NMR Studies of cbEGF-like Domains from Human Fibrillin-1

A thesis submitted in partial fulfilment of the

requirements for the degree of Doctor of Philosophy

by

Rachel Sara Smallridge

St. John's College

Trinity 2000

OCMS/Department of Biochemistry

University of Oxford

South Parks Road

Oxford OX1 3QU

ENGLAND

LIFE SCIENCE LIBRARY OXFORD

26 APR 2001

To my Mum and Dad,
and to Zubin, who waited patiently.

Acknowledgements

This project was funded by the Medical Research Council.

Firstly I would like to acknowledge my supervisors, Kristy Downing and Iain Campbell. Iain plucked me from a group of candidates wishing to become Biochemistry undergraduates at St. John's College, Oxford, in 1992. Little did he know how hard it would be to get rid of me after that point! I have greatly appreciated his support and advice during both my degree and DPhil. With respect to Kristy, well, what can I say? She has been an inspirational Part II and DPhil supervisor, supporting and helping me when required, whilst also ensuring that I was constantly challenged. Her vast knowledge of and enthusiasm for this field of research have been invaluable for my progress, and I realise now that tasks she made me persevere with were indeed "character building". I am also grateful for the opportunities Kristy gave me to travel to international conferences. These events were helpful, interesting and informative, whilst also being extremely enjoyable!

Secondly I would like to thank Pat Whiteman, who set up the expression system for the protein construct that I studied. Pat has always been more than generous with her time and help, and for that I am very grateful. I would also like to thank Penny Handford for advice in relation to the biological aspects of my project.

The IDC group has been a pleasure to work with. Thank you all for your continuous help and support. I'm sure I have asked all of you at least one question during my time here! Special thanks must go to Jörn Werner, without whom this thesis may have been one chapter shorter (is that a good or bad thing!?). His knowledge in relation to NMR spectroscopy and protein dynamics is astounding, and he was an excellent and patient tutor in this field. Thank you also to Jonathan Boyd, Nick Soffe and Christina Redfield for endless advice when setting up NMR experiments.

The AKD lab has recently expanded, so I would like to thank its new members for listening to my moaning over the past months. It was great to have advice from people who have recently experienced this process themselves. Special thanks must go to Xuemei Yuan for continuous help and advice even after leaving the AKD group. The wonder of e-mail!

And now for the F5 posse. Although the AKD group recently migrated to the Biochemistry department so that I no longer physically work in F5, I remain there in spirit! It was my "home" for over three years, and holds some amazing memories, e.g., cheese and wine parties, cheesy Wotsits and Diet Cokes on rough mornings, nice treats and evil treats. It was a great place to work and a great bunch of people to work with, and I will always remember it fondly.

And last, but not least, I would like to thank three wonderful people that I met during my DPhil course, who have always been there when I needed them. To Jo Owen, Lucy Forrest and Beccie Love- here's to more fun and frolics in the future!

ABSTRACT
NMR Studies of cbEGF-like Domains from Human Fibrillin-1

Rachel Smallridge
St. John's College
(Submitted for the degree of Doctor of Philosophy, Trinity 2000)

The calcium binding epidermal growth factor-like (cbEGF) 12-13 domain pair from human fibrillin-1 was the focus of studies for this dissertation. Various nuclear magnetic resonance (NMR) spectroscopy techniques were employed to analyse the calcium binding, structural and dynamic properties of this pair, and to assess the effects of a disease-causing mutation. Fibrillin-1 is a mosaic protein composed mainly of 43 cbEGF domains arranged as multiple, tandem repeats, and mutations within fibrillin-1 have been linked to Marfan syndrome (MFS). 66% of MFS-causing mutations identified thus far are localised to cbEGF domains, emphasising that the native properties of these domains are critical to the functional integrity of this protein. The cbEGF12-13 pair is found within the longest run of cbEGFs in fibrillin-1, and many mutations that cluster in this region are associated with the severe, neonatal form of MFS. It is thought that this region may be important for fibrillin-1 assembly into 10-12nm connective tissue microfibrils.

Calcium binding studies of cbEGF12-13 demonstrated that cbEGF13 contains the highest affinity site thus far investigated from human fibrillin-1. Comparison with previous results showed that fibrillin-1 cbEGF calcium binding affinity can be significantly modulated by the type of domain which is linked to its N-terminus, and also highlighted the high affinity of the "neonatal" region.

The NMR solution structure of cbEGF12-13 is a near-linear, rod-like arrangement of two cbEGF domains, with both exhibiting secondary structure characteristic of this domain type. The rod-like arrangement is stabilised by calcium binding by cbEGF13 and by hydrophobic interdomain packing interactions. This observation supports the hypothesis that all Class I EGF/cbEGF-cbEGF pairs, characterised by a single linker residue, possess this rod-like structure. The structure also exhibits additional packing interactions to those previously observed for cbEGF32-33 from fibrillin-1, which may explain the higher calcium binding affinity of cbEGF13. A model of cbEGF11-15, created based on structural data for cbEGF12-13 and a model of cbEGF32-36, has highlighted a potential protein binding interface, which encompasses all known neonatal MFS mutations, as well as a flexible, unstructured loop region of cbEGF12.

Backbone dynamics data confirmed the extended structure of cbEGF12-13. These data, combined with previous data for cbEGF32-33, highlighted a potential dynamics signature for Class I cbEGF domain pairs. Comparison of data for these pairs also suggested that, in addition to the role of calcium in stabilising rigidity on the pico- to millisecond time-scale, calcium affinity may play a key role in determining the anisotropy of cbEGF pairs. Possible dynamic explanations for the variation in calcium binding affinity of cbEGF domains from human fibrillin-1 were also noted.

The G1127S mutation located in cbEGF13 of fibrillin-1 causes a mild variant of MFS. NMR studies of the G1127S cbEGF12-13 mutant pair showed that cbEGF12 may chaperone folding of mutant cbEGF13, an effect most likely mediated through interdomain packing interactions. These studies have also shown that the effects of this mutation are localised to cbEGF13, suggesting that a "partial" MFS phenotype is the result of altered structural, dynamic and/or calcium binding properties of this domain.

Abbreviations

cb	calcium binding
COSY	correlated spectroscopy
CSA	chemical shift anisotropy
1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
DG	distance geometry
DQF	double quantum filtration
ECM	extracellular matrix
EGF	epidermal growth factor-like
FID	free induction decay
FT	Fourier transform
HMQC	heteronuclear multiple-quantum correlation
HSQC	heteronuclear single-quantum correlation
INEPT	insensitive nuclei enhanced polarisation transfer
MD	molecular dynamics
MFS	Marfan syndrome
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
rf	radio frequency
rMD	restrained molecular dynamics
rmsd	root-mean-square-deviation
SA	simulated annealing
TB	transforming growth factor-β binding protein-like
TOCSY	total correlation spectroscopy

NMR Studies of cbEGF-like Domains from Human Fibrillin-1

Chapter 1 : Introduction **1**
1.1 Modular Proteins 1
1.2 Fibrillin-1, Microfibrils and the Marfan Syndrome 2
1.2.1 Fibrillin-1 2
1.2.2 Fibrillin-1 Assembly into the 10-12nm Connective Tissue Microfibrils 3
1.2.3 The Functions of Fibrillin-1-Containing Microfibrils 6
1.2.4 Other Components of Fibrillin-1-Containing Microfibrils 7
1.2.5 The Marfan Syndrome 8
1.3 Calcium Binding Epidermal Growth Factor-like Domains (cbEGFs) 12
1.4 NMR : A Powerful Tool 16
1.5 Project Aims: Studies of the cbEGF12-13 Domain Pair 18
References 23

Chapter 2 : Methods for Analysis of Protein Structure, Dynamics and Calcium Binding **31**
2.1 Introduction to Intrinsic Protein Fluorescence 31
2.1.1 An Overview of Fluorescence Theory 32
2.1.2 Intrinsic Protein Fluorescence 33
2.1.3 Experimental Considerations 36
2.2 Introduction to Nuclear Magnetic Resonance 37
2.2.1 Qualitative Introduction to One-Dimensional NMR 38
2.2.2 Qualitative Introduction to Two-Dimensional NMR 39
2.2.3 Qualitative Introduction to Three-Dimensional NMR 41
2.2.4 Introduction to the Product Operator Description of NMR 43
2.2.4.1 Product Operators for One Spin 43
2.2.4.2 Product Operators for Two Weakly Coupled Spins 45
2.2.4.3 Product Operators for the Spin-Echo Pulse Sequence 47
2.2.4.3.1 The Single Spin Case 47
2.2.4.3.2 The Two-Spin Homonuclear Case 48
2.2.4.3.3 The Two-Spin Heteronuclear Case 49
2.2.4.4 Product Operators for the INEPT Pulse Sequence 50
2.2.5 Product Operator Calculations for NMR Experiments 52
2.2.5.1 COSY 52
2.2.5.2 TOCSY 56
2.2.5.3 NOESY 59
2.2.5.4 HSQC 63
2.3 Calculating Solution Structures from NMR Data 65
2.3.1 Processing and Analysing NMR Data 66
2.3.1.1 Processing NMR data 66
2.3.1.2 NMR Spectral Assignment 68
2.3.1.2.1 The Sequential Assignment Approach 70
2.3.1.2.2 The Main-Chain Directed Approach and Secondary Structure Identification 72
2.3.2 From NMR Data to Structural Restraints 76
2.3.2.1 From NOE Data to Distance Restraints 76
2.3.2.1.1 NOE Intensity Calibration 76
2.3.2.1.2 Overlapped Crosspeaks and Ambiguous Distance Constraints 78
2.3.2.1.3 Stereospecific Assignments and Floating Chirality 80
2.3.2.2 From HMQC-J Data to ϕ Dihedral Angle Restraints 82
2.3.2.3 From NH-Exchange Data to Hydrogen Bonding Restraints 85
2.3.3 From NMR Data to Structures 87
2.3.3.1 The Development of Structure Calculation Methods 87
2.3.3.2 Dynamical Simulated Annealing 90
2.3.3.3 Structure Calculation Using SA 92

2.3.3.3.1 The Target Function 92
2.3.3.3.2 SA Structure Calculation and Refinement Protocols 94
2.3.4 Evaluation of Structures 98
2.4 Introduction to the Study of Protein Dynamics by NMR 100
2.4.1 T_1, T_2 and heteronuclear NOE 101
2.4.2 The Spectral Density Function 104
2.4.3 Global Dynamics 105
2.4.4 Model-Free Analysis of Internal Dynamics 106
2.4.5 Statistical Analysis 108
2.4.6 Chemical Exchange 110
References 111

Chapter 3 : Expression of cbEGF12-13 Constructs **118**
3.1 Introduction 118
3.2 Expression 119
3.2.1 Growing the Bacteria 119
3.2.2 Lysis of the Bacteria 121
3.3 Purification 121
3.4 Refolding 122
3.5 Removing the His Tag 123
References 124

Chapter 4 : Calcium Binding Studies of cbEGF12-13 ..... **125**
4.1 Introduction 125
4.2 Preparation of cbEGF12-13 Samples for Calcium Titrations 127
4.2.1 Preparation of the NMR Titration Sample 127
4.2.2 Preparation of the Fluorescence Titration Samples 128
4.3 Collection of Calcium Titration Data 128
4.3.1 Data Collection for the NMR Titration 128
4.3.2 Data Collection for the Fluorescence Titrations 129
4.4 Results 132
4.5 Discussion and Conclusions 134
4.5.1 Effects of Pairwise Domain Linkage on cbEGF Affinity 134
4.5.2 Implications of Variable Calcium Affinity Along the Length of Fibrillin-1 138
References 143

Chapter 5 : Determination of the Solution Structure of cbEGF12-13 ..... **146**
5.1 NMR Data Acquisition and Spectral Assignment of cbEGF12-13 146
5.1.1 NMR Sample Preparation 146
5.1.2 Acquisition and Processing of NMR Data for Spectral Assignment and Structure Calculations 147
5.1.3 Spectral Assignment 149
5.1.4 Modelling the Structure of the cbEGF12-13 Domain Pair 154
5.2 Conversion of NMR data to Structural Restraints 157
5.2.1 NOE Data 157
5.2.2 HMQC-J Data 158
5.2.3 NH-Exchange Data 160
5.2.4 Calcium Binding Restraints 160
5.3 Structure Calculation and Refinement 161
5.4 The Structure of the cbEGF12-13 Domain Pair 165
5.4.1 Secondary Structure and Interdomain Packing 165
5.4.2 Tilt and Twist Angles 170
5.4.3 Structural Convergence 172
5.5 Biological Significance of the cbEGF12-13 Structure 176
5.5.1 Implications for Connective Tissue Microfibrils 176
5.5.2 Implications for the Marfan Syndrome 178

References 182

Chapter 6 : Backbone Dynamics of cbEGF12-13 ..... **187**
6.1 Acquisition and Processing of NMR Data for Relaxation Studies 187
6.1.1 Sample Preparation 187
6.1.2 Data Acquisition 187
6.1.3 Data Processing 189
6.2 Analysis and Results of ^{15}N Relaxation Data 189
6.2.1 ^{15}N-T_1, ^{15}N-T_2 and ^{1}H-^{15}N heteronuclear NOE 189
6.2.2 Motion on the Micro- to Millisecond Time-Scale 195
6.2.3 Estimation of the Rotational Diffusion Tensor 198
6.2.4 Model-Free Analysis of Internal Dynamics 203
6.3 Discussion of Results : Biological Significance 209
6.3.1 Correlation of Backbone Flexibility and Calcium Affinity 209
6.3.2 Implications for Connective Tissue Microfibrils and MFS 212
References 215

Chapter 7 : NMR Studies of the cbEGF12-13 G1127S Mutant Domain Pair **219**
7.1 Introduction to the G1127S Mutation 219
7.2 Sample Preparation 221
7.3 Data Acquisition 223
7.4 Results 223
7.5 Discussion 225
7.5.1 Chaperoning of Folding by Tandem Linkage of cbEGF-like Domains 225
7.5.2 Implications for the Pathogenic Effect of the G1127S Mutation 231
References 233

Chapter 8 : Summary of Conclusions **235**
8.1 The Calcium Binding Properties of cbEGF12-13 235
8.2 The Structure of the cbEGF12-13 Domain Pair 236
8.3 Dynamics of the cbEGF12-13 Domain Pair 237
8.4 The G1127S Mutation 239
References 241

Appendix I : Geometric Energy Function Parameters and NMR Structure Calculation Protocols for Proteins 1
Appendix II : Minimal Media Labelling 18
Appendix III : Amino Acid Analysis 19
Appendix IV : Experimental Restraints Used for Structure Calculations of cbEGF12-13 22
Appendix V : ^{1}H and ^{15}N Chemical Shift and J-Coupling Values for Calcium Saturated cbEGF12-13, pH 6.5 and 33°C 64
Appendix VI : ^{15}N-T_1, ^{15}N-T_2 and ^{15}N-Heteronuclear NOE Data for cbEGF12-13 at (a) 500MHz and (b) 750MHz, pH 6.5 and 35°C 72

Chapter 1

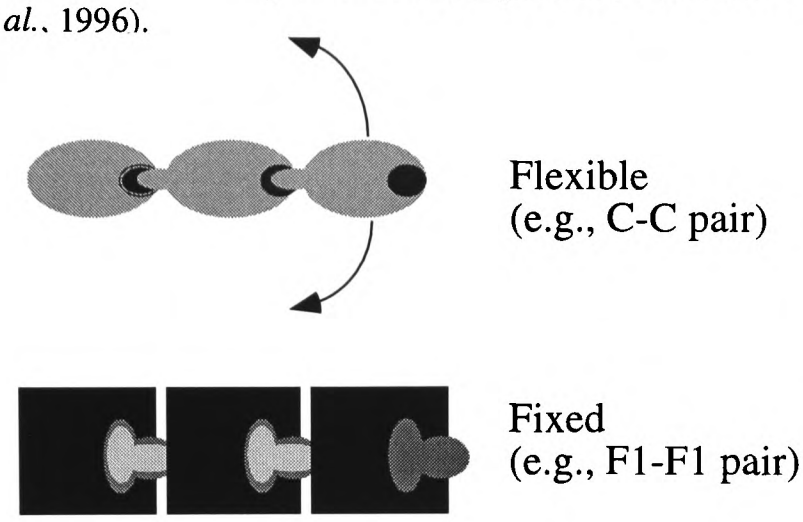
Introduction

1.1 Modular Proteins

Eukaryotic genomes are composed of exons (coding regions) separated by introns (long, non-coding regions), which are reorganised at the RNA level to give a single, coding RNA strand for a protein (Breathnach *et al.*, 1977). In extracellular proteins, single exons often code for a protein unit that folds independently of the rest of the molecule (Blake, 1978). As these domains or modules, as they are called, are usually relatively small (~30-300 residues), they can often be studied by nuclear magnetic resonance (NMR). One application of NMR is determination of the overall structure of a protein, which, for relatively large, modular proteins, can usually be obtained by determining the structure of each domain separately, followed by investigation of the way in which these domains are joined (see Figure 1.1). While this "dissection" approach has increased the rate at which protein structural information is being obtained, care must be taken. Results obtained when investigating the properties of a single domain may differ substantially from those obtained when analysing the domain in its native protein environment. For example, recent NMR structural studies of the ^{6}F1^{1}F2^{2}F2 triple domain fragment from the collagen-binding domain of fibronectin demonstrated that the ^{6}F1 and ^{2}F2 domains interact to form a buried interface of ~850Å^{2}, an interaction not observed in studies of domain pairs from this region (Pickford *et al.*, manuscript in preparation).

Figure 1.1 : Schematic Representation of Types of Domain Linkage

C denotes the CCP domain which is abundant in complement control proteins, whereas F1 denotes a fibronectin type I domain (Bork *et al.*, 1996).



1.2 Fibrillin-1, Microfibrils and the Marfan Syndrome

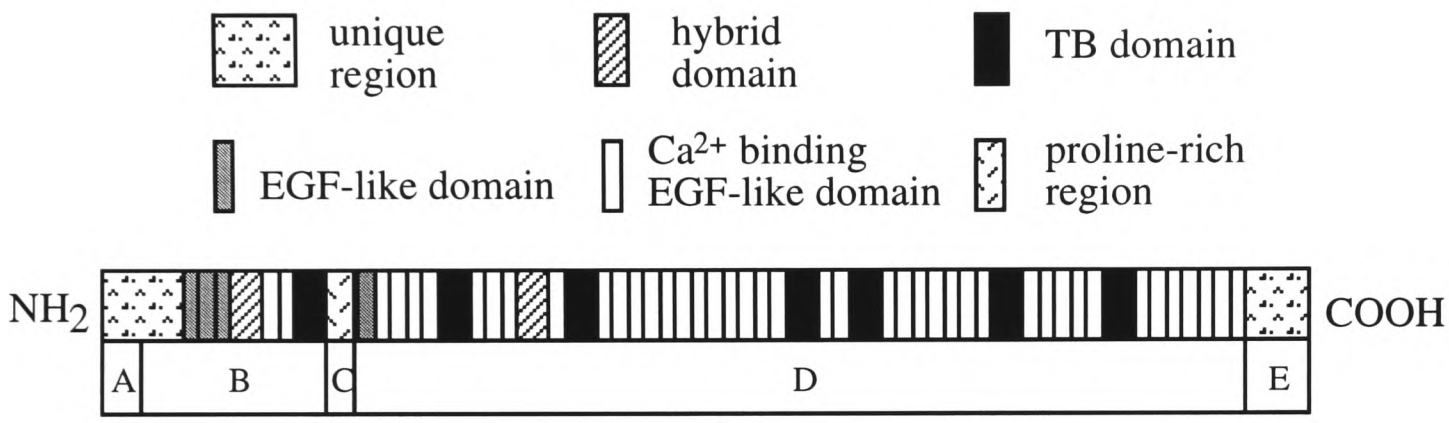
1.2.1 Fibrillin-1

Fibrillin-1 is an extracellular matrix glycoprotein composed of 2872 amino acids (Pereira *et al.*, 1993), and this large size makes it a challenging macromolecule to study. It is, however, a mosaic or modular protein, meaning it can be studied using the "dissection" approach described above. Figure 1.2 shows schematically the domain organisation and the various regions (A-E) of human fibrillin-1, as defined by Pereira *et al.* (1993). This figure clearly demonstrates that human fibrillin-1 is mainly composed of calcium binding (cb) epidermal growth factor-like (EGF) domains and transforming growth factor-β binding protein-like (TB) domains, with 43 cbEGFs domains being separated by seven TB domains. Of the regions described, region C has an unusually high proline content (42% of the residues are prolines), and region D contains a putative cell-attachment (RGD) sequence. By analogy to titin, region C may be important for the elastic properties of fibrillin-1. Titin, a large sarcomeric protein of striated muscle, contains a proline-rich PEVK domain, which is principally responsible for the elasticity of muscle fibres (Linke *et al.*, 1996; 1998). Fibrillin-1 contains 14 potential N-linked glycosylation sites, and can be further

post-translationally modified by glycosylation of O-linked sites and by β-hydroxylation of Asn and/or Asp residues (Glanville *et al.*, 1994).

Figure 1.2 : Schematic Representation of the Domain Organisation of Human Fibrillin-1

Different domain types are represented as differently shaded boxes, with box size being proportional to the number of residues in each domain, and the five regions (A-E) of human fibrillin-1 are highlighted (Pereira *et al.*, 1993). TB represents the transforming growth factor-β binding protein-like domain, and EGF stands for epidermal growth factor.



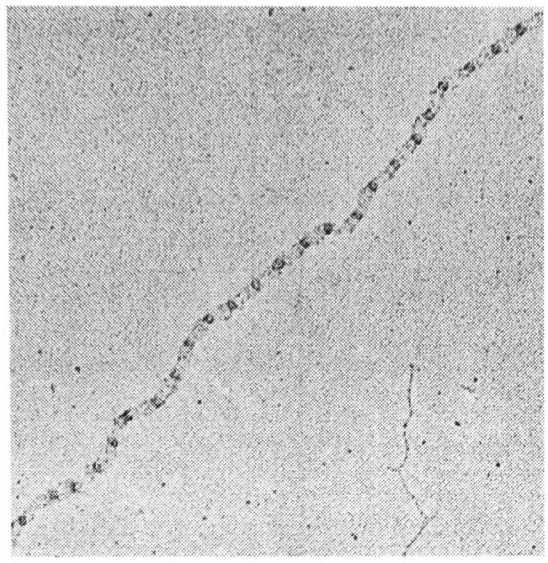
Region A: 29 residues, largely basic.
Region B: 336 residues, cysteine-rich.
Region C: 58 residues, proline-rich.
Region D: 2240 residues, 49 uninterrupted cysteine-rich repeats, i.e., EGF-like, cbEGF-like and TB domains.
Region E: carboxy terminus.

1.2.2 Fibrillin-1 Assembly into the 10-12nm Connective Tissue Microfibrils

Fibrillin-1 monomers assemble repetitively to form the 10-12nm connective tissue microfibrils present in the extracellular matrix (ECM) (Sakai *et al.*, 1986). Figure 1.3 shows the distinctive "beads on a string" appearance of these microfibrils, and illustrates the importance of calcium in the maintenance of their structural architecture (Kielty & Shuttleworth, 1993; Cardy & Handford, 1998; Wess *et al.*, 1998). Little is known regarding the assembly of fibrillin-1 into the 10-12nm microfibrils, but it has recently been observed that the N-terminal region of human fibrillin-1 mediates the rapid formation of disulphide-bonded homodimers and larger aggregates (Trask *et al.*, 1999; Ashworth *et al.*, 1999; Reinhardt *et al.*, 2000a). It is

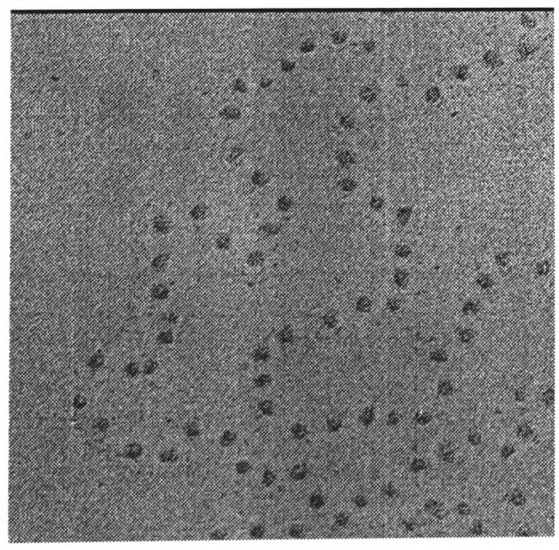
gure 1.3 : Electron Micrographs Demonstrating the Importance of Calcium in Maintenance of the Structural Architecture of the 10-12nm Connective Tissue Microfibrils

These electron micrographs were provided courtesy of Dr. P. A. Handford.



<u>A normal microfibril in the presence of calcium</u>
Note the "beads on a string" appearance.

EDTA addition



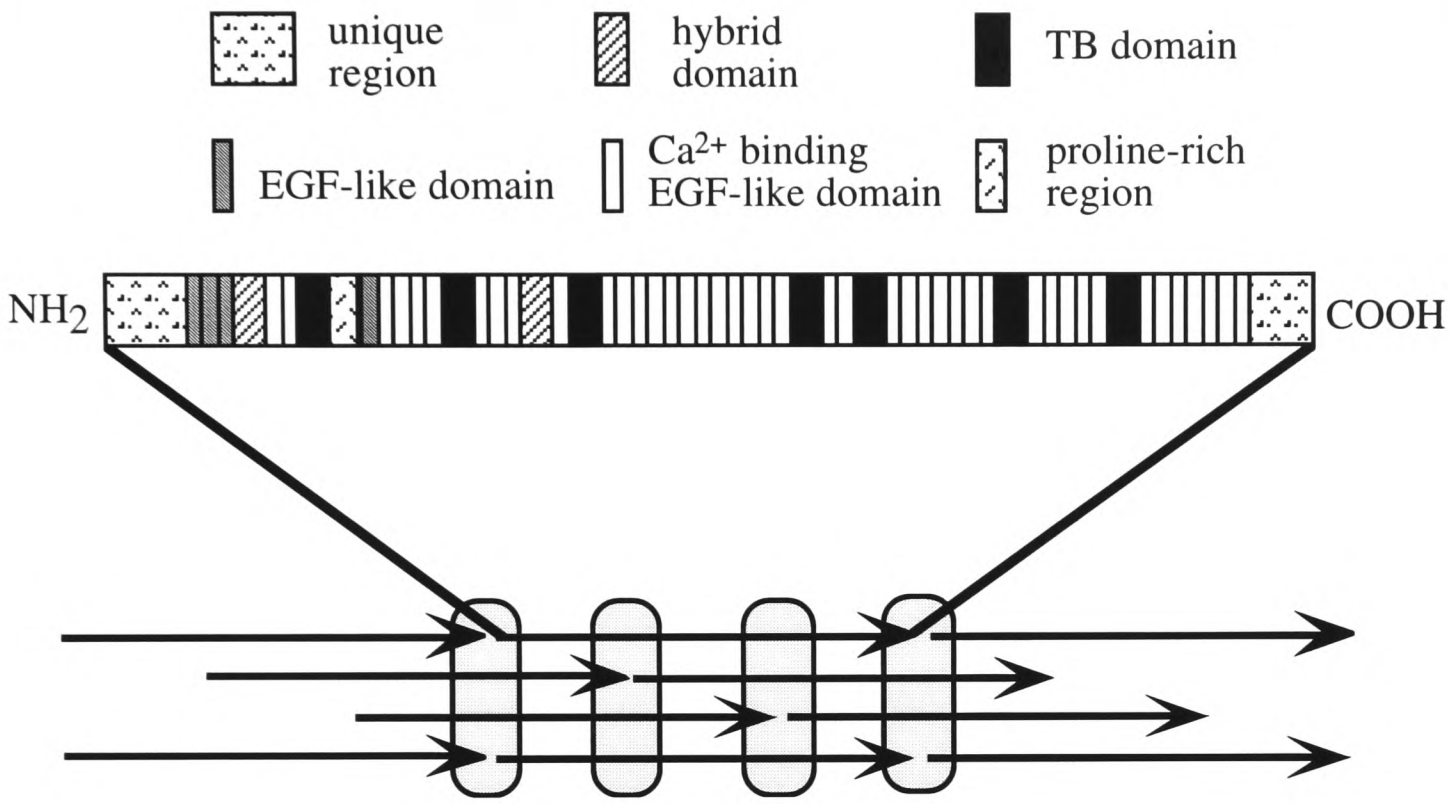
<u>Microfibrils in the absence of calcium</u>
Note the diffuse appearance in the interbead region and loss of linear structure.

thought that this intermolecular disulphide bond formation is an initial step in the assembly process, and a candidate cysteine residue, C204, in the first hybrid domain, has been suggested to participate in these intermolecular disulphide bonds (Reinhardt *et al.*, 2000a). The Pro-rich, C region of fibrillin-1 is thought to be important in determining the specificity of this interaction, since experiments performed on fragments of the fibrillin proteins have shown that fibrillin-1 does not form heterodimers with the other member of the fibrillin family, the homologous fibrillin-2 (Trask *et al.*, 1999; Ashworth *et al.*, 1999). Fibrillin-2 differs most distinctively from fibrillin-1 in that a Gly-rich region replaces the Pro-rich region. The site of formation of these disulphide bonded homodimers/multimers is, however, unclear, with this event being localised to intracellular (Trask *et al.*, 1999) and extracellular compartments (Ashworth *et al.*, 1999; Reinhardt *et al.*, 2000a) in separate studies.

Despite the lack of information regarding microfibrillar assembly, a model for fibrillin-1 organisation into the 10-12nm microfibrils has been proposed (see Figure 1.4) (Downing *et al.*, 1996; Reinhardt *et al.*, 1996; Qian & Glanville, 1997). In this most recent model, the fibrillin-1 monomers are predicted to adopt a parallel, staggered orientation, with each monomer spanning three "interbead" regions. The interbead regions, therefore, contain multiple cbEGF domain repeats. Dimension data used to construct this model include a bead periodicity of ~50-55nm (Kielty & Shuttleworth, 1993; Sakai & Keene, 1994; Handford *et al.*, 1995; Cardy & Handford, 1998), a fibrillin-1 monomer length of ~148nm (Sakai *et al.*, 1991; Sakai & Keene, 1994), and a cbEGF domain span of ~2.9nm, so that 43 cbEGFs span ~125nm (Downing *et al.*, 1996). In this model, it was assumed that the hybrid domains are globular, and the resulting model is consistent with the dimensions given above, with antibody labelling data (Sakai *et al.*, 1991; Reinhardt *et al.*, 1996), and with the position of transglutaminase-derived cross-links, which were detected between fibrillin-1 monomers in the interbead region (Qian & Glanville, 1997).

Figure 1.4 : A Model for the Organisation of Fibrillin-1 Monomers into the 10-12nm Connective Tissue Microfibrils

Each arrow represents a fibrillin-1 monomer, and the shaded regions correspond to the beaded regions of the 10-12nm microfibrils.



1.2.3 The Functions of Fibrillin-1-Containing Microfibrils

In non-elastic tissues, the 10-12nm connective tissue microfibrils have anchoring functions, e.g., in the suspensory ligaments of the ocular lens, whereas in elastic tissues they form the structural system that links elastic fibres to each other and to the ECM, and provide a scaffold upon which elastin is deposited (Mecham & Heuser, 1991). In addition to these structural functions, a gene-targeting experiment in mice has suggested that the main function of fibrillin-1 is in tissue homeostasis rather than elastic matrix assembly (Pereira *et al.*, 1997). Another important physiological role for fibrillin-1-containing microfibrils may include cell-matrix interactions in haemostasis and thrombosis. It is thought that fibrillin-1 monomers in the microfibrils utilise an RGD sequence localised to TB4 to bind to the platelet GPIIb-IIIa integrin receptor under low shear conditions, allowing platelet adhesion and aggregation to occur on the microfibril surface (Ross *et al.*, 1998). In recombinant fibrillin-1 fragments, this RGD sequence has also been shown to be important for cell adhesion and signalling, where it

mediates binding to both αv and β3 integrins (Sakamoto *et al.*, 1996; Pfaff *et al.*, 1996; Darrigo *et al.*, 1998). Microfibrils may also directly influence cell behaviour. They have been shown to support adhesion and spreading of human vascular smooth muscle cells *in vitro* (Kielty *et al.*, 1992), and fibrillin-1 may play a role in establishing cardiac endothelial diversity by defining a subpopulation of cells that will participate in valve formation (Wunsch *et al.*, 1994).

It has been proposed that fibrillin-1 could have cellular growth promoting activity, and, if so, it could influence many morphogenetic programmes (Engel, 1991). Post-translational modifications of fibrillin-1, e.g., glycosylation or β-hydroxylation of Asn/Asp, or fibrillin-1 itself could provide binding sites for proteins that become active, or inactive, when they encounter these sites. Despite the functions discussed here, there still, however, remains a large amount to learn regarding the function of fibrillin-1 and the fibrillin-1-containing microfibrils.

1.2.4 Other Components of Fibrillin-1-Containing Microfibrils

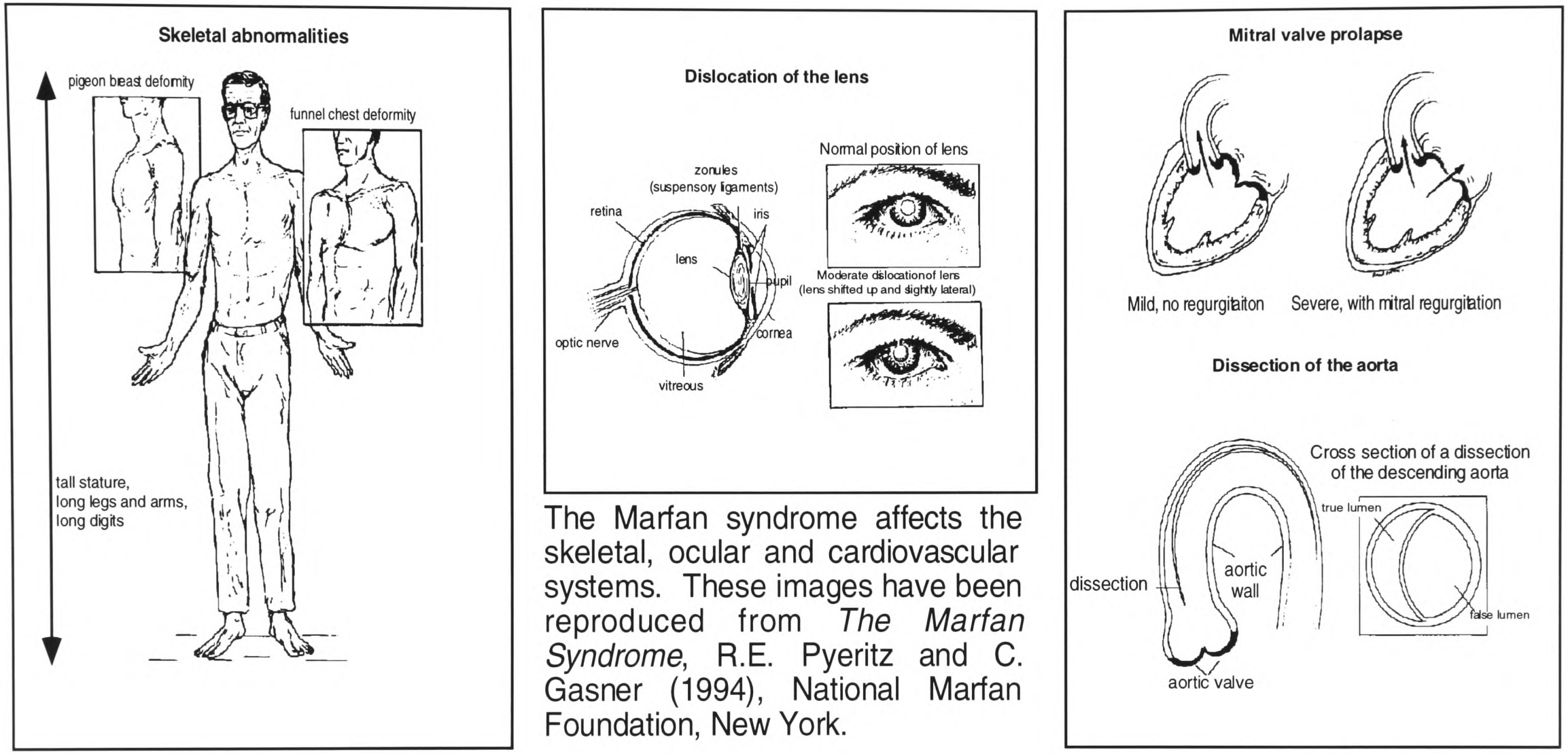
Fibrillin-1 has the potential to interact with many proteins, however, these interactions are not well understood. Immunohistochemical approaches have indicated that a number of proteins may be associated with microfibrils, including amyloid P, fibronectin, thrombospondin, vitronectin and microfibril-associated glycoproteins (MAGPs, MFAPs, emilin) (Kielty & Shuttleworth, 1995). In developing tissues, the proteins MAGP-1 and 2 (Gibson *et al.*, 1986; 1991; 1996; 1998), MFAPs 1,3 and 4 (Yeh *et al.*, 1994; Abrams *et al.*, 1995; Zhao *et al.*, 1995), latent transforming growth factor β-binding protein (LTBP)-1 and -2 (Gibson *et al.*, 1995; Hyytiainen *et al.*, 1998; Dallas *et al.*, 2000) and βig-h3 (Skonier *et al.*, 1992; Gibson *et al.*, 1996) have all been identified in association with fibrillin-1-containing microfibrils. Chondroitin sulphate proteoglycans have also been shown to associate with fibrillin-1

in the beaded regions of microfibrils, and have been shown to be important in microfibrillar organisation (Kielty *et al.*, 1996). More recently, the chondroitin sulphate proteoglycan, decorin, has been shown to bind specifically to an amino-terminal region of fibrillin-1 near the Pro-rich, C region. Decorin can interact with MAGP-1 and fibrillin-1 individually, or with both of these proteins to form a ternary complex (Trask *et al.*, 2000). The functional importance of these interactions is, however, unclear, since knockout mice with a targeted disruption of the decorin gene are viable and appear to have functional elastic tissues (Danielson *et al.*, 1997).

Possible functions of some of these putative microfibrillar components have been suggested. For example, MAGP-1 is present in the beaded region of microfibrils and may utilise intermolecular disulphide bonds to stabilise the arrangement of fibrillin-1 monomers in this region (Henderson *et al.*, 1996), and MAGP-2 may mediate interactions with cell-surfaces at specific stages of development and differentiation (Gibson *et al.*, 1998). LTBP-2 is a component of microfibrils that targets latent forms of TGF-β growth factors to the ECM (Hyytiainen *et al.*, 1998). Despite these insights, there is still much to learn about the structure and function of the components of 10-12nm connective tissue microfibrils.

1.2.5 The Marfan Syndrome

Defects in human fibrillin-1 result in the Marfan syndrome (MFS), an inherited, autosomal dominant disease, which is characterised by ophthalmic, cardiovascular and musculoskeletal symptoms (Figure 1.5) (reviewed in Ramirez *et al.*, 1993; Sakai & Keene, 1994; Dietz & Pyeritz, 1995). The disease affects at least 1 in 5000 in the population, and its occurrence has no relation to ethnic origin or geography (Dietz & Pyeritz, 1995). MFS results from mutations in the fibrillin-1 (*FBN*-1) gene and over 140 different mutations have been reported to date (see Figure 1.6) (Collod-Béroud *et al.*, 1998). These mutations can be divided into the three categories shown in



The Marfan syndrome affects the skeletal, ocular and cardiovascular systems. These images have been reproduced from *The Marfan Syndrome*, R.E. Pyeritz and C. Gasner (1994), National Marfan Foundation, New York.

Figure 1.5 : Symptoms of the Marfan Syndrome

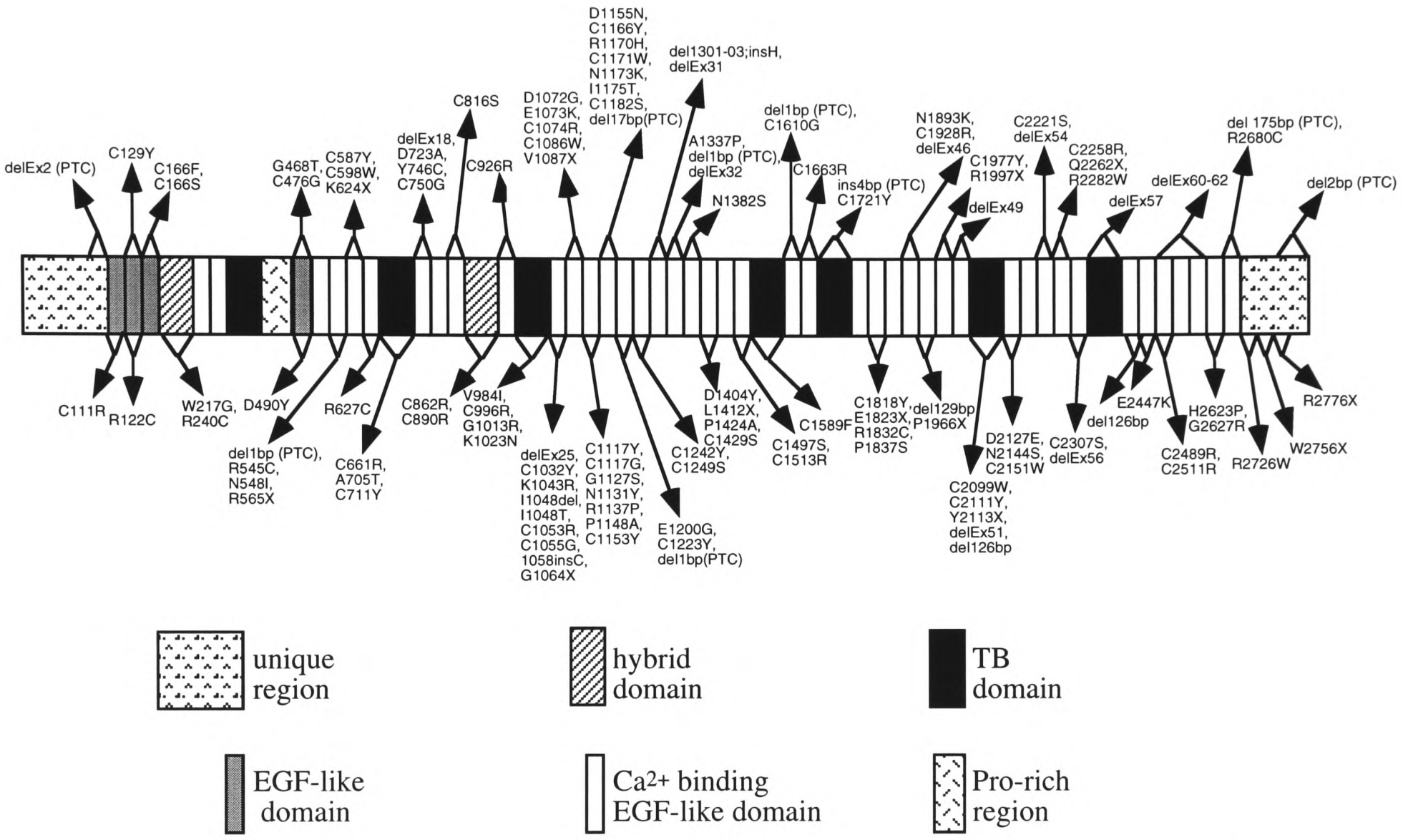


Figure 1.6 : Schematic Representation of Amino Acid Substitutions Caused by Genetic Mutations to the *FBN-1* Gene

This diagram summarises mutations causing the Marfan syndrome, and related disorders, superimposed on the domain organisation of human fibrillin-1 (Pereira *et al.*, 1993). "PTC" denotes mutations that cause a premature termination codon. This figure is modified from Dietz & Pyeritz, 1995.

Table 1.1 (Dietz & Pyeritz, 1995), and the majority of these MFS-causing mutations result in amino acid substitutions in cbEGF domains.

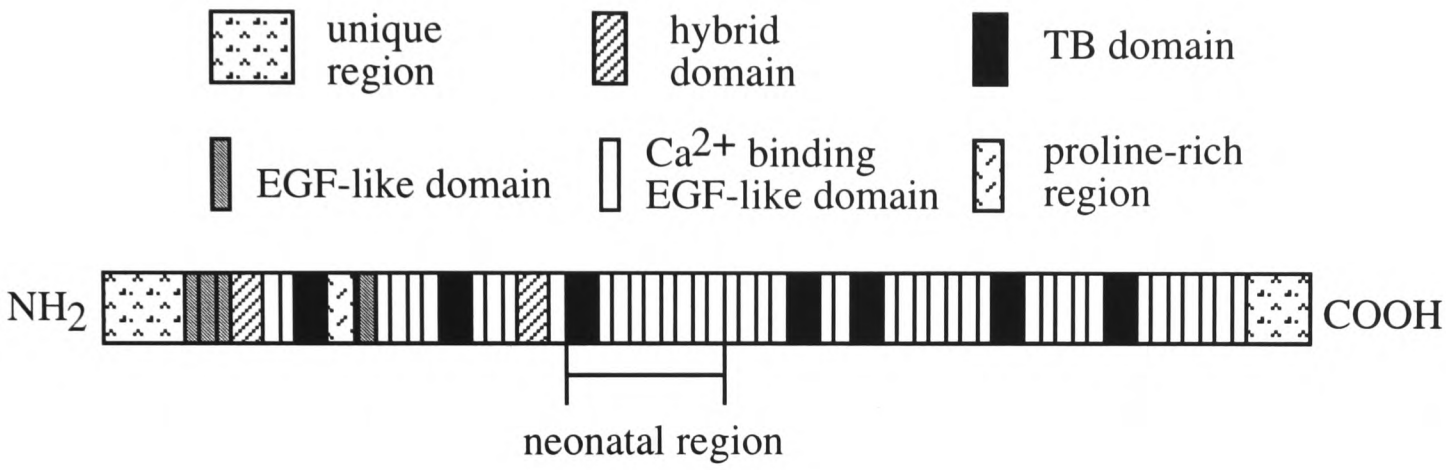
Table 1.1 : The Three Types of Mutation in the *FBN*-1 Gene that Cause MFS

Category of Mutation	Effect of Mutation
1	Missense mutations or small in-frame deletions or insertions.
2	Creation of a premature termination codon, due to nonsense or frameshift mutations.
3	Mutations that result in the loss of the central region of the monomer, due to aberrant splicing events or genomic deletions that preserve the reading frame.

The *FBN*-1 gene mutations shown in Figure 1.6 result in forms of MFS ranging from mild to severe, and thus far it is not understood why these mutations produce such variations in phenotype. To allow molecular diagnosis and optimum treatment of this disease, a genotype/phenotype relationship must be derived. This is an extremely challenging task since most MFS families carry unique mutations, and the fibrillin genotype is not always the sole determinant of the connective tissue phenotype. For example, mutations in three patients with a Marfan-related syndrome, presenting predominantly skeletal symptoms, have been localised to LTBP-2 (Dallas *et al.*, 2000).

Although the genotype/phenotype relationship of MFS remains elusive, it has been observed that a cluster of mutations in the region corresponding to exons 24-32 (encoding TB-3 and cbEGF domains 11-18; Figure 1.7) are associated with the neonatal form of the disease (nMFS) and with another severe type of MFS (Putnam *et al.*, 1996; Collod-Béroud *et al.*, 1997; Blooms *et al.*, 1999). However, mutations which produce a more moderate phenotype are also found in this region (Collod-Béroud *et al.*, 1997; 1998). A cluster of mutations which affect exons 24-34 of fibrillin-2 have been shown to be associated with congenital contractural arachnodactyly (CCA), further emphasising the critical role of this central region of the fibrillin proteins (Park *et al.*, 1998).

Figure 1.7 : Schematic Diagram of the Neonatal Region of Human Fibrillin-1



MFS is suspected to have gone undiagnosed in many famous cases in history, e.g., Abraham Lincoln, Paganini, Rachmaninov (Ramirez *et al.*, 1993). Investigation of fibrillin-1 will assist in early diagnosis of the disease, which may then be treated via surgical intervention if necessary. One day diagnosis of the disease *in utero* may be possible, along with the development of preventative treatments. The elucidation of a genotype/phenotype relationship, combined with identification of mutations within fibrillin-1 that do not produce a disease phenotype, will be invaluable for genetic counselling. Mutations in the *FBN*-1 gene also cause other disorders which require increased understanding at the genome level to optimise treatment, such as ectopia lentis, Shprintzen-Goldberg syndrome, and joint hypermobility (reviewed in Dietz & Pyeritz, 1995). In addition, autoantibodies to fibrillin-1 have recently been identified in patients suffering from the skin disorder, localised scleroderma (Arnett *et al.*, 1999).

1.3 Calcium Binding Epidermal Growth Factor-like Domains (cbEGFs)

Human fibrillin-1 contains 43 cbEGF domains, and these domains are a subgroup of the EGF-like domain family. In addition to the usual consensus sequence exhibited by EGF-like domains, cbEGF domains contain an extra consensus sequence

at the N-terminus of the domain which is associated with calcium binding: D/N-x-D/N-E/Q-x_m-D/N*-x_n-Y/F, where m and n are variable and * indicates possible β-hydroxylation (Figure 1.8) (Rees *et al.*, 1988; Handford *et al.*, 1991; Mayhew *et al.*, 1992; Campbell & Bork, 1993; Bork *et al.*, 1996; Downing *et al.*, 2000). In addition to β-hydroxylation, other post-translational modifications that can occur include O-fucosylation and O-glucosylation (reviewed in Downing *et al.*, 2000). cbEGF domains comprise ~40-45 amino acids, with the disulphide bonding pattern 1-3, 2-4 and 5-6 between the six cysteine residues (Bork *et al.*, 1996). They lack any appreciable hydrophobic core, and are stabilised by a combination of these three disulphide bonds and calcium binding to the N-terminus (Rao *et al.*, 1995; Downing *et al.*, 2000). Differences in loop composition and length between cbEGFs are likely to relate to the specificity of the cbEGF domain interactions (Campbell & Bork, 1993).

The cbEGF domain is a widely distributed domain, and its presence in many different proteins illustrates its importance. Some examples of proteins containing cbEGF domains are shown in Table 1.2, and these proteins have a wide variety of roles in extracellular events including receptor-ligand interactions, adhesion and coagulation (reviewed in Downing *et al.*, 2000). In many of these proteins, e.g., the Notch superfamily, fibrillins and LTBPs, the cbEGF domains are arranged as multiple tandem repeats (Campbell & Bork, 1993; Downing *et al.*, 2000). The structures of cbEGF domains from factor VII (Banner *et al.*, 1996; Muranyi *et al.*, 1998; Kao *et al.*, 1999), factor IX (Baron *et al.*, 1992; Brandstetter *et al.*, 1995; Rao *et al.*, 1995), factor X (Ullner *et al.*, 1992; Selander-Sunnerhagen *et al.*, 1992; Sunnerhagen *et al.*, 1996), Clr (a complement protease) (Bersch *et al.*, 1998), thrombomodulin (Fuentes-Prior *et al.*, 2000) and fibrillin-1 (Downing *et al.*, 1996) have been solved using X-ray crystallography or NMR methods, and the typical secondary structure of a cbEGF domain is shown in Figure 1.9. The structure consists of a major double stranded β-sheet, with the possibility of an additional, C-terminal minor double-stranded β-sheet,

and/or an α-helix between the first two cysteines. The N- and C-termini are located at opposite ends of the molecule, with the calcium binding site positioned near the N-terminus. The bound calcium is coordinated by a pentagonal bipyramidal arrangement of oxygen atoms, six of which are provided by intradomain ligands (Campbell & Bork, 1993; Downing *et al.*, 2000).

Figure 1.8 : Schematic Representation of the Secondary Structure and Consensus Sequence of cbEGF Domains

Each amino acid is shown as a circle. Conserved cysteines are shown in yellow, the predominant conserved calcium binding consensus residues are shown in red, and other conserved residues, likely to be important for folding or interdomain packing of cbEGF domains, are shown in green. * indicates the site of possible β-hydroxylation.

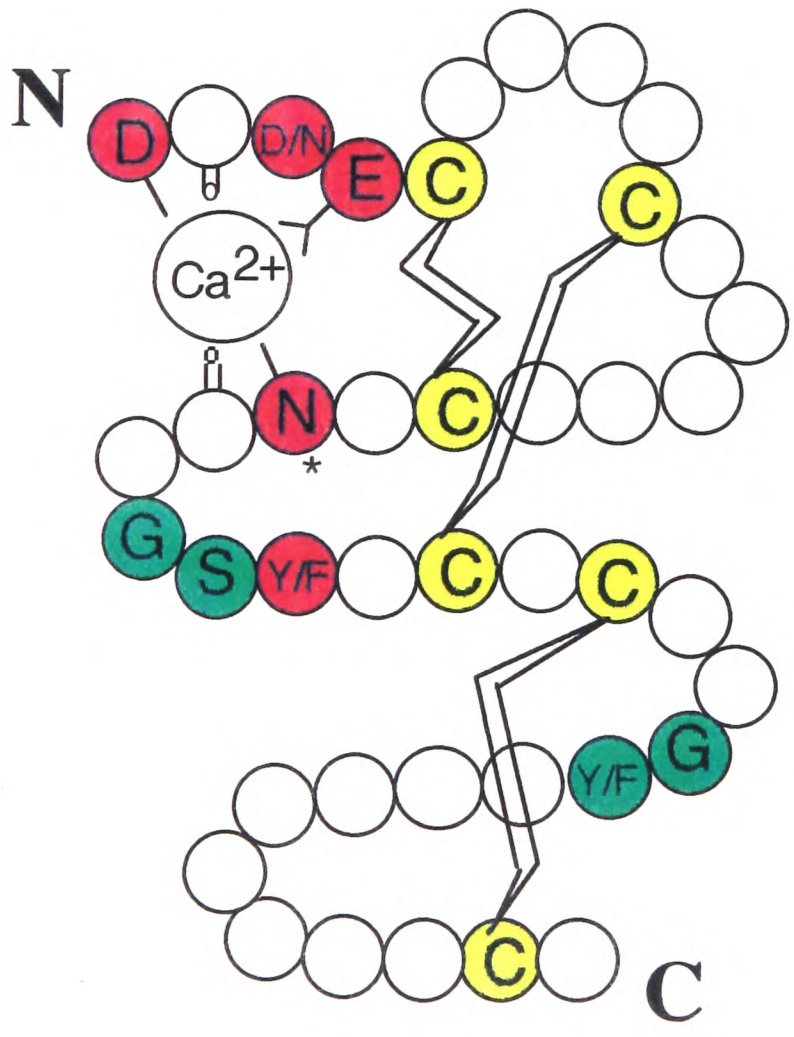
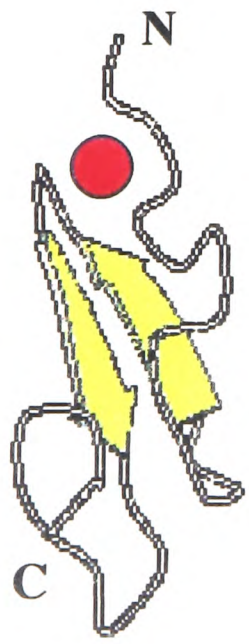


Table 1.2 : The Functions of Some Proteins Containing cbEGF Domains

Protein	Function
Fibrillin-1,2, fibulin-1,2, nidogen, LTBPs	Maintenance of the architecture of the ECM
Notch-3, Jagged-1, *Drosophila* Delta and Serrate	Cell fate determination
Factors VII, IX, X, proteins S and C, thrombomodulin	Control of blood coagulation
Low density lipoprotein receptor (LDLR)	Uptake of cholesterol

The calcium atom bound in the single binding site (see Figures 1.8 and 1.9) stabilises, and restricts the conformational flexibility of, cbEGF domain tertiary structure and the interdomain packing interactions between cbEGF domains (Downing *et al.*, 1996; Werner *et al.*, 2000; Downing *et al.*, 2000). Calcium is also necessary for some protein-protein interactions (Cooke *et al.*, 1987; Handford *et al.*, 1991; Corson *et al.*, 1993; Pereira *et al.*, 1993). It is likely to play a role in the formation of rod-like superstructures, since it has been shown to have a role in orienting domain pairs (Downing *et al.*, 1996; 2000), and it may also protect against proteolysis *in vivo* (Reinhardt *et al.*, 2000b).

Figure 1.9 : Diagrammatic Representation of the Structure of a cbEGF Domain



The major double-stranded β-sheet is represented by yellow arrows. The rest of the peptide is represented by wire. The bound calcium is represented by a red sphere.

The importance of native cbEGF domain function is emphasised by the fact that mutations to this domain type result in a wide variety of inherited disease phenotypes. In addition to MFS being caused by mutations within the cbEGF domains of human fibrillin-1 (Collod-Béroud *et al.*, 1998), genetic mutations that cause amino acid changes within cbEGF domains of Notch-3, protein S, factor IX and the low density lipoprotein (LDL) receptor have been linked to CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy) (Joutel *et al.*, 1997), protein S deficiency (Gandrille *et al.*, 1995), haemophilia B (Giannelli *et al.*, 1996), and familial hypercholesterolaemia (Hobbs *et al.*, 1992), respectively. The mutations that occur in cbEGF domains can be classified into three groups depending on the residues they affect (see Table 1.3) (Downing *et al.*, 1996). The number of inherited diseases caused by missense mutations within cbEGF domains underscores the biological importance of this type of domain.

Table 1.3 : The Potential Effects of Mutations in cbEGF Domains

Mutated Residue	Consequence of Mutation
Cysteines	Alters native disulphide bond formation, which disrupts correct folding.
Calcium binding consensus sequence residues	Leads to reduced calcium affinity, destabilising cbEGF domain linkage.
Other residues	For example, mutation of residues involved in intermolecular interactions could disrupt interdomain packing.

1.4 NMR: A Powerful Tool

Nuclear magnetic resonance is an extremely powerful technique, which can be readily applied to the study of proteins. It can be used for a variety of investigations, including monitoring ligand binding (depending on kinetics), the determination of protein three-dimensional (3D) structure at atomic resolution, and dynamics investigations (see Chapter 2). These studies are invaluable in providing a complete

biophysical picture of a protein with a view to furthering the understanding of its function and mechanism of action.

In simple terms, a NMR experiment begins with the sample of interest in a static, stable magnetic field, at proton frequencies of up to almost a thousand MHz. In this situation the diamagnetic nuclear spins of the protein sample equilibrate with the magnetic field to yield a net equilibrium magnetisation. Signals can then be induced in the sample by the application of radio frequency (rf) pulses at various intervals, which disrupt the equilibrium magnetisation of the sample. The signals obtained as a result of this perturbation can provide information about the interactions between the nuclear spins, and individual signals or resonances in the resultant NMR spectrum can be assigned specifically to particular atoms within the protein. Using the NOESY type of NMR experiment (see section 2.2.5.3), the intensities of the signals obtained can be correlated with the distance between the corresponding atoms in a protein, and it is this information that has formed the basis of structure calculations to date.

NMR has thrived over the past fifteen years. This is in part due to the use of isotopically labelled samples and the formulation of heteronuclear, multi-dimensional NMR experiments, which has raised the limit of protein structure determination by NMR to ~150-300 amino acids (Clore & Gronenborn, 1991). Advances have also been made as a result of improvements to NMR spectrometers and the production of spectrometers with higher proton frequencies, i.e., 800 and 900MHz machines. The study of domains from extracellular proteins using the dissection approach described in section 1.1 has also increased the rate at which protein structural and dynamic information is being obtained.

X-ray crystallography is the alternative, well-established technique applied to protein structure determination, and this method can be applied to much larger proteins than NMR. However, X-ray methods lack the ability to monitor both ligand binding and intrinsic protein dynamics. Time resolved crystallographic methods exist, but are difficult to implement since all molecules must exhibit the same state at the same time.

Photoactivation is therefore sometimes employed. X-ray methods cope poorly with flexible proteins, as flexible proteins often do not form crystals, or, when crystals do form, the flexible regions of the protein often show no electron density. Flexible regions of a protein can also become locked in place by packing interactions, as seen by comparing X-ray and NMR studies of the ^{9}F3^{10}F3 domain pair from fibronectin (Copie *et al.*, 1998). These problems have been a driving force for the improvement of NMR techniques, as evidenced by the increasing number of NMR structures being deposited in the Brookhaven Protein Data Bank (PDB). NMR was the main physical technique used in this thesis.

1.5. Project Aims: Studies of the cbEGF12-13 Domain Pair

Fibrillin-1 contains 47 EGF-like domains, 43 of which are cbEGF domains as previously mentioned. The latter are labelled cbEGF1-43 (Corson *et al.*, 1993; Pereira *et al.*, 1993), and the cbEGF12-13 domain pair from human fibrillin-1 was the main focus of NMR solution studies for this thesis (Figure 1.10). The cbEGF12-13 pair comprises residues 1070-1154 (Figure 1.11) (numbering according to Pereira *et al.*, 1993), and this domain pair contains many mutations that result in both nMFS and other forms of MFS (Table 1.4) (Putnam *et al.*, 1996; Collod-Béroud *et al.*, 1997; 1998; Blooms *et al.*, 1999).

The first investigation undertaken in this project was calcium binding studies of the cbEGF12-13 domain pair (see Chapter 4; Smallridge *et al.*, 1999). This domain pair is centrally positioned within the neonatal region of fibrillin-1 (Figures 1.7 and 1.10), and many mutations that cause the devastating nMFS disease phenotype are found within this pair (Table 1.4) (Putnam *et al.*, 1996; Collod-Béroud *et al.*, 1997; 1998). In fact, there are three mutations to calcium binding consensus residues within the cbEGF12-13 pair, i.e., D1072G, E1073K and N1131Y, which result in nMFS,

Figure 1.10 : Position of the cbEGF12-13 Domain Pair from Human Fibrillin-1

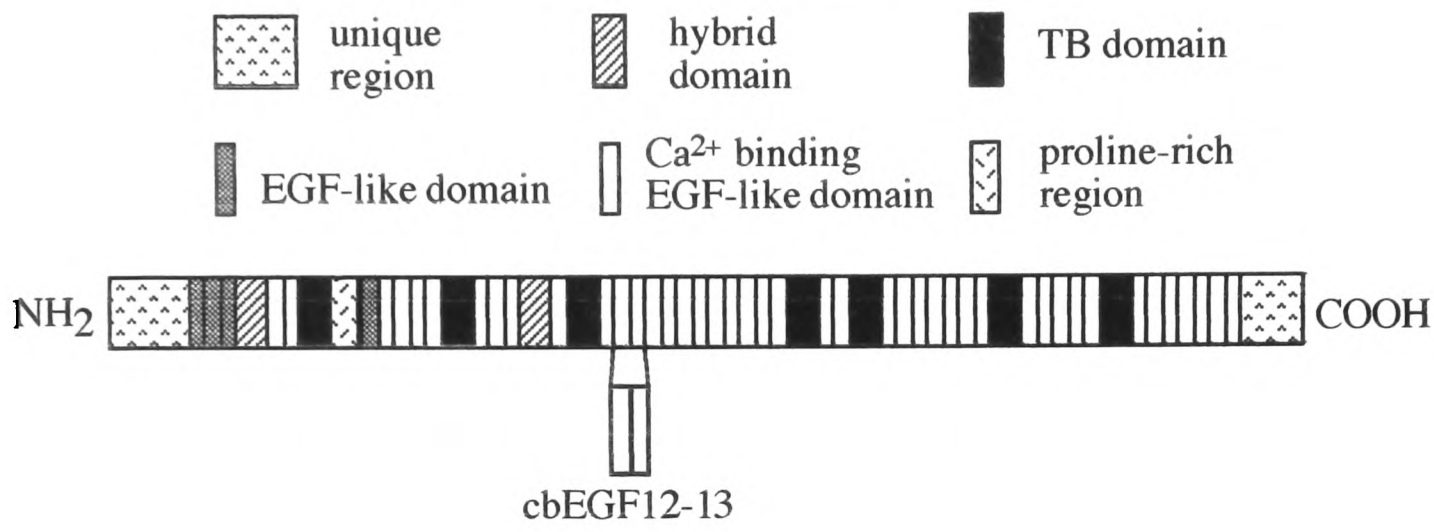
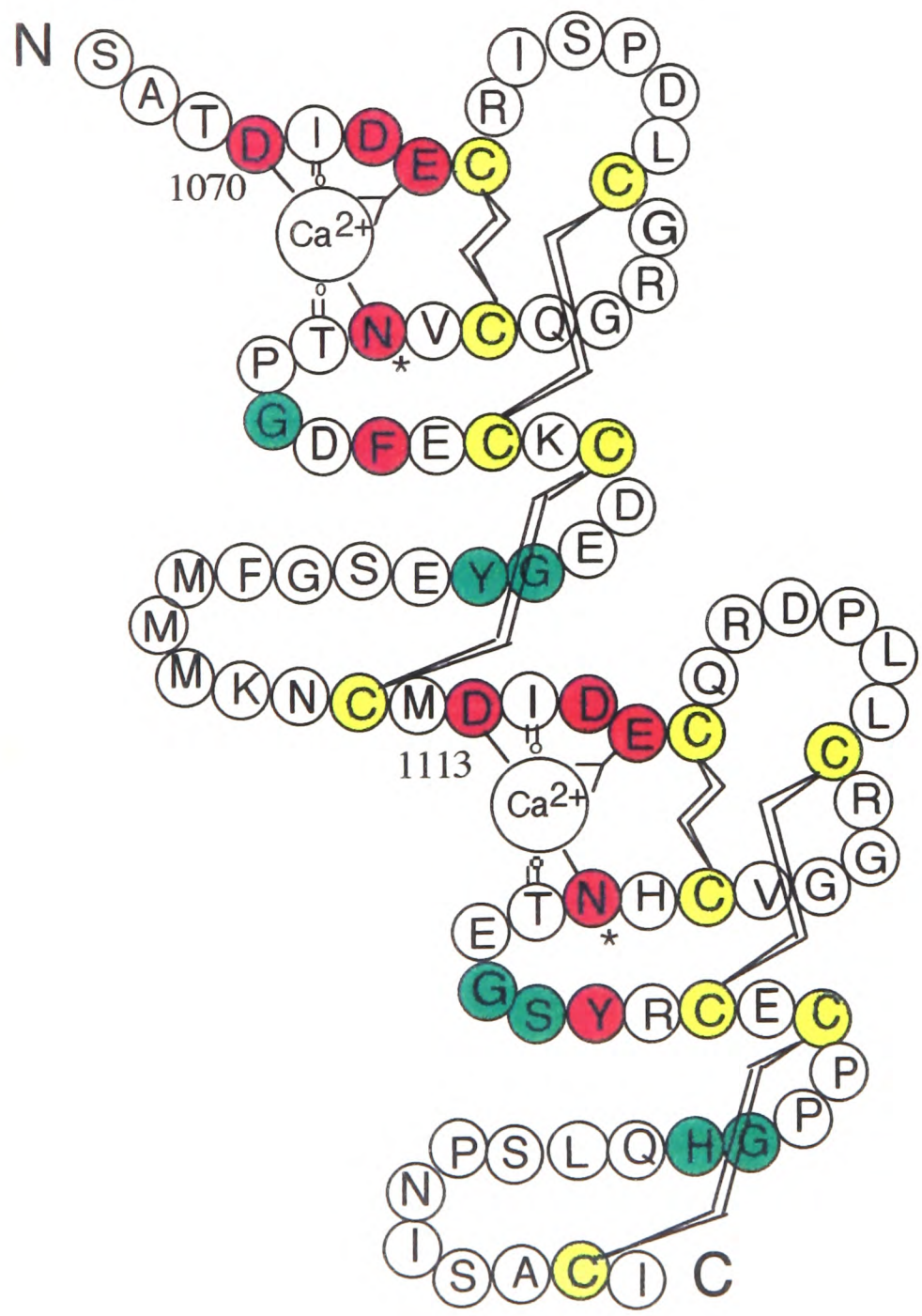


Figure 1.11 : Schematic Diagram of the cbEGF12-13 Domain Pair

The sequence and the predicted schematic secondary structure of the cbEGF12-13 domain pair from human fibrillin-1 are shown. The key is the same as for Figure 1.8. The first residue of each domain, i.e., the first consensus calcium binding residue, is labelled using the numbering of Pereira *et al.*, 1993.



underscoring the biological importance of native calcium binding by this domain pair. Calcium titrations performed on this domain pair resulted in K_d values that could be compared with the results of the previous titrations performed on cbEGF domains from human fibrillin-1 (Glanville *et al.*, 1994; Handford *et al.*, 1995; Knott *et al.*, 1996; Reinhardt *et al.*, 1996; Whiteman *et al.*, 1998; Kettle *et al.*, 1999), which gave invaluable insight into variations in calcium binding affinity of cbEGFs along the length of fibrillin-1 and indicated potentially stable regions of this protein. These calcium binding investigations also helped to improve the understanding of nMFS and other MFS mutations that occur in this region of fibrillin, and allowed determination of the effect of tandem linkage of cbEGF domains on calcium binding affinity.

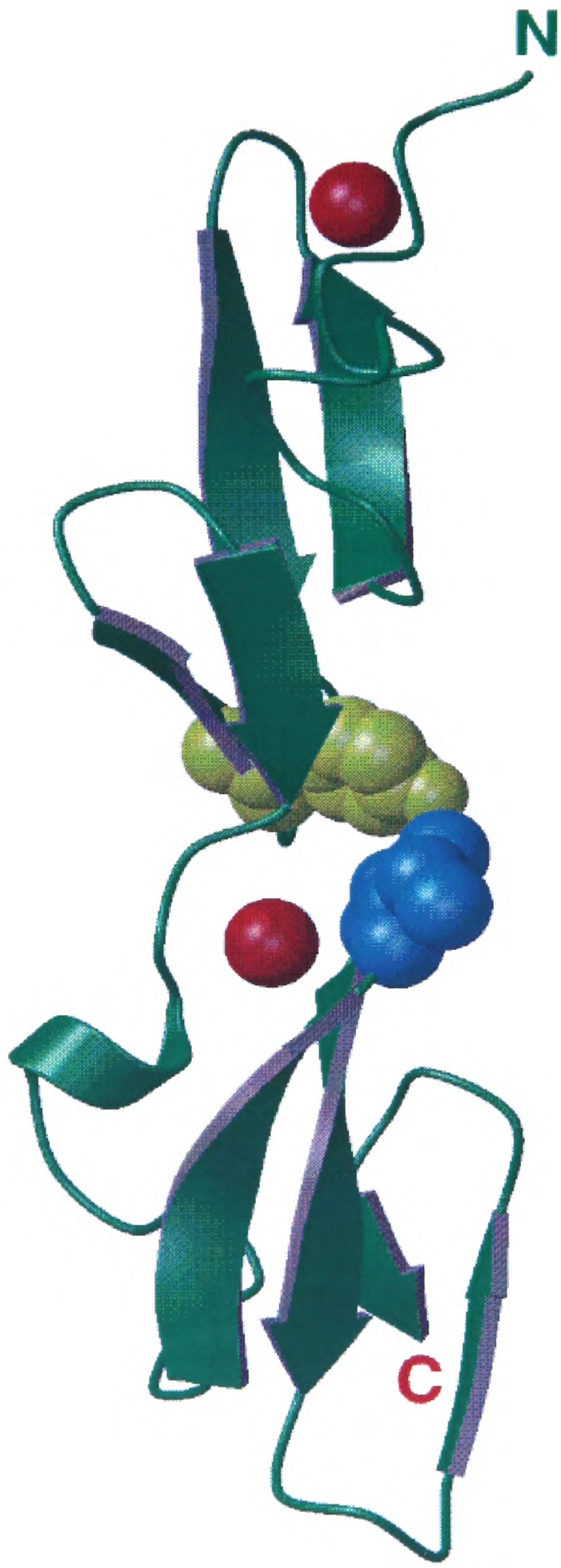
Table 1.4 : Summary of Missense Mutations in the cbEGF12-13 Domain Pair that Cause MFS

Mutation	Effect of Mutation	Phenotype
D1072G	Expected to disrupt calcium binding	Neonatal MFS
E1073K	Expected to disrupt calcium binding	Neonatal MFS
C1074R	Causes disulphide bond disruption	Neonatal MFS
C1086W	Causes disulphide bond disruption	Neonatal MFS
C1086Y	Causes disulphide bond disruption	Neonatal MFS
C1117G	Causes disulphide bond disruption	Neonatal MFS
C1117Y	Causes disulphide bond disruption	Classical MFS
G1127S	Can lead to secondary structure disruption, causing misfolding *in vitro*	Mild form of inherited weakness of elastic tissue. "Partial" MFS phenotype
N1131Y	Expected to disrupt calcium binding	Neonatal MFS
R1137P	Leads to secondary structure disruption, causing misfolding *in vitro*	Classical MFS
C1153Y	Causes disulphide bond disruption	Classical MFS

The second part of this project comprised NMR structural studies of the cbEGF12-13 domain pair (see Chapter 5). Structural studies of the cbEGF12-13 domain pair were performed to enable validation of a hypothesis proposed by Downing *et al.* (1996). The cbEGF12-13 Class I pair is predicted to adopt a conformation analogous to that of the Class I cbEGF32-33 pair (Figure 1.12; Downing *et al.*,

Figure 1.12 : The NMR Solution Structure of the cbEGF32-33 Domain Pair from Human Fibrillin-1

Calcium atoms are shown as red spheres, and β-sheets are represented by arrows. The main hydrophobic interactions include a conserved aromatic residue, Tyr2157 (shown in yellow), at the open end of the minor β-sheet in cbEGF32 which packs against the major β-hairpin of cbEGF33, along with side chain interactions of Tyr2157 with Gly 2186 (blue) and the two methyls of Ile2185 (blue) (Downing *et al.*, 1996).



1996), i.e., a conformation thought to be exhibited by all Class I EGF-cbEGF/cbEGF-cbEGF pairs. It has been suggested that Class I pairs, containing only one residue in their interdomain linker region, adopt a different conformation from Class II domain pairs, which have two residues in their linker. The length of the loop connecting cyteines 5 and 6 significantly differs between the Class I and Class II pairs, which may be important in determining the different pairwise interactions between the domains. In Class II pairs a calcium binding ligand may possibly be donated from the N- to C-terminal domain, resulting in a structure where the two domains are folded onto one another, in contrast to the rigid, rod-like structure observed for the Class I cbEGF32-33 domain pair (see Figure 1.12; Downing *et al.*, 1996). The hypothesis for Class II pair structure arises from the observation that a conserved carboxylate-carboxyamide side chain is observed at the position of the ligand donated from the N- to C-terminal domain in the crystal structure of the cbEGF domain from factor IX (Rao *et al.*, 1995).

These NMR structural studies also enabled the structural consequences of MFS mutations to be assessed, furthering the understanding of the genotype/phenotype relationship of MFS. The structure of cbEGF32-36 from human fibrillin-1 has previously been modelled, using the NMR solution structure of the cbEGF32-33 domain pair, which showed that stretches of cbEGF domains are likely to adopt an extended, rigid, rod-like structure (Downing *et al.*, 1996). The run of twelve cbEGF domains (cbEGF11-22) containing the cbEGF12-13 pair and the neonatal region of fibrillin-1 are therefore likely to adopt an extended, rod-like structure, a prediction supported by dimension data discussed in section 1.2.2. This extended region could therefore be important for fibrillin-1 assembly into the 10-12nm connective tissue microfibrils. This hypothesis is supported by immunofluorescence studies of neonatal MFS fibrils, which has shown them to be short, fragmented and frayed relative to wild-type microfibrils (Superti-Furga *et al.*, 1992; Godfrey *et al.*, 1995), although the specific mutation(s) giving rise to this morphology has, however, not been identified.

In the third part of this project, NMR backbone relaxation measurements were used to obtain an insight into the intrinsic dynamics of the cbEGF12-13 pair (see Chapter 6). A complete picture of a protein cannot be obtained purely from structural studies, as proteins are not generally static. Dynamics studies of the cbEGF12-13 domain pair from human fibrillin-1 highlighted regions of this domain pair with increased flexibility, which may indicate their involvement in protein-protein interactions. The results from cbEGF12-13 could also be compared with previous results obtained from backbone dynamics measurements of the cbEGF32-33 domain pair (Werner *et al.*, 2000), providing information regarding variations in intrinsic dynamics of cbEGF domains along the length of human fibrillin-1. This comparison also allowed further verification of the role of calcium in stabilising the rigidity of cbEGF domains.

The final part of this thesis involved NMR investigations of the consequences of the G1127S mutation localised to cbEGF13 in the cbEGF12-13 domain pair (see Chapter 7). This mutation produces a mild variant of the MFS phenotype, predisposing to ascending aortic aneurysm and dissection in later life (Franke *et al.*, 1995), and it was investigated to further the understanding the genotype/phenotype relationship of the MFS. The G1127S mutation has previously been shown to cause domain misfolding *in vitro* in the cbEGF13 single domain and the cbEGF13-14 domain pair (see Table 1.4; Whiteman *et al.*, 1998). Initial data obtained here, however, suggest that folding of domain 13 in the presence of the mutation may be chaperoned by N-terminal linkage of domain 12.

References

Abrams, W. R., Ma, R.-I., Kucich, U., Bashir, M. M., Decker, S., Tsipouras, P., McPherson, J. D., Wasmuth, J. J. & Rosenbloom, J. (1995). Molecular cloning of the microfibrillar protein MFAP3 and assignment of the gene to chromosome 5q32-q33.2. *Genomics*. **26**, 47-54.

Arnett, F. C., Tan, F. K., Uziel, Y., Laxer, R. M., Krafchik, B. R., Antohi, S. & Bona, C. (1999). Autoantibodies to the extracellular matrix microfibrillar protein, fibrillin 1, in patients with localized scleroderma. *Arthritis Rheum.* **42**, 2656-2659.

Ashworth, J. L., Kelly, V., Wilson, R., Shuttleworth, C. A. & Kielty, C. M (1999). Fibrillin assembly: dimer formation mediated by amino-terminal sequences. *J. Cell. Sci.* **112**, 3549-3558.

Banner, D. W., D'Arcy, A., Chene, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y. & Kirchhofer, D. (1996). The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature.* **380**, 41-46.

Baron, M., Norman, D. G., Harvey, T. S., Handford, P. A., Mayhew, M., Tse, A. G. D., Brownlee, G. G. & Campbell, I. D. (1992). The three-dimensional structure of the first EGF-like module of human factor IX: comparison with EGF and TGF-α. *Prot. Sci.* **1**, 81-90.

Bersch, B., Hernandez, J.-F., Marion, D. & Arlaud, G. J. (1998). Solution structure of the epidermal growth factor (EGF)-like module of human complement protease C1r, an atypical member of the EGF family. *Biochemistry.* **37**, 1204-1214.

Blake, C. C. F. (1978). Do genes-in-pieces imply proteins-in-pieces ? *Nature*. **273**, 267-268.

Blooms, P., Cisler, J., Mathews, K. R., Godfrey, M., Tiecke, F., Kaufmann, U. C., Vetter, U., Hagemeier, C. & Robinson, P. N. (1999). Novel exon skipping mutation in the fibrillin-1 gene: Two "hot spots" for the neonatal Marfan syndrome. *Clin. Genet.* **55**, 110-117.

Bork, P., Downing, A. K., Kieffer, B. & Campbell, I. D. (1996). Structure and distribution of modules in extracellular proteins. *Quart. Rev. Biophys.* **29**, 119-169.

Brandstetter, H., Bauer, M., Huber, R., Lollar, P. & Bode, W. (1995). X-ray structure of clotting factor IXa: Active site and module structure related to Xase activity and hemophilia B. *Proc. Natl. Acad. Sci. USA.* **92**, 9796-9800.

Breathnach, R., Mandel, J. L. & Chambon P. (1977). Ovalbumin gene is split in chicken DNA. *Nature*. **270,** 314-319.

Campbell, I. D. & Bork, P. (1993). Epidermal growth factor-like modules. *Curr. Op. Struct. Biol.* **3**, 385-392

Cardy, C. M. & Handford, P. A. (1998). Metal ion dependency of microfibrils supports a rod-like conformation for fibrillin-1 calcium binding epidermal growth factor-like domains. *J. Mol. Biol.* **276**, 855-860.

Clore, G. M. & Gronenborn, A. M. (1991). Structures of lager proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. *Science*. **252**, 1390-1399.

Collod-Béroud, G., Béroud, C., Adès, L., Black, C., Boxer, M., Brock, D. J., Godfrey, M., Hayward, C., Karttunen, L., Milewicz, D., Peltonen, L., Richards, R. I., Wang, M., Junien, C. & Boileau, C. (1997). Marfan database (second edition): software and database for the analysis of mutations in the human *FBN*-1 gene. *Nucl. Acids Res*. **25**, 147-150.

Collod-Béroud, G., Béroud, C., Adès, L., Black, C., Boxer, M., Brock, D. J. H., Holman, K. J., de Paepe, A., Francke, U., Grau, U., Hayward, C., Klein, H. G., Liu, W., Nuytinck, L., Peltonen, L., Perez, A. B., Rantamäki, T., Junien, C. &

Boileau, C. (1998). Marfan database (third edition): new mutations and new routines for the software. *Nucl. Acids Res.* **26**, 229-233.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H. & Sheard, B. (1987). The solution structure of human epidermal growth factor. *Nature*. **327**, 339-341.

Copie, V., Tomita, Y., Akiyama, S. K., Aota, S., Yamada, K. M., Venable, R. M., Pastor, R. W., Krueger, S. & Torchia, D. A. (1998). Solution structure and dynamics of linked cell attachment modules of mouse fibronectin containing the RGD and synergy regions: Comparisons with the human fibronectin crystal structure. *J. Mol. Biol.* **277**, 663-682.

Corson, G.M., Chalberg, S. C., Dietz, H. C., Charbonneau, N. & Sakai, L.Y. (1993). Fibrillin binds calcium and is encoded by cDNAs that reveal a multidomain structure and alternatively spliced exons at the 5' end. *Genomics*. **17**, 468-475.

Dallas, S. L., Keene, D. R., Bruder, S. P., Saharinen, J., Sakai, L. Y., Mundy, G. R. & Bonewald, L. F. (2000). Role of the latent transforming growth factor β binding protein 1 in fibrillin-containing microfibrils in bone cells *in vitro* and *in vivo*. *J. Bone Min. Res.* **15**, 68-81.

Danielson, K. G., Baribault, H., Holmes, D. F., Graham, H., Kadler, K. E. & Iozzo, R. V. (1997). Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J. Cell Biol.* **136**, 729-743.

Darrigo, C., Burl, S., Withers, A. P., Dobson, H., Black, C. & Boxer, M. (1998). TGF-β1 binding protein-like modules of fibrillin-1 and -2 mediate integrin-dependent cell adhesion. *Connect. Tissue Res.* **37**, 29-52.

Dietz, H. C. & Pyeritz, R. E. (1995). Mutations in the human gene for fibrillin-1 (*FBN*-1) in the Marfan syndrome and related disorders. *Hum. Mol. Genet.* **4**, 1799-1809.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Downing, A. K., Campbell, I. D. & Handford, P. A. (2000). Chapter 4 : Calcium binding EGF-like domains. In Calcium Homeostasis. Carafoli, Krebs, eds. *Topics Biol. Inorg. Chem.*, volume 3.

Engel, J. (1991). EGF-like domains in extracellular matrix proteins: localised signals for growth and differentiation? *FEBS Lett.* **251**, 1-7.

Franke, U., Berg, M. A., Tynan, K., Brenn, T., Liu, W., Aoyama, T., Gasner, C. Miller, D. C. & Furthmayr, H. (1995). A Gly1127Ser mutation in an EGF-like domain of the fibrillin-1 gene is a risk factor for ascending aortic aneurysm and dissection. *Am. J. Hum. Genet.* **56**, 1287-1296.

Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D. R & Bode, W. (2000). Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. *Nature*. **404**, 518-525.

Gandrille, S., Borgel, D., Eschwege-Gufflet, V., Aillaud, Mf., Dreyfus, M., Matheron, C., Gaussem, P., Abgrall, J. F., Jude, B., Sie, P., Toulon, P. & Aiach, M. (1995). Identification of 15 different candidate casual point mutations and three polymorphisms in 19 patients with protein S deficiency using a scanning method for the analysis of the protein S active gene. *Blood.* **85**, 130-138.

Giannelli, F., Green, P. M., Sommer, S. S., Poon, M.-C., Ludwig, M., Schwaab, R., Reitsma, P. H., Goossens, M., Yoshioka, A. & Brownlee, G. G. (1996). Haemophilia B (sixth edition): a database of point mutations and short additions and deletions. *Nucl. Acids Res.* **24**, 103-118.

Gibson, M. A., Hughes, J. L., Fanning, J. C. & Cleary, E. G. (1986). The major antigen of the elastin-associated microfibrils is a 31kDa glycoprotein. *J. Biol. Chem.* **261**, 11429-11436.

Gibson, M. A., Kumaratilake, J. S. & Cleary, E. G. (1989). The protein components of the 12-nanometer microfibrils of elastic and non-elastic tissues. *J. Biol. Chem.* **264**, 4590-4598.

Gibson, M. A., Sandberg, L. B., Gross, L. E. & Cleary, F. G. (1991). Complementary DNA cloning establishes microfibril-associated glycoprotein (MAGP) to be a discrete component of elastin-associated microfibrils. *J. Biol. Chem.* **266**, 7596-7601.

Gibson, M. A., Hatzinikolas, G., Davis, E. C., Baker, E., Sutherland, G. R. & Mecham, R. P. (1995). Bovine latent transforming growth factor β1-binding protein-2: molecular cloning, identification of tissue isoforms, and immunolocalization to elastin-associated microfibrils. *Mol. Cell. Biol.* **15**, 6932-6942.

Gibson, M. A., Hatzinikolas, G., Kumaratilake, J. S., Sandberg, L. B., Nichol, J. K., Sutherland, G. R. & Cleary, E. G. (1996). Further characterization of proteins associated with elastic fiber microfibrils including the molecular cloning of MAGP-2 (MP25). *J. Biol. Chem.* **271**, 1096-1103.

Gibson, M. A., Finnis, M. L., Kumaratilake, J. S. & Cleary, E. G. (1998). Microfibril-associated glycoprotein-2 (MAGP-2) is specifically associated with fibrillin-containing microfibrils but exhibits more restricted patterns of tissue localization and developmental expression than its structural relative MAGP-1. J. *Histochem. Cytochem.* **46**, 871-885.

Glanville, R. W., Qian, R., McClure, D. W. & Maslen C. L. (1994). Calcium binding, hydroxylation, and glycosylation of the precursor epidermal growth factor-like domains of fibrillin-1, the Marfan gene protein. *J. Biol. Chem.* **269**, 26630-26634.

Godfrey, M., Raghunath, M., Cisler, J., Bevins, C. L., de Paepe, A., Di Rocco, M., Gregoritch, J., Imaizumi, K., Kaplan, P., Kuroki, Y., Silberbach, M., Superti-Furga, A., Van Thienen, M. N., Vetter, U. & Steinmann, B. (1995). Abnormal morphology of fibrillin microfibrils in fibroblast cultures from patients with neonatal Marfan syndrome. *Am. J. Pathol.* **146**, 1414-1421.

Handford, P. A., Mayhew, M., Baron, M., Winship, P. R., Campbell, I. D. & Brownlee, G. G. (1991). Key residues involved in calcium binding motifs in EGF-like domains. *Nature.* **351**, 164-167.

Handford, P. A., Downing, A. K., Rao, Z., Hewett, D. R., Sykes, B. C. & Kielty, C. M (1995). The calcium binding properties and molecular organisation of EGF-like domains in human fibrillin-1. *J. Biol. Chem.* **270**, 6751-6756.

Henderson, M., Polewski, R., Fanning, J. C. & Gibson, M. A. (1996). Microfibril-associated glycoprotein-1 (MAGP-1) is specifically located on the beads of the beaded-filament structure for fibrillin-containing microfibrils as visualised by the rotary shadowing technique. *J. Histochem. Cytochem.* **44**, 1389-1397.

Hobbs, H. H., Brown, M. S. & Goldstein, J. L. (1992). Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum. Mutat.* **1**, 445-466.

Hyytiainen, M., Taipale, J., Heldin, C.-H. & Keski-Oja, J. (1998). Recombinant latent transforming growth factor β-binding protein 2 assembles to fibroblast extracellular matrix and is susceptible to proteolytic processing and release. *J. Biol. Chem.* **273**, 20669-20676.

Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cecillion, M., Marechal, E., Maciazek, J., Vayssiere, C., Cruaud, C., Cabanis, E. A., Ruchoux, M. M., Weissenbach, J., Bach, J. F., Bousser, M. G. & TournierLasserve, E. (1997). Notch3 mutations in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), a mendelian condition causing stroke and vascular dementia. *Annals Of the New York Academy Of Sciences.* **826**, 213-217.

Kao, Y.-H., Lee, G. F., Wang, Y., Starovasnik, M. A., Kelley, R. F., Spellman, M. W. & Lerner, L. (1999). The effect of O-fucosylation on the first EGF-like domain from human blood coagulation Factor VII. *Biochemistry.* **38**, 7097-7110.

Kettle, S., Yuan, X., Grundy, G., Knott, V., Downing, A. K. & Handford, P. A. (1999). Defective calcium binding to fibrillin-1: consequence of an N2144S change for fibrillin-1 structure and function. *J. Mol. Biol.* **285**, 1277-1287.

Kielty, C. M., Whittaker, S. P., Grant, M. E. & Shuttleworth, C. A. (1992). Attachment of human vascular smooth-muscle cells to intact microfibrillar assemblies of collagen VI and fibrillin. *J. Cell. Sci.* **103**, 445-451.

Kielty, C. M. & Shuttleworth, C. A. (1993). The role of calcium in the organisation of fibrillin molecules. *FEBS Lett.* **336**, 323-326.

Kielty, C. M. & Shuttleworth, C. A. (1995). Fibrillin-containing microfibrils - structure and function in health and disease. *Int. J. Biochem. Cell. Biol.* **27**, 747-760.

Kielty, C. M., Whittaker, S. P. & Shuttleworth, C. A. (1996). Fibrillin: evidence that chondroitin sulphate proteoglycans are components of microfibrils and associate with newly synthesised monomers. *FEBS Lett.* **386**, 169-173.

Knott, V., Downing, A. K., Cardy, C. M. & Handford, P. (1996). Calcium binding properties of an EGF-like domain pair from human fibrillin-1. *J. Mol. Biol.* **255**, 22-27.

Linke, W. A., Ivemeyer, M., Olivieri, N., Kolmerer, B., Rüegg, J. C. & Labeit, S. (1996). Towards a molecular understanding of the elasticity of titin. *J. Mol. Biol.* **261**, 62-71.

Linke, W. A., Ivemeyer, M., Mundel, P., Stockmeier, M. R. & Kolmerer, B. (1998). Nature of PEVK-titin elasticity in skeletal muscle. *Proc. Natl. Acad. Sci. USA*. **95**, 8052-8057.

Mayhew, M., Handford, P. A., Baron, M., Tse, A. G., Campbell, I. D. & Brownlee, G. G. (1992). Ligand requirements for Ca^{2+} binding to EGF-like domains. *Protein. Eng*. **5**, 489-494.

Mecham, R. P. & Heuser, J. E. (1991). The elastic fiber. In Cell Biology of the Extracellular Matrix, 2nd edition. E. D. Hay, ed., Plenum Publishing Co., New York, pp. 79-109.

Muranyi, A., Finn, B. E., Gippert, G. P., Forsen, S., Stenflo, J. & Drakenberg, T. (1998). Solution structure of the N-terminal EGF-like domain from human factor VII. *Biochemistry*. **37**, 10605-10615.

Park, E. S., Putnam, E. A., Chitayat, D., Child, A. & Milewicz, D. M. (1998). Clustering of *FBN2* mutations in patients with congenital contractural arachnodactyly indicates an important role of the domains encoded by exons 24 through 34 during human development. *Am. J. Med. Genet.* **78**, 350-355.

Pereira, L., D'Alessio, M., Ramirez, F., Lynch, J. R., Sykes, B., Pangilinan, T. & Bonadio, J. (1993). Genomic organisation of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Hum. Mol. Genet.* **2**, 961-968.

Pereira, L., Andrikopoulos, K., Tian, J., Lee, S. Y., Keene, D. R., Ono, R., Reinhardt, D. P., Sakai, L. Y., Biery, N. J., Bunton, T., Dietz, H. C. & Ramirez, F. (1997). Targetting of the gene encoding fibrillin-1 recapitulates the vascular aspect of Marfan syndrome. *Nat. Genet.* **17**, 218-222.

Pfaff, M., Reinhardt, D. P., Sakai, L. Y. & Timpl, R. (1996). Cell adhesion and integrin binding to recombinant human fibrillin-1. *FEBS Letts*. **384**, 247-250.

Pickford, A. R., Smith, S. P., Staunton, D., Boyd, J. & Campbell, I. D.. Structure and function of a three-module fragment from the gelatin binding domain of human fibronectin. Manuscript in preparation.

Putnam, E. A., Cho, M., Zinn, A. B., Towbin, J. A., Byers, P. H. & Milewicz, D. M. (1996). Delineation of the Marfan phenotype associated with mutations in exons 23-32 of the *FBN*-1 gene. *Am. J. Med. Genet.* **62**, 233-242.

Qian, R. Q. & Glanville, R. W. (1997). Alignment of fibrillin molecules in elastic microfibrils is defined by transglutaminase-derived cross-links. *Biochemistry*. **36**, 15841-15847.

Ramirez, F., Pereira, L., Zhang, H. & Lee, B. (1993). The fibrillin-Marfan syndrome connection. *Bioessays*. **15**, 589-594.

Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G. & Stuart, D. (1995). The structure of a Ca^{2+}-binding epidermal growth factor-like domain: Its role in protein-protein interactions. *Cell.* **82**, 131-141.

Rees, D. J. G., Jones, I. M., Handford, P. A., Walter, S. J., Esnouf, M. P., Smith K. J. & Brownlee G. G. (1988). The role of β-hydroxyaspartate and adjacent

carboxylate residues in the first EGF domain of human factor IX. *EMBO. J.* **7**, 2053-2061.

Reinhardt, D. P., Keene, D. R., Corson, G. M., Pöschl, E., Bächinger, H. P., Gambee, J. E. & Sakai, L. Y. (1996). Fibrillin-1: organisation in microfibrils and structural properties. *J. Mol. Biol.* **258**, 104-116.

Reinhardt, D. P., Gambee, J. E., Ono, R. N., Bachinger, H. P. & Sakai, L. Y. (2000a). Initial steps in the assembly of microfibrils: Formation of disulphide-cross-linked multimers containing fibrillin-1. *J. Biol. Chem.* **275**, 2205-2210.

Reinhardt, D. P., Ono, R. N., Notbohm, H., Muller, P. K., Bachinger, H. P. & Sakai, L. Y. (2000b). Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis: A potential disease-causing mechanism in Marfan syndrome. *J. Mol. Biol.* **275**, 12339-12345.

Ross, J. M., McIntire, L. V., Moake, J. L., Kuo, H.-J., Qian, R.-Q., Glanville, R. W., Schwartz, E. & Rand, J. H. (1998). Fibrillin containing elastic microfibrils support platelet adhesion under dynamic sheer conditions. *Thromb. Haemost.* **79**, 155-161.

Sakai, L. Y., Keene, D. R. & Envall, E. (1986). Fibrillin, a new 350kD glycoprotein is a component of extracellular microfibrils. *J. Cell. Biol.* **103**, 2499-2509.

Sakai, L. Y., Keene, D. R., Glanville, R. W. & Bachinger, H. P. (1991). Purification and partial characterisation of fibrillin, a cys-rich structural component of connective tissue microfibrils. *J. Biol. Chem.* **266**, 14763-14770.

Sakai, L. Y. & Keene, D. R. (1994). Fibrillin: Monomers and microfibrils. *Meth. Enzym.* **245**, 29-52.

Sakamoto, H., Broekelmann, T., Cheresh, D. A., Ramirez, F., Rosenbloom, J. & Mecham, R. P. (1996). Cell-type specific recognition of RGD- and non-RGD-containing cell binding domains in fibrillin-1. *J. Biol. Chem.* **271**, 4916-4922.

Selander-Sunnerhagen, M., Ullner, M., Persson, E., Teleman, O., Stenflo, J. & Drakenberg, T. (1992). How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the NH_2-terminal EGF-like domain in coagulation factor X. *J. Biol. Chem.* **267**, 19642-19649.

Skonier, J., Neubauer, M., Madisen, L., Bennett, K., Plowman, G. & Purchio, A. (1992). cDNA cloning and sequence analysis of βig-h3, a novel gene induced in a human adenocarcinoma cell-line after treatment with transforming growth factor-β. *DNA Cell Biol.* **11**, 511-522.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol.* **286**, 661-668.

Sunnerhagen, M., Olah, G. A., Stenflo, J., Forsen, S., Drakenberg, T. & Trewhella, J. (1996). The relative orientation of Gla and EGF domains in coagulation- factor-X is altered by Ca^{2+} binding to the first EGF domain - a combined NMR small-angle X-ray-scattering study. *Biochemistry.* **35**, 11547-11559.

Superti-Furga, A., Raghunath, M. & Willems, P. J. (1992). Deficiencies of fibrillin and decorin in fibroblast cultures of a patient with neonatal Marfan syndrome. *J. Med. Genet.* **29**, 875-878.

Trask, T. M., Ritty, T. M., Broekelmann, T., Tisdale, C. & Mecham, R. P. (1999). N-terminal domains of fibrillin 1 and fibrillin 2 direct the formation of homodimers: a possible first step in microfibril assembly. *Biochem. J.* **340**, 693-701.

Trask, B. C., Trask, T. M., Broekelmann, T. & Mecham, R. P. (2000). The microfibrillar proteins MAGP-1 and fibrillin-1 form a ternary complex with the chondroitin sulphate proteoglycan decorin. *Mol. Biol. Cell.* **11**, 1499-1507.

Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T. & Teleman, O. (1992). Three-dimensional structure of the Apo form of the N-terminal EGF-like module of blood coagulation factor X as determined by NMR spectroscopy and simulated folding. *Biochemistry.* **31**, 5974-5983

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (2000). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296**, 1065-1078.

Wess, T. J., Purslow, P. P., Sherratt, M. J., Ashworth, J., Shuttleworth, C. A. & Kielty, C. M. (1998). Calcium determines the supramolecular organisation of fibrillin-rich microfibrils. *J. Cell Biol.* **141**, 829-837.

Whiteman, P., Downing, A. K., Smallridge, R., Winship, P. R. & Handford, P. A. (1998). A Gly to Ser change causes defective folding *in vitro* of calcium binding epidermal growth factor-like domains from factor IX and fibrillin-1. *J. Biol. Chem.* **273**, 7807-7813.

Wunsch, A. M., Little, C. D. & Markwald, R. R. (1994). Cardiac endothelial heterogeneity defines valvular development as demonstrated by the diverse expression of JB3, an antigen of the endocardial cushion tissue. *Devel. Biol.* **165**, 585-601.

Yeh, H., Chow, M.,. Abrams, W. R., Fan, J., Foster, J., Mitchell, H., Muenke, M. & Rosenbloom, J. (1994). Structure of the human gene encoding the associated microfibrillar protein (MFAP1) and localization to chromosome 15q15-q21. *Genomics.* **23**, 443-449.

Zhao, Z., Lee, C. C., Jiralerspong, S., Juyal, R. C., Lu, F., Baldini, A., Greenberg, F., Caskey, C. T. & Patel, P. I. (1995). The gene for a human microfibril-associated glycoprotein is commonly deleted in Smith-Magenis patients. *Hum. Mol. Genet.* **4**, 589-597.

Chapter 2

Methods for the Analysis of Protein Structure, Dynamics and Calcium Binding

2.1 Introduction to Intrinsic Protein Fluorescence

When a molecule absorbs a photon of light and subsequently emits light of a longer wavelength (lower energy), the emission is called fluorescence. Fluorescence emission results when excited electrons revert from their first excited state back to their ground state. The lifetime of the excited state is such that it can be influenced by many interactions and perturbations, making fluorescence emission extremely sensitive to environment. This sensitivity means that only relatively small amounts of sample are required for fluorescence spectroscopy measurements, i.e., nM-lower μM range. Fluorescence measurements, on only a small amount of sample, can therefore provide a wealth of biochemical information, including details concerning conformation, flexibility, ligand binding, intermolecular distances and solvent interactions.

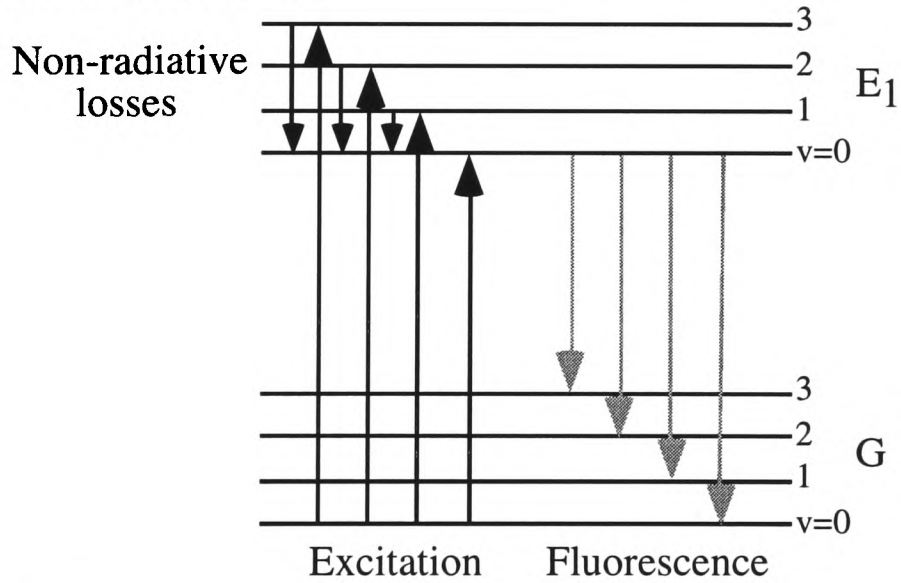
Initially in this section, an introduction to the basic theory of fluorescence will be presented based on the works of Freifelder (1982), Schmid (1997) and Lakowicz (1999). Intrinsic protein fluorescence will then be discussed, followed by an overview of the experimental considerations required when performing fluorescence experiments.

2.1.1 An Overview of Fluorescence Theory

The simplest way of understanding fluorescence is by considering the energy-level diagram shown in Figure 2.1. This figure shows two electronic states of a molecule which is capable of fluorescence, the ground state (G) and the first excited state (E_1). Within each of these states are a number of sub-levels called vibrational energy levels (v). Absorption of a photon induces electronic shifts from the G state to the E_1 state, i.e., the molecule has been "excited" and has moved from a lower to a higher energy state. Which vibrational level (v) is reached within the E_1 state depends upon the amount of energy absorbed.

Figure 2.1 : Energy-level Diagram to Illustrate the Origin of Fluorescence Emission

This diagram is a simplified Jablonski diagram (see Lakowicz, 1999), and was modified from Freifelder, 1982. Vibrational levels (v) are shown as horizontal lines. Absorption of a photon (excitation) is shown as black, upward arrows. Non-radiative losses are shown as small, black, downward arrows. Fluorescence emission is represented as grey, downward arrows. E_1 and G represent the first, excited and ground state of a molecule, respectively.



After excitation, the molecule then begins to lose energy and returns to its ground state. Initially non-radiative losses occur as the molecule relaxes to the lowest vibrational level of the E_1 state. This process is called internal conversion. Subsequently, energy can then be lost either by non-radiative processes (discussed later) or by fluorescence emission. In Figure 2.1, the latter is a transition from the lowest vibrational energy level of E_1, v = 0, to one of the vibrational energy levels in G,

$v \geq 0$. As energy is lost in dropping to the lowest vibrational state of E_1, the emitted light has a lower energy than that of the absorbed light, and will therefore be observed at a longer wavelength. Also, because the fluorescence emission occurs from E_1, $v = 0$ to G, $v \geq 0$, the emitted light can have a range of energies.

Fluorescence intensity is generally described in terms of quantum yield, Q,

$$Q = \frac{\text{number of photons emitted}}{\text{number of photons absorbed}}$$

where a photon is a unit of light having the energy $E = hv$, where h is Planck's constant and v is the light wave frequency. Determination of Q is difficult, because, as mentioned previously, fluorescence emission occurs at a range of energies/wavelengths. Therefore in most biochemical analyses the fluorescence intensity measured is usually the total fluorescence energy produced by the whole system. This is generally sufficient for obtaining the information desired from fluorescence spectroscopy (Freifelder, 1982).

Both molecular properties and environmental factors determine fluorescence emission. For example, in the case of molecular properties, if the upper vibrational levels of the G state overlap with the lower vibrational levels of the E_1 state, non-radiative losses can occur to the ground state, meaning no fluorescence is emitted. This is often the case for flexible molecules. In terms of environment, fluorescence emission can be reduced by quenching. Quenching usually occurs as a result of collision or resonance energy transfer. The former can be either simple collision with energy exchange or the result of a chemical reaction, while the latter is a long-range radiative transfer process. The latter will not be described further here (see Freifelder, 1982; Lakowicz, 1999).

2.1.2 Intrinsic Protein Fluorescence

In proteins, only three of the twenty naturally occurring amino acids are capable of fluorescence emission. These residues are tryptophan, tyrosine and

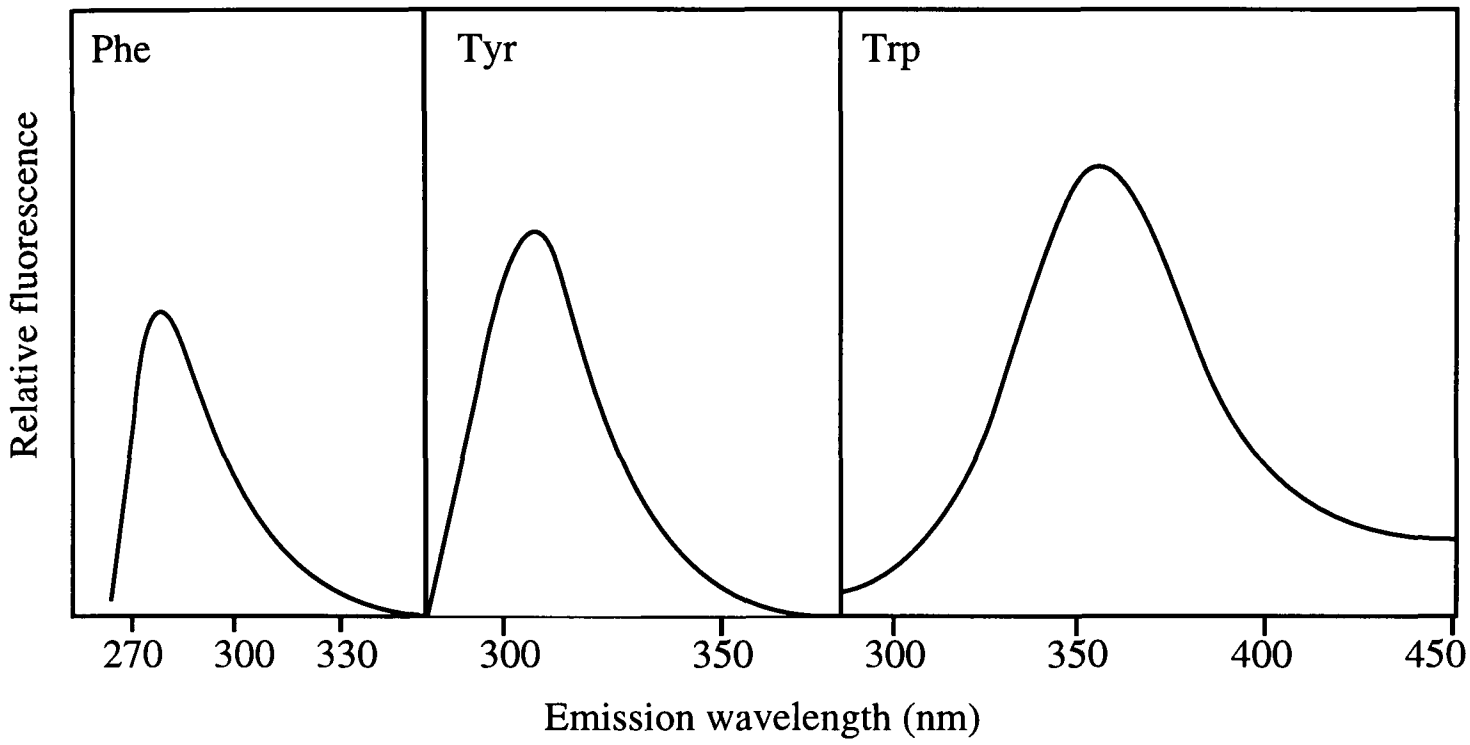
phenylalanine, i.e., non-flexible, aromatic residues. These amino acids have the advantage of appearing infrequently in proteins, i.e., a protein may contain only one or a few fluorescent residues. Fluorescence emission of proteins can therefore be of great value, since the resultant spectra can often be easily interpreted.

The emission spectra of Trp, Tyr and Phe are shown in Figure 2.2, and their absorbance and emission properties are shown in Table 2.1. From these data it is clear that when proteins contain Trp residues, these residues contribute dominantly to fluorescence emission. The difference in fluorescence potential of the three fluorescent residues is highlighted by the sensitivity parameter (Table 2.1), which shows that Trp is > 3 times more sensitive than Tyr and > 180 times more sensitive than Phe. In proteins containing Trp residues, it is unlikely that fluorescence will be observed from either Phe or Tyr. Phe fluorescence is likely to be quenched by energy transfer to the other two aromatic amino acids, and Tyr fluorescence is usually also transferred to Trp or is undetectable due to the strength of Trp fluorescence (Schmid, 1997; Lakowicz, 1999).

So in proteins devoid of Trp, is fluorescence no longer an experimental option? From the sensitivity values (Table 2.1), the next best residue to exploit is tyrosine, but tyrosine fluorescence is often weak due to quenching. Tyrosine fluorescence can be quenched if it is near an amino group, carboxyl group or tryptophan. However, in some situations it can be detected after excitation at ~275nm. Proteins of interest should therefore always have their fluorescence capability assessed, even when no tryptophan residues are present. In this project, intrinsic tyrosine fluorescence was used to monitor calcium binding to the high affinity site in cbEGF13 of the cbEGF12-13 domain pair, as this domain pair does not contain tryptophan residues (see Chapter 4). This method proved successful, despite the inherently low sensitivity of tyrosine, which resulted in its subsequent use in monitoring calcium binding to another cbEGF domain pair devoid of tryptophan residues, i.e., the AB pair from the LDL receptor (Malby *et al.*, manuscript in preparation).

Figure 2.2 : Fluorescence Emission Spectra of the Aromatic Amino Acids

This diagram was reproduced from Schmid, 1997. The spectra were recorded in 0.01M potassium phosphate buffer pH 7.0 at 25°C. 100μM Phe, 6μM Tyr and 1μM Trp were used, with the excitation wavelengths being 257nm, 274nm and 278nm, respectively.



The main reason to study intrinsic protein fluorescence is to acquire information about conformation. Both tryptophan and tyrosine fluorescence are extremely sensitive to their environment, so changes in fluorescence are indicative of an altered environment around these residues. Monitoring intrinsic protein fluorescence can therefore provide information concerning protein structure, dynamics, folding and ligand binding.

Table 2.1 : Absorbance and Fluorescence Properties of the Aromatic Amino Acids

This table was reproduced from Schmid, 1997. The properties were observed in water at neutral pH.

Amino acid	**Absorbance**		**Fluorescence**		**Sensitivity**
	λ_{max} (nm)	ε_{max} ($M^{-1}cm^{-1}$)	λ_{max} (nm)	Quantum yield, Q	ε_{max} x Q ($M^{-1}cm^{-1}$)
Tryptophan	280	5600	355	0.13	730
Tyrosine	275	1400	304	0.14	200
Phenylalanine	258	200	282	0.02	4

2.1.3 Experimental Considerations

As fluorescence spectroscopy is a highly sensitive technique, it is essential to optimise experimental conditions. For example, fluorescence intensity is greatly affected by temperature, with intensity usually decreasing with increasing temperature. Tyrosine emission can decrease by more than 1% for every degree rise in temperature (Schmid, 1997). It is therefore essential to maintain a constant optimum temperature throughout fluorescence experiments, and in this project thermostatted cuvette holders were used for this purpose.

Choice of solvent is also essential. The ideal solvent is minimalistic, i.e., deionised water, since this reduces the likelihood of fluorescent impurities. However, the inclusion of other components is often unavoidable, e.g., buffers to minimise pH fluctuations. Any solvent or buffer used must therefore not absorb in the emission and excitation spectral regions of the fluorophore to be observed. Cuvettes used must be cleaned thoroughly and free from solid particles, and the samples used should also be filtered to remove solid particles.

In a fluorescence spectrum of protein in water, a Raman scattering peak is observed resulting from the water molecules. This scattering peak often overlaps with the fluorescence emission spectrum of the protein, and is located at a fixed energy difference from the incident radiation. The influence of the Raman peak can be removed in two ways. The first method involves subtraction of a blank spectrum from that of the sample spectrum. Alternatively, the excitation wavelength can be shifted to remove Raman peak interference. In this project, the former subtraction method was employed.

The choice of slit width used for excitation is another consideration. A compromise must be reached between signal-to-noise and possibility of sample damage. The lower the protein concentration and the less sensitive the fluorescent residue, the wider the slit width required. In general, the narrowest slit width possible should be utilised, and the excitation slit should always be closed between

measurements to minimise sample damage due to exposure to intense light. Choice of excitation wavelength is reasonably flexible, and can be designed to minimise contributions from the Raman peak or impurities. A variety of wavelengths and emission slits widths are utilised to monitor protein emission spectra.

In summary, fluorescence is a highly sensitive technique that can provide a wealth of information. This sensitivity, however, means that great care must be taken when performing fluorescence experiments, to ensure the most accurate results possible are obtained.

2.2 Introduction to Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) is a technique that detects the reorientation of nuclear spins subjected to applied magnetic fields. It is a very powerful technique, which can be used in the investigation of the biological properties of proteins. NMR is one of only two techniques available for determining the three-dimensional structures of macromolecules at atomic resolution, and NMR also has many other applications, e.g., monitoring ligand binding (Craik & Wilce, 1997) and the study of backbone dynamics (Stone *et al.*, 1992). In this project, NMR was utilised to characterise the cbEGF12-13 domain pair from human fibrillin-1, in terms of calcium binding (Smallridge *et al.*, 1999), structure, backbone dynamics and the consequences of a MFS-causing mutation (Whiteman *et al.*, manuscript submitted).

This section presents an introduction to NMR theory. This theory has been described extensively elsewhere, e.g., in Abragam, 1961; Ernst *et al.*, 1987; Cavanagh *et al.*, 1996, so this description will focus on introducing key experiments used in this project. Firstly, a general introduction to one- (1D), two- (2D) and three-dimensional (3D) NMR will be presented. 3D heteronuclear NMR experiments were used in this project to both increase the amount of NMR data available for the cbEGF12-13 domain pair and to improve the resolution of the data. Use of 3D data resulted in greater

confidence in spectral assignment, as well as an increased number of restraints for structure calculations. A quantitative description of 3D NMR pulse sequences, however, lies beyond the scope of this thesis, therefore a qualitative overview of 3D NMR will be given. The product operator formalism will then be introduced. Product operators provide a method of precisely analysing the time evolution of a nuclear spin's magnetisation during the course of a NMR experiment, and, although they are inherently quantum mechanical in nature, their use requires no knowledge of quantum mechanics.

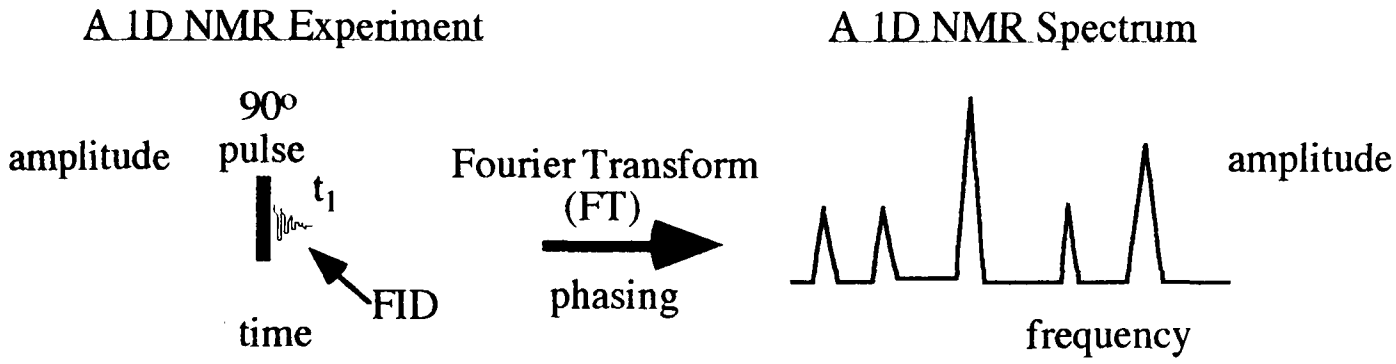
2.2.1 Qualitative Introduction to One-Dimensional NMR

A 1D NMR experiment, shown schematically in Figure 2.3, involves the application of a 90° radio frequency (rf) pulse to a sample in a static, magnetic B_o field, and the length of time for which a rf pulse is applied is known as the pulse width. The pulse width determines the range of frequencies excited in the sample, and usually has a value in the μs range. Following this pulse, the free induction decay (FID) is detected and recorded during the t_1 time period. This FID is the change in the amplitude of the response over time, i.e., the change in the signal as the bulk sample magnetisation returns to its equilibrium state. For a given spin within a macromolecule, the rate at which the induced magnetisation relaxes depends upon its local environment. The FID is processed using Fourier transformation (FT), which converts time-domain data into a frequency-domain spectrum. The resulting data are then phased (see section 2.3.1.1), so that the lineshapes in the spectra have an absorptive character (Figure 2.3).

1D NMR experiments can be successfully applied to the study of small biological molecules. For example, for proteins less than ~5KDa, 1D NMR can be used to monitor ligand binding, and this method was used to determine a K_d value for calcium binding to the isolated cbEGF13 domain from human fibrillin-1 (Whiteman *et al.*, 1998). However, as molecular weight increases, significant line-broadening causes

1D spectra to become too overlapped to allow the identification of specific resonances. 1D NMR experiments are, however, still invaluable for both initial checks of sample quality, and when setting up NMR experiments, e.g., for measuring 90° pulse widths.

Figure 2.3 : Schematic Representation of a 1D NMR Experiment



2.2.2 Qualitative Introduction to Two-Dimensional NMR

There are many types of 2D NMR experiments, e.g., COSY, TOCSY, NOESY, which are all made up of different arrangements of 90° pulses during four basic time periods (Figure 2.4). The first time period is preparation, where the

Figure 2.4 : Schematic Representation of the Four Basic Time Periods of a 2D NMR Experiment

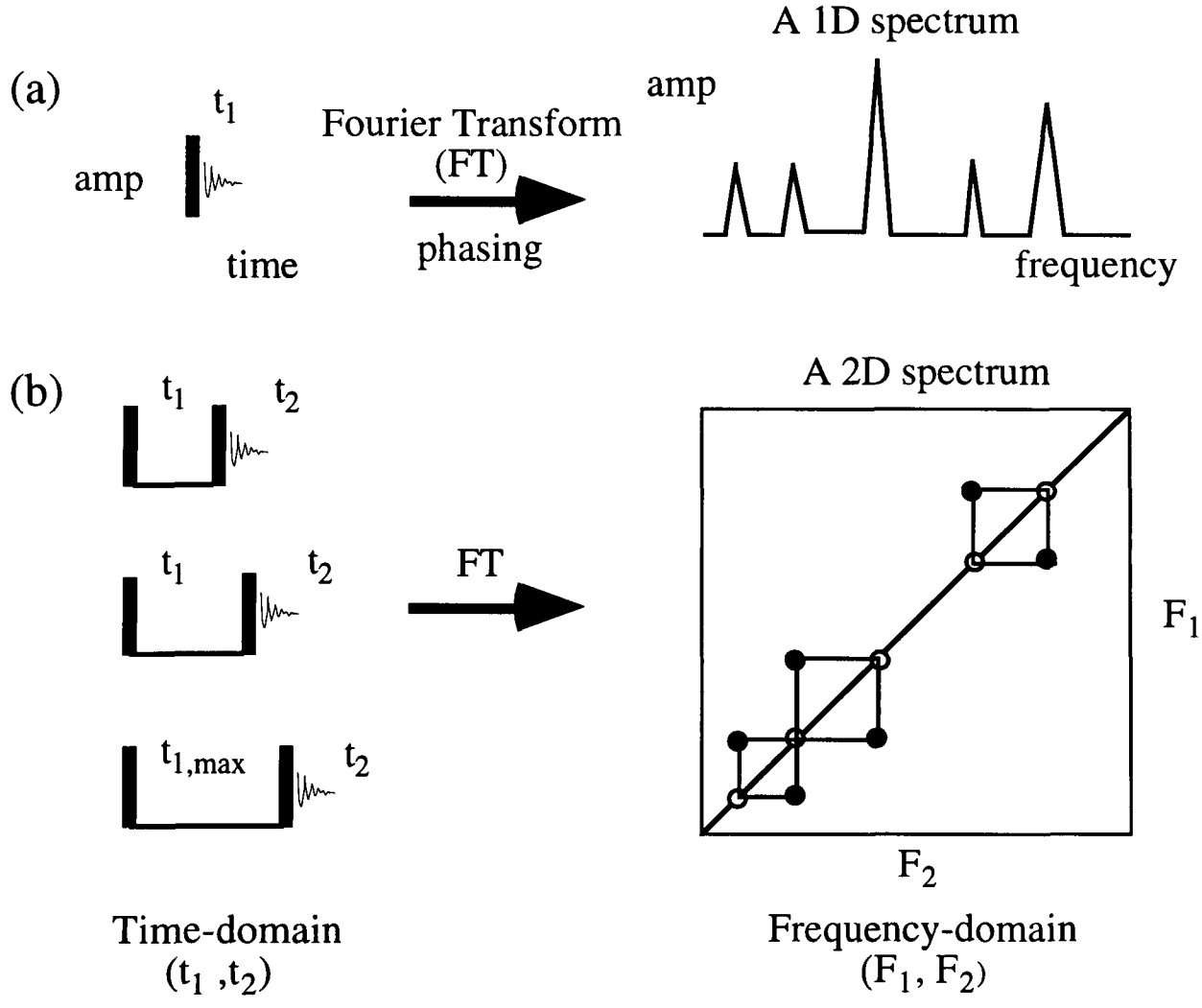
Preparation	Evolution (t_1)	Mixing	Acquisition(t_2)

equilibrium magnetisation of a spin system, a result of the static B_0 magnetic field, is converted to the desired nonequilibrium state. This is usually achieved with a 90° pulse, which generates transverse magnetisation. This nonequilibrium magnetisation then evolves during the second time period called evolution. Often this evolution period is simply a period of free precession, although evolution can be under applied rf pulses. During this evolution period, the detected magnetisation produces signals whose frequencies correspond to the F_1 (D_2) dimension of the resulting 2D spectrum, a process known as F_1 frequency labelling of magnetisation. The evolution time, t_1, is

incremented in discrete steps up to a maximum value known as $t_{1,max}$ (Figure 2.5). The third time period is the mixing period, where magnetisation is transferred from one spin to another. This mixing period establishes the correlation between the two dimensions of a 2D NMR spectrum, and dictates the information present in the resultant spectrum. Finally there is the acquisition period (t_2), where the FID is detected. Phase cycling and pulsed field-gradient methods can be used to select a particular magnetisation transfer pathway at this stage, if more than one pathway is possible. For a detailed description of the latter two methods the reader is referred to Keeler *et al.* (1994).

Figure 2.5 : Schematic Comparison of (a) 1D and (b) 2D NMR Basic Experimental Schemes and Spectra

The black bars denotes 90° pulses. The label "amp" stands for amplitude. Fourier transform (FT) is used to convert digitised time-domain signals into frequency-domain NMR spectra. Diagonal peaks in the 2D spectrum are represented as hollow circles, whereas cross-peaks are represented as filled circles.



To summarise, 2D NMR experiments are purely an extension of 1D NMR experiments, i.e., the signal modulated as a function of t_1 in 1D NMR is , in 2D NMR,

later detected as a function of t_2. The duration of t_1 is varied over the course of the experiment to resolve the data in two dimensions (Figure 2.5). If the magnetisation evolves with the same frequency in t_1 and t_2, only diagonal peaks will be seen in the 2D spectrum. The diagonal corresponds exactly to the 1D spectrum of the sample, meaning that no extra information is obtained if the frequency evolution is the same in t_1 and t_2. Hence, in 2D experiments the magnetisation is labelled so that it evolves differently in t_1 and t_2, which results in cross peaks in the spectrum at the intersection of the chemical shifts of the interacting spins.

2.2.3 Qualitative Introduction to Three-Dimensional NMR

As 2D NMR experiments are modular in nature, new higher-dimensional NMR experiments can be constructed from combinations of 2D NMR experiments, e.g., two 2D pulse sequences can be combined to make a 3D pulse sequence (Figure 2.6). In 3D NMR experiments, the acquisition period is now denoted as t_3, and the signal recorded here will have evolved during two evolution time periods (t_1 and t_2). These time periods are incremented independently from one 2D experiment to the next. Three-dimensional experiments can allow a combination of magnetisation transfer by both through-bond (dipolar) and through-space (scalar) coupling interactions within a single pulse sequence.

The result of a 3D NMR experiment is a three-dimensional time-domain matrix, which can be Fourier transformed to produce a three-dimensional frequency-domain matrix (F_1, F_2, F_3). A 3D spectrum can be imagined as a cube, but to simplify analysis 2D slices of this cube are taken (Figure 2.7). Each 2D slice is like a 2D spectrum, except that the two dimensions visualised are edited by coupling to a third nucleus, which is often a heteronucleus, e.g., ^{15}N. The 3D NMR experiments used in this project were heteronuclear experiments, i.e., one dimension of the resultant spectrum was ^{15}N and two were ^{1}H.

Figure 2.6 : Schematic Diagram to Show the Production of a Three-Dimensional NMR Experiment from Two Two-Dimensional NMR Experiments

This figure is modified from Cavanagh *et al.*, 1996.

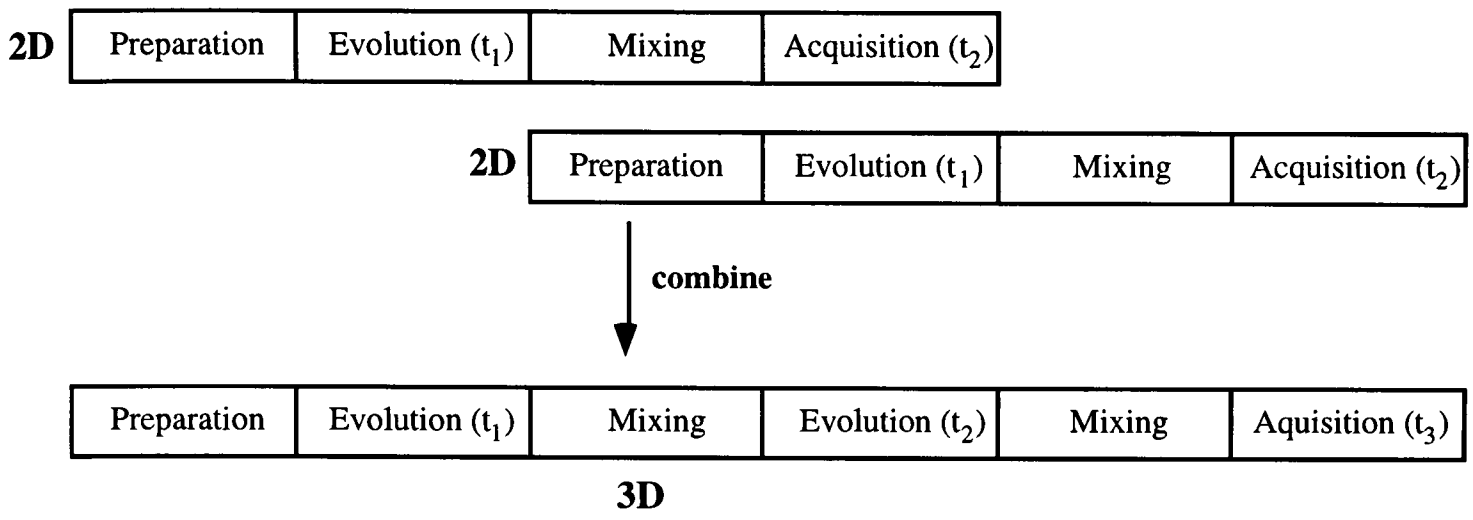
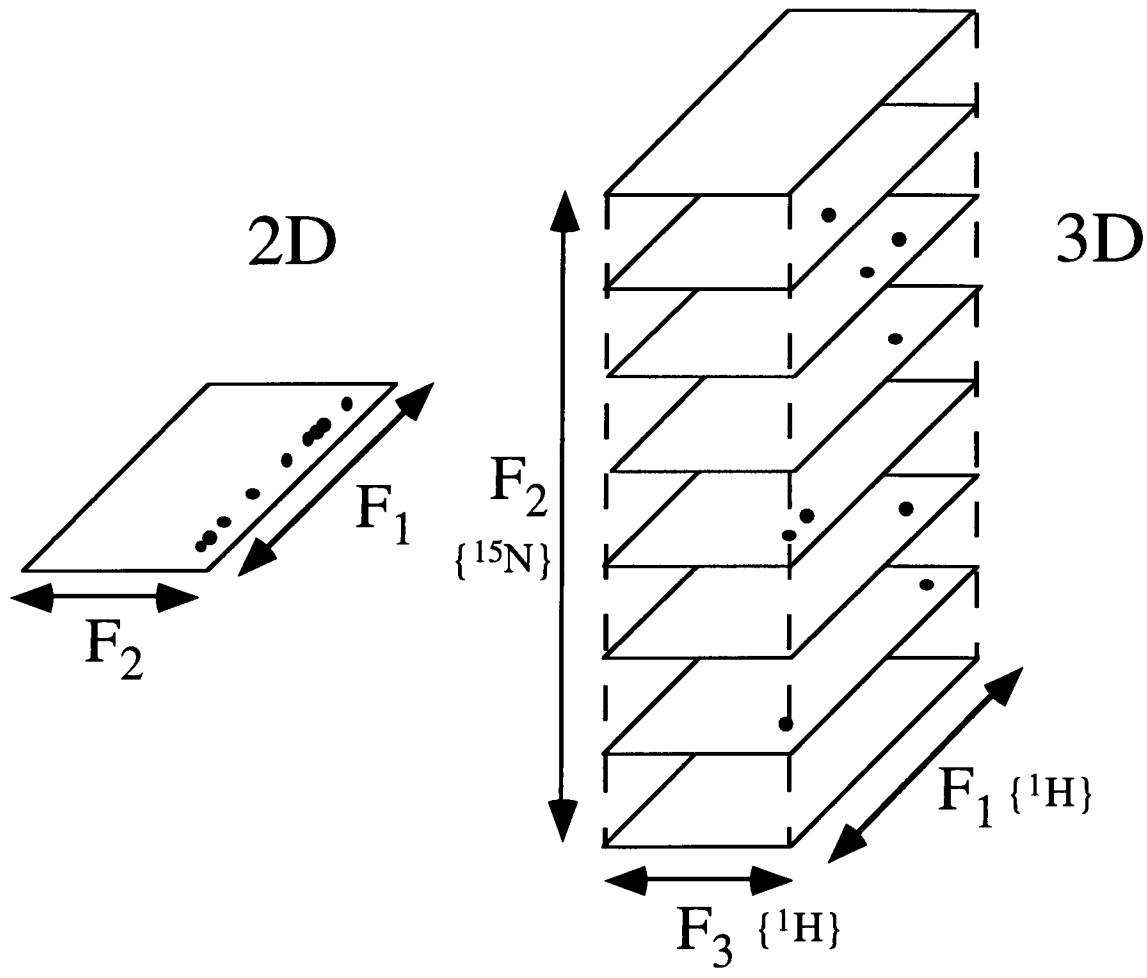


Figure 2.7 : Schematic Representation of 2D and 3D NMR Spectra

This diagram illustrates how increasing the number of dimensions can improve spectral resolution and clarity. This figure is modified from Clore & Gronenborn, 1991, and shows the dimension labels relevant to this project.



2.2.4 Introduction to the Product Operator Description of NMR

Product operators are a convenient way of following the magnetisation evolution of a spin system during application of a pulse sequence. They are used to represent the magnetisation states of a spin system under the influence of an external magnetic field. During NMR experiments, the magnetisation state of a spin system is changed constantly by a series of applied pulses and/or delays specific to a particular pulse sequence, and product operators allow visualisation of the magnetisation state of a spin system at each stage of an experiment. They provide an exact analysis under conditions of weak J coupling, and are useful in predicting the outcome of complex NMR experiments. The latter is possible because the time evolution of the spin system is determined by the time evolution of the product operators (Sørensen *et al.*, 1983).

This section provides an overview of the product operator formalism (for more detail see Sørensen *et al.*, 1983; Cavanagh *et al.*, 1996). Firstly the formalism will be introduced with respect to a single spin, two weakly coupled spins, the spin-echo pulse sequence and the INEPT pulse sequence. Subsequently, application of product operator calculations to the 2D experiments used in this project will be described.

2.2.4.1 Product Operators for One Spin

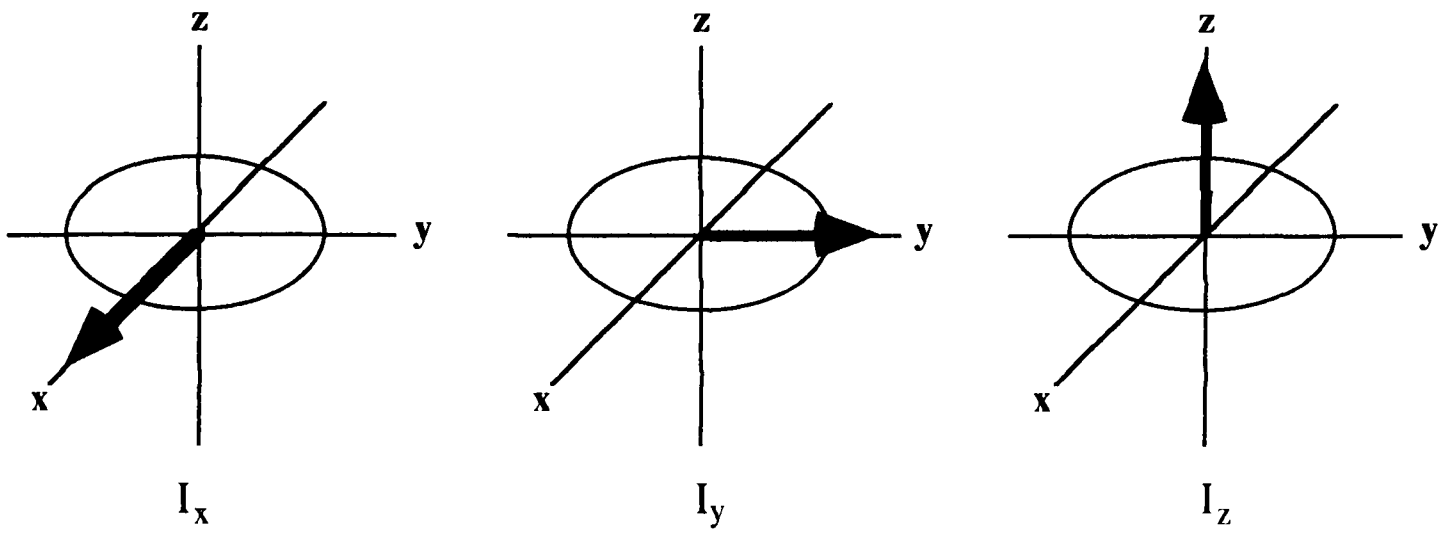
Only four product operators are needed to describe NMR experiments on a system containing only one $I = {}^1/_2$ spin,

$$ {}^1/_2E \qquad I_x \qquad I_y \qquad I_z $$

where E is the identity operator, and I_x, I_y and I_z correspond to x, y and z magnetisation of a single I spin in the rotating frame vector model (Figure 2.8). Using this vector model and the right-hand rule, the effect of 90° x and y pulses on these three operators can be easily predicted,

$$I_x \xrightarrow{90^\circ_x} I_x \qquad I_x \xrightarrow{90^\circ} -I_z$$
$$I_y \xrightarrow{90^\circ_x} I_z \qquad I_y \xrightarrow{90^\circ_y} I_y$$
$$I_z \xrightarrow{90^\circ_x} -I_y \qquad I_z \xrightarrow{90^\circ_y} I_x$$

Figure 2.8 : The Vector Model of a Single I Spin Illustrating x, y and z Magnetisation



I_x I_y I_z

As described in sections 2.2.1-3, 90° pulses are often used during the first time period of NMR experiments to prepare nonequilibrium magnetisation, and the above illustrates how product operators can allow visualisation of this effect. To generalise the previous set of equations, the effect of α_x and α_y pulses on the above operators can be written as,

$$I_x \xrightarrow{\alpha_x} I_x \qquad I_x \xrightarrow{\alpha_y} I_x\cos\alpha - I_z\sin\alpha$$
$$I_y \xrightarrow{\alpha_x} I_y\cos\alpha + I_z\sin\alpha \qquad I_y \xrightarrow{\alpha_y} I_y$$
$$I_z \xrightarrow{\alpha_x} I_z\cos\alpha - I_y\sin\alpha \qquad I_z \xrightarrow{\alpha_y} I_z\cos\alpha + I_x\sin\alpha$$

Similarly product operators can be used to describe the evolution of magnetisation of an I spin during a period of free precession, for example, in the evolution period (t_1) of a general 2D experiment. Under free precession for a time t and

at a resonance offset Ω (where $\Omega = \omega - \omega_o$, where ω_o is the Larmor frequency), the product operators evolve as,

$$I_x \xrightarrow{\Omega t} I_x\cos\Omega t + I_y\sin\Omega t$$
$$I_y \xrightarrow{\Omega t} I_y\cos\Omega t - I_x\sin\Omega t$$
$$I_z \xrightarrow{\Omega t} I_z$$

The above results can be visualised by looking at the vector model diagram (Figure 2.8) and imagining anticlockwise precession of magnetisation in each plane.

2.2.4.2 Product Operators for Two Weakly Coupled Spins

For a system of two weakly coupled spin-$^1/_2$ nuclei, I and S, there are sixteen product operators to be considered,

		$^1/_2E$	S_x	S_y	S_z
	$^1/_2E$	$^1/_2E$	S_x	S_y	S_z
2x	I_x	I_x	$2I_xS_x$	$2I_xS_y$	$2I_xS_z$
	I_y	I_y	$2I_yS_x$	$2I_yS_y$	$2I_yS_z$
	I_z	I_z	$2I_zS_x$	$2I_zS_y$	$2I_zS_z$

The factor of two is included for normalisation. The four product operators in the centre of the table, highlighted by the dashed box, represent multiple-quantum coherences. Multiple-quantum coherences are represented by product operators containing more than one transverse operator, and, as they cannot be directly observed by NMR, they will not be discussed further. For a two-spin system, only product operators containing a single transverse operator are observable by NMR.

Evolution of magnetisation as the result of an α_x or α_y pulse for a weakly coupled two-spin system is the same as for a single spin system, however, differences

arise when considering evolution of magnetisation during periods of free precession. Free precession for a two-spin system, I and S, under a scalar coupling of J_{IS} for a time t can be represented as,

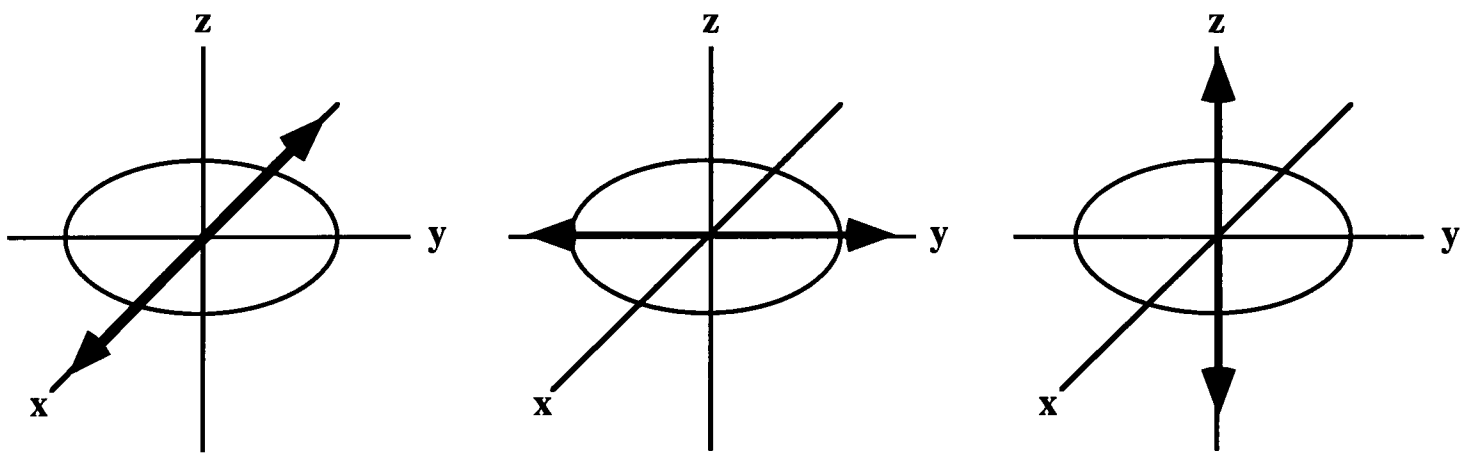
$$I_x \xrightarrow{\pi J_{IS} t} I_x\cos(\pi J_{IS}t) + 2I_yS_z\sin(\pi J_{IS}t)$$

$$I_y \xrightarrow{\pi J_{IS} t} I_y\cos(\pi J_{IS}t) - 2I_xS_z\sin(\pi J_{IS}t)$$

$$I_z \xrightarrow{\pi J_{IS} t} I_z$$

The first two equations demonstrate how in-phase magnetisation develops into anti-phase magnetisation under the influence of the weak scalar coupling (Figure 2.9).

Figure 2.9 : Vector Model to Illustrate Anti-Phase Magnetisation of the Two-Spin Product Operators $2I_xS_z$, $2I_yS_z$ and $2I_zS_z$



$2I_xS_z$
Antiphase magnetisation along x

$2I_yS_z$
Antiphase magnetisation along y

$2I_zS_z$
Antiphase magnetisation along z

To further explain, S_z can be represented in terms of energy level populations, i.e., $S_z = {}^1/_2 (\alpha - \beta)$, where $S_\alpha + S_\beta = 1$. Consequently, $2I_xS_z = I_xS_\alpha - I_xS_\beta$, i.e., antiphase magnetisation along x for spin I.

Similarly the evolution of two-spin operators under the conditions described above is given by,

$$2I_xS_z \xrightarrow{\pi J_{IS} t} 2I_xS_z\cos(\pi J_{IS}t) + I_y\sin(\pi J_{IS}t)$$

$$2I_yS_z \xrightarrow{\pi J_{IS} t} 2I_yS_z\cos(\pi J_{IS}t) - I_x\sin(\pi J_{IS}t)$$

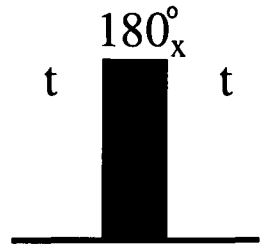
$$2I_zS_z \xrightarrow{\pi J_{IS}t} 2I_zS_z$$

Evolution of the S spin in the above cases can be obtained simply by exchanging the I and S labels.

2.2.4.3 Product Operators for the Spin-Echo Pulse Sequence

The spin-echo pulse sequence, shown in Figure 2.10, is an essential part of many NMR experiments. It consists of a single 180° pulse between two time periods, t, and the effect of this pulse sequence will now be examined in three different cases.

Figure 2.10 : The Spin-Echo Pulse Sequence



2.2.4.3.1 The Single Spin Case

Consider a single spin I. Before the application of the spin-echo pulse sequence, initial equilibrium I_z magnetisation is converted to $-I_y$ magnetisation by a $90°_x$ pulse,

$$I_z \xrightarrow{90°_x} -I_y$$

During the subsequent spin-echo pulse sequence, the chemical shift evolution of the I spin can be written as,

$$-I_y \xrightarrow{\Omega t} -I_y\cos\Omega t + I_x\sin\Omega t$$
$$\xrightarrow{180°_x} I_y\cos\Omega t + I_x\sin\Omega t$$

$$\xrightarrow{\Omega t} I_y\cos^2\Omega t - I_x\cos\Omega t\sin\Omega t + I_x\sin\Omega t\cos\Omega t + I_y\sin^2\Omega t$$

Using the trigonometric identity $\cos^2\theta + \sin^2\theta = 1$, the above result simplifies to,

$$\longrightarrow I_y$$

Therefore, the effect of the spin-echo on a single spin I was to take $-I_y$ magnetisation and evolve it to the final state I_y, i.e., no net evolution under chemical shift occurred during the spin-echo pulse sequence. Evolution under chemical shift was refocused, explaining why the spin-echo 180°_x pulse is often referred to as a refocusing pulse.

2.2.4.3.2 The Two-Spin Homonuclear Case

Consider application of the same spin-echo pulse sequence to a pair of scalar-coupled spins, I and S, where I and S are both protons, i.e., the homonuclear case. The pulses applied are assumed to be non-selective and will therefore affect both spins equally. In the previous section, it was shown that for a single spin there is no net evolution under chemical shift during the spin-echo pulse sequence. Consequently, the chemical shift evolution during the spin-echo can be ignored for each spin of the scalar-coupled pair. Considering I magnetisation, an initial 90°_x pulse is first used to generate $-I_y$ magnetisation. The evolution during the spin-echo for the IS system, under the scalar-coupling J_{IS}, can then be written as,

$$-I_y \xrightarrow{\pi J_{IS}t} -I_y\cos(\pi J_{IS}t) + 2I_xS_z\sin(\pi J_{IS}t)$$
$$\xrightarrow{(180^\circ_x)_I} I_y\cos(\pi J_{IS}t) + 2I_xS_z\sin(\pi J_{IS}t)$$
$$\xrightarrow{(180^\circ_x)_S} I_y\cos(\pi J_{IS}t) - 2I_xS_z\sin(\pi J_{IS}t)$$
$$\xrightarrow{\pi J_{IS}t} I_y\cos^2(\pi J_{IS}t) - 2I_xS_z\cos(\pi J_{IS}t)\sin(\pi J_{IS}t) - 2I_xS_z\sin(\pi J_{IS}t)\cos(\pi J_{IS}t) - I_y\sin^2(\pi J_{IS}t)$$

Therefore, using the trigonometric identities $\cos(2\theta) = \cos^2\theta - \sin^2\theta$ and $\sin(2\theta) = 2\sin\theta\cos\theta$, the above result simplifies to,

$$\longrightarrow I_y\cos(2\pi J_{IS}t) - 2I_xS_z\sin(2\pi J_{IS}t)$$

The result for the S spin can be obtained simply by interchanging the I and S labels. Setting $t = 1/(4J_{IS})$ results in generation of the anti-phase term $2I_xS_z$, whereas setting $t = 1/(2J_{IS})$ results in generation of the in-phase term $-I_y$. The spin-echo pulse sequence is routinely used to generate the anti-phase state in many NMR pulse sequences.

2.2.4.3.3 The Two-Spin Heteronuclear Case

If the two spins involved in a scalar coupling interaction are of a different nuclear type, e.g., ^{15}N and ^{1}H, then the rf pulses of a spin-echo sequence will no longer affect both spins equally. The rf pulses in this sequence can therefore only be applied to the I or S spin in the heteronuclear IS scalar-coupled pair. Chemical-shift evolution can again be ignored, so considering I magnetisation evolution in this system, after a 90°_x pulse has generated $-I_y$ magnetisation,

$$-I_y \xrightarrow{\pi J_{IS}t} -I_y\cos(\pi J_{IS}t) + 2I_xS_z\sin(\pi J_{IS}t)$$
$$\xrightarrow{(180^\circ_x)_I} I_y\cos(\pi J_{IS}t) + 2I_xS_z\sin(\pi J_{IS}t)$$

i.e., in this case the S spin remains unaffected by the 180°_x pulse. The second delay then produces,

$$\xrightarrow{\pi J_{IS}t} I_y\cos^2(\pi J_{IS}t) - 2I_xS_z\cos(\pi J_{IS}t)\sin(\pi J_{IS}t)$$
$$+ 2I_xS_z\sin(\pi J_{IS}t)\cos(\pi J_{IS}t) + I_y\sin^2(\pi J_{IS}t)$$

Therefore, using the trigonometric identities noted above, this result can be reduced to,

$$\longrightarrow I_y$$

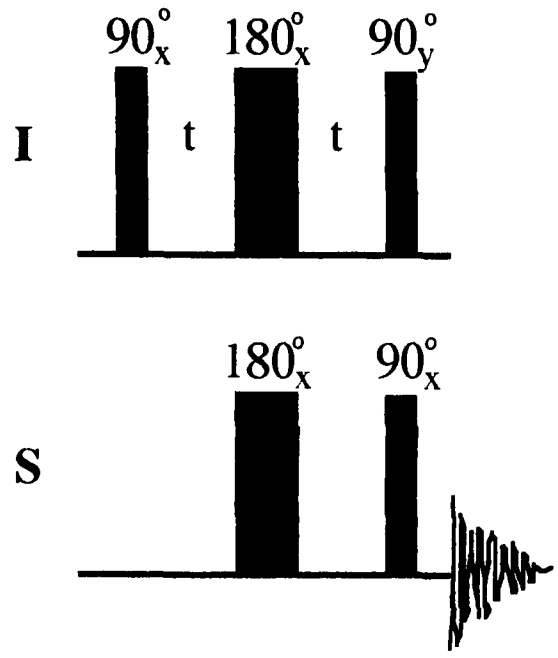
Consequently, for a heteronuclear two-spin system during a spin-echo pulse sequence, evolution under the scalar coupling, as well as the chemical shift, is refocused. Essentially, use of the spin-echo pulse sequence in this case has decoupled the S spin from the I spin, and this decoupling method is essential for many NMR experiments, including the INEPT pulse sequence which will now be described.

2.2.4.4 Product Operators for the INEPT Pulse Sequence

The "insensitive nuclei enhanced by polarisation transfer" (INEPT; Morris and Freeman, 1979) pulse sequence is essential for heteronuclear NMR experiments. The signal detected from heteronuclei used in protein NMR spectroscopy, e.g., ^{15}N or ^{13}C, is weak in comparison to signal detected from ^{1}H and therefore needs to be enhanced by the INEPT sequence. The INEPT sequence transfers magnetisation via a scalar coupling interaction from a sensitive nucleus with a high gyromagnetic ratio, e.g., ^{1}H, to a less sensitive nucleus with a lower gyromagnetic ratio, e.g., ^{15}N or ^{13}C, resulting in increased signal detected on the heteronucleus.

The INEPT pulse sequence, shown in Figure 2.11, is constructed from the spin-echo pulse sequence. In the previous section it was shown that, for a heteronuclear two-spin scalar-coupled system, the spin-echo sequence serves to refocus evolution under both the chemical shift and the scalar-coupling interaction. However, for the INEPT pulse sequence to work, evolution under the heteronuclear scalar coupling interaction is desired. For this to happen, the spin-echo pulse sequence must be applied to both the I and S spins simultaneously, as shown in Figure 2.11. During this spin-echo sequence the chemical shift is still refocused.

Figure 2.11 : The INEPT Pulse Sequence



For the INEPT product operator calculation, consider the IS scalar-coupled heteronuclear system, where I is the sensitive ^{1}H nucleus and S is the less sensitive ^{15}N nucleus. Starting with the equilibrium magnetisation $k_I I_z + k_S S_z$, where k_i is the intensity factor which is proportional to the gyromagnetic ratio of each spin, the evolution of this equilibrium magnetisation can be written as,

$$k_I I_z + k_S S_z \xrightarrow{90^\circ_x} -k_I I_y + k_S S_z$$

$$\xrightarrow{\pi J_{IS} t} k_I\{-I_y\cos(\pi J_{IS}t) + 2I_x S_z \sin(\pi J_{IS}t)\} + k_S S_z$$

$$\xrightarrow{(180^\circ_x)_I} k_I\{ I_y\cos(\pi J_{IS}t) + 2I_x S_z \sin(\pi J_{IS}t)\} + k_S S_z$$

$$\xrightarrow{(180^\circ_x)_S} k_I\{ I_y\cos(\pi J_{IS}t) - 2I_x S_z \sin(\pi J_{IS}t)\} - k_S S_z$$

$$\xrightarrow{\pi J_{IS} t} k_I\{ I_y\cos^2(\pi J_{IS}t) - 2I_x S_z \cos(\pi J_{IS}t)\sin(\pi J_{IS}t)$$
$$- 2I_x S_z \sin(\pi J_{IS}t)\cos(\pi J_{IS}t) - I_y \sin^2(\pi J_{IS}t)\} - k_S S_z$$

The latter result can be simplified using the trigonometric relationships previously noted,

$$\longrightarrow k_I\{I_y\cos(2\pi J_{IS}t) - 2I_x S_z \sin(2\pi J_{IS}t)\} - k_S S_z$$

The result of the final 90° x and y pulses is then,

$$\xrightarrow{(90^\circ_y)_I + (90^\circ_x)_S} k_I\{I_y\cos(2\pi J_{IS}t) - 2I_z S_y \sin(2\pi J_{IS}t)\} + k_S S_y$$

If t is set to $1/(4J_{IS})$, the above result simplifies to $k_I\{-2I_zS_y\} + k_SS_y$, where both terms correspond to signals that will be detected on the S spin. The term k_SS_y represents in-phase S spin y magnetisation, which has not been enhanced by magnetisation transfer. The intensity factor in this term is proportional to the lower gyromagnetic ratio of the S spin, i.e., of the ^{15}N nucleus in this example. However, the $k_I\{-2I_zS_y\}$ term represents anti-phase S spin magnetisation, which has the intensity factor of the I spin, i.e., of the ^{1}H nucleus in this example. This demonstrates that the INEPT pulse sequence can result in transfer of I spin magnetisation to the S spin via the heteronuclear scalar coupling interaction. This effect can be extremely significant. For example, if I is ^{1}H and S is ^{13}C, then the intensity of the anti-phase term is approximately four times greater than that of the in-phase term. For the ^{1}H and ^{15}N example above, the effect is even more pronounced, with the anti-phase term being approximately ten times more intense than the in-phase term.

2.2.5 Product Operator Calculations for NMR Experiments

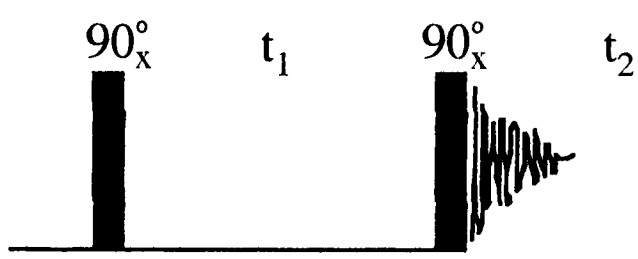
In the following four sections, product operators will be used to describe the evolution of magnetisation during 2D experiments used in this project. Two-dimensional COSY, TOCSY (HOHAHA), NOESY and HSQC spectra were recorded on cbEGF12-13 for use in spectral assignment and structure calculations, and 2D NOESY spectra were also used to monitor calcium binding to this domain pair and for studies of mutant cbEGF12-13 constructs.

2.2.5.1 COSY

Correlation spectroscopy (COSY; Aue *et al.*, 1976) is utilised to correlate J-coupled spins for use in spectral assignment. It allows visualisation of connectivities between two spin-$^{1}/_{2}$ nuclei, e.g., I and S, that are separated by three

bonds or less and have a scalar coupling of J_{IS}. The pulse sequence for the COSY experiment can be seen in Figure 2.12. It is the most basic 2D NMR experiment, and consists of two 90° pulses separated by the evolution time, t_1.

Figure 2.12 : The Pulse Sequence of a COSY Experiment



Consider a homonuclear two-spin IS system, i.e., where I and S are both protons. Evolution of magnetisation for both of these spins will be identical throughout the course of the experiment, so the product operator calculations below will only consider the I spin. To obtain the product operators for the evolution of S spin magnetisation, the I and S labels can be interchanged. Firstly, a $90°_x$ pulse is applied to the I and S spins, but considering the I spin only,

$$I_z \xrightarrow{(90°_x)_{IS}} -I_y$$

The I spin magnetisation then evolves during a period of free precession, time t_1, which must be considered in terms of both the influence of the resonance offset for the I spin, Ω_I, and in terms of the scalar coupling between the I and S spin, J_{IS},

$$-I_y \xrightarrow{\Omega_I t_1} -I_y\cos\Omega_I t_1 + I_x\sin\Omega_I t_1$$

$$\xrightarrow{\pi J_{IS} t_1} -I_y\cos\Omega_I t_1\cos(\pi J_{IS} t_1) + 2I_xS_z\cos\Omega_I t_1\sin(\pi J_{IS} t_1)$$
$$+I_x\sin\Omega_I t_1\cos(\pi J_{IS} t_1) + 2I_yS_z\sin\Omega_I t_1\sin(\pi J_{IS} t_1)$$

The two evolution effects above occur simultaneously during the free precession period and were only separated here to improve clarity. Finally, a $90°_x$ pulse is applied to the I and S spins,

$$\xrightarrow{(90^{o}_{x})_{IS}} -I_z\cos\Omega_I t_1\cos(\pi J_{IS}t_1) - 2I_xS_y\cos\Omega_I t_1\sin(\pi J_{IS}t_1)$$
$$+I_x\sin\Omega_I t_1\cos(\pi J_{IS}t_1) - 2I_zS_y\sin\Omega_I t_1\sin(\pi J_{IS}t_1)$$

In the final result above, the I_z and $2I_xS_y$ terms are not observable, as neither contains a single transverse operator. These terms can therefore be ignored. Of the latter two product operator terms, the I_x term represents I spin magnetisation that has evolved at the frequency Ω_I during the t_1 evolution period and will again evolve at the frequency Ω_I during the t_2 acquisition period. This produces a diagonal peak at Ω_I in t_1 and Ω_I in t_2 in the 2D COSY spectrum. The $2I_zS_y$ term, however, represents S spin magnetisation that has evolved at the frequency Ω_I during the t_1 evolution period, but will then evolve at the frequency Ω_S during the t_2 acquisition period. This produces a cross-peak at Ω_I in t_1 and Ω_S in t_2 in the 2D COSY spectrum (Figure 2.13a).

COSY peaks have a unique multiplet structure, which can be explained by trigonometric manipulation of the I_x and $2I_zS_y$ terms,

$$+I_x\sin\Omega_I t_1\cos(\pi J_{IS}t_1) \longrightarrow +{}^1/_2I_x[\sin(\Omega_I t_1 + \pi J_{IS}t_1) + \sin(\Omega_I t_1 - \pi J_{IS}t_1)]$$
$$-2I_zS_y\sin\Omega_I t_1\sin(\pi J_{IS}t_1) \longrightarrow -I_zS_y[\cos(\Omega_I t_1 - \pi J_{IS}t_1) - \cos(\Omega_I t_1 + \pi J_{IS}t_1)]$$

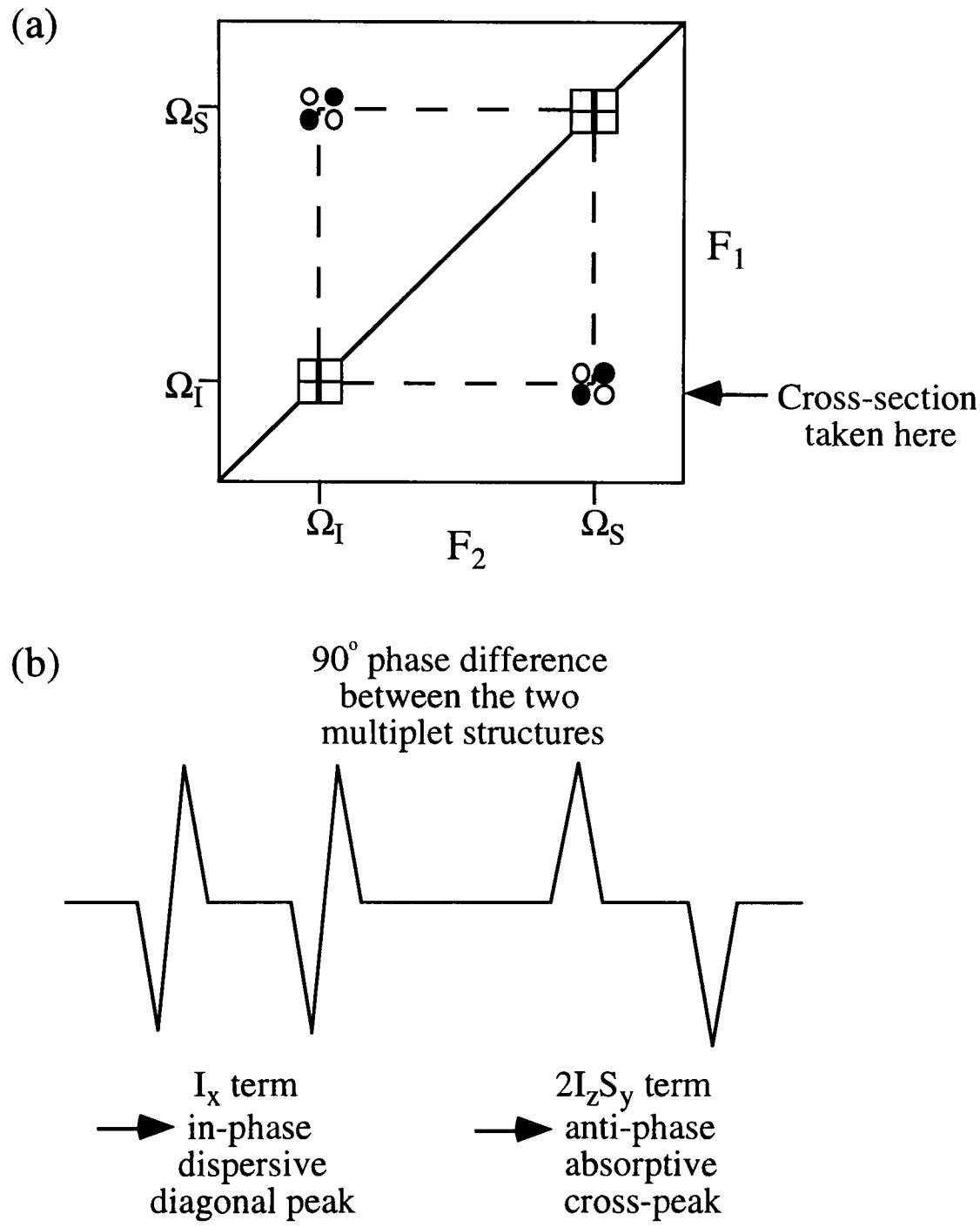
The first line indicates that the I spin diagonal peak centred at (Ω_I, Ω_I) is in-phase in F_1, as both sine modulations have the same sign, and that its two multiplet components are positioned at $\Omega_I \pm \pi J_{IS}$. The second term indicates that the cross-peak centred at (Ω_I, Ω_S) is anti-phase in F_1, as both cosine modulations have the opposite sign, and that its two multiplet components are also positioned at $\Omega_I \pm \pi J_{IS}$. Evolution of the I_x and $2I_zS_y$ terms during t_2 will be the same as during t_1, meaning that identical peak patterns will be seen in the F_1 and F_2 dimensions (Figure 2.13a).

In 2D COSY spectra there is a 90° phase difference between the diagonal and cross-peaks, meaning that when, for example, the cross-peaks are phased to be absorptive, the diagonal peaks will be dispersive (Figure 2.13). This 90° phase difference between the two multiplet peak structures results because the I_x term is cosine modulated for the scalar coupling interaction, whereas the $2I_zS_y$ term is sine modulated. Double quantum filtration (DQF) can be used in COSY experiments (DQF-COSY) to produce spectra which can be phased to produce both absorptive diagonal and cross-peaks (Rance *et al.*, 1983).

The most useful region of a 2D COSY spectrum is the fingerprint region, which shows correlations between $^1H^N$ and $^1H^\alpha$ protons (F_2 ~10.5ppm-6.5ppm and F_1 ~6.0-2.5ppm). The number of peaks in this region should correspond to the amino acid composition of the protein under study. Some amino acid spin systems can easily be identified in 2D COSY spectra, namely glycine, alanine, threonine and aromatic ring spin systems (see section 2.3.1.2). The two H^α protons coupled to the H^N proton in the glycine spin system produces a distinctive cross-peak composed of eight antiphase components. The H^α-H^β and H^β-H^γ cross-peaks of alanine and threonine, respectively, are easily distinguished from other H^α-H^β cross-peaks in the same spectral region based on their strong intensity and narrow shape in the F_2 dimension (Redfield, 1993). Complete assignment of these spin systems can subsequently be made using, for example, a 2D TOCSY experiment, which will now be described.

Figure 2.13 : Schematic Diagram to Show the Peak Appearance in a Two-Dimensional COSY Spectrum

In (a) a simplified schematic representation of a two-dimensional COSY spectrum is shown. White squares represent in-phase dispersive diagonal peaks, white circles represent the negatively phased part of the absorptive cross-peaks, and black circles represent the positively phased part of the absorptive cross-peak. Together the white and black circles form anti-phase absorptive cross-peaks. In (b) a F_2 cross-section is shown to highlight dispersive and absorptive peak appearance. This cross-section is taken at the point marked in (a). This diagram is modified from Wimperis *et al.* (1998).



2.2.5.2 TOCSY

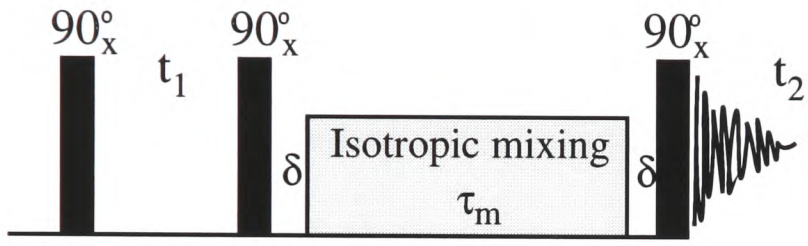
Total correlation spectroscopy (TOCSY), which is also known as homonuclear Hartmann-Hahn (HOHAHA) spectroscopy, is an extension of the COSY through-bond experiment. The TOCSY experiment is still a through-bond experiment,

but addition of an isotropic mixing period means it can correlate both directly connected spins, as well as more remotely connected spins, i.e., spins connected by a chain of spin-spin couplings. The TOCSY experiment can therefore be used to map out entire amino acid spin systems, making it an invaluable asset to aid spectral assignment.

The pulse sequence of a TOCSY experiment is shown in Figure 2.14.

Figure 2.14 : The Pulse Sequence of a TOCSY Experiment

The delays signified by the two deltas are long enough to allow the transmitter power to be altered, but are short enough to prevent NOE build-up, i.e., δ ~20μs.



For a homonuclear IS spin system, where the signal is detected from the I spin, the mixing period in the above scheme removes the chemical-shift difference between the I and S spins, i.e., the chemical-shift interaction approximates to zero. This is known as broad-band decoupling (Hartmann & Hahn, 1962; Braunschweiler & Ernst, 1983; Bax & Davis, 1985), and means that magnetisation only evolves under the strong scalar coupling interaction during this mixing period.

Qualitatively, the TOCSY pulse sequence starts by converting z magnetisation into transverse magnetisation using a $90°_x$ pulse. This magnetisation is then frequency-labelled during the t_1 evolution period, before being flipped back to the z axis for the isotropic mixing period. The length of this mixing period (τ_m) determines how far along a chain of coupled spins the magnetisation will be transferred, with scalar connectivities between protons five bonds apart routinely resolvable. The final $90°_x$ pulse converts the z magnetisation back into the transverse plane for detection.

As the TOCSY experiment results in transfer of magnetisation throughout a spin system, i.e., involving more than two spins, the product operator notation needs to be

adapted. Instead of using the IS notation, consider the I_1 magnetisation of a two-spin scalar-coupled system, I_1I_2, where I_1 and I_2 are both protons. The product operator calculations are initially the same as for the COSY experiment,

$$I_{1z} \xrightarrow{(90^\circ_x)_{12}\text{-}t_1\text{-}(90^\circ_x)_{12}} -I_{1z}\cos\Omega_1 t_1\cos(\pi J_{12}t_1) - 2I_{1x}I_{2y}\cos\Omega_1 t_1\sin(\pi J_{12}t_1) + I_{1x}\sin\Omega_1 t_1\cos(\pi J_{12}t_1) - 2I_{1z}I_{2y}\sin\Omega_1 t_1\sin(\pi J_{12}t_1)$$

Of these four terms, the I_{1x} and $2I_{1z}I_{2y}$ terms will be eliminated by phase cycling methods. The $2I_{1x}I_{2y}$ multiple-quantum coherence term will dephase during the isotropic mixing period due to the inhomogeneity of the rf field, and hence can also be ignored. Therefore, focusing only on the I_{1z} term, magnetisation proportional to I_{1z} is transferred throughout the spin system via strong scalar coupling interactions during the isotropic mixing period. This can be shown as,

$$-I_{1z}\cos\Omega_1 t_1\cos(\pi J_{12}t_1) \xrightarrow{\text{isotropic mixing } (\tau_m)} -\sum_{k=1}^{N} I_{kz}a_{1k}(\tau_m)\cos\Omega_1 t_1\cos(\pi J_{12}t_1)$$

where $a_{1k}(\tau_m)$ are mixing coefficients for magnetisation transfer through a spin system from spin I_1 to spin I_k. The final 90°_x pulse then follows to bring magnetisation back into the transverse plane for detection,

$$\xrightarrow{(90^\circ_x)_{12}} \sum_{k=1}^{N} I_{ky}a_{1k}(\tau_m)\cos\Omega_1 t_1\cos(\pi J_{12}t_1)$$

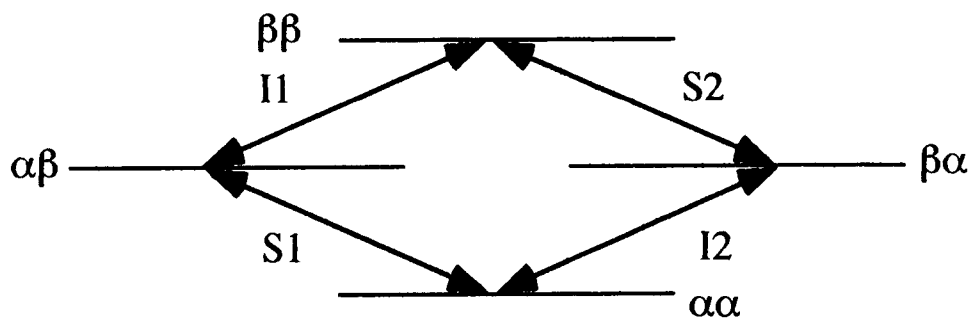
During acquisition (t_2), the above terms evolve under the effects of both the chemical-shift and scalar-coupling interactions. In the resultant 2D TOCSY spectrum, the diagonal peak is represented by the term proportional to I_{1y}, whereas the cross-peaks are represented by the term proportional to I_{ky} for $k > 1$. The TOCSY experiment produces cross-peaks that are in-phase, an aesthetic advantage of the TOCSY over the COSY experiment. Another advantage of the TOCSY experiment is that the cross-peaks

produced are less affected than COSY cross-peaks by increased linewidths resulting from proteins of high molecular mass.

2.2.5.3 NOESY

In contrast to the through-bond magnetisation transfer experiments previously described, nuclear Overhauser effect (NOE; Overhauser, 1953; Solomon, 1955; Jeener *et al.*, 1979; Macura *et al.*, 1981) spectroscopy (NOESY) exploits through-space, dipolar transfer of magnetisation. The discovery of the NOE was essential for the development of solution structure determination by NMR, since this effect allows correlation of protons which are ≤ 5Å apart in space. Distance restraints for structure calculations are derived from NOESY cross-peak intensities, since these intensities are inversely proportional to the distance between the corresponding nuclei. Nuclear Overhauser effect spectroscopy relies on the fact that cross-relaxation can occur between two spin-$^1/_2$ nuclei, and that the rate of this dipolar relaxation is dependent on both the rate of molecular reorientation, parameterised by τ_c (molecular correlation time), and the distance between the spins (r).

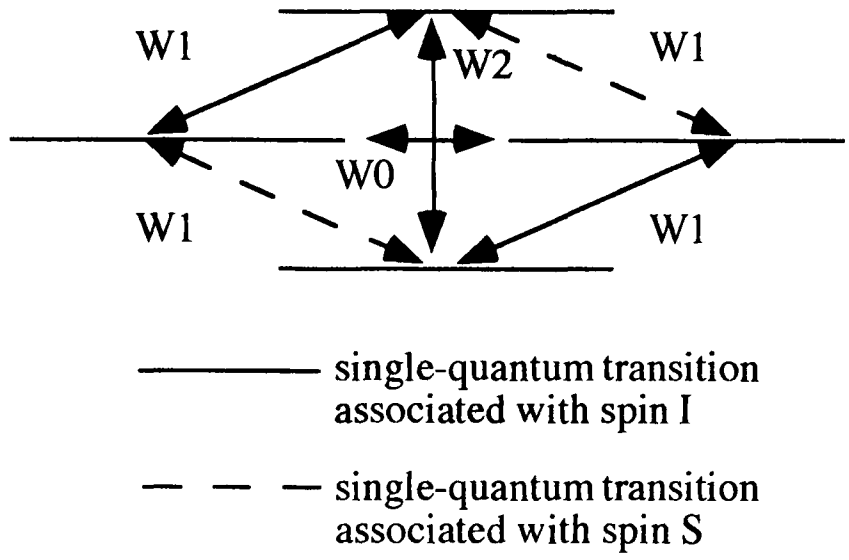
Consider an energy-level diagram for a two spin-$^1/_2$, homonuclear IS system,



IS spin system

where I1 and I2 represent I spin transitions, S1 and S2 represent S spin transitions, and αα, αβ, βα and ββ represent the four possible energy levels. The transition

probabilities for zero, single and double quantum coherence (W_0, W_1 and W_2), which are all a function of τ_c and $1/r^6$, can be shown as,



W_2 and W_0 in this diagram represent cross-relaxation. If this system is forced into a nonequilibrium state, i.e., the population difference across one spin is not in equilibrium, the system as a whole will act to restore equilibrium. The NOE is defined as the fractional change in the intensity of the NMR resonance of a dipolar coupled spin caused by cross-relaxation, when the populations of the energy levels are not at equilibrium. For homonuclear spins, the maximum observable NOE is given by,

$$\eta = \frac{W_2 - W_0}{2W_1 + W_2 + W_0}$$

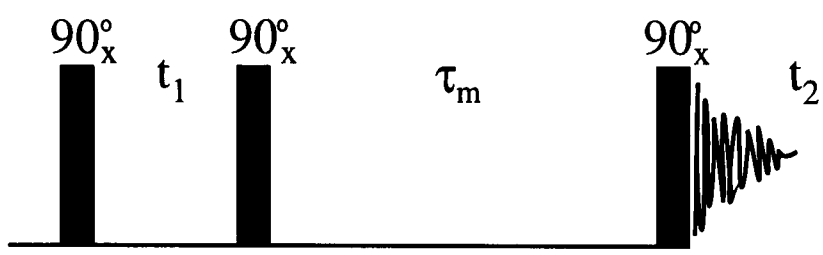
For proteins, which are relatively large with long correlation times, W_0 is greater than W_2 and W_1 (Derome, 1987). W_0 is therefore the predominant cross-relaxation mechanism for proteins.

More formally, the nuclear Overhauser effect is the time-dependent change in the intensities of resonance signals that are dipolar coupled, when the populations of the stationary states within this system are perturbed. The cross-relaxation constants for this process are inversely proportional to r^6 (Cavanagh *et al.*, 1996).

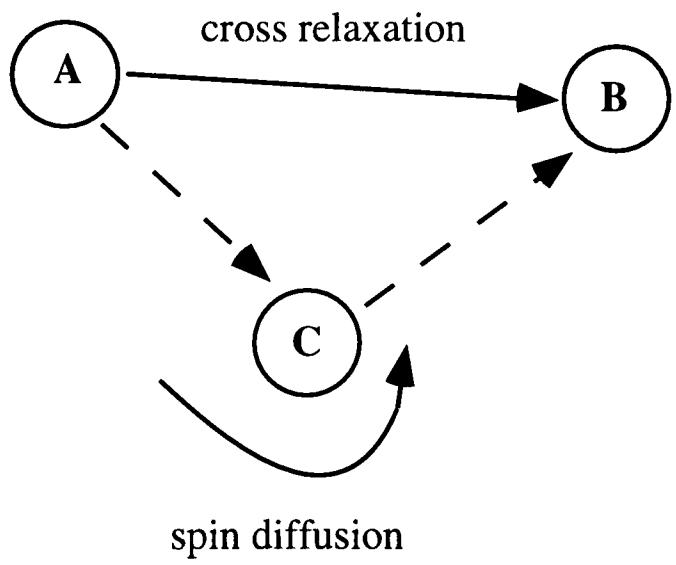
The pulse sequence of a NOESY experiment is shown in Figure 2.15.

Figure 2.15 : The Pulse Sequence of a NOESY Experiment

The term τ_m denotes the mixing time of the NOESY experiment.



Qualitatively considering a dipolar-coupled two-spin system, the NOESY pulse sequence starts with a $90°_x$-t_1-$90°_x$ sequence which serves to create a nonequilibrium population difference across the energy levels of the dipolar-coupled spins. This causes a NOE to be produced, and the mixing time, τ_m, then determines how long this NOE is allowed to build-up. The shorter the value of τ_m the more accurately the NOE will relate to interproton distance, but the longer the value of τ_m, the greater the signal-to-noise ratio and the greater the NOE intensity. Care must be taken, however, that the mixing time is not too long, as this can result in spin diffusion, i.e., relaxation that occurs along an undesired pathway (two or more steps) :



The τ_m value is generally set to a value of the order of $1/R_1$, where R_1 is the longitudinal relaxation time for the protein of interest. Spin diffusion does not cause a problem, and can even be advantageous when assigning NOESY spectra, but it can lead to incorporation of incorrect distance restraints into structure calculations.

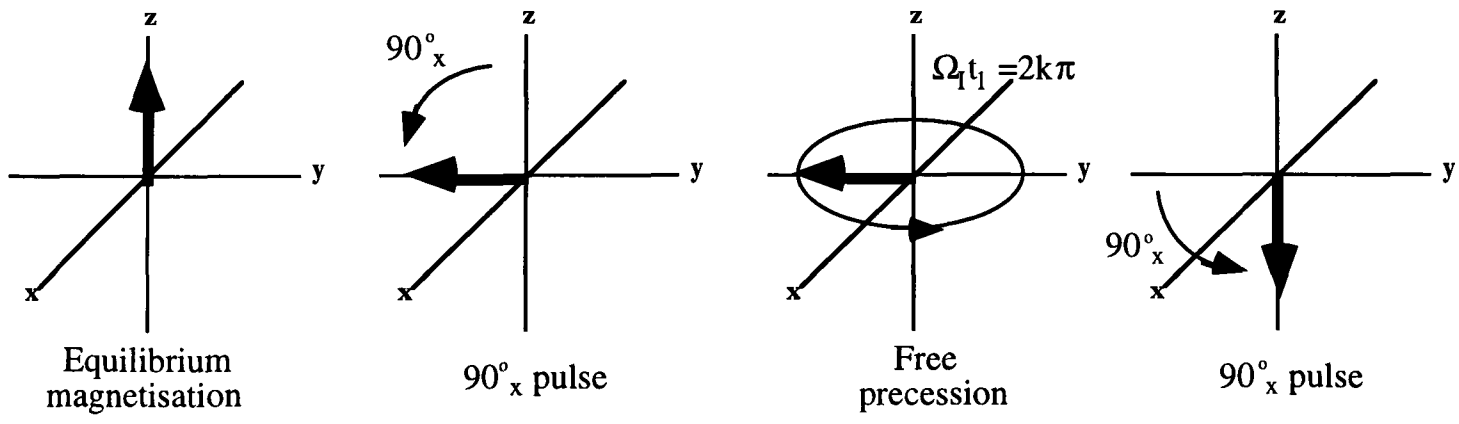
Consider a pair of dipolar-coupled spins, I and S, where I and S are not scalar coupled. Focusing only on the I spin, the product operator evolution for the first part of the NOESY experiment can be written as,

$$I_z \xrightarrow{(90^\circ_x)_{IS}} -I_y \xrightarrow{\Omega_I t_1} -I_y\cos\Omega_I t_1 + I_x\sin\Omega_I t_1 \xrightarrow{(90^\circ_x)_{IS}} -I_z\cos\Omega_I t_1 + I_x\sin\Omega_I t_1$$

Of the two terms produced after the second 90° pulse, the I_x term is not desired and can be suppressed by phase cycling or pulsed-field gradient techniques. The term of interest is the $-I_z\cos\Omega_I t_1$ term, which results from a perturbation of the I spin magnetisation from its original, equilibrium $+I_z$ state. The exact perturbation depends on the length of the t_1 evolution period. For example, whenever $\Omega_I t_1 = 2k\pi$ and k is an integer, the populations across the transitions of the I spins are inverted (Figure 2.16).

Figure 2.16 : Schematic Diagram to Show How I Spin Magnetisation Can Be Inverted During a NOESY Experiment

Inversion of I spin magnetisation occurs whenever k is an integer for $\Omega_I t_1 = 2k\pi$.



Therefore, during the mixing period, τ_m, a NOE will be induced on the S spin as a result of the dipolar coupling between the I and S spin. The cross-relaxation effects that occur during the mixing time can be written as,

$$-I_z\cos\Omega_I t_1 \xrightarrow{\text{mixing period } (\tau_m)} -I_z\cos\Omega_I t_1 a_{II}(\tau_m) - S_z\cos\Omega_I t_1 a_{IS}(\tau_m)$$

where $a_{II}(\tau_m)$ represents the proportion of the original magnetisation remaining on the I spin, and $a_{IS}(\tau_m)$ represents the proportion of the original magnetisation that is transferred from the I to S spin during τ_m by dipolar cross-relaxation. The final 90° pulse produces observable, transverse magnetisation for detection,

$$\xrightarrow{(90^\circ_x)_{IS}} I_y \cos\Omega_I t_1 a_{II}(\tau_m) + S_y \cos\Omega_I t_1 a_{IS}(\tau_m)$$

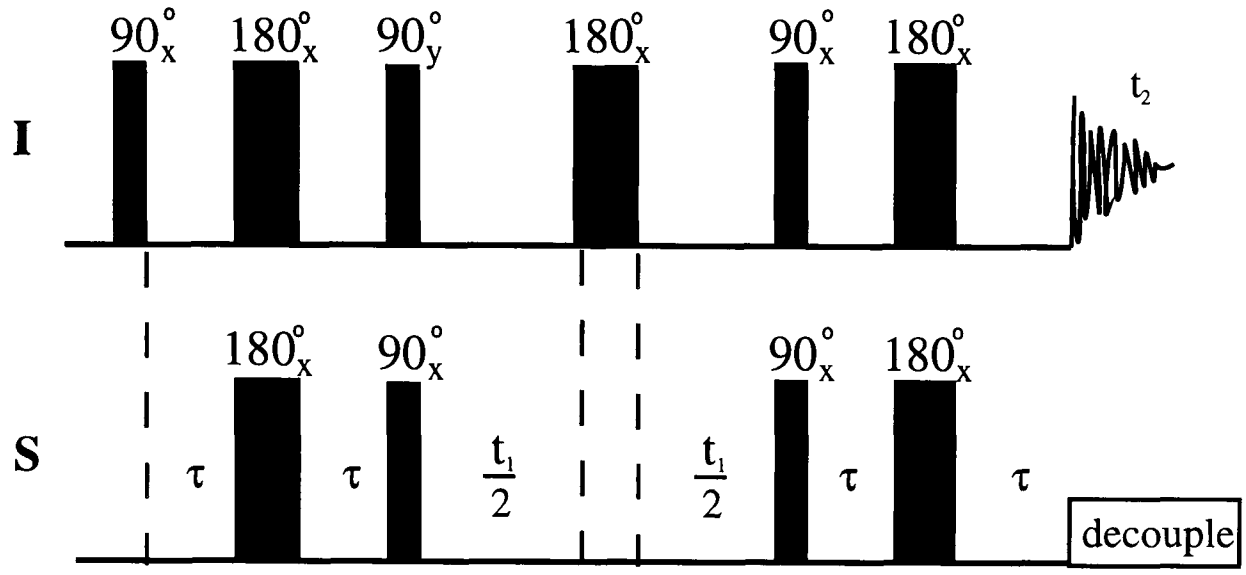
Fourier transformation of the above time-domain data yields a 2D frequency-domain spectrum with a cross-peak centred at Ω_I in F_1 and Ω_S in F_2. Any magnetisation remaining on the I spin will produce a diagonal peak centred at Ω_I in F_1 and Ω_I in F_2. Magnetisation can also be transferred from the S to I spin in an identical manner, leading to a diagonal peak at Ω_S in F_1 and Ω_S in F_2, and a cross-peak at Ω_S in F_1 and Ω_I in F_2. Both diagonal and cross-peaks are in-phase and can be phased to be absorptive in both dimensions.

2.2.5.4 HSQC

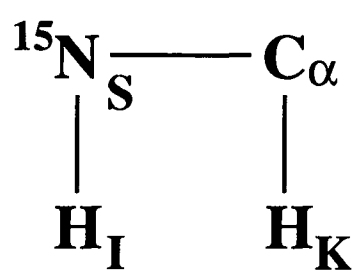
The three 2D NMR experiments described above are homonuclear experiments, and, since ^{15}N-labelling of the cbEGF12-13 pair was used in this project for dynamic and structural studies, the final part of this NMR section will describe a heteronuclear 2D experiment. This heteronuclear experiment may be combined with a 2D experiment as outlined in section 2.2.3 for isotopic editing. The basic pulse sequence for the heteronuclear single-quantum correlation (HSQC) experiment (Bodenhausen and Ruben, 1980) will be described using the product operator formalism. This pulse sequence, shown in Figure 2.17, is constructed from two INEPT pulse sequences (see section 2.2.4.4) separated by a t_1 evolution period.

Figure 2.17 : The Basic Pulse Sequence of a HSQC Experiment

The phase cycling scheme is not shown.



Consider an amide proton, I, attached to a ^{15}N-nucleus, S, where I is also scalar-coupled to an alpha-carbon proton, K,



The HSQC pulse sequence first uses INEPT to convert I_z magnetisation into the enhanced, anti-phase heteronuclear single-quantum coherence, $2I_zS_y$, seen at the end of section 2.2.4.4. This single-quantum coherence is subsequently frequency-labelled during the t_1 evolution period, before being converted back to observable proton magnetisation by a final INEPT sequence. In product operator terms, ignoring evolution under the homonuclear scalar-coupling interaction, J_{IK}, the preceding sequence can be written as,

$$I_z \xrightarrow{(90^\circ_x)_{IK}\text{-}\tau\text{-}(180^\circ_x)_{IK},(180^\circ_x)_S\text{-}\tau\text{-}(90^\circ_y)_{IK},(90^\circ_x)_S} -2I_zS_y$$

$$\xrightarrow{t_1/_2\text{-}(180^\circ_x)_{IK}\text{-}t_1/_2} 2I_zS_y\cos(\Omega_S t_1) - 2I_zS_x\sin\Omega_S t_1$$

$$\xrightarrow{(90^\circ_x)_{IK},(90^\circ_x)_S\text{-}\tau\text{-}(180^\circ_x)_{IK},(180^\circ_x)_S\text{-}\tau} -I_x\cos\Omega_S t_1 - 2I_yS_x\sin\Omega_S t_1$$

The delay (τ) is set to $1/(4J_{IS})$, and the $(180°_x)_I$ pulse in the centre of the t_1 evolution period serves to refocus evolution under the heteronuclear (^1H-^{15}N) scalar coupling interaction. The homonuclear scalar coupling interaction in this system, J_{IK}, evolves independently of the $2I_zS_y$ term during t_1, so the resultant F_1 lineshape after Fourier transformation does not contain contributions from the J_{IK} interaction. Of the two final terms obtained, the latter $2I_yS_x$ term represents multiple-quantum coherence, which is not observable. The former I_x term, however, is observable, and, after Fourier transformation, produces an in-phase, absorptive cross-peak in a 2D HSQC spectrum at Ω_I in F_1 and Ω_S in F_2.

In the basic HSQC experiment described above, only the first of the two single-quantum coherence terms obtained after the t_1 evolution period is converted back into observable magnetisation for detection. This experiment does not, therefore, reach its maximum potential in terms of sensitivity. This basic scheme has subsequently been adapted to enhance sensitivity, i.e., new schemes have been devised that transform both single-quantum coherence terms present after the t_1 period back to observable magnetisation for detection. Examples of such experiments include the sensitivity-enhanced HSQC (Palmer *et al.*, 1991) and the gradient-enhanced HSQC (Kay *et al.*, 1992), which have been used in the research described in this thesis, but will not be discussed further here.

2.3 Calculating Solution Structures from NMR Data

As previously stated, NMR is one of only two methods available for determining the three-dimensional structure of proteins at atomic resolution. After a complete set of NMR data for structure determination has been acquired and processed, the spectra are analysed until the majority of peaks have been identified. This assignment process requires the use of a number of different spectra (see below). Once

assignment is complete, NOE data and scalar coupling data can be extracted. These data are the principal NMR data currently used for solution structure determination. The NOE and scalar coupling data are converted into interproton distance restraints and dihedral angle restraints, respectively, and it is these restraints that form the basis of structure calculations.

In this section, the general approach for calculating a three-dimensional protein structure will be described. The description will begin with acquired data, and will progress step-by-step through to the refinement and evaluation of NMR solution structures.

2.3.1 Processing and Analysing NMR Data

2.3.1.1 Processing NMR Data

Once NMR data has been acquired, it is processed to allow the FID to be more conveniently visualised. As previously mentioned, Fourier transformation is the most basic processing technique, which is used to convert digitised time-domain signals into frequency-domain NMR spectra. However, this is not the only processing technique available and use of Fourier transformation alone will rarely produce the best possible NMR spectrum. Other processing methods exploited to optimise a spectrum's appearance include zero-filling, apodisation, linear prediction and phasing. Use of a combination of these processing methods can greatly improve a spectrum's signal-to-noise ratio and resolution.

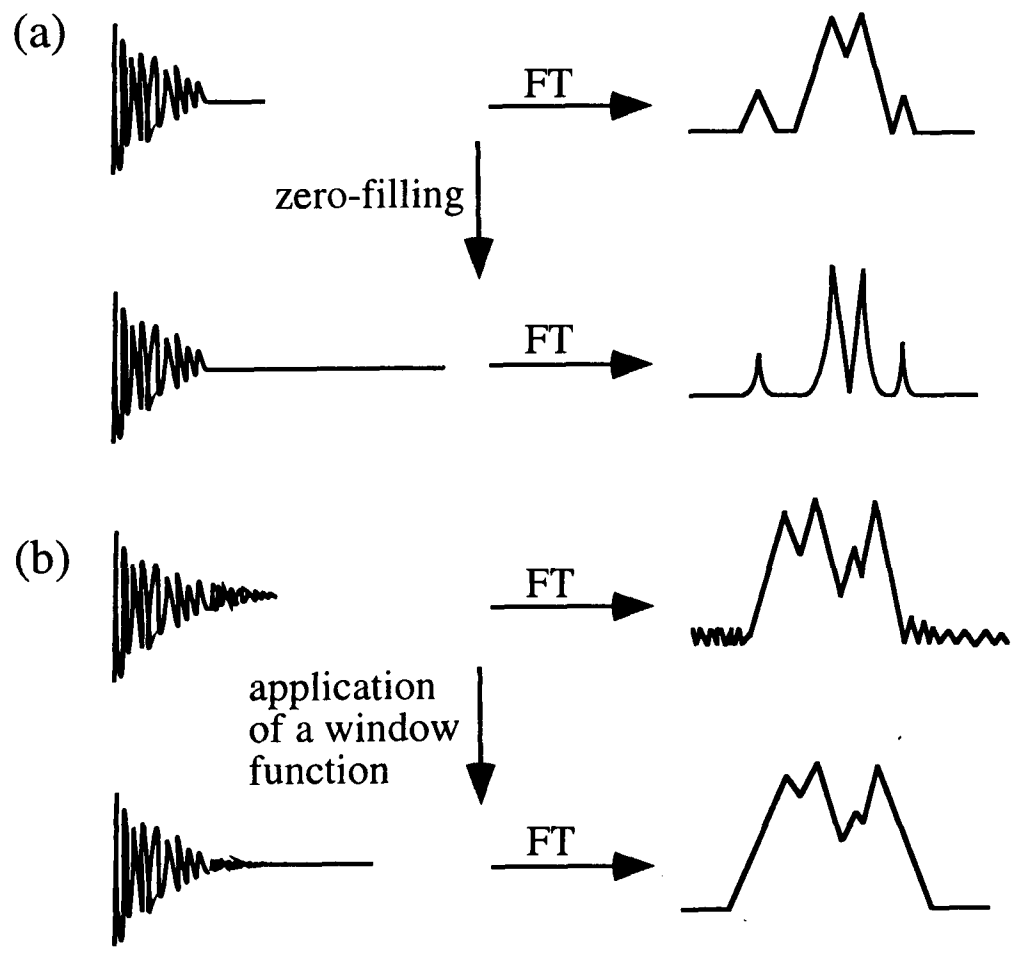
Zero-filling prior to Fourier transformation can greatly increase the digital resolution of a spectrum. This method extends the acquired data set by adding on fictitious data points that have zero intensity, and, generally, at least double the number of real data points are appended. If a FID decays to zero by the end of the acquisition period, zero-filling is equivalent to increasing the acquisition time. It interpolates

between data points and results in the attainment of spectra with apparently higher digital resolution (Figure 2.18a).

Apodisation is the application of a window function to a FID prior to Fourier transformation. This involves multiplying the FID by a function, and the function selected depends upon the desired effect. Window functions can be selected to decrease the effect of truncation artefacts, increase signal-to-noise and/or improve peak appearance of a spectrum. For example, as shown in Figure 2.18b, a function can be chosen to decrease the contribution of the tail end of the FID, where the signal-to-noise is lower, hence improving the overall signal-to-noise of the resultant spectrum.

Figure 2.18 : Schematic Representation of the Effect of (a) Zero-Filling and (b) Apodisation on Spectral Appearance

The diagram is schematically modified from Derome, 1987.



Linear prediction is a technique used to predict an extension to acquired NMR data prior to Fourier transformation (Barkhuijsen *et al.*, 1985). This method can reduce the effect of truncation artefacts and is particularly useful when processing multi-

dimensional, i.e., 3D and 4D, NMR datasets, because in these datasets the indirectly detected dimensions are usually seriously truncated due to time constraints.

The final processing method to be discussed here is phasing. NMR data are phased after Fourier transformation, and both zero-order and first-order empirical phase corrections can be applied to a NMR spectrum. These corrections are used to reduce baseline distortions and to optimise peak shape.

In summary, there are several different techniques to exploit when processing NMR data. Obtaining optimal spectra for a particular sample requires trial and error experimentation with the above methods. This experimentation is essential to maximise and optimise the spectral information available.

2.3.1.2 NMR Spectral Assignment

Assignment of NMR spectra is akin to puzzle solving, and the aim of the puzzle is to identify all the peaks in a NMR spectrum. NMR spectra of proteins contain three sets of information that are exploited for spectral assignment. These are through-bond information resulting from scalar coupling interactions (e.g., in COSY and TOCSY spectra), through-space information resulting from dipolar coupling interactions (e.g., in NOESY spectra), and information about chemical environment (obtained from the chemical shift position of each peak). Every cross-peak resonance in a protein NMR spectrum results from a connectivity between two atoms either within the same amino acid or between amino acids closely positioned in three-dimensional space. There are two, well-established techniques for the assignment of protein NMR spectra that take advantage of the latter condition. The first is the sequential assignment approach (Wüthrich, 1986; Redfield 1993), which is used to link amino acid spin systems together in the linear order of the protein sequence. The second is the main-chain directed approach (Englander & Wand, 1987), which links together amino acid spin systems in a cyclic fashion relating to secondary structural elements. Both of these

assignment methods were used to complete spectral assignment for the cbEGF12-13 domain pair, and these methods will be discussed further in sections 2.3.1.2.1 and 2.3.1.2.2.

Spectral assignment by both methods is achieved by comparing various NMR spectra in H_2O and ^2H_2O, and these spectral comparisons can be made using simple paper plots and/or computational methods, e.g., XEASY (Bartels *et al.*, 1995) or NMRView (Johnson & Blevins, 1994). Two different solvents are used to maximise the information obtained. In H_2O, signals are obtained from labile protons, i.e., most backbone amide protons, but the large signal obtained from H_2O itself can greatly obscure signals from protons at or around the water resonance. Also, signals for protons that exchange with H_2O can be significantly reduced by selective saturation of the H_2O resonance, a method sometimes used for solvent suppression. In ^2H_2O, signals from labile protons are lost, but the proton signals at or around the water resonance and in the aromatic region can be clearly visualised.

The set of spectra required for assignment depends on the size of the protein of interest. For small proteins, < 10kDa, spectral assignment can be accomplished using only homonuclear experiments recorded in H_2O and ^2H_2O, i.e., 2D COSY, TOCSY and NOESY spectra. Spectral overlap can be resolved by recording spectra at different temperatures and/or pH values, remembering that data from the same temperature and pH must ultimately be used for structure determination. In addition to 2D data, use of 3D heteronuclear-edited experiments, e.g., ^{15}N-edited TOCSY- and NOESY-HSQC spectra, can greatly reduce spectral overlap and simplify spectral assignment (Figure 2.7). Use of these 3D heteronuclear spectra pushes the molecular weight limit of structure determination by NMR up to ~20kDa. Triple resonance experiments (^1H, ^{15}N and ^{13}C) can also be used if a ^{15}N, ^{13}C-double-labelled sample is available. These experiments greatly facilitate spectral assignment, since, for example, inter-residue correlations can be directly observed. This increases the accuracy and speed with which

spectral assignment can be completed, and also further raises the molecular weight limit for protein structure determination by NMR.

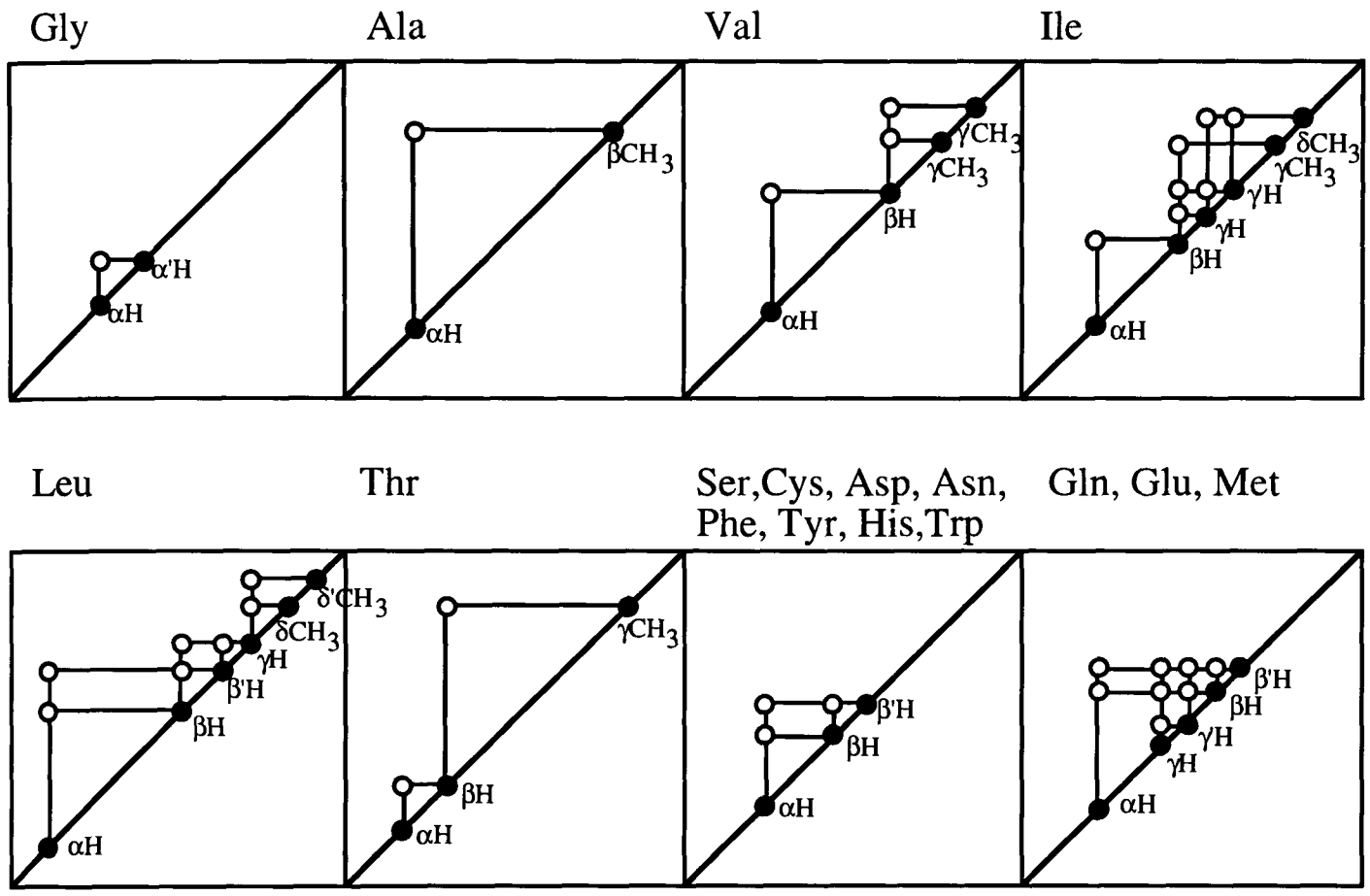
2.3.1.2.1 The Sequential Assignment Approach

In the first stage of sequential assignment, resonances belonging to the same amino acid spin system, i.e., intraresidue peaks, are identified. Spectra containing through-bond information, e.g., COSY and TOCSY spectra, are used to link up sets of HN, Hα and aliphatic side-chain peaks for every amino acid in the protein sequence. The twenty naturally occurring amino acids give rise to ten distinct spin system patterns, which are shown in Figure 2.19, plus four extra patterns for the aromatic ring protons. These patterns were determined based on random coil chemical shifts, and the exact random coil chemical shift values for these protons were determined in a short, denatured peptide (GGXA) (Wüthrich, 1986). These chemical shift values, combined with the characteristic peak patterns, are an invaluable reference when assigning intraresidue peaks. However, it must be noted that formation of higher order structure can result in significant differences between the actual and random coil chemical shift position.

In the second stage of sequential assignment, once the intraresidue peaks have been correlated, the separate amino acid spin systems are linked together linearly using both through-space, dipolar coupling data obtained from NOESY data and amino acid sequence data. This is most commonly accomplished by looking for intense NOEs between the HN, Hα and/or Hβ protons of one amino acid spin system and the HN of the next. As intense NOE peaks can also result from residues close to each other in three-dimensional space rather than in the protein sequence (see below), this is very much a trial and error process. However, as more spin systems are connected, the likelihood increases that a particular spin system sequence will match a particular region

Figure 2.19 : Schematic Representation of the Characteristic Spin System Patterns for the Twenty Amino Acids

Open circles indicate COSY cross-peaks, while closed circles represent diagonal peaks for the protons of each amino acid. The corresponding cross-peak displays observed in TOCSY and NOESY spectra depend upon both the geometry of the side-chain and the mixing time used. This diagram is modified from Wüthrich, 1986.



of the protein's primary amino acid sequence. This process of elimination can help resolve inconsistencies, and ultimately allows sequence-specific assignments to be made.

Once sequence-specific assignment is complete, chemical shift values for all the observable protons in each amino acid can be obtained from ^2H_2O and H_2O NOESY spectra. Additional information can also be obtained from COSY and TOCSY spectra in both solvents. These chemical shift values are essential to allow identification of the remaining through-space interactions, i.e., the inter-residue proton interactions, in the NOESY spectra.

To summarise, in the sequential assignment approach intraresidue peaks are first identified and linked up using through-bond spectral data. These assigned amino acid spin systems are subsequently linked together in a linear fashion using through-space spectral data combined with a knowledge of the amino acid sequence of the protein of interest (Wüthrich, 1986; Redfield 1993). Assignment is finally complete when the remaining NOE peaks are identified using knowledge of the chemical shift values for the protons in each assigned amino acid.

2.3.1.2.2 The Main-Chain Directed Approach and Secondary Structure Identification

The second method for spectral assignment is the main-chain directed approach (Englander & Wand, 1987), which uses the same spectra and sequence information as the sequential assignment process. However, rather than linearly connecting up spin systems in order of amino acid sequence, the main-chain directed approach is based on recognition of the NOE patterns of particular units of secondary structure. It is a method which connects up amino acids that are close to one another in three-dimensional space. In α-helices, the main-chain directed and sequential assignment approaches are identical, since the residues are close in both sequence and

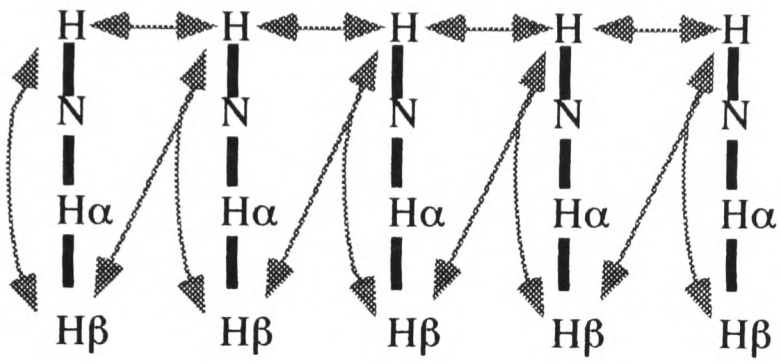
three-dimensional space. This is not the case, however, for β-sheet structures, where residues can be only 2.3Å apart in three-dimensional space, but can be far apart in the primary protein sequence (Figure 2.20). The main-chain directed method for spectral assignment is particularly useful when there is already some knowledge of the expected secondary structure, especially where β-sheet structures are predicted. It can prevent the misassignment of residues with strong NOE peaks between their NH groups as being sequential, when they are actually located opposite one another on separate β-strands.

In addition to its use in spectral assignment, knowledge of protein secondary structure greatly benefits structure determination. If the main-chain directed approach is not used for assignment purposes, secondary structure is usually identified once sequential assignment is complete and the $^{3}J_{HN\text{-}H\alpha}$ coupling constants have been measured (see section 2.3.2.2). This identification is done using the parameters shown in Table 2.2 (Barsukuv & Lian, 1993). For example, observation of strong $d_{NN(i,i+1)}$ NOEs, in combination with $^{3}J_{HN\text{-}H\alpha}$ values less than 4Hz, indicates a helical or turn region of the polypeptide chain. On the other hand, observation of strong $d_{\alpha N(i,i+1)}$ NOEs, with weak $d_{NN(i,i+1)}$ and $d_{\beta N(i,i+1)}$ NOEs and $^{3}J_{HN\text{-}H\alpha}$ values greater than 9Hz, indicates extended β-structures. Once secondary structure has been identified, peak intensity information from these regions can be reliably used to calibrate NOE cross-peak intensities into strong (≤ 2.8Å), medium (≤ 3.5Å) and weak (≤ 5Å) interproton distance categories for incorporation into structure calculations (see section 2.3.2.1.1). Hydrogen bonding data obtained from slow amide proton exchange experiments can also be incorporated into the structure calculation process if it is consistent with the identified secondary structure. This incorporation can only be performed when hydrogen bonding partners for these amide groups have been identified through structure calculations in the absence of these data (see section 2.3.2.3).

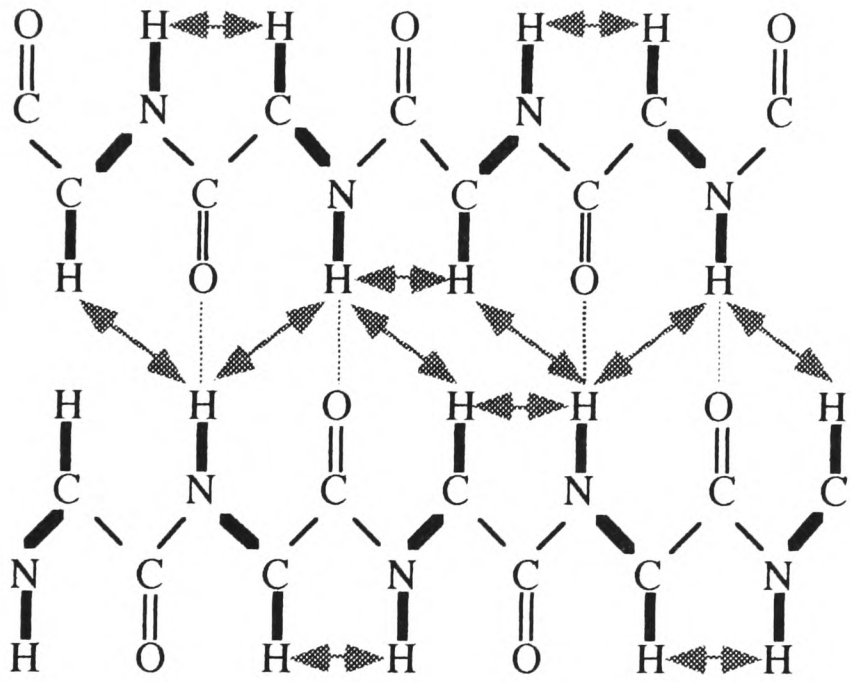
Figure 2.20 : Schematic Representation of the Proton-Proton Interactions Observed in Protein Secondary Structure

Some useful patterns for both main-chain directed assignments, as well as secondary structure identification in general, are highlighted. Protons are linked by both through-bond connectivities (shown as thick, black lines) and by short-range, through-space interactions (grey arrows). Hydrogen bonds are shown as dashed grey lines. This diagram is modified from Englander & Wand, 1987.

Helix



Anti-parallel β-sheet



Parallel β-sheet

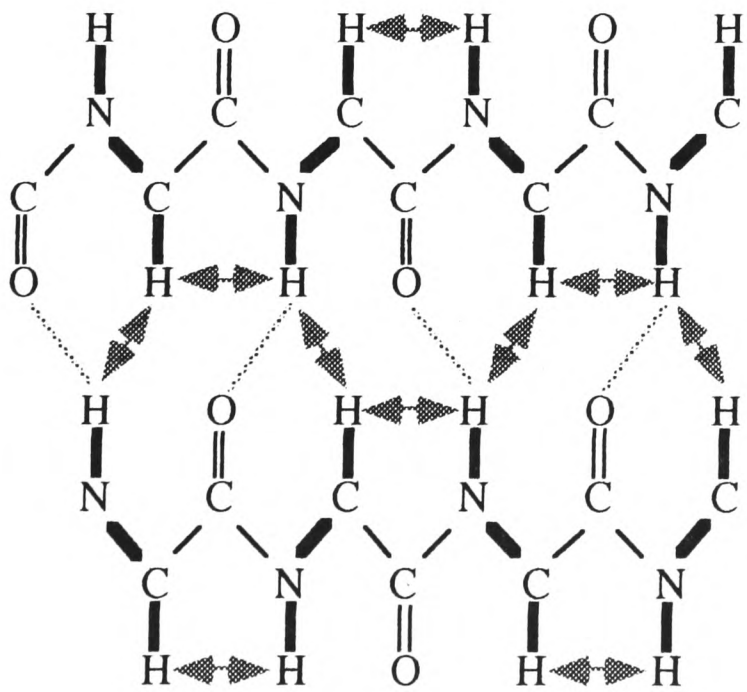


Table 2.2 : Characteristic Parameters Observed in Protein Secondary Structure

This table is reproduced from Barsukov & Lian, 1993.

Parameters	α-Helix	3_{10}-Helix	β-Antiparallel[b]	β-Parallel[b]
$d_{\alpha N}(i,i)$ (Å)	2.6	2.6	2.8	2.8
$d_{\alpha N}(i,i+1)$ (Å)	3.5	3.4	**2.2**[a]	**2.2**[a]
$d_{\alpha N}(i,i+2)$ (Å)	4.4	**3.8**[a]		
$d_{\alpha N}(i,i+3)$ (Å)	**3.4**[a]	3.3		
$d_{\alpha N}(i,i+4)$ (Å)	4.2	(> 4.5)		
$d_{NN}(i,i+1)$ (Å)	**2.8**[a]	**2.6**[a]	4.3	4.2
$d_{NN}(i,i+2)$ (Å)	4.2	4.1		
$d_{\beta N}(i,i+1)$ (Å)	2.5-4.1	2.9-4.4	3.2-4.5	3.7-4.7
$d_{\alpha\beta}(i,i+3)$ (Å)	2.5-4.4	3.1-5.1		
$d_{\alpha\alpha}(i,j)$ (Å)			**2.3**[a]	4.8
$d_{\alpha N}(i,j)$ (Å)			3.2	3.0
$d_{NN}(i,j)$ (Å)			**3.3**[a]	4.0
$^3J_{HN\text{-}H\alpha}$ (Hz)	(< 4)	(< 4)	(> 9)	(> 9)
NH exchange rate[c]	**slow**[a,d]	**slow**[a,d]	**slow**[a,e]	**slow**[a,e]

[a] The most readily available parameters are highlighted in bold. These highlighted parameters are often used to calibrate NOE cross-peak intensities into strong (≤ 2.8Å), medium (≤ 3.5Å) and weak (≤ 5Å) interproton distance categories (see section 2.3.2.1.1).

[b] $d_{\alpha\alpha}(i,j)$, $d_{\alpha N}(i,j)$ and $d_{NN}(i,j)$ refer to inter-strand distances.

[c] Amide protons still observable 24 hours after being dissolved in ^2H_2O are considered to be "slowly exchanging" (Barsukov & Lian, 1993).

[d] The first four residues in α-helices and the first three residues in 3_{10}-helices will generally have fast amide proton exchange rates.

[e] Every second residue in the flanking strand will generally have fast amide proton exchange rates.

2.3.2 From NMR Data to Structural Restraints

2.3.2.1 From NOE Data to Distance Restraints

2.3.2.1.1 NOE Intensity Calibration

As previously described in section 2.2.5.3, the intensities of NOE cross-peaks are inversely proportional to the distance between the corresponding nuclei. Interproton NOE cross-peaks are produced when dipolar cross-relaxation occurs between longitudinal magnetisations, and the cross-relaxation rate for this process is inversely proportional to r^6, where r is the distance between the two dipolar-coupled spins. If mixing times are kept short to prevent spin diffusion, NOE cross-peak intensities are proportional to the cross-relaxation rate constants, and NOE calibration can be performed using the equation,

$$r_{ij} = r_{ref}\left(\frac{V_{ref}}{V_{ij}}\right)^{1/6}$$

where r_{ij} and V_{ij} are the interproton distance and the NOE cross-peak intensity to be calibrated, respectively, and r_{ref} and V_{ref} are the corresponding reference values (Barsukov & Lian, 1993). Reference intensities and distances are usually obtained from the averages of well-resolved NOEs for interproton interactions in regions of regular secondary structure (Figure 2.20 and Table 2.2), although values can also be obtained from chemically fixed distances where the covalent geometry is known, e.g., the distance between geminal or aromatic ring protons (Barsukov & Lian, 1993). The former referencing is generally preferred, as it provides a greater range of interproton reference distances, which is particularly useful when calibrating NOEs into more generous categories (see below).

A problem arises, however, because of the possibility of spin diffusion. If the mixing time for a NOESY experiment is set longer than the optimum value for the

above equation, which is often done in practise to maximise signal-to-noise, the calibration method described can produce inaccurate results. If spin diffusion occurs, cross-peak intensities for pairs of closely positioned protons can be weakened by "leakage" of magnetisation to other protons. This leads to an overestimation of these interproton distances, i.e., the values calculated will be longer than actual distances between the protons. Correspondingly, the cross-peak intensities for pairs of protons which are far apart in space can increase as a result of spin diffusion, leading to an underestimation of the interproton distances (Clore & Gronenborn, 1985). This effect can be extremely problematic, as it can lead to incorporation of incorrect distance restraints into structure calculations. In addition, internal motions, frequently observed in macromolecules, can further complicate use of this calibration equation.

Consequently, instead of calculating an exact interproton distance for each NOE cross-peak intensity, the intensities are usually calibrated into three groups, each with an upper distance limit. These interproton distance categories are strong ($\leq$ 2.8Å), medium ($\leq$ 3.5Å) and weak ($\leq$ 5Å). The upper NOE intensity boundary for each distance category is calculated in the same way as the reference intensities and distances are calculated for the above equation. The lower distance limit is determined from the sum of the van der Waals' radii of the two interacting protons (Sutcliffe, 1993). However, this is not the case in structure calculations, where this lower distance limit is set to 0Å during simulated annealing (SA) (see below). This allows protons to move more freely and even pass through one another during the high temperature SA process, allowing them to search all areas of conformational space to find the global energy minimum. Towards the end of the structure calculation process, the size of each atom is gradually increased to its real value (see section 2.3.3).

For methyl groups, the $\langle r^{-6} \rangle$ averaging method described above is not appropriate for NOE intensity calibration, since the three protons of a methyl group are constantly in a state of fast rotation. $\langle r^{-3} \rangle$ averaging may therefore be used, since the results obtained by this method, rather than the $\langle r^{-6} \rangle$ method, are closer to the mean

distance (Tropp, 1980). This $\langle r^{-3} \rangle$ averaging is, however, incompatible with ambiguous distance restraints (see below) and so was not used in this project.

2.3.2.1.2 Overlapped Cross-Peaks and Ambiguous Distance Restraints

As the molecular weight of a protein increases, the possibility of cross-peak overlap in a NMR spectrum also increases. For proteins of ~20kDa, even with multi-dimensional NMR techniques, it is unlikely that all peaks in the resultant spectra will be unambiguously assigned. Even for the ~10kDa cbEGF12-13 domain pair studied in this project, a considerable number of peaks were overlapped due to sequence and structure repetition between the two domains. So how is peak overlap resolved in the structure calculation process?

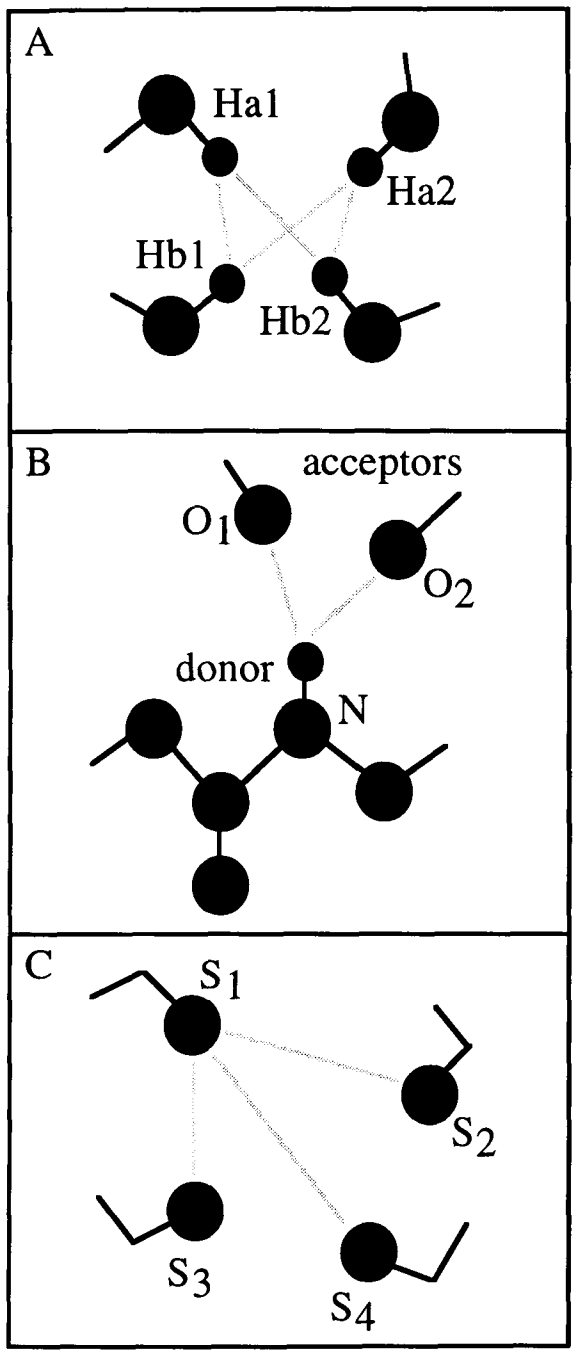
Use of ambiguous distance restraints is the most recent and most successful method to date that allows incorporation of the interproton distance information contained within overlapped, NOE cross-peaks into structure calculations (Nilges, 1995; 1997). These restraints are based on the fact that an overlapped, ambiguous NOE cross-peak at the chemical shift position (F_1, F_2) contains contributions from all the interproton interactions that have the same chemical shift assignment. The ambiguous NOE depends on the sum of the inverse sixth powers of the individual distances of the interproton interactions involved (Nilges, 1995). Sum-averaging is therefore used,

$$\overline{D} = \left(\sum_{n} D_{ij}^{-6} \right)^{-1/6}$$

where $\overline{D}$ is defined as the "r^{-6}-summed" average distance of all the interproton distances contributing a particular NOE, and is the distance calibrated from the overlapped NOE cross-peak intensity. As well as being applied to ambiguous interproton restraints resulting from NOE cross-peak overlap, e.g., this method can be used to deal with the equivalence of Tyr and Phe Hδ1/2 and Hε1/2, this method can also be successfully applied to hydrogen bond and disulphide bond ambiguous connections (Figure 2.21).

Figure 2.21 : Schematic Diagram to Show Examples of Ambiguous Distance Restraints

In A, an ambiguous NOE between two α protons and two β protons with degenerate chemical shifts leads to four contributions. In B, a protected amide hydrogen can have two possible oxygen acceptors. In C, a disulphide bond has an unknown connectivity. This diagram is modified from Nilges, 1997.



$$\bar{D} = (D^{-6}_{Ha1\text{-}Hb1} + D^{-6}_{Ha1\text{-}Hb2} + D^{-6}_{Ha2\text{-}Hb1} + D^{-6}_{Ha2\text{-}Hb2})^{-1/6}$$

$$(D^{-6}_{N\text{-}O1} + D^{-6}_{N\text{-}O2})^{-1/6} < 3\text{Å}$$

$$(D^{-6}_{S1\text{-}S2} + D^{-6}_{S1\text{-}S3} + D^{-6}_{S1\text{-}S4})^{-1/6} = 2.02\text{Å}$$

This latest method has several advantages over techniques previously employed to deal with overlapped peaks. In the above method, the average distance calculated is always less than the smallest, contributing, interproton distance, consistent with the intensity of the overlapped NOE cross-peak being the sum of all the contributing NOE intensities. In the $\langle r^{-6} \rangle^{-1/6}$ averaging method previously used to treat methyl and methylene groups (Brünger *et al.*, 1986), the calculated average distance was always between the smallest and the largest of the contributing interproton distances. Although these two methods are very similar, the difference between them is a scale factor of $n^{1/6}$,

where n is the number of interproton interactions contributing to the overlapped peak. This scale factor is not negligible, as several interproton distances can often contribute to a particular overlapped peak (Nilges, 1995). For example, consider three interproton distances contributing to a particular overlapped NOE peak, 2.8Å, 3.2Å and 3.5Å. The sum-average is 2.56Å, whereas the $<r^{-6}>^{-1/6}$ average is 3.07Å, a scale factor difference of $3^{1/6}$.

The sum-averaging method is also an improvement over use of pseudoatom corrections (Wüthrich *et al.*, 1983). In this method a pseudoatom is placed at the mean position of the atoms in an equivalent group when stereospecific assignments cannot be made. For example, this can be performed for the protons of a methylene groups or for the individual methyl groups in an isopropyl fragment (Wüthrich *et al.*, 1983). As the distance restraints are applied to the pseudoatom, rather than to an actual atom, the restraints are no longer directly related to the protein structure or to the original NOE data. The sum-averaging scheme, however, converts all cross-peak intensities to distance restraints without corrections, meaning the structure more closely relates to the experimentally-derived data.

2.3.2.1.3 Stereospecific Assignment and Floating Chirality

It has long been recognised that the accuracy and precision of an ensemble of NMR structures can be improved by knowledge of stereospecific assignments for both methylene protons and isopropyl groups (Driscoll *et al.*, 1989). Stereospecific assignment of prochiral β-methylene protons usually requires a P.E.COSY spectrum (Mueller, 1987), which is used to measure the $^3J_{\alpha\beta}$ scalar coupling constants. This coupling constant data, combined with analysis of the intensities of the $d_{\alpha\beta(i,i)}$ and $d_{N\beta(i,i)}$ intraresidue, NOE cross-peaks, can provide information about χ_1 side chain torsion angles (Driscoll *et al.*, 1989). A problem arises, however, because the

difference in cross-peak intensity expected between the HN-Hβ2 and HN-Hβ3 NOEs can become insignificant as a result of spin diffusion, and it is this difference that is fundamental in allowing stereospecific assignment. Problems can also arise when studying larger proteins, since increased linewidth and overlap in the P.E.COSY spectrum can hinder measurement of $^3J_{\alpha\beta}$ scalar coupling constants. Stereospecific assignment of the protons in methylene and isopropyl groups is therefore not always possible.

Previously, if these assignments were missing, the $\langle r^{-6} \rangle^{-1/6}$ averaging method or the pseudoatom method (see previous section) were employed (Brünger *et al.*, 1986; Wüthrich *et al.*, 1983). Neither of these methods were ideal, and for the latter the corresponding experimental distance restraints had to be widened to correct for the difference between the position of the pseudoatom and the position of the protons for which the NOE had been measured. This effect was significant, because the corrections used were generally based on worst-case geometries (Folmer *et al.*, 1997).

Subsequently, the floating chirality method was developed to treat non-stereospecifically assigned methylene or isopropyl groups (Folmer *et al.*, 1997). It is a simple and effective computational approach that requires no additional NMR data, and allows incorporation of these non-stereospecifically assigned groups into SA structure calculations (see section 2.3.3). Floating chirality assignment can be applied to all non-stereospecifically assigned methylene and isopropyl groups, as well as glycine alpha protons, that have non-degenerate chemical shift positions. The NOE peaks are first labelled according to their chemical shift positions, e.g., for glycine, Hα1 would simply be defined as the high-field resonance and Hα2 the low-field resonance. Then, during various stages of the SA structure calculation process, these symmetrically positioned protons or methyl groups are explicitly, but not simultaneously, swapped. The NOE energy term is evaluated for all possible conformations, and the conformation accepted is the one giving rise to the lowest energy term.

Floating chirality assignment is an invaluable tool in NMR structure calculations, since it allows distance restraints to be extracted directly from the measured peak intensity, without the need for averaging or pseudoatom correction. This means the structures calculated are closer to the NOE data, and are of a quality almost comparable to that obtained when stereospecific assignments are available (Folmer *et al.*, 1997).

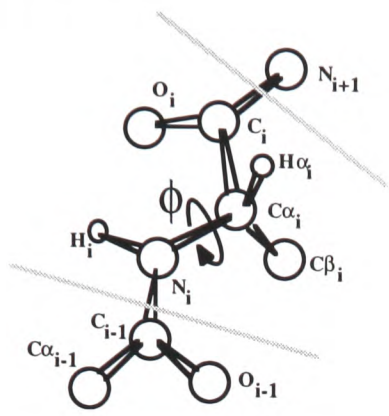
2.3.2.2 From HMQC-J Data to ϕ Dihedral Angle Restraints

Vicinal spin-spin coupling constants can provide useful additional geometric information which can be incorporated into structure calculations, and $^{3}J_{HN\text{-}H\alpha}$ scalar coupling constant information can be easily obtained from 2D ^{1}H, ^{15}N-heteronuclear multiple-quantum correlation (HMQC) spectra (Kay & Bax, 1990). This scalar coupling data can be converted into main chain ϕ dihedral angle information (Figure 2.22), which is the most readily available torsion angle information extensively used in structure calculations.

HMQC-J data are processed twice using two variants of a window function in the ^{15}N (F_1) dimension, e.g., 2 and 5Hz exponential line narrowing (Kay & Bax, 1990). J-coupling data are then extracted using non-linear least squares fitting to the lineshapes of the resonances arising from residues other than glycine. This method can only be applied where resonances show line splitting or significant line-broadening (Norwood *et al.*, 1992; Dr. P. C. Driscoll, unpublished results). The window function used in F_1 is selected to increase the resolution of the peak splittings by accentuating the information obtained from the tail end of the FID, that contains the "high frequency" components of the signal. Use of two different processing methods can resolve ambiguity, and can increase result confidence because the J-coupling constants obtained from both should agree.

Figure 2.22 : Schematic Diagram to Show the Position of the Backbone ϕ Dihedral Angle in a Polypeptide Chain

The limits of a residue are indicated by grey lines. This diagram is modified from IUPAC-IUB Commission on Biochemical Nomenclature, 1970.



$^3J_{HN\text{-}H\alpha}$ scalar coupling constants can be converted into ϕ dihedral angle values using the Karplus equation (Karplus, 1959),

$$^3J_{HN\text{-}H\alpha} = 6.4\cos^2(\phi - 60^\circ) - 1.4\cos(\phi - 60^\circ) + 1.9$$

which is shown schematically in Figure 2.23. The parameters most commonly used in the Karplus equation are those according to a study of bovine pancreatic trypsin inhibitor, where the $^3J_{HN\text{-}H\alpha}$ coupling constants in solution were correlated with the backbone ϕ dihedral angles measured from the crystal structure (Pardi *et al.*, 1984).

From Figure 2.23, it is clear that in many instances there is ambiguity as to which ϕ angle corresponds to a particular $^3J_{HN\text{-}H\alpha}$ coupling constant. This problem was overcome by an analysis of the range of ϕ values observed in known protein structures, which showed that, apart from Asn, Asp and Gly, all amino acids usually adopt negative ϕ values which concentrate in the -30° to -180° range (Richardson, 1981). Furthermore, within this range there are particular ϕ angle values, and consequently $^3J_{HN\text{-}H\alpha}$ values, that are expected for regions of regular secondary structure (Table 2.3).

Figure 2.23 : Schematic Representation of the Karplus Equation Showing the Relationship Between $^3J_{HN\text{-}H\alpha}$ Coupling Constants and Backbone ϕ Angles

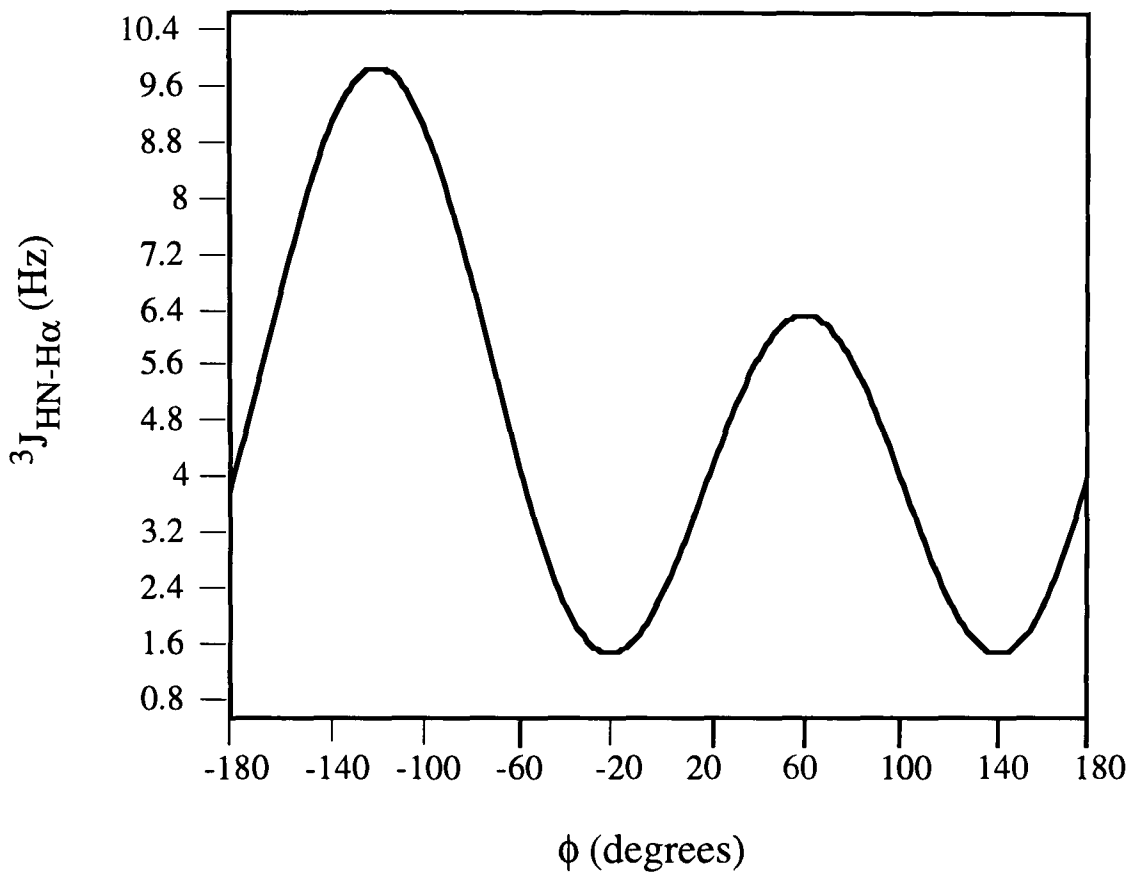


Table 2.3 : Approximate Backbone ϕ Dihedral Angle and $^3J_{HN\text{-}H\alpha}$ Values For Regions of Regular Secondary Structure

This table is modified from Sutcliffe, 1993.

Secondary Structure	ϕ (degrees)	$^3J_{HN\text{-}H\alpha}$ (Hz)
α-helix	-57[a]	3.9[c]
3_{10}-helix	-49[b]	3.0[c]
Antiparallel β-sheet	-139[a]	8.9[c]
Parallel β-sheet	-119[a]	9.7[c]

[a] From IUPAC-IUB Commission on Biochemical Nomenclature, 1970.
[b] From Ramachandran & Sasisekharan, 1968.
[c] Calculated using the Karplus equation, 1959.

From Table 2.3, it can be clearly seen that $^3J_{HN\text{-}H\alpha}$ values can be used to identify residues in the two main types of secondary structure. $^3J_{HN\text{-}H\alpha}$ values <5Hz indicate helical structures, although it is impossible to define the exact helix type with $^3J_{HN\text{-}H\alpha}$ information alone. Correspondingly, $^3J_{HN\text{-}H\alpha}$ values >8.5Hz indicate β-sheet structures, although it is impossible to define whether a sheet is parallel or anti-parallel with $^3J_{HN\text{-}H\alpha}$ information alone. These values are slightly relaxed when compared to those in Tables 2.2 and 2.3 to compensate for slight deviations from regular secondary structure and experimental inaccuracies. $^3J_{HN\text{-}H\alpha}$ values between 5-8.5Hz tend not to be used for the derivation of ϕ angle constraints, since extensive backbone motions produce average $^3J_{HN\text{-}H\alpha}$ values of ~7Hz for a range of different backbone conformations. However, values <5Hz are incorporated into structure calculations as the ϕ angle restraint -60° ± 30°, and values >8.5Hz are incorporated as -120° ± 30°. The inclusion of a ± 30° range compensates for motional averaging effects, that can systematically increase and decrease the magnitude of small and large $^3J_{HN\text{-}H\alpha}$ values, respectively. These ϕ angle restraints are, however, only incorporated into the structure calculation process if they are compatible with initial structures calculated in their absence.

2.3.2.3 From NH-Exchange Data to Hydrogen Bonding Restraints

If amide protons show signs of slow exchange, this usually indicates that these protected protons are involved in the formation of hydrogen bonds, and if main chain amide protons are involved in hydrogen bonding, this usually indicates that they are located in regions of secondary structure (Baker & Hubbard, 1984). In α-

helices and 3_{10}-helices, these hydrogen bonds are formed between $CO_{(i)}$-$HN_{(i+4)}$ and $CO_{(i)}$-$HN_{(i+3)}$, respectively. In β-sheet structures, there are a large number of CO-HN hydrogen bonds formed between adjacent strands (Figure 2.20). In fact, all amide protons in secondary structural elements are involved in hydrogen bonds except for the first four residues in an α-helix, the first three residues of a 3_{10}-helix, and every second residue in the flanking strands of β-sheets (Sutcliffe, 1993).

Amide proton exchange rates can be visualised by monitoring the change in NH cross-peak intensity as a function of time after dissolving a fully protonated protein sample in ^2H_2O. This can be performed by recording a series of COSY or TOCSY spectra starting as soon as possible after the protein was dissolved. Typical time periods used for these experiments are 4 hours, 8 hours, 24 hours, 48 hours up to several days, keeping the data acquisition time to a minimum for the initial data points. The availability of a ^1H, ^{15}N-labelled sample is a great advantage here, since, instead of the homonuclear experiments mentioned above, a series of HSQC spectra can be recorded. This decreases the acquisition time required for each experiment meaning that changes in the sample can be visualised more frequently, e.g., after every hour. The use of HSQC spectra also increases the resolution and clarity of the data obtained. Amide protons still observed in the resultant spectra 24 hours after the protein was dissolved in ^2H_2O are considered "slowly exchanging" (Barsukov & Lian, 1993), and it is these amide protons that are likely to be involved in hydrogen bonding interactions.

Recently, a more rapid experimental method has been developed to visualise the exchange rates of labile protons, which observes exchange on the millisecond time-scale (Böckmann & Guittet, 1997). It is a new pulse sequence which maintains the principle of determining amide exchange rates by diffusion, but has the important advantage of being much less time-consuming. It employs the use of a sensitive, highly selective and easy to implement water excitation scheme, and eliminates the laborious solvent exchange requirements of the previous method. As this method observes amide

protons that exchange with H_2O, those observed here are unlikely to be involved in hydrogen bond formation. In summary, this method is quicker and less laborious, and can visualise labile proton exchange on a different time-scale to the method described above.

As it is generally accepted that there is a correlation between "slow" amide proton exchange and the involvement of these protons in hydrogen bonds, the data obtained from the experiments above can be incorporated into structure calculations as hydrogen bonding restraints. A problem arises, however, since amide proton exchange can also be slowed due to proton inaccessibility to solvent. Also, the fact that a proton does not exchange on the millisecond time-scale only suggests that it could be involved in hydrogen bonding. Amide exchange data are therefore not discriminating enough to be used alone, and must be used in conjunction with NOE and $^3J_{HN\text{-}H\alpha}$ data. Consequently, hydrogen bonding data are only incorporated into structure calculations once they agree with the secondary structure observed in initial structures calculated in their absence. These initial structures are also beneficial for identification of hydrogen bonding partners, since NH exchange data give no indication as to which atoms are the hydrogen bond acceptors. Once the hydrogen bonds have been verified, they are defined as distance restraints of 1.8-2.3Å and 2.5-3.3Å for the H-O and N-O atom pairs, respectively (Barsukov & Lian, 1993; Sutcliffe, 1993).

2.3.3 From NMR Data to Structures

2.3.3.1 The Development of Structure Calculation Methods

Calculations used to determine NMR solution structures aim to search all areas of protein conformational space to find the global energy minimum that satisfies both a certain target energy function and a list of experimentally-derived restraints. In

this section, a brief overview of the development of methods for calculating three-dimensional protein structures from NMR data will be presented.

The first methods employed to calculate three-dimensional protein structures from interproton distance data were distance geometry (DG) techniques, and two DG procedures were proposed. The first operates in distance/Cartesian space and is frequently called "metric matrix distance geometry", because it converts the initial set of distance restraints into a set of Cartesian coordinates using the metric matrix (Crippen & Havel, 1978). The second, more recent, method operates in torsional space, and uses least-squares minimisation of a target function which comprises interproton distance estimates and weak van der Waals' repulsion terms (Braun & Gö, 1985).

Structures obtained using DG methods generally satisfy all the experimentally-derived distance restraints, but, when compared to available X-ray crystal structures, the DG-calculated structures appear slightly expanded with relatively large distortions in local backbone structure (Clore *et al.*, 1986). These distortions, together with close non-bonded contacts, result in large potential energy values, since these energy values are not taken into consideration in these calculations (Sutcliffe, 1993). Two hybrid methods were therefore developed, which combined DG methods with one of two further algorithms, i.e., restrained molecular dynamics (rMD; Brünger *et al.*, 1986; Clore *et al.*, 1986) or simulated annealing (SA; Nilges *et al.*, 1988a). The latter techniques were introduced to refine the structures obtained from DG. They were designed to eliminate unfavourable conformational distortions and to reduce potential energies, therefore bringing the DG-calculated structures closer to their global energy minimum.

The target function usually used in the rMD technique contains the potential energy terms,

$$\begin{aligned} F_{\text{tot}} &= (F_{\text{covalent}}) + (F_{\text{non-bonded}}) \\ &= (F_{\text{bond}} + F_{\text{angle}} + F_{\text{improper}}) + (F_{\text{vdW}} + F_{\text{electrostatic}} + F_{\text{H-bond}} + F_{\text{experimental}}) \end{aligned}$$

where F_{tot} represents the total potential energy of a system defined by the hybrid energy function. The first six terms in the second line represent the force field usually used in MD, and the extra $F_{experimental}$ term is incorporated to define the potential energy relating to the experimentally-determined restraints (Brünger *et al.*, 1986; Clore *et al.*, 1986). Theoretically, rMD has the potential to allow sampling of all areas of protein conformational space, however, due to the computational expense of this process, this is rarely the case.

In the SA technique, the target function is simplified, and the three F_{vdW}, $F_{electrostatic}$ and $F_{H\text{-}bond}$ terms in the rMD target function equation above are replaced by a simple van der Waals' repulsive force (see section 2.3.3.3.1). This means that SA requires considerably less computational time than rMD. SA involves heating the system to a high temperature followed by slow cooling, which allows DG-calculated structures to overcome potential energy barriers and maximises conformational space sampling on the pathway to the global energy minimum (Nilges *et al.*, 1988a; 1991). This SA process is therefore generally preferred to the rMD method.

After the development of the DG-rMD and DG-SA hybrid calculation methods, the great potential of the rMD and SA methods was realised and these methods were further developed into independent structure calculations. Structure calculations using rMD are, however, less popular when compared to SA, because rMD techniques use considerably more CPU time. SA has therefore developed into the major method for the determination of three-dimensional protein structures from NMR data (Nilges *et al.*, 1988b).

SA is computationally efficient, it thoroughly samples conformational space, and it reliably locates the global energy minimum of the target function (Nilges *et al.*, 1991). SA structure calculations were therefore used in this project for structural studies of the cbEGF12-13 domain pair. SA can be performed in both Cartesian space (Brünger, 1992) and torsion angle space (Güntert *et al.*, 1997), with the latter requiring fewer independent variables as it has less degrees of freedom. Covalent structure

parameters (bond lengths and angles, planarities and chiralities) are kept fixed at their optimal values during SA in torsion angle space, so the strong potentials required to maintain covalent structure in Cartesian space SA are absent, as are the concomitant high-frequency bond and angle vibrations. This means longer time-steps and/or higher temperatures can be used in torsion angle space SA (Güntert *et al.*, 1997). Torsion angle space is therefore often the method of choice since it is more computationally efficient. Cartesian space, however, is more amenable to the incorporation of the ambiguous distance restraints described in section 2.3.2.1.2, so the SA structure calculations in this project were performed in Cartesian space. Structure calculations using SA, and SA in general, will now be described in more detail.

2.3.3.2 Dynamical Simulated Annealing

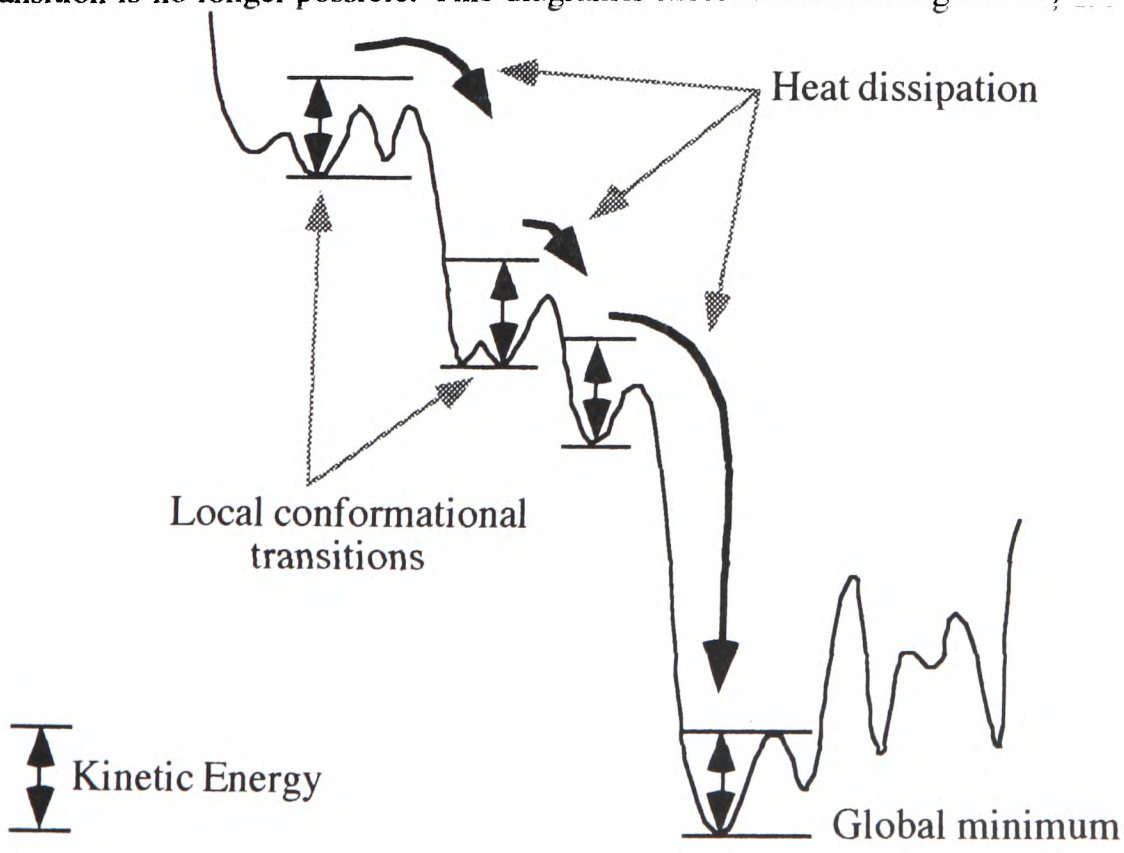
Simulated annealing was first introduced by Kirkpatrick *et al.* (1983) as a method to simulate the physical process of annealing using the Monte Carlo algorithm (Metropolis *et al.*, 1953), where a solid is heated until all the atoms move randomly and then cooled so that these particles settle in low energy states. The goal and strength of SA in structure calculations is that it searches for the conformation(s) of the molecule that best fits the experimental data whilst also maintaining reasonable covalent and non-covalent interactions (Brünger *et al.*, 1997). This process, which has been proven successful (Nilges *et al.*, 1991), is dominated by the need the satisfy the experimentally-derived restraints, and it is based on the assumption that these restraints accurately determine the fold of the molecule.

As well as being computationally efficient, SA structure calculations and refinement procedures produce structures with an excellent convergence rate and have a much larger radius of convergence than previous methods (Brünger *et al.*, 1997). This is because SA can find a lower energy minimum of the target function than is found simply by, for example, gradient descent optimisation techniques (Press *et al.*, 1986).

An advantage of SA is that it requires no initial structures, it simply starts with an arbitrary extended template with ideal stereochemistry and topology, and it has the ability to overcome barriers in the target energy function (Figure 2.24; Brünger *et al.*, 1997). During each phase in the SA procedure a fixed temperature is used, but temperature fluctuations arise from local conformational changes which tend to lower the target energy value. This leads to an increase in kinetic energy, and hence the temperature, of the system. Once the temperature coupling has dissipated the excess kinetic energy as heat, the reverse transition is unlikely to occur. This temperature coupling effect therefore maintains enough kinetic energy for local conformation corrections, but does not provide enough energy to allow escape from the global energy minimum once it has been reached (Brünger *et al.*, 1997). Therefore, during the high temperature phase of SA, many highly unstable conformations with large potential energy values can be adopted, and it is this thorough sampling of conformational space that ensures that the most stable state will be reached at the end of the SA process.

Figure 2.24 : Schematic Representation of the Search for the Global Energy Minimum During Simulated Annealing

Kinetic energy can allow local conformational transitions where the barriers are smaller than the kinetic energy, but if a larger energy drop is encountered, excess kinetic energy is dissipated as heat and the reverse transition is no longer possible. This diagram is modified from Brünger *et al.*, 1997.



2.3.3.3 Structure Calculations Using SA

XPLOR (Brünger, 1992) and DYANA (Güntert *et al.*, 1997) are the two most commonly used computational method for implementing structure calculations using SA in combination with NMR data. XPLOR works in Cartesian space, whilst DYANA works in torsion angle space, so the cbEGF12-13 domain pair structures calculated in this project (see Chapter 5) were calculated using XPLOR-version 3.851, since DYANA does not implement the use of ambiguous distance restraints. The algorithm employed by XPLOR will now be described in detail.

2.3.3.3.1 The Target Function

The XPLOR target function can be written as,

$$F_{tot} = (F_{covalent}) + (F_{non\text{-}bonded})$$
$$= (F_{bond} + F_{angle} + F_{improper}) + (F_{repel} + F_{NOE} + F_{dihedral})$$

where F_{tot} represents the total potential energy of all the atoms in a system at a certain temperature during a dynamic process. F_{tot} is determined by solving Newton's equations of motion for all n atoms of the system,

$$\frac{\delta^2 x_i}{\delta t^2} = -\frac{1}{m_i}\frac{\delta}{\delta x_i} F_{tot}(x_1, x_2, \ldots\ldots x_n)$$

where x_i is the movement of atom i, m_i is the atomic mass of atom i, and t is the duration of the motion (Nilges *et al.*, 1988a; 1988b).

The $F_{covalent}$ term in the target function above maintains correct bond lengths, angles, chirality and planes, and is given by,

$$F_{covalent} = \sum_{bonds} k_b(r - r_o)2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{impropers} k_\omega(\omega - \omega_0)^2$$

The force constants of the energy terms for bonds (k_b), angles (k_θ) and improper torsion angles (k_ω) are set to uniform high values to ensure that each amino acid has

near perfect stereochemistry throughout the structure calculation. The improper torsion term maintains appropriate planarity and chirality, and peptide bonds are assumed to be planar and *trans* (Nilges *et al.*, 1988a). If *cis*-proline residues are present, their conformation must be manually edited (see Appendix I - generate_fib1213.inp). In XPLOR, the standard values for bond lengths, bond angles, correct chirality and planarity of aromatic rings and peptide bonds, as defined by known covalent protein structures, are specifically defined in two files, i.e., the parameter file (parallhdg.pro) and the topology file (topallhdg.pro) (see Appendix I).

As mentioned in section 2.3.3.1, in SA structure calculations three of the non-bonded energy terms of the rMD hybrid energy function, i.e., F_{vdW}, $F_{electrostatic}$ and $F_{H\text{-}bond}$ terms, are replaced by a simple van der Waals' repulsive force, F_{repel}, which significantly reduces the computational time required for these calculations. The non-bonded interactions are described by this single repulsive term to prevent forbidden close non-bonded contacts (Nilges *et al.*, 1988a). F_{repel} is given as,

$$F_{repel} = \begin{cases} 0 & , \text{ if } r \geq r_{min} \\ k_{rep}\ (s^2 \cdot r_{min}^2 - r^2)^2 & , \text{ if } r < r_{min} \end{cases}$$

where the values of r_{min} are the standard values for van der Waals' radii represented by the Lennard-Jones potential used in the CHARMM empirical energy function (Brooks *et al.*, 1983). s is a van der Waals' radius scale factor and k_{rep} is the van der Waals' repulsive force constant (Nilges *et al.*, 1988a; 1988b).

The final two energy terms, that combine with F_{repel} to make up the non-bonded contribution to the total potential energy of the system, are the terms relating to the experimentally-derived restraints. The potential energies of the NOE-derived distance restraints and dihedral angle restraints, F_{NOE} and $F_{dihedral}$, respectively, are represented as square-well potential functions.

F_{NOE} can be written as,

$$F_{\text{NOE}} = \begin{cases} k_{NOE}(d_{ij} - d_{ij}^{u})^2, & \text{if } d_{ij}^{u} < d_{ij} \le d_{ij}^{u} + d_{sw} \\ 0, & \text{if } d_{ij}^{l} \le d_{ij} \le d_{ij}^{u} \\ k_{NOE}(d_{ij} - d_{ij}^{l})^2, & \text{if } d_{ij} < d_{ij}^{l} \end{cases}$$

where k_{NOE} is a variable NOE force constant, d_{ij} is the NOE-derived distance between atoms i and j, and d_{ij}^{u} and d_{ij}^{l} are the upper and lower limits of the target distance restraints, respectively (Nilges *et al.*, 1988a; 1988b). An extra term, d_{sw}, is also introduced to allow small and large NOE violations to be distinguished. The value of d_{sw} is set arbitrarily to 1.0Å. When $d_{ij} > d_{ij}^{u} + d_{sw}$, the F_{NOE} term is changed to a softsquare potential, so it becomes linear for large NOE violations,

$$F_{\text{NOE}} = a + \frac{b}{d_{ij} - d_{ij}^{u}} + c(d_{ij} - d_{ij}^{u})$$

The values of a and b are selected so that F_{NOE} gives the same value for $d_{ij} > d_{ij}^{u} + d_{sw}$ for both the square-well and softsquare potentials, and c is the asymptote which is set to a constant value during structure calculations.

F_{dihedral} can be written as,

$$F_{\text{dihedral}} = \begin{cases} k_{dihedral}(\phi - \phi^{u})^2, & \text{if } \phi > \phi^{u} \\ 0, & \text{if } \phi^{l} \le \phi \le \phi^{u} \\ k_{dihedral}(\phi - \phi^{l})^2, & \text{if } \phi < \phi^{1} \end{cases}$$

where ϕ is the dihedral angle value obtained from $^{3}J_{\text{HN-H}\alpha}$ data (see section 2.3.2.2), ϕ^{u} and ϕ^{l} are the upper and lower limits of the dihedral angle restraints, respectively, and k_{dihedral} is the dihedral angle force constant.

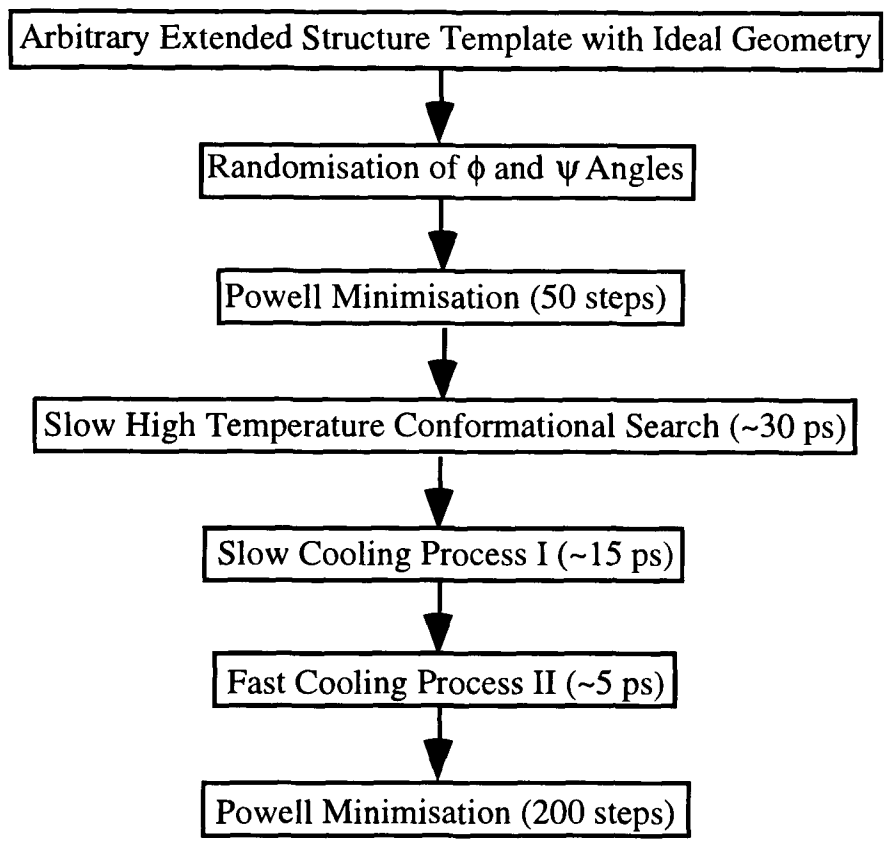
2.3.3.3.2 SA Structure Calculations and Refinement Protocols

The outline of the SA structure calculation method is shown schematically in Figure 2.25, and the computational protocols required for structure calculations and refinement, i.e., sa.inp and refine.inp, respectively, are shown in Appendix I. These protocols were written by M. Nilges (EMBL, Germany), and have

been modified from those originally given in the XPLOR manual (version 3.1; Brünger, 1992), to allow the incorporation of both ambiguous distance restraints (see section 2.3.2.1.2) and floating chirality (see section 2.3.2.1.3). Floating chirality assignment is carried out before each dynamics/minimisation procedure for initial structures. The SA structure calculation procedure will now be described in detail.

Figure 2.25 : Schematic Outline of the Simulated Annealing Structure Calculation Protocol

This figure was provided courtesy of Dr. X. M. Yuan.



SA structure calculations start with an extended template structure with ideal geometry, where the ϕ and ψ angles of this template structure are subsequently randomised to ensure thorough sampling of conformational space (Nilges *et al.*, 1988b). The SA structure calculation process then begins with 50 steps of Powell (1977) minimisation, to eliminate any unfavourable, close non-bonded contacts caused by the ϕ and ψ angle randomisation (Nilges *et al.*, 1988b).

The next stage involves a slow, high temperature conformational search (~30ps), where the NOE, dihedral and covalent force constants are increased. During this process, k_{NOE} is the most dominant of these forces and becomes the main driving force in folding the extended template. The value of the van der Waals' repulsive force constant, k_{rep}, is kept low during this stage (only 0.2% of k_{NOE}), which makes it sufficient to define and maintain global structure, but allows atoms to move freely, get very close to one another, and even pass through one another to improve local structure. The atoms are only weakly coupled in this phase and can essentially move independently to satisfy the restraints, therefore avoiding problems associated with protein folding. Although this method is not the "natural" method of protein folding, it has been proven successful. As all the force constant values are relatively low during this high temperature conformational search phase, energy barriers between different folds of the protein can be easily overcome and the global energy minimum of the target function is reliably located (Figure 2.24). At the end of this phase, it is likely that the resultant conformation will be similar to the fold seen in the final structures (Nilges *et al.*, 1988b; 1991).

The SA structure calculation process then has a slow cooling phase (~15ps). During this phase the k_{rep} terms is gradually increased from 0.003kcalmol^{-1}.Å^{-4} to its final value of 4kcalmol^{-1}.Å^{-4}, the $k_{dihedral}$ value is set to its final high value, and the van der Waals' radius scale factor is decreased from 0.9 to a final value of 0.75. This restores the hardness of the atoms and also improves the local geometry of the calculated structures (Nilges *et al.*, 1988b).

A fast cooling process (~5ps) is then performed with all parameters set to their final values. This is done purely to increase the likelihood of reaching the global energy minimum. This is finally followed by 200 steps of Powell minimisation (1977), to remove any further unfavourable, close non-bonded contacts. The latter may have been introduced during the high temperature phase where the k_{NOE} and $k_{dihedral}$ force constants were set relatively high to ensure that the experimentally-derived distance restraints

were satisfied (Nilges *et al.*, 1988b). When using this SA structure calculation protocol, 50-100 structures are usually calculated for each round, and a random seed generator is used to ensure that each cycle within a calculation round produces a different conformation.

Raising the force constants as described above is akin to decreasing the temperature of the system. However, a simple lowering of the temperature would not allow the force constants to be altered differentially, and this differential change is highly advantageous. This is because, for example, increasing F_{NOE} at a faster rate than F_{repel} has been shown to be more efficient in the SA procedure than changing both at the same rate (Nilges *et al.*, 1988b).

In XPLOR structure calculations, cysteine residues predicted to be involved in disulphide bond formation are initially "patched" together using NOE restraints (see Appendix I). The hydrogen atoms of the -SH groups are deleted, the size of the sulphur atoms is adjusted, and a normal NOE distance restraint is implemented, with a k_{NOE} value the same as for other restraints. This is performed so that initial structures are defined by the experimentally-derived NOE data, rather than by the inclusion of rigid disulphide bonds. Covalent disulphide bonds are introduced into the structures in the refinement process.

The force constant values described above are designed for a protein of up to ~10kDa, and are therefore the values that were used in the structure calculations of the cbEGF12-13 pair from human fibrillin-1. For studies of proteins of a different molecular mass, these values may need to be optimised. In fact several parameters may need to be optimised for a protein depending on its molecular mass and dynamics, including the initial high temperature used, the duration of the dynamic simulations and the steps of energy minimisation. The optimum parameters can be judged by following the changes in the potential energy of the system as these parameters are adjusted.

Multiple rounds of the SA structure calculation technique are performed in combination with modification of violated restraints. If restraints are violated, the

corresponding spectral data are reanalysed with the result that these restraints are left unaltered, adjusted or corrected. Possible adjustments include relaxing distance restraints by one distance class, or, in extreme cases, by two classes, e.g., for interactions involving methyl protons which are in a state of fast rotation. After these multiple rounds, a set of ~100 satisfactory structures will be obtained. These structures can then be further improved by refinement, which improves both close non-bonded contacts and local geometry (refine.inp -see Appendix I). This refinement process uses the structures calculated by SA structure calculations, introduces covalent disulphide bonds, and simply comprises a cooling minimisation sequence analogous to the slow cooling process described above. In refinement the starting temperature for this slow cooling process is, however, lower, because the structures used by refinement already have a low potential energy and require only local geometry improvements. Refinement improves convergence rates, and one to three rounds of refinement are usually performed on the SA structures calculated. In the end, a final family of ~20-25 NMR solution structures may be selected from these one to three rounds of refinement, taking care not to duplicate structures. This selection process is based various factors, including low potential energy values and maximum agreement with experimentally-derived data.

2.3.4 Evaluation of NMR Structures

There are several well-established methods for evaluating the quality of a final ensemble of NMR structures (see Laskowski *et al.*, 1996), and it must noted that each of these methods is less efficient when used alone. True insight into a structure's quality can only be obtained when a broad range of factors are considered, since these factors are not independent and one quality may be optimised at the expense of another.

The first method for evaluating NMR structures is via consideration of the convergence of the ensemble. The overall precision of a NMR ensemble is usually

expressed in terms of either an average pairwise root-mean-square-deviation (rmsd) of the coordinates across the structures in the final family, or a rmsd of the structures relative to the mean coordinates. The fact that all the structures in a final ensemble are not uniform is due to the conformational variation allowed in structure calculations using solution NMR data. The rmsd values can be calculated over the whole molecule or only over "well-defined" segments of the structure, which can be defined by dynamics data. It is well documented that secondary structural elements in a set of final structures overlay better than loop regions, with rmsd values of ~0.5Å from the mean coordinates. The greater rmsd variation seen in loop regions is usually consistent with their dynamic properties in solution, and can sometimes be related to function. Using these rmsd values alone as an indication of structural quality is not, however, wise, as low rmsd values may indicate a precise result rather than an accurate one.

In addition to structural convergence, agreement with experimental data can also be used to evaluate a final family of structures. This is usually expressed in terms of the number of violated restraints and the magnitude of these violations. The total or average restraint violation is often quoted, combined with the magnitude of the largest violation.

The quality of the ensemble can also be measured in terms of the number of restraints per residue, which is a very significant determinant of the quality of NMR structures. Structure calculations with an average of 16 restraints per residue usually give structures with a rmsd value of ~0.5Å for the backbone atoms in secondary structural regions.

Another method for evaluating NMR structures is based on analysis of stereochemical quality. This method is independent of experimental data, and involves analysing the geometric properties of the protein, e.g., bond lengths, bond angles and dihedral angles. This analysis is performed using computational methods, which make comparisons between the ensemble structures and what is already known about standard protein structure and geometry from available high-resolution X-ray crystal structures. The programs used for this analysis are PROCHECK-NMR and AQUA

(Laskowski *et al.*, 1996), which give a full list of the geometrical properties of each amino acid residue in all the structures of the ensemble, and the degree of agreement between the structures and the experimental data, respectively. These programs can also be beneficially used during the structure calculation and refinement processes, as they can highlight problems arising from incorrect or inaccurate restraints.

In summary, a combination of the methods described above should be used to assess the quality of structures determined using NMR-based structure calculations, and the greater the number of evaluation methods employed, the greater the confidence in the accuracy/quality of a final ensemble of structures. Where possible the accuracy of the fold should also be verified with biological data, for example, assessing whether an enzyme active site is capable of supporting its biological function. This biological data, however, is not always available.

2.4 Introduction to the Study of Protein Dynamics by NMR

The single, well-defined three-dimensional protein structures determined by X-ray crystallography are often considered higher resolution than the family of structures calculated using NMR methods. However, as previously mentioned, the latter can give insight into the intrinsic dynamics of a macromolecule. This insight is less easily obtained using X-ray crystallography methods, and, as biochemical processes in general are extremely dynamic, there is no reason to believe that biological macromolecules maintain a static structure *in vivo*. Conformational flexibility can be essential for the activity and biomechanical properties of these macromolecules (recently reviewed in relation to NMR studies of protein dynamics by Kay, 1998). Regions of a structural ensemble with poor convergence can indicate areas of increased flexibility, which, in turn, can be an indicator of function. For example, highly unstructured loop regions are frequently thought to be involved in interactions with other macromolecules. Increased rmsd values alone, however, are an unreliable measure of intrinsic protein

dynamics, as poor convergence can purely indicate of a lack of experimental data for a structured region, e.g., for a loop projecting into solution.

Information regarding intrinsic protein dynamics in solution can be obtained by studying the relaxation rates of ^{15}N or ^{13}C nuclei using NMR. These NMR relaxation measurements probe motions on the picosecond (ps) to nanosecond (ns) and microsecond (μs) to millisecond (ms) time-scale, providing detailed information concerning protein dynamics. These experiments enable a more complete description of a macromolecule to be obtained and are therefore now used routinely. Since backbone dynamics measurements were obtained for the cbEGF12-13 domain pair from human fibrillin-1 (Chapter 6), a brief overview of relaxation theory will be presented here. For a more detailed description of the theory of NMR relaxation with respect to protein dynamics the reader is referred to Cavanagh *et al.*, 1996.

2.4.1 T_1, T_2 and heteronuclear NOE

Information about protein backbone dynamics in solution is usually obtained by measuring the relaxation rates of ^{15}N or ^{13}C nuclei in isotopically labelled samples. The primary mechanism by which these nuclei "relax" is through dipolar interactions with directly bonded protons (DD). Chemical shift anisotropy (CSA) provides an additional relaxation mechanism at high magnetic field strengths (Abragam, 1961). These two interactions are not the only ones capable of affecting relaxation rates, as, for example, chemical-exchange processes can also significantly influence transverse relaxation rates (see below; Kaplan & Fraenkel, 1980). The three parameters measured in the NMR relaxation study described in Chapter 6 are T_1, T_2 and heteronuclear NOE, and it is these parameters, combined with detailed statistical analysis, that allow characterisation of protein backbone dynamics on the ps to ns and μs to ms time-scale.

T_1 is known as the spin-lattice or longitudinal relaxation time constant, and $T_1 = 1/R_1$, where R_1 is the spin-lattice relaxation rate constant. T_1 describes the time for the recovery of longitudinal magnetisation to thermal equilibrium after this equilibrium has been perturbed, and was first formulated in the Bloch equations (Bloch *et al.*, 1946; Cavanagh *et al.*, 1996). In solution NMR, T_1 values are sensitive to high frequency motions on the ps to ns time-scale (Abragam, 1961). Considering relaxation of protonated ^{15}N nuclei, i.e., the nuclei considered in the NMR relaxation study of cbEGF12-13 (Chapter 6), and ignoring cross-correlation effects, T_1 is given by,

$$1/T_1 = 1/T_1^{DD} + 1/T_1^{CSA}$$

where T_1^{DD} and T_1^{CSA} are the dipolar and CSA longitudinal relaxation time constants, respectively.

T_2 is defined as the spin-spin or transverse relaxation time constant, and $T_2 = 1/R_2$, where R_2 is the spin-spin relaxation rate constant. T_2 describes the decay of transverse magnetisation after this magnetisation has been produced by a perturbation of the equilibrium magnetisation. This relaxation time constant, together with T_1, provides the essential phenomenological description of NMR (Bloch *et al.*, 1946; Cavanagh *et al.*, 1996). T_2 values, in addition to being sensitive to the same high frequency motions as T_1, are also a function of lower frequency processes on the μs to ms time-scale. Combining T_1 and T_2 data therefore provides information about protein backbone dynamics over a large time-scale range (Kay *et al.*, 1989). Considering relaxation of protonated ^{15}N nuclei, ignoring cross-correlation effects, T_2 is given by,

$$1/T_2 = 1/T_2^{DD} + 1/T_2^{CSA} + R_{ex}$$

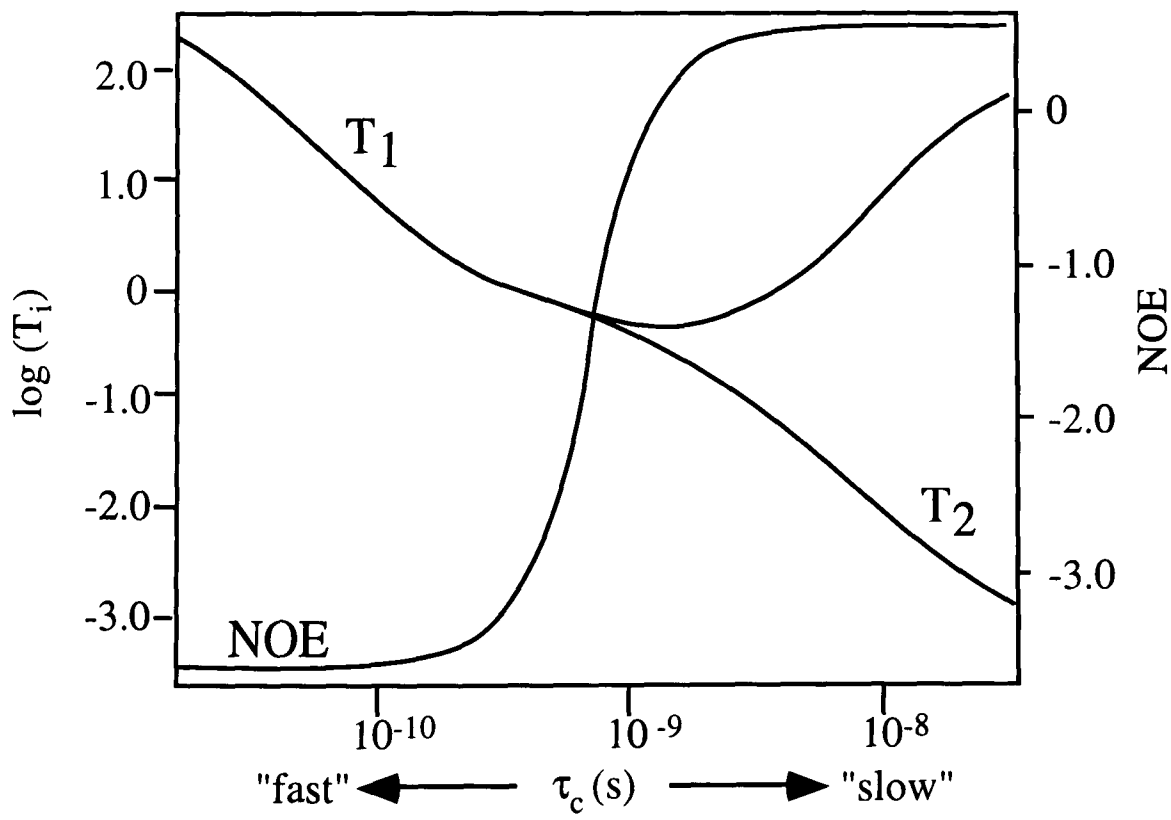
where T_2^{DD} and T_2^{CSA} are the dipolar and CSA transverse relaxation time constants, respectively, and R_{ex} is the sum of rate constants for other pseudo-first-order processes contributing to transverse relaxation, for example, ^{15}N conformational exchange on the μs to ms time-scale (Bloom *et al.*, 1965).

The homonuclear nuclear Overhauser effect (NOE) has been described in section 2.2.5.3 using the phenomenological approach of Solomon (1955), and is defined as the fractional change in the intensity of a dipolar coupled spin's NMR resonance when cross-relaxation occurs to restore equilibrium populations. The heteronuclear NOE is produced analogously to the homonuclear NOE, except that in the former magnetisation transfer occurs via the heteronuclear dipolar coupling (Cavanagh *et al.*, 1996). Heteronuclear NOEs, like T_1 values, are sensitive to high frequency motions on the ps to sub-ns time-scale (10^{-12} - 10^{-8}s).

Figure 2.26 shows a plot of the dependence of the three relaxation parameters discussed above on τ_c, the rotational correlation time of a N-H spin pair. ^{15}N-T_1 and T_2 relaxation rates and ^{1}H-^{15}N heteronuclear NOE are plotted against τ_c at a frequency of 50.7MHz according to the equations in section 2.4.2 (see below), assuming both isotropic motion and the absence of internal motions.

Figure 2.26 : Plots of Log (T_i) (i=1,2) and ^{1}H-^{15}N NOE Versus τ_c for ^{15}N Nuclei Relaxed by Dipolar and CSA Mechanisms at the Nitrogen Frequency 50.7MHz

For "fast" values of τ_c, T_1 and T_2 are essentially the same, whereas for "slow" values of τ_c T_1 values increase whilst T_2 values decrease. This figure was provided courtesy of Dr. A. K. Downing.



2.4.2 The Spectral Density Function

A quantum mechanical description of relaxation is required in order to relate the relaxation time constants and NOE to correlation times. The dipolar and CSA relaxation time constants described above, together with the NOE, are all dependent on the spectral density function, $J(\omega)$, and can be written as (Abragam, 1961),

$$1/T_1^{DD} = (1/4)d_{NH}^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)]$$
$$1/T_2^{DD} = (1/8)d_{NH}^2[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)]$$
$$1/T_1^{CSA} = (1/3)\omega_N^2(\sigma_{\parallel} - \sigma_{\perp})^2J(\omega_N)$$
$$1/T_2^{CSA} = (1/18)\omega_N^2(\sigma_{\parallel} - \sigma_{\perp})^2[4J(0) + 3J(\omega_N)]$$
$$NOE = 1 + (1/4)T_1d_{NH}^2(\gamma_H/\gamma_N)[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]$$

where $d_{NH} = (h/2\pi)(\mu_o/4\pi)\gamma_H\gamma_N\langle r_{NH}^{-3}\rangle$, h is Plank's constant (6.626×10^{-34} J s), μ_o is the permeability of free space ($4\pi \times 10^{-7}$ T m A^{-1}), γ_H and γ_N are the gyromagnetic ratios of ^1H and ^{15}N (2.6752×10^8 and -2.7108×10^7 s^{-1} T^{-1}, respectively), ω_H and ω_N are the Larmor frequencies of ^1H and ^{15}N ($2\pi \times 500.13$ and $-2\pi \times 50.68$ MHz, respectively), and r_{NH} is the N-H bond length (here set to 0.102nm). The CSA tensor was assumed to be axially symmetric, with the parallel and perpendicular components, $\sigma_{\parallel}$ and $\sigma_{\perp}$. In this project, the value of $\sigma_{\parallel} - \sigma_{\perp}$ was kept constant at -170ppm (Tjandra *et al.*, 1996).

The spectral density function, $J(\omega)$, in the above equations is the Fourier transform of the orientational correlation function of a unit vector along the N-H bond and incorporates the fact that relaxation rates depend on molecular motion (Abragam, 1961). For a protein in solution, $J(\omega)$ is dependent on both the motion of the macromolecule as a whole, i.e., global dynamics, and on the internal motions of the ^1H-^{15}N bond vector (Lipari & Szabo, 1982a). To gain information about these motions, $J(\omega)$ needs to be determined using a combination of the above equations and the

experimental T_1, T_2 and heteronuclear NOE data. A problem arises, however, as there are usually insufficient data to allow $J(\omega)$ to be determined uniquely. The spectral density function is therefore usually approximated using the model-free formalism which will be described in section 2.4.4 (Lipari & Szabo, 1982a; 1982b).

2.4.3 Global Dynamics

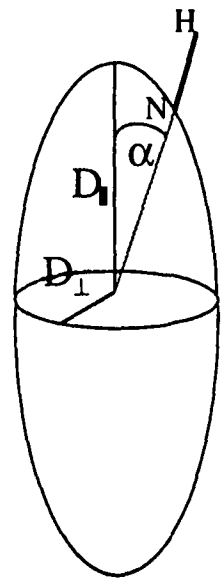
The cbEGF12-13 pair was predicted to fit an anisotropic, symmetric top model from homology to the cbEGF32-33 pair from human fibrillin-1 (Downing *et al.*, 1996; Werner *et al.*, 2000a), hence both isotropic and anisotropic diffusion models were used to describe global dynamics when fitting experimental relaxation data for this pair (Chapter 6). Figure 2.27 shows a schematic diagram of an anisotropic, symmetric top model with principal axes $D_{\parallel}$ and $D_{\perp}$, where α describes the angle between a N-H bond vector and $D_{\parallel}$. For this anisotropic diffusion model, global dynamics are given by,

$$J(\omega) = \frac{2}{5}\left[\sum_{i}^{3} \frac{A_i \tau_i}{1 + (\tau_i \omega)^2} \right]$$

where $A_1 = (1.5\cos^2\alpha - 0.5)^2$, $A_2 = 3\sin^2\alpha\cos^2\alpha$, $A_3 = 0.75\sin^4\alpha$ and the correlation times $\tau_1 = (6D_{\perp})^{-1}$, $\tau_2 = (D_{\parallel} + 5D_{\perp})^{-1}$ and $\tau_3 = (4D_{\parallel} + 2D_{\perp})^{-1}$ (Woessner, 1962). In the case where $D_{\parallel}$ equals $D_{\perp}$, the model becomes isotropic, i.e., a perfect sphere. For a spherical diffusion tensor the equation above is simplified and depends only on the overall rotational molecular correlation time, τ_m , where $\tau_m = \tau_1 = \tau_2 = \tau_3 = (6D)^{-1}$ with $A_1 + A_2 + A_3 = 1$.

Figure 2.27 : Schematic Diagram of an Anisotropic, Symmetric Top Model With Principal Axes $D_{\parallel}$ and $D_{\perp}$

The angle of the N-H bond vector with $D_{\parallel}$ is labelled α.



2.4.4 Model-Free Analysis of Internal Dynamics

For a complete description of protein dynamics, internal dynamics must be considered in addition to global dynamics, since, as mentioned in section 2.4.2, $J(\omega)$ is dependent on both the global dynamics of a macromolecule and on the internal motions of a N-H bond vector (Lipari & Szabo, 1982a). Amplitudes and motions of protein intramolecular dynamics are determined from the model-free formalism designed by Lipari & Szabo (1982a; 1982b), which is an extension of the equation shown in section 2.4.3 (Woessner, 1962). For a symmetric top anisotropic diffusion model, the spectral density function is assumed to be,

$$J(\omega) = \frac{2}{5}\left[S^2 \sum_{i}^{3} \frac{A_i \tau_i}{1+(\tau_i \omega)^2} + (1-S^2)\frac{\tau_c}{1+(\tau_c \omega)^2} \right]$$

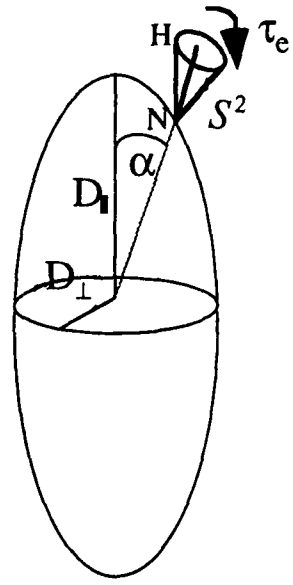
where A_1, A_2, A_3, τ_1, τ_2 and τ_3 are as described previously, $1/\tau_c = 1/\tau_m + 1/\tau_e$, where τ_e is an effective internal correlation time describing rapid internal motions, and S^2 is a generalised order parameter (Figure 2.28). Assumptions implicit in this model are that global and internal motions are independent of one another, and that internal motions

can be described by a single effective correlation time. This model was subsequently extended by Clore *et al.* (1990a; 1990b) (see below).

The generalised order parameter, S^2, characterises the amplitudes of internal motions on the ps to ns time-scale. S^2 measures the degree of spatial restriction of the unit bond vector, and values of S^2 range from 0 to 1, where, for a freely diffusing N-H bond vector, $S^2 = 1$ indicates completely restricted internal motion and $S^2 = 0$ generally indicates the absence of restriction on internal motion (Figure 2.28; Stone *et al.*, 1992; Mandel *et al.*, 1995). This time-scale range can be further broken down using the model-free formalism extension, $S^2 = S^2_f S^2_s$, where, within this ps to ns range, S^2_f and S^2_s represent motions on the fast and slow time-scales, respectively (Clore *et al.*, 1990b).

Figure 2.28 : Schematic Diagram of an Anisotropic, Symmetric Top Model Highlighting Internal Motions

The principal axes are labelled $D_{\parallel}$ and $D_{\perp}$. The angle of the N-H bond vector with $D_{\parallel}$ is labelled α. S^2 and τ_e are shown in a diffusion-in-a-cone interpretation.



In the basic model-free formalism, $1/\tau_c = 1/\tau_m + 1/\tau_e$, where τ_e is an effective internal correlation time describing rapid internal motions (Figure 2.28). Motions characterised by τ_e represent internal motions on the ps time-scale. Again, this time-scale can be broken down using the model-free formalism extension, where $\tau_e = \tau_f \tau_s$.

Here, τ_f represents an effective correlation time for internal motions on a fast time-scale (<100-200ps), and τ_s represents an effective correlation time for internal motions on a slower time-scale ($\tau_f < \tau_s < \tau_m$) (Clore *et al.*, 1990b).

R_{ex} is an adjustable parameter that can be added to the model-free formalism when data cannot be fit by S^2 and τ_e. Motions characterised by a R_{ex} term are due to ^{15}N conformational exchange processes. It is used to account for the ^{15}N conformational exchange that can occur on the time-scale of the delays in the spin-echo pulse sequences used in NMR relaxation experiments, i.e., on the μs to ms time-scale (Stone *et al.*, 1992; Mandel *et al.*, 1995).

Including the above extensions, there are five possible parameters that can be fit to the three experimentally-determined data points T_1, T_2 and NOE, i.e., ${S^2}_s$ and ${S^2}_f$ (S^2), τ_s and τ_f (τ_e), and R_{ex} (Lipari & Szabo, 1982a; 1982b; Clore *et al.*, 1990b; Mandel *et al.*, 1995). τ_f, however, is not generally used since it is usually too fast to be observed (Clore *et al.*, 1990b). No more than three of these model-free parameters, in addition to τ_m, can be can be fitted to the three relaxation parameters with statistical significance, and the original model-free formalism of Lipari & Szabo (1982a; 1982b) is retained unless motions within S^2 can be resolved.

2.4.5 Statistical Analysis

Quantitative analysis of relaxation data employs similar statistical reasoning at three stages, i.e., when deriving relaxation time constants and their associated uncertainties, when determining a global diffusion model using a set of T_1 and T_2 values, and when selecting a model of internal dynamics for each residue

(Chapter 6; Werner *et al.*, 2000b). This statistical analysis will now be described in relation to model selection for internal dynamics.

For each amino acid of the protein of interest, the combination of parameters that best fit the experimental data is determined, and the nomenclature and selection process for the five models of internal dynamics follow Mandel *et al.* (1995; Table 2.4). Various dynamic models of increasing complexity (Table 2.4) are chosen for each residue using χ^2 values and F-statistics (see below; Mandel *et al.*, 1995; Werner *et al.*, 2000b). Models are compared using the F-test, followed by calculation of the probability, Q, that this F could have been obtained by chance (Press *et al.*, 1990).

Table 2.4 : Models of Internal Dynamics for the Analysis of ^{15}N Relaxation Data

Model	Values of fixed parameters	Optimised parameters
1	$S^2_s = 1, \tau_e = 0, R_{ex} = 0$	S^2_f
2	$S^2_s = 1, R_{ex} = 0$	S^2_f, τ_e
3	$S^2_s = 1, \tau_e = 0$	S^2_f, R_{ex}
4	$S^2_s = 1$	S^2_f, τ_e, R_{ex}
5	$R_{ex} = 0$	S^2_f, S^2_s, τ_e

Calculation of model-free parameters for each residue from T_1, T_2 and NOE values is performed by minimisation of the target function,

$$\chi^2 = (T_1^{exp} - T_1^{th})^2/\sigma_1^2 + (T_2^{exp} - T_2^{th})^2/\sigma_2^2 + (NOE^{exp} - NOE^{th})^2/\sigma_{NOE}^2$$

where T_1^{exp}, T_2^{exp} and NOE^{exp} are the experimentally-determined relaxation parameters, σ_1, σ_2 and σ_{NOE} are the uncertainties in the experimentally-determined relaxation parameters, and T_1^{th}, T_2^{th} and NOE^{th} are the theoretical values calculated using the equations in section 2.4.2. The quality of the fit between the experimental data and the theoretical model can be assessed by comparing the optimal value of χ^2 with the 95% critical value of the distribution of χ^2 determined from a set of simulations. The critical

value is the χ^2 value that is exceeded by a specified fraction of the simulated data (Mandel *et al.*, 1995).

Models of increasing complexity are chosen only if simpler models are not adequate. If a more complex model is found to be adequate, the F-test is used to determine the statistical significance of the improvement (Mandel *et al.*, 1995). The F of a n- and m-parameter model with χ^2 values of χ^2_1 and χ^2_2, respectively, is given by,

$$F = \frac{\nu_2(\chi_1^2 - \chi_2^2)}{(\nu_1 - \nu_2)\chi_2^2}$$

where ν_1 and ν_2 denote the degrees of freedom of a model calculated as M-n and M-m, respectively, where M is the number of measured values and n and m are the number of free parameters in each model. The significance of an improvement in χ^2 on moving to a more complex model is tested by the probability $Q(\nu_1, \nu_2, F)$ of acquiring a F value as large as that obtained by chance (Press *et al.*, 1990). Small values of Q indicate that a more complex model is justified.

In summary, NMR experiments can be used to measure T_1 and T_2 relaxation time constants, as well as the steady-state heteronuclear NOE relaxation parameter. Since a unique spectral density function cannot generally be obtained from these data, the three experimentally-derived parameters are analysed using a model-free formalism. Characterisation of motions by these model-free parameters allows a full description of protein dynamic events on the ps to ns and μs to ms time-scale, both in terms of the overall molecular motion and the intrinsic, internal motions.

2.4.6 Chemical Exchange

For relaxation measurements performed at a single field strength, modelling R_{ex} into the above statistical analysis can provide misleading information. Several

approaches have therefore been developed to allow more direct quantification of the ^{15}N conformational exchange processes, including monitoring the quadratic dependence of exchange line broadening with B_1, the rotating frame field (Szyperski *et al.*, 1993; Zinnjustin *et al.*, 1997; Akke *et al.*, 1998), or B_0 (Peng & Wagner, 1995; Farrow *et al.*, 1995; Phan *et al.*, 1996), or by determining the R_2/R_1 ratio resulting from ^1H-^{15}N dipole and CSA relaxation interference (Kroenke *et al.*, 1998).

In this project, R_{ex} terms were estimated from the field dependence of R_2-$R_1/2$ (Phan *et al.*, 1996) in a model-independent way, as well as by treatment as adjustable parameters within the framework of the model-free analysis at a single field. Provided the spectral density at the proton frequency is small compared to the spectral density at zero frequency, it has been shown that R_2-$R_1/2$ is a linear function of the square of the spectrometer field strength (Phan *et al.*, 1996). The slope depends on the CSA value, assumed here to be -170ppm (Tjandra *et al.*, 1996), and on the R_{ex} terms. The advantage of this method is that it is independent of the estimation of the overall diffusion properties. The precision of the R_{ex} terms is, however, limited by the accuracy of the relaxation rate constants and by the range of spectrometer field strengths available. In addition to allowing R_{ex} terms to be directly determined, recording relaxation data at multiple field strengths also increases the amount of data available for statistical analyses, and therefore increases result confidence.

References

Abragam, A. (1961). Principles of nuclear magnetism. Clarendon Press, Oxford.

Akke, M., Liu, J., Cavanagh, J., Erickson, H. P. & Palmer, A. G. (1998). Pervasive conformational fluctuations on microsecond time scales in a fibronectin type III domain. *Nat. Struct. Biol.* **5**, 55-59.

Aue, W. P., Bartholdi, E. & Ernst, R. R. (1976). Two-dimensional spectroscopy: application to nuclear magnetic resonance. *J. Chem. Phys.* **64**, 2229-2246.

Baker, E. N. & Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Molec. Biol.* **44**, 97-179.

Barkhuijsen, H., De Beer, R., Bovée, W. M. M. J. & Van Ormondt, D. (1985). Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure. *J. Magn. Reson.* **61**, 465-481.

Barsukov, I. L. & Lian, L.-Y. (1993). Structure determination from NMR data I. Analysis of NMR data. In NMR of macromolecules. G. C. K. Roberts, ed., Oxford University Press Inc., New York, pp. 315-357.

Bartels, C., Xia, T.-H., Billeter, M., Güntert P. & Wüthrich K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR.* **5**, 1-10.

Bax, A. & Davis, D. G. (1985). MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **65**, 355-360.

Bloch, F., Hansen, W. W. & Packard, M. (1946). The nuclear induction experiment. *Phys. Rev.* **70**, 474-485.

Bloom, M., Reeves, L. W. & Wells, E. J. (1965). Spin echoes and chemical exchange. *J. Chem. Phys.* **42**, 1615-1624.

Böckmann, A. & Guittet, E. (1997). Determination of fast proton exchange rates of biomolecules by NMR using water selective diffusion experiments. *FEBS Lett.* **418**, 127-130.

Bodenhausen, G. & Ruben, D. J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185-189.

Braun, W. & Gö, N. (1985). Calculation of protein conformations by proton-proton distance constraints. A new efficient algorithm. *J. Mol. Biol.* **186**, 611-626.

Braunschweiler, L. & Ernst, R. R. (1983). Coherence transfer by isotropic mixing - application to proton correlation spectroscopy. *J. Magn. Reson.* **53**, 521-528.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187-217.

Brünger, A. T., Clore, G. M., Gronenborn, A. M. & Karplus, M. (1986). Three-dimensional structure of proteins determined by molecular dynamics with interproton distance restraints: Application to crambin. *Proc. Natl. Acad. Sci.* **83**, 3801-3805.

Brünger, A. T. (1992). X-PLOR (version 3.1): A system for X-ray crystallography and NMR. Yale University, New Heaven, CT, USA.

Brünger, A. T., Adams, P. D. & Rice, L. M. (1997). New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure.* **5**, 325-336.

Cavanagh, J., Fairbrother, W. J., Palmer III, A. G. & Skelton, N. J. (1996). Protein NMR spectroscopy. Academic Press, Inc., San Diego.

Clore, G. M & Gronenborn, A. M. (1985). Assessment of errors involved in the determination of interproton distance ratios and distances by means of one- and two-dimensional NOE measurements. *J. Magn. Reson.* **61**, 158-164.

Clore, G. M., Brünger, A. T., Karplus, M. & Gronenborn, A. M. (1986). Application of molecular dynamics with interproton distance restraints to three-dimensional protein structure determination. A model study of crambin. *J. Mol. Biol.* **191**, 523-551.

Clore, G. M., Driscoll, P. C., Wingfield, P. T. & Gronenborn, A. M. (1990a). Analysis of the backbone dynamics of interleukin-1β using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy. *Biochemistry.* **29**, 7387-7401.

Clore, G. M., Szabo, A., Bax A., Kay, L. E., Driscoll, P. C. & Gronenborn, A. M. (1990b). Deviations from the simple two parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Amer. Chem. Soc.* **112**, 4989-4991.

Clore, G. M. & Gronenborn, A. M. (1991). Structures of lager proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. *Science.* **252**, 1390-1399.

Craik, D. J. & Wilce, J. A. (1997). Studies of protein-ligand interactions by NMR. In Protein NMR techniques. D. G. Reid, ed., Humana Press Inc., Totowa, New Jersey, pp. 195-232.

Crippen, G. M. & Havel, T. F. (1978). Stable calculation of coordinates from distance information. *Acta Cryst.* **A34**, 282-284.

Derome, A. E. (1987). Modern NMR techniques for chemistry research. Pergamon Press Ltd.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Driscoll, P. C., Gronenborn, A. M. & Clore, G. M. (1989). The influence of stereospecific assignments on the determination of three-dimensional structures of proteins by nuclear magnetic resonance spectroscopy. Application to the sea anemone protein BDS-I. *FEBS Lett.* **243**, 223-233.

Englander, S. W. & Wand, A. J. (1987). Main-chain-directed strategy for the assignment of ^{1}H NMR spectra of proteins. *Biochemistry.* **26**, 5953-5958.

Ernst, R. R., Bodenhausen, G. & Wokaun, A. (1987). Principles of nuclear magnetic resonance in one and two dimensions. Clarendon Press, Oxford.

Farrow, N. A., Zhang, O. W., Szabo, A., Torchia, D. A. & Kay, L. E. (1995). Spectral density-function mapping using ^{15}N relaxation data exclusively. *J. Biomol. NMR.* **6**, 153-162.

Folmer, R. H. A., Hilbers, C. W., Konings, R. N. H. & Nilges, M. (1997). Floating stereospecific assignment revisited: Application to an 18 kDa protein and comparison with J-coupling data. *J. Biomol. NMR.* **9**, 245-258.

Freifelder, D. M. (1982). Principles of physical chemistry, with applications to the biological sciences. Jones and Bartlett Publishers, Inc., Boston.

Güntert P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283-298.

Hartmann, S. R. & Hahn, E. L. (1962). Nuclear double resonance in the rotating frame. *Phys. Rev.* **128**, 2042-2053.

IUPAC-IUB Commission on Biochemical Nomenclature (1970). Abbreviations and symbols for the description of the conformation of polypeptide chains. *Biochemistry.* **9**, 3471-3479.

Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546-4553.

Johnson, B. A. & Blevins, R. A. (1994). NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR.* **4**, 603-614.

Kaplan, J. I. & Fraenkel, G. (1980). NMR of chemically exchanging systems. Academic Press, New York.

Karplus, M. (1959). Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.* **30**, 11-15.

Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry.* **28**, 8972-8979.

Kay, L. E. & Bax, A. (1990). New methods for the measurement of HN-CαH coupling constants in ^{15}N-labeled proteins. *J. Magn. Reson.* **86**, 110-126.

Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient-enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665.

Kay, L. E. (1998). Protein dynamics from NMR. *Biochem. Cell. Biol.* **76**, 145-152.

Keeler, J., Clowes, R. T., Davis, A. L. & Laue, E. D. (1994). Pulsed-field gradients - theory and practice. *Methods Enzymol.* **239**, 145-207.

Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science.* **220**, 671-680.

Kroenke, C. D., Loria, J. P., Lee, L. K., Rance, M. & Palmer, A. G. (1998). Longitudinal and transverse ^{1}H-^{15}N dipolar ^{15}N chemical shift anisotropy relaxation interference : Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. *J. Am. Chem. Soc.* **120**, 7905-7915.

Lakowicz, J. R. (1999). Principles of fluorescence spectroscopy. Kluwer Academy/Plenum Press, New York.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486.

Lipari, G. & Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559.

Lipari, G. & Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559-4570.

Macura, S., Huang, Y., Suter, D. & Ernst, R. R. (1981). Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.* **43**, 259-281.

Malby S., Pickering, R., Saha, S., Smallridge, R., Lindse, S. & Downing, A. K.. The first EGF domain of the LDL receptor contains a non-canonical calcium binding site. Manuscript in preparation.

Mandel, A. M., Akke, M. & Palmer, A. G. (1995). Backbone dynamics of *Escherichia coli* ribonuclease Hi: Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144-163.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087-1092.

Morris, G. A. & Freeman, R. (1979). Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **101**, 760-762.

Mueller, L. (1987). P.E.COSY, a simple alternative to E.COSY. *J. Magn. Reson.* **72**, 191-196.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988a). Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* **229**, 317-324.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988b). Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. Circumventing problems associated with folding. *FEBS Lett.* **239**, 129-136.

Nilges, M., Kuszewski, J. & Brünger, A. T. (1991). Sampling properties of simulated annealing and distance geometry. In Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy. J. C. Hoch *et al.*, eds., Plenum Press, New York, pp. 451-455.

Nilges, M. (1995). Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J. Mol. Biol.* **245**, 645-660.

Nilges, M. (1997). Ambiguous distance data in the calculation of NMR structures. *Folding & Design.* **2**, S53-S57.

Norwood, T. J., Crawford, D. A., Steventon, M. E., Driscoll, P. C. & Campbell, I. D. (1992). Heteronuclear ^{1}H-^{15}N nuclear magnetic resonance studies of the c-subunit of the *Escherichia coli* F_1F_0ATP synthase -assignment and secondary structure. *Biochemistry.* **31**, 6285-6290.

Overhauser, A. W. (1953). Polarization of nuclei in metals. *Phys. Rev.* **92**, 411-415.

Palmer, A. G., Cavanagh, J., Wright, P. E. & Rance, M. (1991). Sensitivity improvement in proton-detected 2-dimensional heteronuclear correlation nmr-spectroscopy. *J. Magn. Reson.* **93**, 151-170.

Pardi, A., Billeter, M. & Wüthrich, K. (1984). Calibration of the angular dependence of the amide proton-C^{α} proton coupling constants, $^{3}J_{HN\alpha}$, in a globular protein: Use of $^{3}J_{HN\alpha}$ for identification of helical secondary structure. *J. Mol. Biol.* **180**, 741-751.

Peng, J. W. & Wagner, G. (1995). Frequency spectrum of NH bonds in eglin c from spectral density mapping at multiple field strengths. *Biochemistry*. **34**, 16733-16752.

Phan, I. Q. H., Boyd, J. & Campbell, I. D. (1996). Dynamics studies of a fibronectin type I module pair at three frequencies : Anisotropic modelling and direct determination of conformational exchange. *J. Biomol. NMR*. **8**, 369-378.

Powell, M. J. D. (1977). Restart procedures for the conjugate gradient method. *Math. Program*. **12**, 241-254.

Press, W. H., Flannery, B. P., Teukolosky, S. A. & Vetterling, W. T. (1986). Modelling of data. In Numerical Recipes. Cambridge University Press, Cambridge, pp. 498-546.

Press, W. H., Flannery, B. P., Teukosky, S. A. & Vetterling, W. T. (1990). In Numerical Recipes. Second edition, Cambridge University Press, New York.

Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem*. **23**, 283-437.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983). Improved spectral resolution in COSY ^{1}H-NMR of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun*. **117**, 479-485.

Redfield, C. (1993). Resonance assignment strategies for small proteins. In NMR of macromolecules, G. C. K. Roberts, ed., Oxford University Press Inc., New York, pp. 71-99.

Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem*. **34**, 167-339.

Schmid, F. X. (1997). Optical spectroscopy to characterize protein conformation and conformational changes. In protein structure : a practical approach. T. E. Creighton, ed., Oxford University Press, Oxford, pp. 261-297.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol*. **286**, 661-668.

Solomon, I. (1955). Relaxation processes in a system of two spins. *Phys. Rev*. **99**, 559-565.

Sørensen, O. W., Eich, G. W., Levitt, M. H., Bodenhausen, G. & Ernst, R. R. (1983). Product operator-formalism for the description of NMR pulse experiments. *Prog. Nucl. Magn. Reson. Spec*. **16**, 163-192.

Stone, M. J., Fairbrother, W. J., Palmer, A. G. III, Reizer, J., Saier, M. H. Jr. & Wright, P. E. (1992). Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from ^{15}N NMR relaxation measurements. *Biochemistry*. **31**, 4394-4406.

Sutcliffe, M. J. (1993). Structure determination from NMR data II. Computational approaches. In NMR of macromolecules, G. C. K. Roberts, ed., Oxford University Press Inc., New York, pp. 359-390.

Szyperski, T., Luginbuhl, P., Otting, G., Güntert, P. & Wüthrich, K. (1993). Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J. Biomol. NMR*. **3**, 151-164.

Tjandra, N., Szabo, A. & Bax, A. (1996). Protein backbone dynamics and ^{15}N chemical-shift anisotropy from quantitative measurement of relaxation interference effects. *J. Am. Chem. Soc.* **118**, 6986-6991.

Tropp, J. (1980). Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: The effect of fluctuating internuclear distances. *J. Chem. Phys.* **72**. 6035-6043.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (2000a). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296**, 1065-1078.

Werner, J. M., Campbell, I. D. & Downing, A. K. (2000b). Shape and dynamics of calcium binding proteins investigated by nitrogen-15 relaxation. In Calcium binding protein protocols, H. J. Vogel, ed., Meth. Mol. Biol. Humana Press, in press.

Whiteman, P., Downing, A. K., Smallridge, R., Winship, P. R. & Handford, P. A. (1998). A Gly to Ser change causes defective folding *in vitro* of calcium binding epidermal growth factor-like domains from factor IX and fibrillin-1. *J. Biol. Chem.* **273**, 7807-7813.

Whiteman, P., Smallridge, R. S., Knott, V., Cordle, J. J., Downing, A. K. & Handford, P. A.. A glycine to serine amino acid change in calcium-binding epidermal growth factor-like domain 13 of human fibrillin-1 causes short range conformational effects. Manuscript submitted.

Wimperis, S., Jones, J. A. & Hore P. J. (1998). Eight lectures on NMR spectroscopy. Physical and Theoretical Chemistry Laboratory, University of Oxford.

Woessner, D. E. (1962). Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion. *J. Chem. Phys.* **37**, 647-654.

Wüthrich, K., Billeter, M. & Braun, W. (1983). Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949-961.

Wüthrich, K. (1986). NMR of proteins and nucleic acids. John Wiley & Sons, Inc., New York.

Zinnjustin, S., Berthault, P., Guenneugues, M. & Desvaux, H. (1997). Off-resonance rf fields in heteronuclear NMR : Application to the study of slow motions. *J. Biomol. NMR*. **10**, 363-372.

Chapter 3

Expression of cbEGF12-13 Constructs

3.1 Introduction

The expression system for the cbEGF12-13 constructs used during the course of this DPhil project was developed by Dr. Pat Whiteman (Whiteman *et al.*, 1998), based on methods established in the laboratory of Dr. Penny A. Handford, Biochemistry Department, University of Oxford (Knott *et al.*, 1996). The method below describes the production of ^{15}N-labelled cbEGF12-13 wild-type sample using minimal media, after transformation into *Escherichia coli* BL21 competent cells. Production of unlabelled cbEGF12-13 was first performed to test the expression system, using a 1l preparation and minimal media containing unlabelled NH_4Cl. The expression system was found to be successful, i.e., ~13mg of protein were obtained, therefore ^{15}N-labelled sample was expressed. The purity and homogeneity of all samples produced by this method were confirmed by Dr. Whiteman using assays for free thiols and mass spectrometry.

The unlabelled cbEGF12-13 wild-type sample produced was utilised for NMR and fluorescence calcium binding studies, and also for solution structure determination by NMR. The ^{15}N-labelled wild-type sample was used for solution structure determination, as well as for backbone dynamic investigations of cbEGF12-13.

An unlabelled cbEGF12-13 G1127S mutant sample was also produced by the method described below, to allow investigation into the effect of this mutation on the

calcium binding and structural properties of this domain pair (see Chapter 7; Whiteman *et al.*, manuscript submitted). Subsequently, a ^{15}N-labelled mutant sample was produced for further future investigations into the structural and dynamic consequences of this mutation.

3.2 Expression

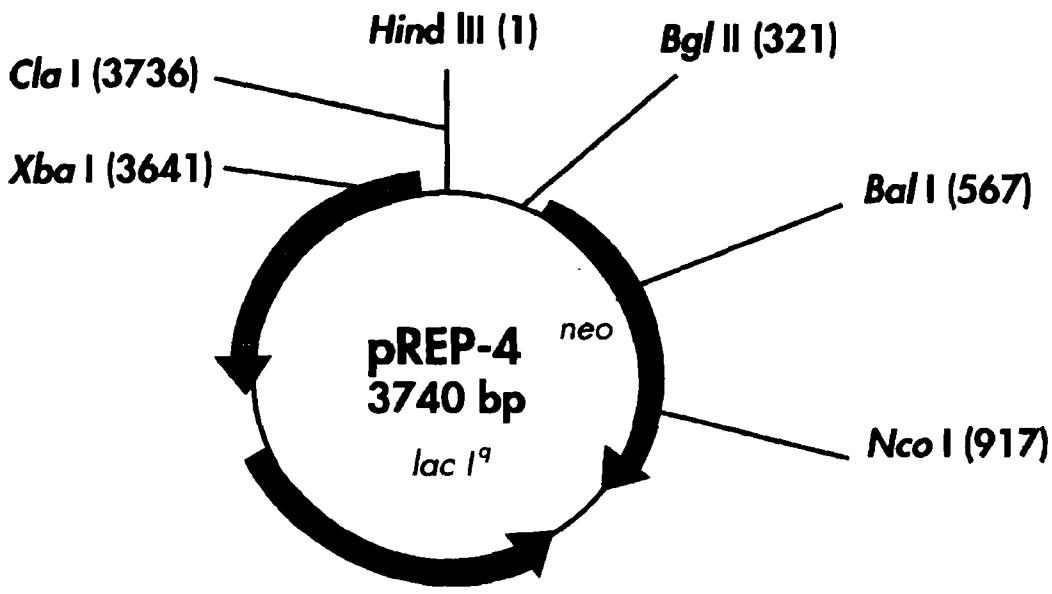
3.2.1 Growth of the Bacteria

E. coli BL21 [pREP4] competent cells were transformed with the pQE-30 plasmid containing the cbEGF12-13 construct. The resultant BL21 cells contained two plasmids; pREP4, which contained the *lac* repressor and kanamycin resistance genes, and pQE-30, which contained the ampicillin resistance genes, the cbEGF12-13 construct, the His-tag and the *lac* promoter and operator (Figure 3.1). 200µl of these transformed cells were plated onto a kan/amp minimal media plate (see Appendix II) and the plate was left at 37˚C for ~24 hours. A single colony was then used to inoculate a flask containing 170ml minimal media plus 340µl 50mg/ml ampicillin and 85µl 50mg/ml kanamycin. This flask was incubated at 37˚C for a further 24 hours.

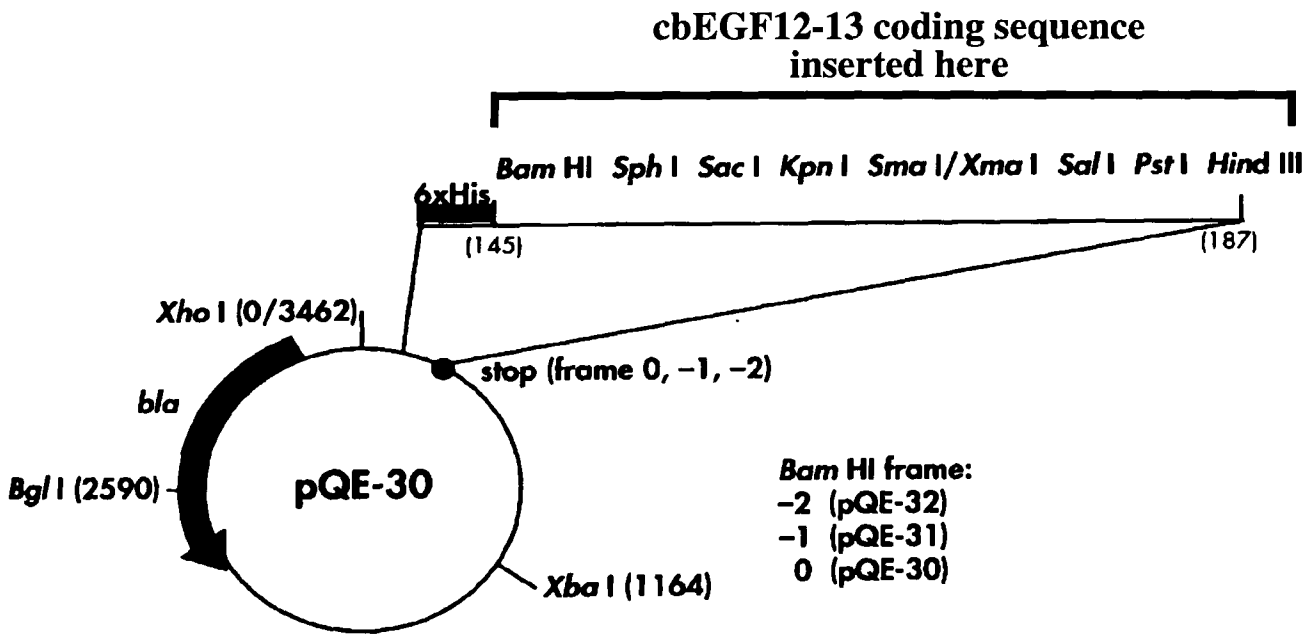
Four 2l flasks were prepared each containing 500ml of minimal media plus 1ml ampicillin (50mg/ml) and 0.25ml of kanamycin (50mg/ml). Each flask was then inoculated with 40ml of the culture prepared above, and the cultures were left to grow at 37˚C until they reached an $O.D._{600}$ of 0.8-0.9 (~2-3 hours). The cultures were induced to express cbEGF12-13 protein by addition of 1ml 1M isopropyl-β-D-thiogalactoside (IPTG, 2mM final concentration), and left to grow at 37°C for a further hour. The bacteria were subsequently harvested using a Beckman JA10 rotor at 5,000rpm for 20 minutes at 4˚C.

Figure 3.1 : Schematic Diagrams of the Plasmids Used in the Production of the cbEGF12-13 Constructs

(a) pREP4



(b) pQE-30



3.2.2 Lysis of the Bacteria

The supernatant was discarded, and the pellets were soaked in 6M guanidine buffered in 50mM sodium phosphate pH 7.4 and 5mM 2-mercaptoethanol ~pH 6 for ~1 hour at room temperature. The sample was also sonicated to reduce viscosity, and then centrifuged at 40,000rpm for 30 minutes at 25°C.

3.3 Purification

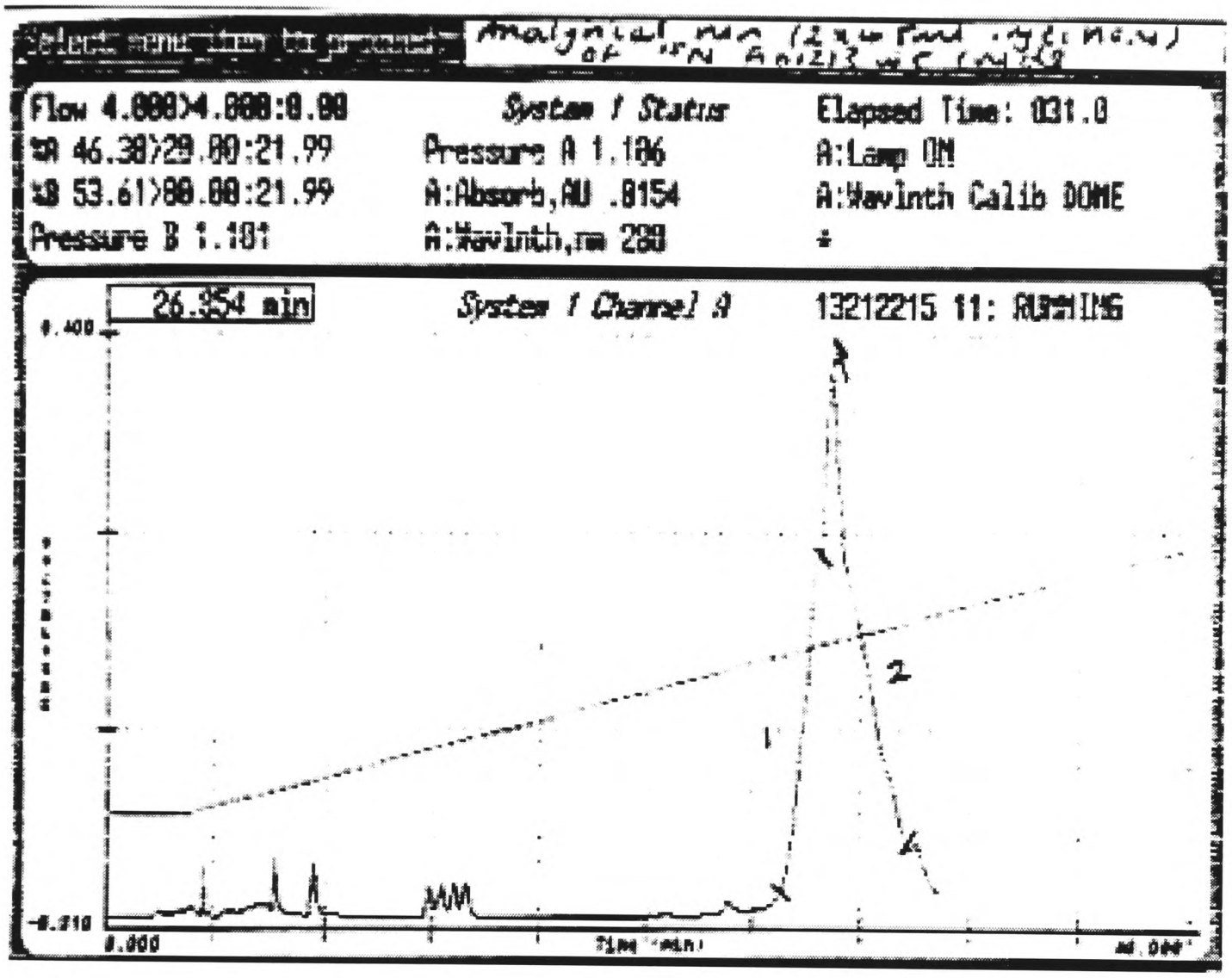
The cleared lysate was loaded onto a Ni^{2+} chelating Sepharose fast flow (Pharmacia) column which had been washed with water, loaded with 1M $NiSO_4$ solution, washed again with water, and then equilibrated with the 6M guanidine buffer solution described above. The effluent was collected as a safeguard.

The protein was eluted from the column using 20mM EDTA in the guanidine buffer solution, and 1.5ml fractions were collected. SDS-PAGE analysis was then performed by Dr. Whiteman on the effluent, the original cleared lysate, and on each of the fractions collected. This indicated that fractions 1-14 should be pooled, and that the effluent could be discarded since it was devoid of cbEGF12-13.

The pooled fractions were reduced for ~1 hour by adjusting to 0.1M dithiothreitol (DTT) and 0.1M Tris•HCl (pH 8.3). The sample was then dialysed in 0.1% trifluoroacetic acid (TFA) for 36-48 hours, changing the dialysis buffer twice, before undergoing high performance liquid chromatography (HPLC, see Figure 3.2). The sample was then freeze-dried before further treatment.

Figure 3.2 : Results of HPLC on the Purified ^{15}N-labelled cbEGF12-13 Domain Pair

In this example, two 4.5ml injections of purified cbEGF12-13 were loaded onto a C8 semi-prep column (Anachem) and the arrow indicates the peak collected.

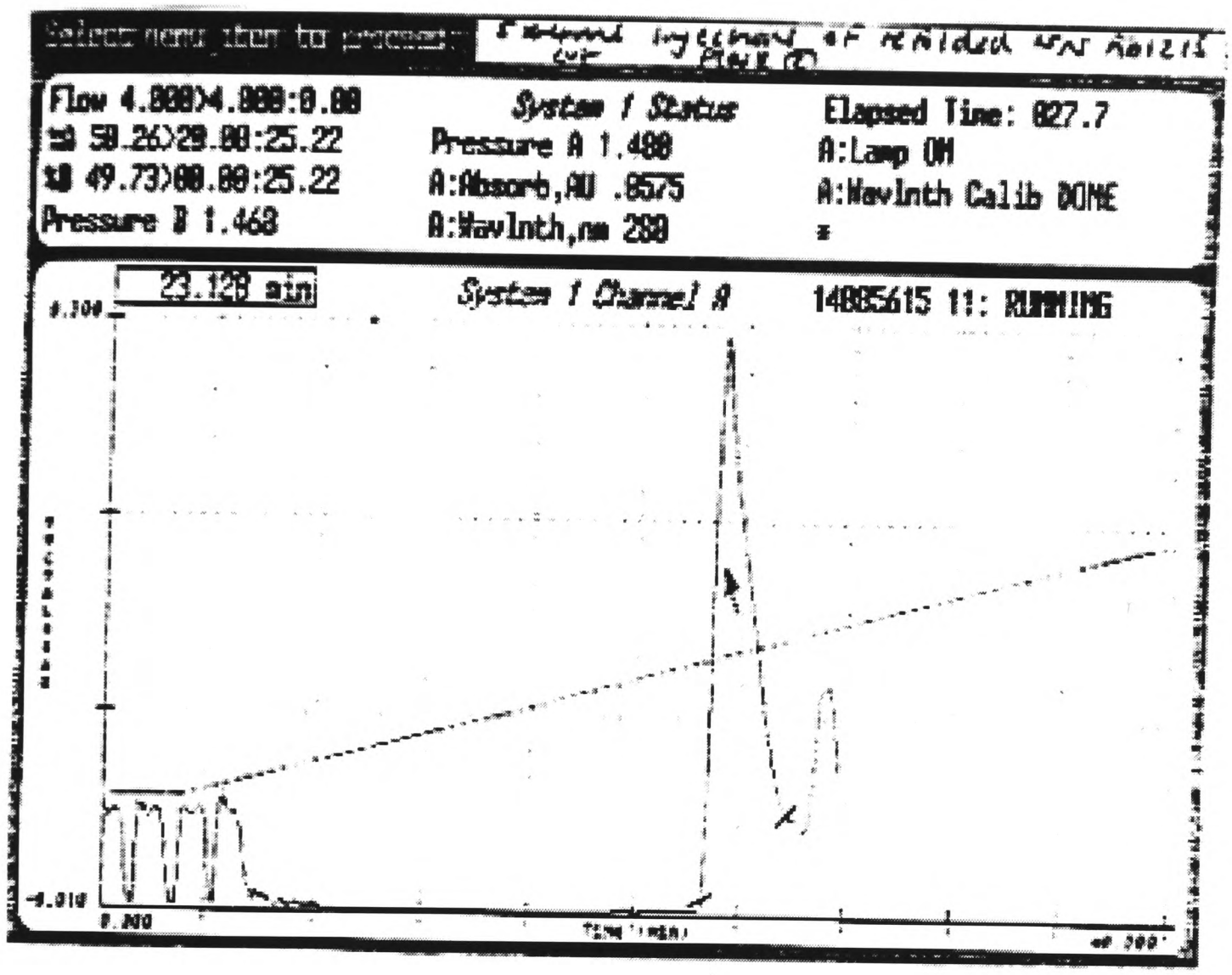


3.4 Refolding

The sample was refolded in 50mM $CaCl_2$, 0.1M Tris•HCl pH 8.3, 3mM cysteine and 0.3mM cystine ensuring a final protein concentration of 0.5 mg/ml (determined by measuring the $O.D._{280}$). This was left for ~24 hours at room temperature. After the pH was adjusted to 2-3 using 0.1% TFA and/or 1M HCl, the sample again underwent HPLC (Figure 3.3) and freeze-drying.

Figure 3.3 : Results of HPLC on the Refolded ^{15}N-labelled cbEGF12-13 Domain Pair

In this example, five 4ml injections of refolded cbEGF12-13 were loaded onto a C8 semi-prep column (Anachem) and the arrow indicates the peak collected.

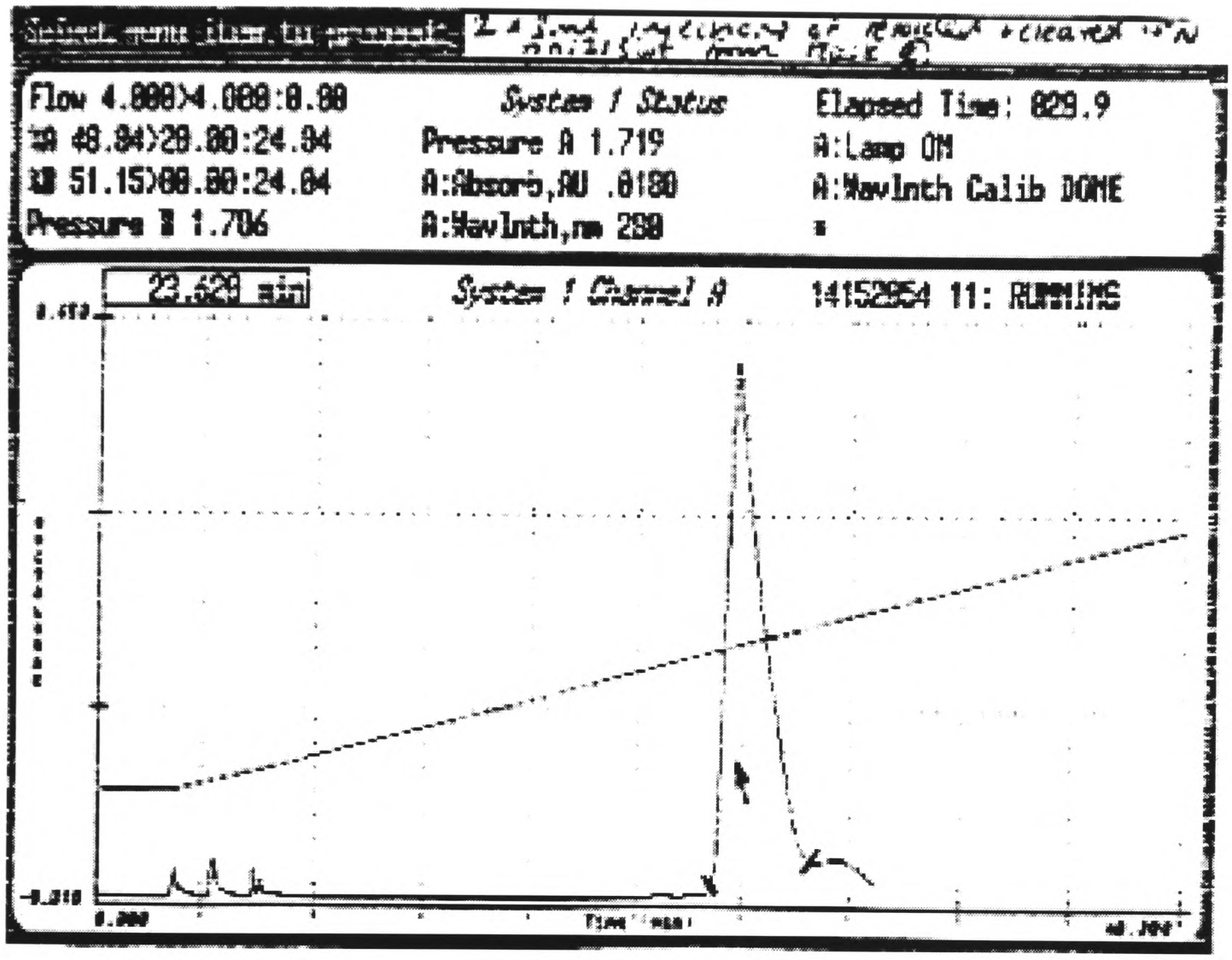


3.5 Removal of the His Tag

The protein was dissolved to give a final peptide concentration of 5mg/ml in 50mM Tris•HCl pH 7.5, 0.1M NaCl, and 5mM $CaCl_2$. Factor Xa was added to cleave the protein, and 1µg was added per mg of protein present. This was incubated at 37˚C for ~24 hours. After pH adjustment to 2-3, as described above, the sample again underwent HPLC (Figure 3.4) and freeze-drying. The sample was then ready for use in NMR experiments.

Figure 3.4 : Results of HPLC on the Refolded and Cleaved ^{15}N-labelled cbEGF12-13 Domain Pair

In this example, two 3ml injections of refolded and cleaved cbEGF12-13 were loaded onto a C8 semi-prep column (Anachem) and the arrow indicates the peak collected.



References

Knott, V., Downing, A. K., Cardy, C. M. & Handford, P. (1996). Calcium binding properties of an EGF-like domain pair from human fibrillin-1. *J. Mol. Biol.* **255**, 22-27.

Whiteman, P., Downing, A. K. & Handford, P. A. (1998). NMR analysis of cbEGF domains gives new insights into the structural consequences of a P1148A substitution in fibrillin-1. *Protein Eng.* **11**, 957-959.

Whiteman, P., Smallridge, R. S., Knott, V., Cordle, J. J., Downing, A. K. & Handford, P. A.. A glycine to serine amino acid change in calcium-binding epidermal growth factor-like domain 13 of human fibrillin-1 causes short range conformational effects. Manuscript submitted.

Chapter 4

Calcium Binding Studies of cbEGF12-13

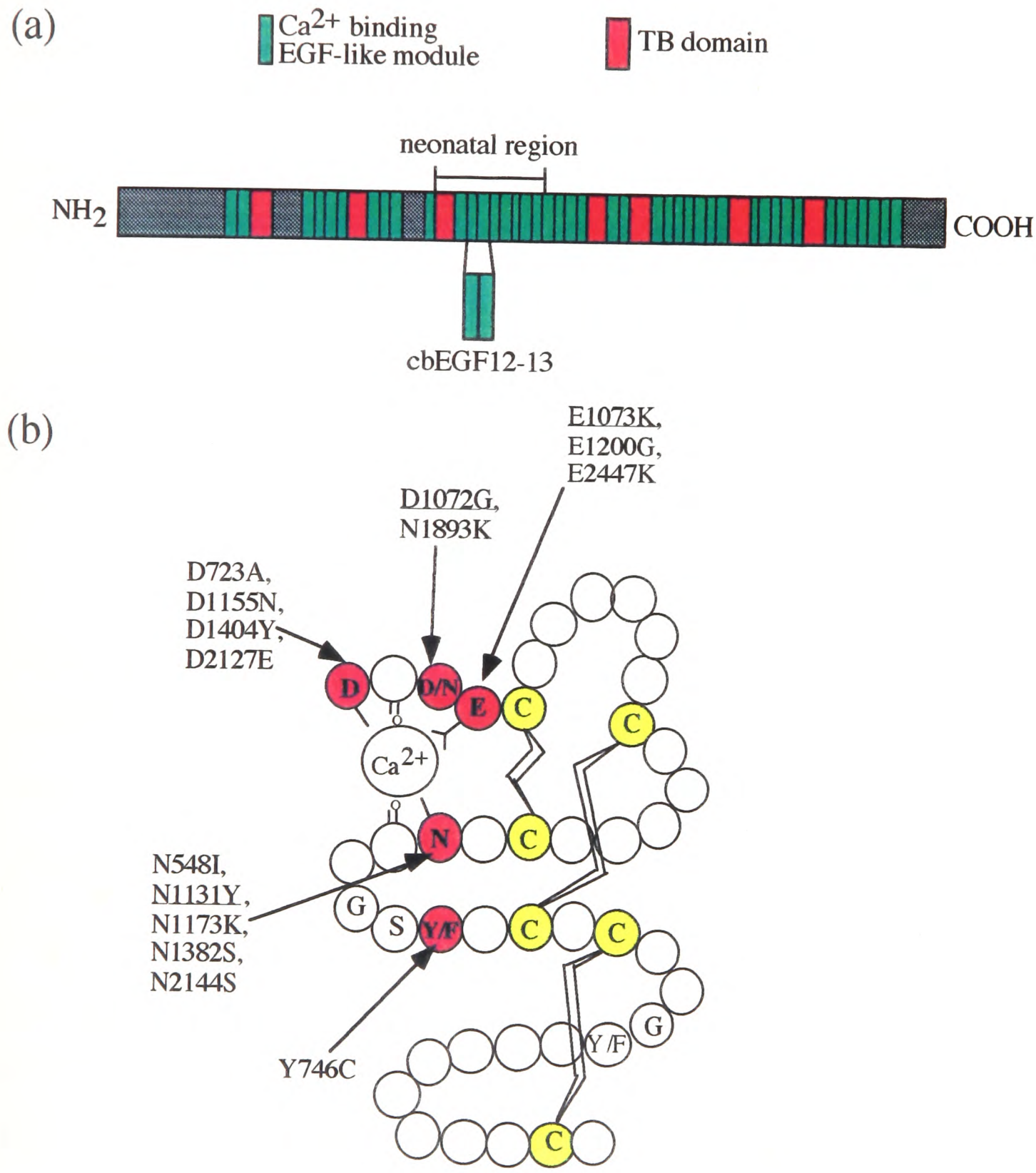
4.1 Introduction

As described in Chapter 1, human fibrillin-1 is mainly comprised of multiple, tandem repeats of cbEGF domains (Figure 4.1a). Electron micrographs recorded in the presence and absence of calcium have revealed that calcium plays a key role in maintaining the architecture of fibrillin-1-containing microfibrils (Kielty & Shuttleworth, 1993; Cardy & Handford, 1998; Wess *et al.*, 1998; Figure 1.3). Of the 140 MFS-causing mutations to fibrillin-1 that have been described (Collod-Béroud *et al.*, 1998), ~66% of these mutations are localised to cbEGF domains. A number of these mutations directly alter the calcium binding consensus sequence (Figure 4.1b), emphasising the critical functional role of calcium in human fibrillin-1.

Methods which have successfully been used to investigate calcium binding by cbEGF domains include NMR (Persson *et al.*, 1989; Handford *et al.*, 1991; Sunnerhagen *et al.*, 1993; Knott *et al.*, 1996; Stenberg *et al.*, 1997a; 1997b; 1999), equilibrium dialysis (Glanville *et al.*, 1994; Reinhardt *et al.*, 1996), intrinsic protein fluorescence (Astermark *et al.*, 1992; Hughes *et al.*, 1993) and chromophoric chelator methods (Rand *et al.*, 1997; Stenberg *et al.*, 1997a; 1997b; 1999). Of these, only NMR allows definitive assignment of K_d values to specific domains when studying constructs containing multiple binding sites.

Figure 4.1 : Schematic Illustration of (a) the Distribution of cbEGF Domains in Human Fibrillin-1 and (b) the Secondary Structure and Consensus Sequence of the cbEGF Domain

In (a) only the cbEGF and TB domains are shown, and the position of the neonatal region (as defined so far by mutation studies) and the construct used in this study are highlighted. In (b) conserved cysteine residues are shown in yellow and the consensus calcium binding sequence is shown in red. Point mutations in the calcium binding consensus sequence associated with MFS are highlighted, and underlined mutations are known to cause the neonatal MFS. Mutation data were obtained from the Marfan syndrome database on the World Wide Web (Collod-Béroud *et al.*, 1997; 1998).



The aim of this study was to investigate the calcium binding properties of the cbEGF12-13 domain pair from human fibrillin-1 (Figure 4.1a). This pair is located within the longest contiguous section of cbEGFs in fibrillin-1, and a number of mutations in this region are associated with the most severe neonatal form of MFS (Putnam *et al.*, 1996; Collod-Béroud *et al.*, 1997; 1998; Figure 4.1). In addition to providing further information on the calcium binding properties of this "neonatal" region of human fibrillin-1, the results obtained could be compared with previous studies of other cbEGF domain-containing constructs. This enabled elucidation of the general role of pairwise domain interactions in determining calcium affinity in fibrillin-1 (Smallridge *et al.*, 1999).

4.2 Preparation of cbEGF12-13 Samples for Calcium Titrations

4.2.1 Preparation of the NMR Titration Sample

The cbEGF12-13 domain pair from human fibrillin-1 was expressed, refolded and purified as described in Chapter 3 (Whiteman *et al.*, 1998b). ~1.63mg of unlabelled cbEGF12-13, as determined by $O.D._{280}$ measurements, were dissolved in 500µl 165mM NaCl ~pH 6.5 and 50µl 50mM Tris•HCl in ^2H_2O ~pH 6.5. This achieved a final concentration of 150mM NaCl to approximate physiological ionic strength, and Tris•HCl was added to simplify maintenance of a constant pH. Before data collection commenced, the pH of the sample was adjusted to 6.5 using 1µl aliquots of 0.1M NaOH or 0.1M HCl, and the sample was maintained at pH 6.5 throughout the experiment (see Table 4.1). This pH value was selected to allow comparison with calcium binding results previously obtained for other constructs from human fibrillin-1. 10µl of the sample were reserved for amino acid analysis to allow protein concentration to be determined accurately (see Appendix III).

4.2.2 Preparation of the Fluorescence Titration Samples

A 500μM protein stock solution was prepared by dissolving 1.5mg of unlabelled cbEGF12-13 in 311μl H_2O. Five ~2ml samples were then prepared which contained 20μl protein stock, 182μl 50mM Tris•HCl in ^2H_2O ~pH 6.5 and 1.82ml 165mM NaCl pH ~6.5, i.e., a final concentration of ~5μM cbEGF12-13, 4.55mM Tris•HCl and 150mM NaCl. The samples were then adjusted to pH 6.5 using 1μl aliquots of 0.1M NaOH or 0.1M HCl. Blanks were also prepared for background correction that contained 20μl H_2O substituted for protein.

4.3 Collection of Calcium Titration Data

4.3.1 Data Collection for the NMR Titration

A 2D NOESY spectrum was recorded at each titration point. All the spectra were recorded with a mixing time of 150ms at 33°C on a home-built/GE Omega spectrometer at 600MHz. 33°C was selected to allow comparison with calcium binding results previously obtained for other constructs from human fibrillin-1. For each titration point ~700 t_1 increments of 2K complex points were acquired, with a spectral width of 8000Hz. Water suppression was achieved using field gradients (Hwang & Shaka, 1995). Data were processed using Felix 2.3 (Biosym, Inc.). The spectra were referenced with respect to the H^2HO resonance, treated with a Gaussian window function of -15 0.15 (which increased signal to noise) and zero-filled in the F_2 dimension to produce spectra with a digital resolution of 1.95Hz/pt. The chemical shift changes of the Hδ* resonances of Phe1093 and Tyr1136, consensus calcium binding

aromatic residues for the N and C-terminal domains respectively, were used to monitor calcium binding in the 2D NOESY spectra (see Figure 4.2). The calcium free chemical shift positions of these resonances were confirmed by a duplicate titration.

Spectra were recorded at the calcium concentrations shown in Table 4.1, and small aliquots of $CaCl_2$ stock solutions (1.5-15µl) were added to prevent large fluctuations in total sample volume. The protein concentration was monitored by amino acid analysis during the course of the titration. The initial concentration of cbEGF12-13 in the NMR sample was found to be 0.131mM (see Appendix III), and the amino acid analysis data were incorporated into calculation of K_d values.

Table 4.1 : NMR Titration Calcium Additions

Experiment	Ca^{2+} added (µl)	[ ] of $CaCl_2$ solution	$[Ca^{2+}]$ total (mM)	pH before NMR	pH after NMR
0Ca	0	0	0	6.512	6.600
1Ca	2	50mM	0.18	6.530	6.531
2Ca	1.5	50mM	0.32	6.492	6.513
3Ca	1.5	50mM	0.45	6.565	6.561
4Ca	10	50mM	1.36	6.506	6.576
5Ca	10	200mM	4.99	6.554	6.585
6Ca	6	1M	15.90	6.599	6.530
7Ca	14	1M	41.35	6.599	6.507
8Ca	7	2.5M	73.17	6.481	6.516
9Ca	7.5	5.0M	141.35	6.506	6.520
10Ca	15	5.0M	277.72	6.473	6.448

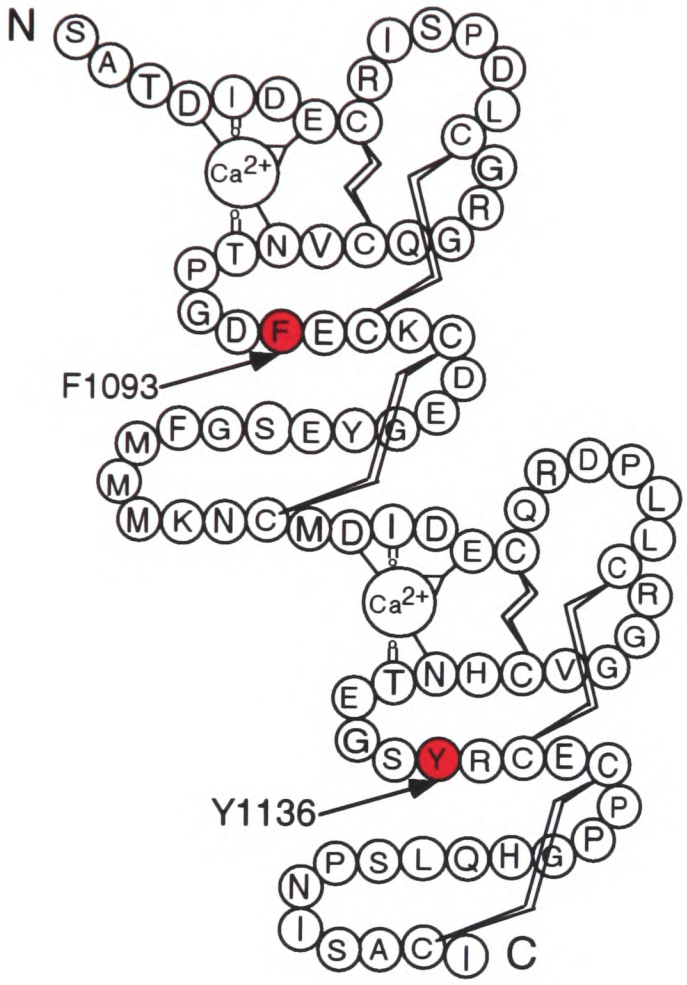
4.3.2 Data Collection for the Fluorescence Titrations

Fluorescence emission spectra were recorded at 33°C between 285 and 450nm. An excitation wavelength of 277nm was utilised, and the excitation and emission bandwidths were 5nm and 10nm, respectively. Calcium-induced fluorescence intensity changes were observed at an emission wavelength corresponding to the maximum fluorescence intensity measured in the absence of calcium (308nm, Figure 4.3). The calcium additions used are shown in Table 4.2, and two stocks solutions

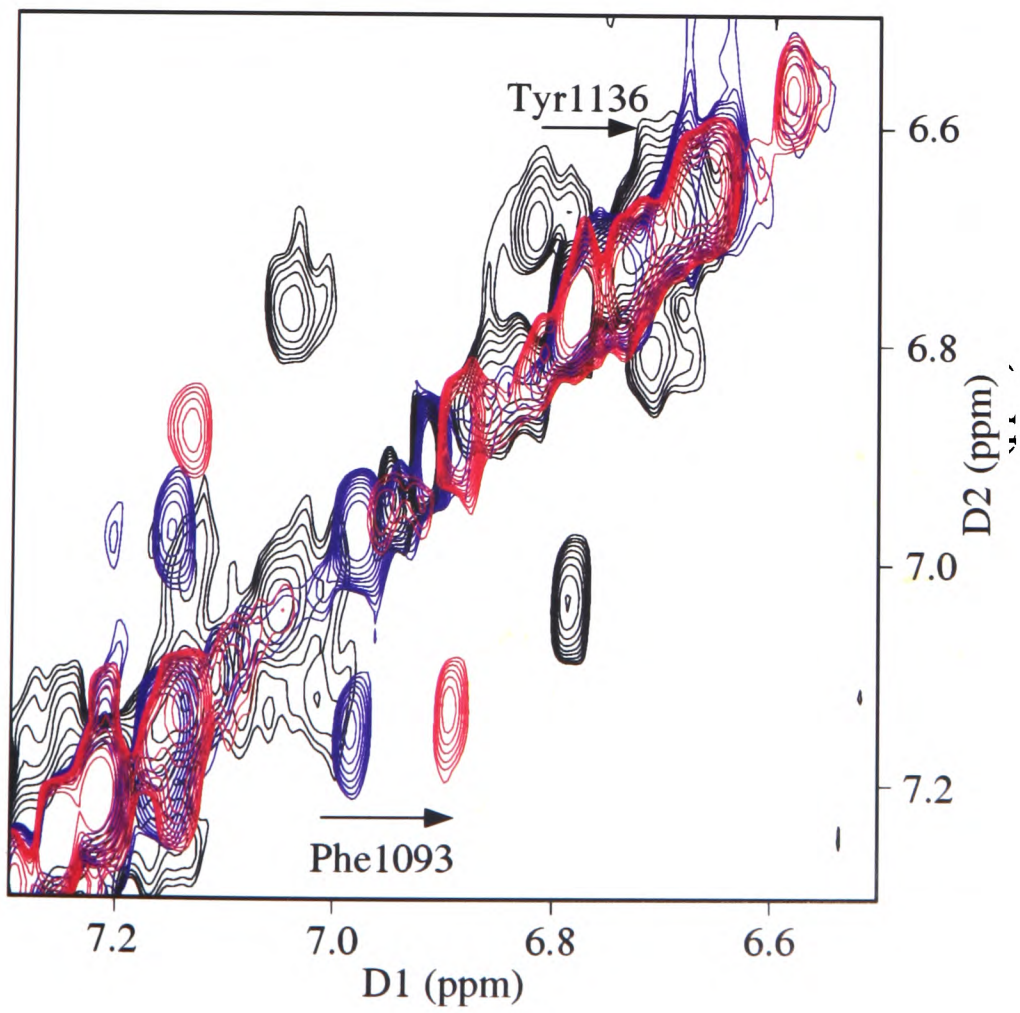
Figure 4.2 : (a) Schematic Diagram of the cbEGF12-13 Domain Pair, and (b) Superposition of 2D NOESY Spectra of cbEGF12-13 at the 0Ca (black), 3Ca (450μM, blue) and 10Ca (277.7mM, red) Calcium Titration Points

In (a) the consensus aromatic residues used to monitor calcium binding by NMR are highlighted, and in (b) arrows highlight the change of the Hδ* chemical shift for these residues. Note that by the 3Ca point the Tyr1136 crosspeak has reached its maximal displacement, whereas the Phe1093 crosspeak has only moved slightly. Although complete chemical shift assignments for the calcium-free form of cbEGF12-13 have not been made, the other obvious peak in the spectra below, that is only clearly resolved in the 0Ca spectrum, is likely to result from one of the consensus asparagine residues involved in calcium binding. The position of this peak appears to shift dramatically upon the first calcium addition, but from Figure 7.3 it is clear that this peak purely becomes broader and has a reduced intensity in the presence of calcium.

(a)



(b)



of $CaCl_2$ were used for the titration (5mM and 10mM). The total volume change during the course of the titration was <2.5% of the original sample volume. Fluorescence spectra were recorded, in duplicate, five minutes after each addition to allow temperature equilibration. The spectra were corrected for background at each point. An identical magnesium titration was performed, in duplicate, to confirm that the fluorescence intensity increases observed were calcium-dependent.

At the end of every titration, three 20µl aliquots from each cuvette were submitted for amino acid analysis and the mean protein concentration was found to be 3.9 ± 0.4µM for the five titrations performed. Lower protein concentration samples were assayed for the fluorescence titration (e.g., 1µM), but the signal to noise ratio was inadequate due to the inherently low intensity of tyrosine fluorescence.

Figure 4.3 : Emission Spectra for the cbEGF12-13 Domain Pair at 0 and 2.5mM $CaCl_2$

These spectra were corrected for a background of water and $CaCl_2$. The results indicated an emission maxima at 308nm, and the fluorescence intensity change upon addition of calcium was significant enough to allow calcium titrations to be performed using this method.

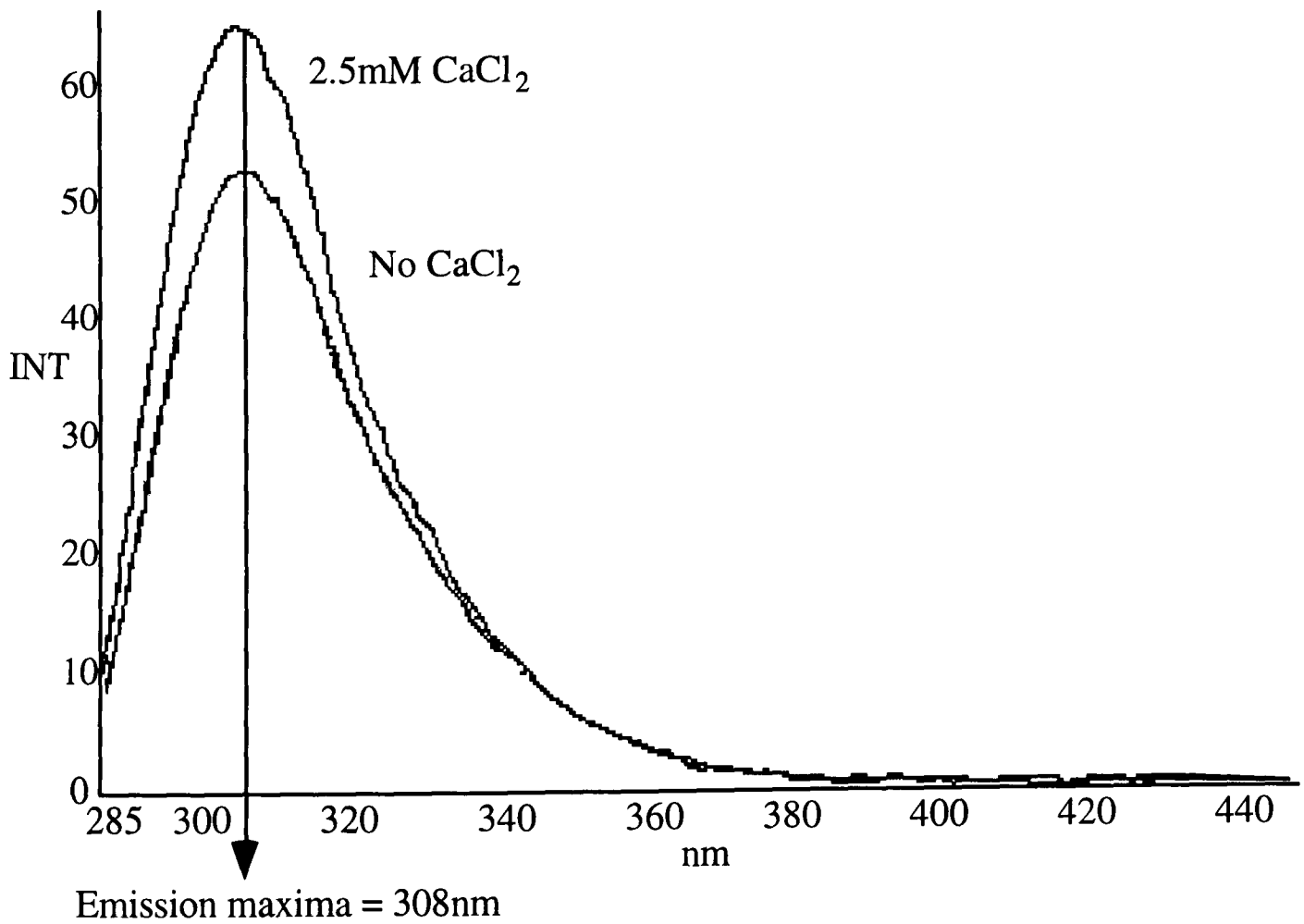


Table 4.2 : Fluorescence Titration Calcium Additions

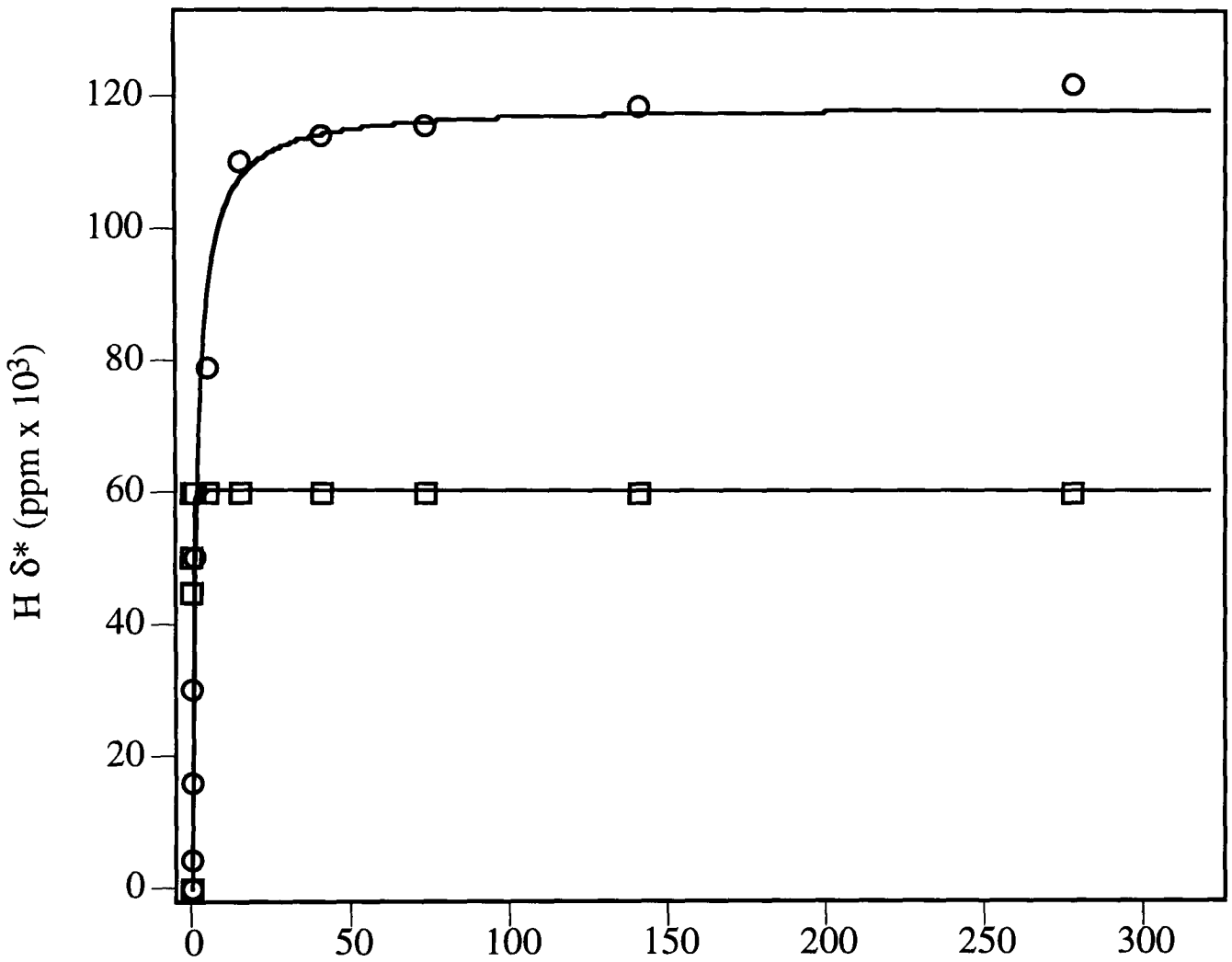
Experiment	Ca^{2+} added (µl)	[ ] of $CaCl_2$ solution (mM)	$[Ca^{2+}]$ total (µM)
0Ca	0	0	0
1Ca	1.6	5	4.0
2Ca	1.0	5	6.5
3Ca	1.0	5	9.0
4Ca	1.8	5	13.5
5Ca	1.8	5	18.0
6Ca	3.6	10	36.0
7Ca	3.8	10	55.0
8Ca	3.6	10	73.0
9Ca	3.6	10	91.0
10Ca	3.6	10	109.0
11Ca	7.2	10	145.0
12Ca	7.2	10	181.0
13Ca	7.2	10	217.0

4.4 Results

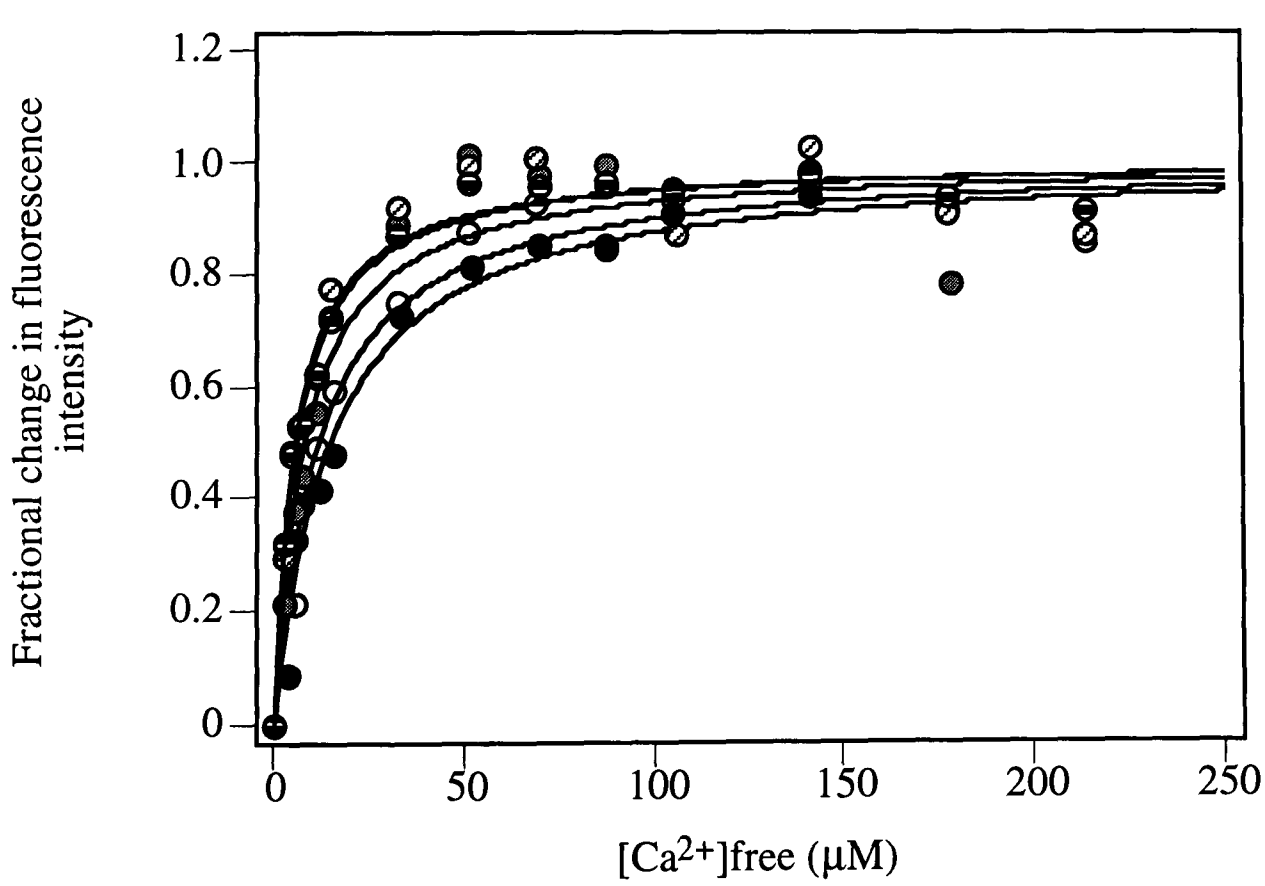
A plot of the change in chemical shift of the Hδ* resonances of Phe1093 and Tyr1136 versus the concentration of free calcium is shown in Figure 4.4a. The concentration of free calcium was computed as $[Ca^{2+}]_{free} = [Ca^{2+}]_{total} - \sum_{i=1}^{n} \left(\frac{\Delta}{\Delta_o}\right)_i [protein]$, where n is the number of calcium binding sites and (Δ/Δ_o) is the fractional change in the Hδ* chemical shift for the consensus aromatic residues (Knott *et al.*, 1996). This equation assumes independent binding of calcium by the two sites. Standard regression analysis was used to curve fit the data for each site using the equation $\Delta = \Delta_o [Ca^{2+}]_{free} / (K_d + [Ca^{2+}]_{free})$ (Knott *et al.*, 1996). From the curve fits to these data, the K_d values of the cbEGF12 and cbEGF13 sites were calculated to be 1.61 ± 0.24mM and 27 ± 4.4µM, respectively. The accuracy of the calcium dissociation constant for domain 13 is reduced, because the lower limit of K_d values which can be measured accurately using two-dimensional NMR is set by the need for a high enough sample concentration to observe the cross-peak(s) of interest.

Figure 4.4 : (a) The Change in Chemical Shift of the Hδ* Resonances of Phe1093 (○) and Tyr1136 (□) *versus* $[Ca^{2+}]_{free}$, and (b) the Fractional Change in Fluorescence Intensity of cbEGF12-13 *versus* $[Ca^{2+}]_{free}$ for the Five Titrations Performed

(a)



(b)



Intrinsic protein fluorescence (see Chapter 2), which has the potential to accurately measure smaller K_d values than NMR, was therefore utilised in the study of this pair. Since the cbEGF12-13 sequence does not contain any tryptophan residues, changes in tyrosine fluorescence were monitored. The cbEGF12-13 domain pair contains two tyrosine residues, one in the binding site of domain 13 and one which is predicted to be engaged in calcium dependent pairwise domain interactions based on homology to the cbEGF32-33 pair (Downing *et al.*, 1996; Figure 4.5). Hence calcium dependent changes in tyrosine fluorescence for this pair would be expected to relate solely to binding by domain 13. In addition, use of low concentration calcium additions in the fluorescence titration enabled visualisation of calcium binding by the cbEGF13 site only, as, from NMR, the cbEGF12 site shows no significant binding until $[Ca^{2+}]_{total}$ = 450μM (Figure 4.2b). A plot of the fractional change in fluorescence intensity versus the concentration of free calcium for the five titrations performed is shown in Figure 4.4b. Results were calculated as for the NMR titration data. From the curve fits to these data, the $K_{d,\,mean}$ of the cbEGF13 site was calculated to be 9.1 ± 3.6μM. Together the NMR and fluorescence data indicate that the K_d for cbEGF13 is <30μM, the highest affinity thus far observed for a cbEGF domain from human fibrillin-1.

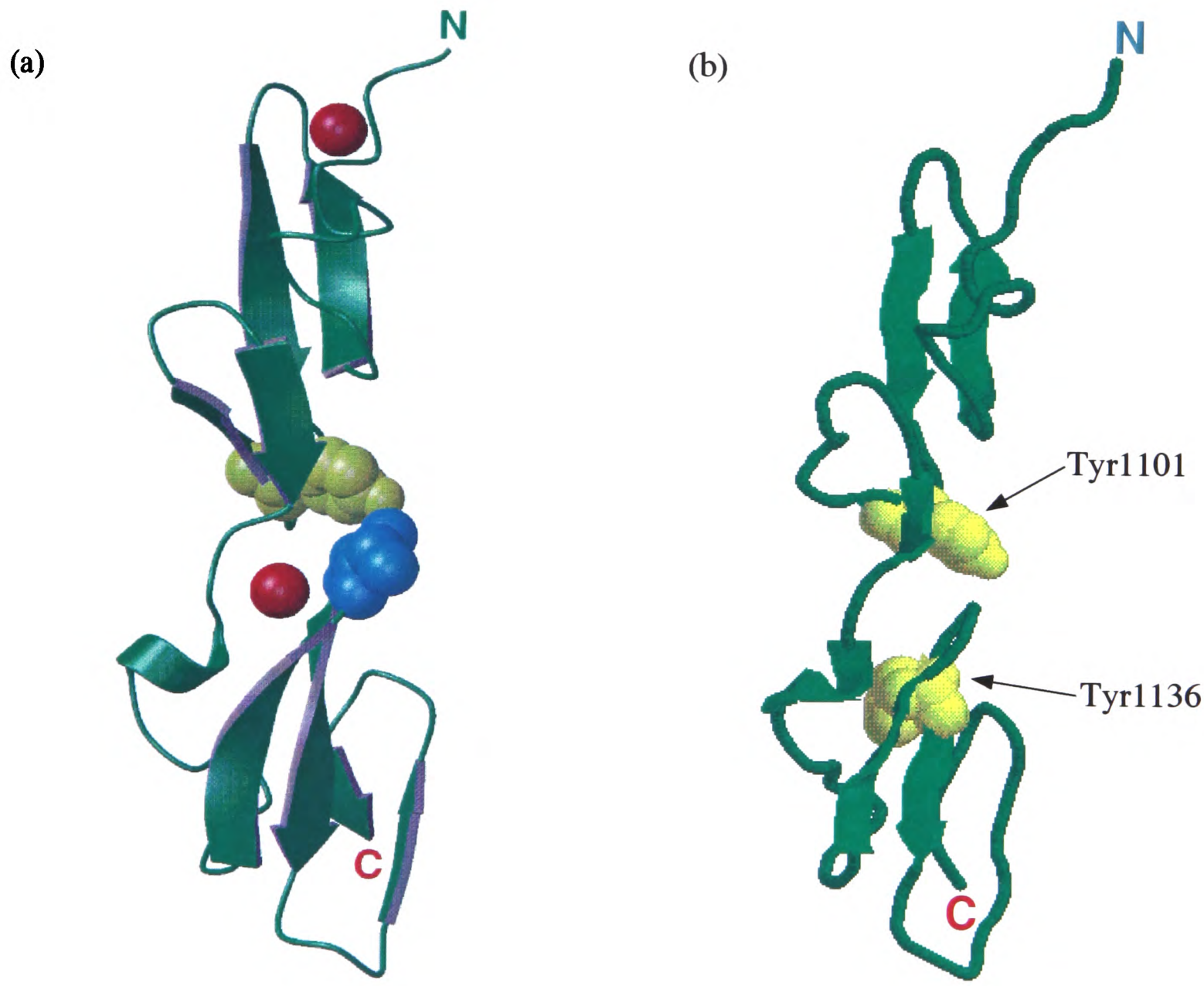
4.5 Discussion and Conclusions

4.5.1 Effects of Pairwise Domain Linkage on cbEGF Affinity

Dissociation constants for fibrillin-1 constructs in the two regions cbEGF12-14 and TB6-cbEGF33 are summarised in Figure 4.6. Comparing the K_d values obtained for cbEGF12-13 and cbEGF13 (Whiteman *et al.*, 1998a) shows that addition of the N-terminal cbEGF12 domain to cbEGF13 is associated with at least a 70-fold increase in the affinity of cbEGF13. In contrast, addition of the non-homologous TB6 domain to cbEGF32 only increases its affinity ~2.7-fold (Kettle *et al.*

Figure 4.5 : (a) The Structure of the cbEGF32-33 Domain Pair, and (b) A Model of the Structure of the cbEGF12-13 Pair

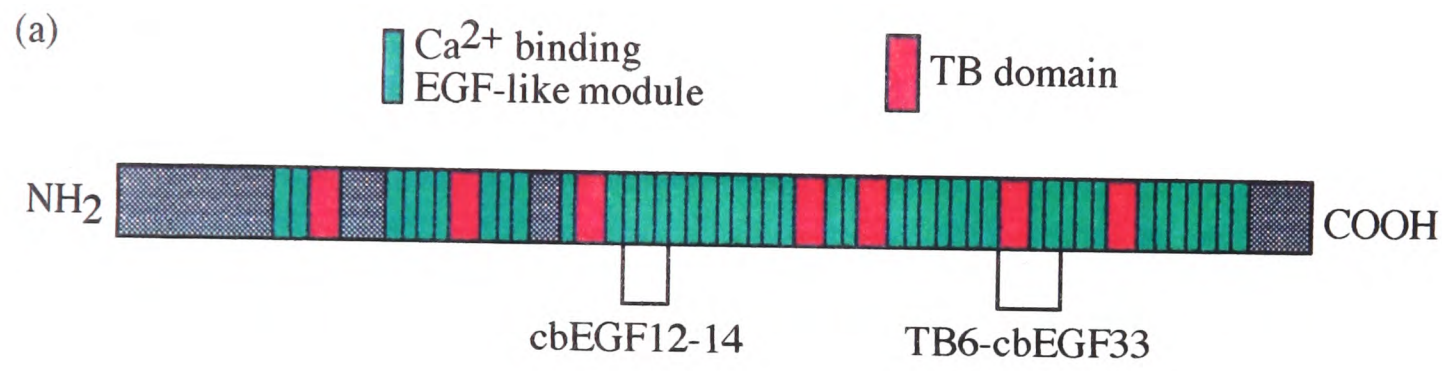
The structure of the cbEGF12-13 domain pair was modelled on the coordinates of cbEGF32-33 (Downing *et al.*, 1996) using Modeller4 (Šali & Blundell, 1993; see Chapter 5, section 5.1.4). Tyrosine residues in cbEGF12-13 are highlighted in yellow. In cbEGF32-33, the main hydrophobic interactions include a conserved aromatic residue (Tyr2157 (yellow)) at the open end of cbEGF32's minor β-sheet which packs against cbEGF33's Gly2186 (blue) and the two methyls of Ile2185 (blue). Tyr 2157 is equivalent to Tyr 1101 in cbEGF12, therefore this residue was predicted to be engaged in calcium dependent pairwise domain interactions in cbEGF12-13. Calcium atoms are shown as red spheres in (a).

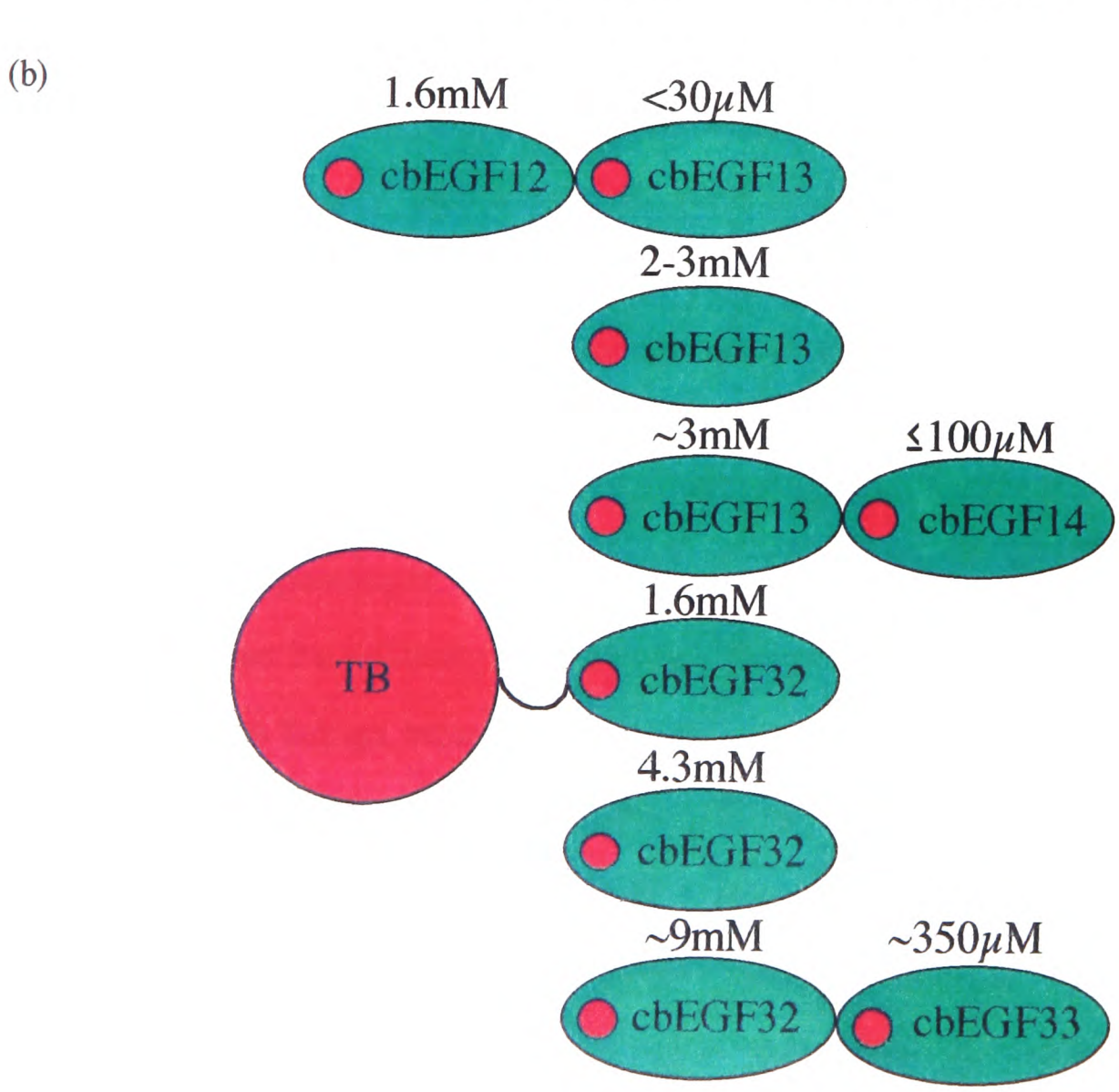


1999). These results suggest that specific pairwise domain interactions modulate the calcium binding affinity of cbEGF domains.

Figure 4.6 : Summary of K_d Values Obtained from Calcium Binding Experiments on Human Fibrillin-1 Constructs Comprising Domain(s) from the Regions cbEGF12-14 and TB6-cbEGF33

The position of the constructs whose calcium binding properties have been previously studied is highlighted in (a). In (b) calcium atoms are shown as small, red circles. Data were obtained from cbEGF12-13 (this study; Smallridge *et al.*, 1999), cbEGF13 and cbEGF13-14 (Whiteman *et al.*, 1998a), TB6-cbEGF32 (Kettle *et al.*, 1999), cbEGF32 (Handford *et al.*, 1995) and cbEGF32-33 (Knott *et al.*, 1996).





The structure of cbEGF32-33 from human fibrillin-1 (Downing *et al.*, 1996; Figure 4.5a) revealed that the two domains are arranged in a rod-like conformation with the calcium atom bound by the C-terminal domain localised to the domain interface. It appears that the presence of the N-terminal domain facilitates the formation of a correctly presented, rigid C-terminal calcium binding site, whereas the exposed N-terminal domain has a less well-defined site, which may account for the consistently lower affinities observed for N-terminal domains (Figure 4.6). Backbone relaxation measurements for the *holo* versus *apo* forms of the cbEGF32-33 pair suggest that calcium binding controls the rigidity of the interdomain linkage (Werner *et al.*, 2000 and manuscript in preparation). Hence increased rigidity of the interdomain linker region appears to be correlated with increased calcium binding. This theory is consistent with electron micrographs showing the effect of EDTA on microfibril appearance. Removal of calcium causes the rod-like superstructure of microfibrils to break down to yield a more diffuse, curved structure, implying that calcium affects the flexibility of fibrillin-1 (Kielty & Shuttleworth, 1993; Cardy & Handford, 1998). In addition, velocity sedimentation and rotary shadowing experiments on purified fibrillin-1 have shown that this protein has a stable, extended, rigid structure in the presence of calcium, in contrast to the flexible structure observed in the absence of calcium (Reinhardt *et al.*, 1997). Analytical ultracentrifugation, rotary shadowing and circular dichroism experiments have also shown the same effect for a recombinant fragment of fibrillin-1 (Reinhardt *et al.*, 1997). Correspondingly, based on the structure and dynamics of TB6 (Yuan *et al.*, 1997; 1998), and the relatively low calcium affinity of the TB6-cbEGF32 pair, it was predicted that the linkage between these two domains is flexible (Kettle *et al.*, 1999). The latter prediction has now been confirmed (Yuan *et al.*, manuscript in preparation).

Calcium binding studies of cbEGF domains from vitamin-K-dependent protein S, using chromophoric chelator methods, have implied that both N to C- and C to N-terminal affinity enhancement effects occur in this protein (Stenberg *et al.*, 1997a; 1997b; 1999). The results summarised in Figure 4.6, however, indicate that C-terminal

domain linkage does not enhance cbEGF domain calcium affinity in these two regions of fibrillin-1. Structural data are not currently available for the cbEGF domains from protein S, and it is possible that the relative orientation of protein S cbEGF domains may be different than for cbEGF domains in fibrillin-1. Also, studies of fibrillin-1 were performed on constructs containing a maximum of two linked cbEGF domains, while the protein S work was performed on up to four tandemly linked cbEGF domains, introducing the possibility of long-range (i.e., i, i+2 or i, i+3) domain interactions (Stenberg *et al.*, 1997a; 1997b).

In calcium binding studies of protein S and human Notch-1 cbEGF domains, attempts have been made to correlate charge distribution in the vicinity of the binding site with calcium affinities (Stenberg *et al.*, 1997a, b; Rand *et al.*, 1997). However, the distribution of charged amino acids and net charges of the constructs shown in Figure 4.6 show no apparent correlation with the relative affinities of the domains (Table 4.3 and Figure 4.7). Specific tandem cbEGF pairs from protein S and Notch have been predicted to adopt an alternative orientation to cbEGF pairs from fibrillin-1, with an additional calcium ligand donated from the N to C-terminal domain (Rao *et al.*, 1995; Downing *et al.*, 1996). Further structural and calcium binding studies will be necessary to reveal whether or not there exists a correlation between cbEGF domain pair structure and affinity.

4.5.2 Implications of Variable Calcium Affinity Along the Length of Fibrillin-1

The calcium binding properties of several regions of human fibrillin-1 containing one or more cbEGF domains have been studied to date (Table 4.4): cbEGF 12-14 (this study, Smallridge *et al.*, 1999; Whiteman *et al.*, 1998a), the region spanning cbEGF22 to within TB6 (Reinhardt *et al.*, 1996), cbEGF25-31 (Glanville *et al.*, 1994) and TB6-cbEGF33 (Handford *et al.*, 1995; Knott *et al.*, 1996; Kettle *et al.*, 1999). Of these, cbEGF domains 13 and 14 have significantly higher affinities

Table 4.3 : Table to Show the Distribution of Charge and the Isoelectric Points for the cbEGF12-13, cbEGF13-14 and cbEGF32-33 Domain Pairs

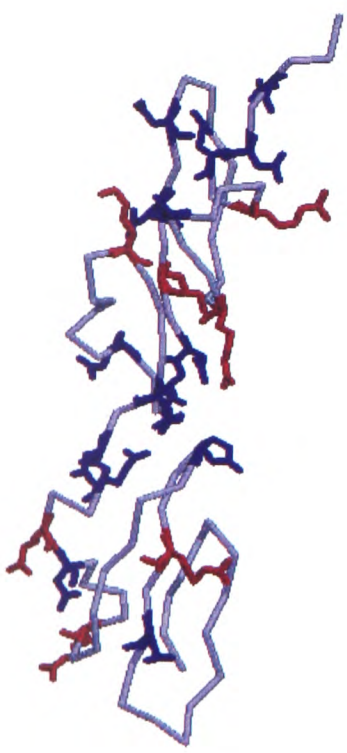
Histidine residues are assumed to be uncharged *in vivo*.

Construct	Overall charge of N-terminal domain	Overall charge of C-terminal domain	Overall charge	Isoelectric point (pI)
cbEGF12-13	**-5**	**-3**	**-8**	**4.27**
cbEGF13-14	**-3**	**-1**	**-4**	**5.51**
cbEGF32-33	**-5**	**-7**	**-12**	**3.81**

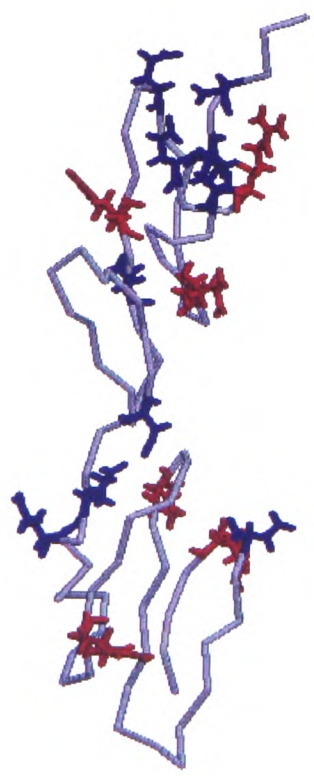
Figure 4.7 : Comparison of the Charge Distribution of the cbEGF12-13, cbEGF13-14 and cbEGF32-33 Domain Pairs

The cbEGF12-13 and cbEGF13-14 models were calculated using the coordinates of cbEGF32-33 (Downing *et al.*, 1996). The cbEGF12-13 model was created using Modeller4 (Šali and Blundell, 1993), and the cbEGF13-14 model was created using Insight v2.3 (Biosym, Inc.). In all domain pairs, glutamate and aspartate residues are shown in blue, with lysine and arginine residues shown in red. Calcium atoms have been omitted to increase clarity.

cbEGF12-13



cbEGF13-14



cbEGF32-33

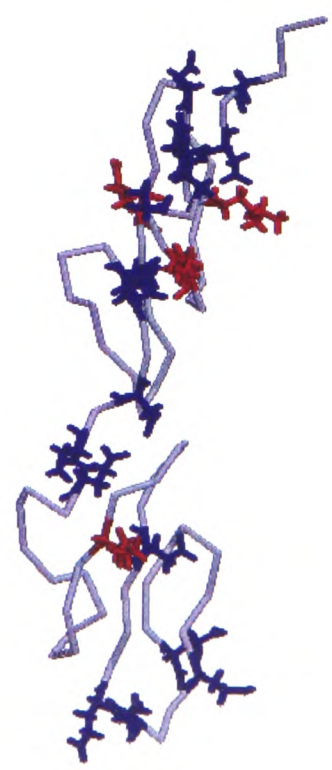


Table 4.4 : The Results of Calcium Binding Studies on cbEGF Domain(s) from Human Fibrillin-1

The diagram and table below highlight the regions of human fibrillin-1 that have been investigated by calcium binding studies. The methods used and the results obtained for these studies are shown.



cbEGF domain(s) studied	Experimental method(s) used	Kd value(s) obtained	Reference
cbEGF12-13	2D NMR and intrinsic protein fluorescence	12: 1.6mM, 13: <30μM	Smallridge *et al.*, *J. Mol. Biol.*, 1999, 286, 661-668.
cbEGF13	1D NMR on a deuterated sample	2-3mM	Whiteman *et al.*, *J. Biol. Chem.*, 1998a, 273, 7807-7813.
cbEGF13-14	2D NMR	13: ~3mM, 14: ⩽100μM	Whiteman *et al.*, *J. Biol. Chem.*, 1998a, 273, 7807-7813.
cbEGF22-23	Equilibrium dialysis	~400μM (n=2)	Reinhardt *et al.*, *J. Mol. Biol.*, 1996, 258, 104-116.
cbEGF22-TB6	Equilibrium dialysis	~500μM (n=10)	Reinhardt *et al.*, *J. Mol. Biol.*, 1996, 258, 104-116.
cbEGF25-31	Equilibrium dialysis	250μM (n=6-7)	Glanville *et al.*, *J. Biol. Chem.*, 1994, 269, 26630-26634.
TB6-cbEGF32	2D NMR	1.6mM	Kettle *et al.*, *J. Mol. Biol.*, 1999, 285, 1277-1287.
cbEGF32	1D NMR on a deuterated sample	4.3mM	Handford *et al.*, *J. Biol. Chem.*, 1995, 270, 6751-6756.
cbEGF32-33	2D NMR	32: ~9mM, 33: ~350μM	Knott *et al.*, *J. Mol. Biol.*, 1996, 255, 22-27.

(2.5 to >16-fold) than cbEGF domains located in the other three regions. Although the physiological relevance of these measurements is unclear, this variation in calcium binding affinity of cbEGF domains along the length of fibrillin-1 suggests that calcium binding sites may be differentially occupied *in vivo*. Using the equation $[Ca^{2+}]_{bound} = [Ca^{2+}]_{free} / (K_d + [Ca^{2+}]_{free})$, with $[Ca^{2+}]_{free} = 1.4mM$, it was estimated that while the cbEGF13 domain calcium binding site will be 98% saturated *in vivo*, the cbEGF33 site will only be 80% saturated. These values were calculated assuming that $[Ca^{2+}]_{free}$ in the extracellular matrix is the same as in blood, and that the K_d values *in vivo* are comparable to those calculated *in vitro* for the cbEGF domain pairs. This means that linkages between domains with lower affinity sites, which may not always be fully occupied with calcium, have the potential to be more flexible than those with higher affinity sites. The variable rigidity of cbEGF domains within fibrillin may also contribute to the extensibility and/or elasticity of the molecule. The importance of ions (including calcium) for the biomechanical properties of microfibrils has been analysed by studies of microfibril-rich sea cucumber dermis, which have shown that in deionised water the network is much more compliant than when tested in artificial sea water (Thurmond & Trotter, 1996).

cbEGFs 12-14 are within the first four domains in the longest run of cbEGFs in human fibrillin-1 (Figure 4.6a). This region contains a cluster of mutations associated with neonatal MFS and another severe form of MFS (Putnam *et al.*, 1996; Collod-Béroud *et al.*, 1997). The "neonatal" region has been localised to exons 24-32 of human fibrillin-1, which encode TB3 and cbEGF domains 11-18 (Putnam *et al.*, 1996; Collod-Béroud *et al.*, 1997; Figure 4.1a). Similarly, a cluster of mutations which affect exons 24-32 of fibrillin-2 have been found to be associated with congenital contractural arachnodactyly, further emphasising the critical role of this central region of the fibrillin proteins (Park *et al.*, 1998). The rigidity of the neonatal region of fibrillin-1 "imposed" by a high calcium affinity may be essential for assembly of connective tissue microfibrils. This is consistent with immunofluorescence studies of neonatal MFS

fibrils which have shown that these are short, fragmented and frayed relative to wild-type microfibrils (Superti-Furga *et al.*, 1992; Godfrey *et al.*, 1995). The specific mutation(s) giving rise to this morphology has not, however, been identified and may not affect calcium binding directly.

The potential importance of this high affinity region in microfibril assembly could partially explain why a number of mutations in this region result in the most severe forms of the MFS. For example, mutation of a calcium ligand in cbEGF12 (E1073K) produces a lethal phenotype (Nijbroek *et al.*, 1995; Putnam *et al.*, 1996), whereas an equivalent mutation in cbEGF38 (E2447K) produces ectopia lentis (EL), a milder fibrillin disorder (Lönnqvist *et al.*, 1994). Both of these mutations occur in the second cbEGF domain in a series of tandem cbEGF repeats. However, similar calcium ligand mutations within the neonatal region can also produce phenotypic variations (Figure 4.1b and Table 4.5). Thus additional factors such as the influence of the mutation on protein-protein interactions and/or the fraction of mutant fibrillin incorporated into the microfibril must also affect phenotype.

Table 4.5 : MFS and Related Phenotypes for Mutations Affecting the Same Calcium Ligand in Different cbEGF Domains of Fibrillin-1

Amino Acid Change	**Domain number**	**Disease Phenotype**
E1073K	cbEGF12	neonatal MFS
E1200G	cbEGF15	classical MFS
E2447K	cbEGF38	ectopia lentis
N1131Y	cbEGF13	neonatal MFS
N1173K	cbEGF14	incomplete MFS

References

Astermark, J., Hogg, P. J., Björk, I. & Stenflo, J. (1992). Effects of γ-carboxyglutamic acid and epidermal growth factor-like modules of factor IX on factor X activation. *J. Biol. Chem.* **267**, 3249-3256.

Cardy, C. M. & Handford, P. A. (1998). Metal ion dependency of microfibrils supports a rod-like conformation for fibrillin-1 calcium binding epidermal growth factor-like domains. *J. Mol. Biol.* **276**, 855-860.

Collod-Béroud, G., Béroud, C., Adès, L., Black, C., Boxer, M., Brock, D. J., Godfrey, M., Hayward, C., Karttunen, L., Milewicz, D., Peltonen, L., Richards, R. I., Wang, M., Junien, C. & Boileau, C. (1997). Marfan database (second edition): software and database for the analysis of mutations in the human FBN-1 gene. *Nucl. Acids Res.* **25**, 147-150.

Collod-Béroud, G., Béroud, C., Adès, L., Black, C., Boxer, M., Brock, D. J. H., Holman, K. J., de Paepe, A., Francke, U., Grau, U., Hayward, C., Klein, H. G., Liu, W., Nuytinck, L., Peltonen, L., Perez, A. B., Rantamäki, T., Junien, C. & Boileau, C. (1998). Marfan database (third edition): new mutations and new routines for the software. *Nucl. Acids Res.* **26**, 229-233.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Glanville, R. W., Qian, R., McClure, D. W. & Maslen C. L. (1994). Calcium binding, hydroxylation, and glycosylation of the precursor epidermal growth factor-like domains of fibrillin-1, the Marfan gene protein. *J. Biol. Chem.* **269**, 26630-26634.

Godfrey, M., Raghunath, M., Cisler, J., Bevins, C. L., de Paepe, A., Di Rocco, M., Gregoritch, J., Imaizumi, K., Kaplan, P., Kuroki, Y., Silberbach, M., Superti-Furga, A., Van Thienen, M. N., Vetter, U. & Steinmann, B. (1995). Abnormal morphology of fibrillin microfibrils in fibroblast cultures from patients with neonatal Marfan syndrome. *Am. J. Pathol.* **146**, 1414-1421.

Handford, P. A., Mayhew, M., Baron, M., Winship, P. R., Campbell, I. D. & Brownlee, G. G. (1991). Key residues involved in calcium binding motifs in EGF-like domains. *Nature.* **351**, 164-167.

Handford, P. A., Downing, A. K., Rao, Z., Hewett, D. R., Sykes, B. C. & Kielty, C. M (1995). The calcium binding properties and molecular organisation of EGF-like domains in human fibrillin-1. *J. Biol. Chem.* **270**, 6751-6756.

Hughes, P. E., Morgan, G., Rooney, E. K., Brownlee, G. G. & Handford, P. (1993). Tyrosine 69 of the first epidermal growth factor-like domain of human factor IX is essential for clotting activity. *J. Biol. Chem.* **268**, 17727-17733.

Hwang, T. L. & Shaka, A. J. (1995). Water suppression that works; excitation sculpting using arbitrary wave forms and pulsed-field gradients. *J. Magn. Reson.* **112**, 275-279.

Kettle, S., Yuan, X., Grundy, G., Knott, V., Downing, A. K. & Handford, P. A. (1999). Defective calcium binding to fibrillin-1: consequence of an N2144S change for fibrillin-1 structure and function. *J. Mol. Biol.* **285**, 1277-1287.

Kielty, C. M. & Shuttleworth, C. A. (1993). The role of calcium in the organisation of fibrillin molecules. *FEBS Letters.* **336**, 323-326.

Knott, V., Downing, A. K., Cardy, C. M. & Handford, P. (1996). Calcium binding properties of an EGF-like domain pair from human fibrillin-1. *J. Mol. Biol.* **255**, 22-27.

Lönnqvist, L., Child, A., Kainulainen, K., Davidson, R., Puhakka, L. & Peltonen, L. (1994). A novel mutation of the fibrillin gene causing ectopia lentis. *Genomics.* **19**, 573-576.

Nijbroek, G., Sood, S., McIntosh, I., Francomano, C. A., Bull, E., Pereira, L., Ramirez, F., Pyeritz, R. E. & Dietz, H. C. (1995). Fifteen novel FBN-1 mutations causing Marfan syndrome detected by heteroduplex analysis of genomic amplicons. *Am. J. Hum. Genet.* **57**, 8-21.

Park, E. S., Putnam, E. A., Chitayat, D., Child, A. & Milewicz, D. M. (1998). Clustering of FBN2 mutations in patients with congenital contractural arachnodactyly indicates an important role of the domains encoded by exons 24 through 34 during human development. *Am. J. Med. Genet.* **78**, 350-355.

Persson, E., Selander, M., Linse, S., Drakenberg, T., Ohlin, A-K. & Stenflo, J. (1989). Calcium binding to the isolated β-hydroxyaspartic acid-containing epidermal growth factor-like domain of bovine factor X. *J. Biol. Chem.* **264**, 16897-16904.

Putnam, E. A., Cho, M., Zinn, A. B., Towbin, J. A., Byers, P. H. & Milewicz, D. M. (1996). Delineation of the Marfan phenotype associated with mutations in exons 23-32 of the FBN-1 gene. *Am. J. Med. Genet.* **62**, 233-242.

Rand, M. D., Lindblom, A., Carlson, J., Villoutreix, B. O. & Stenflo, J. (1997). Calcium binding to tandem repeats of cbEGF-like modules. Expression and characterisation of the EGF-like modules of human Notch-1 implicated in receptor-ligand interactions. *Protein Sci.* **6**, 2059-2071.

Rao, Z., Handford, P. A., Mayhew, M., Knott, V., Brownlee, G. G. & Stuart, D. (1995). The structure of a Ca^{2+} binding epidermal growth factor-like domain: its role in protein-protein interactions. *Cell.* **82**, 131-141.

Reinhardt, D. P., Keene, D. R., Corson, G. M., Pöschl, E., Bächinger, H. P., Gambee, J. E. & Sakai, L. Y. (1996). Fibrillin-1: organisation in microfibrils and structural properties. *J. Mol. Biol.* **258**, 104-116.

Reinhardt, D. P., Mechling, D. E., Boswell, B. A., Keene, D. R., Sakai, L. Y. & Bachinger, H. P. (1997). Calcium determines the shape of fibrillin. *J. Biol. Chem.* **272**, 7368-7373.

Sali A. & Blundell T. L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol.* **286**, 661-668.

Stenberg, Y., Linse, S., Drakenberg, T. & Stenflo J. (1997a). The high affinity calcium-binding sites in the epidermal growth factor module region of vitamin K-dependent protein S. *J. Biol. Chem.* **272**, 23255-23260.

Stenberg, Y., Julenius, K., Dahlqvist, I., Drakenberg, T. & Stenflo, J. (1997b). Calcium-binding properties of the third and fourth epidermal-growth-factor-like modules in vitamin-K-dependent protein S. *Eur. J. Biochem.* **248**, 163-170.

Stenberg, Y., Muranyi, A., Steen, C., Thulin, E., Drakenberg, T. & Stenflo J. (1999). EGF-like module pair 3-4 in vitamin K-dependent protein S: Modulation of calcium affinity of module 4 by module 3, and interaction with factor X. *J. Mol. Biol.* **293**, 653-665.

Sunnerhagen, M. S., Persson, E., Dahlqvist, I., Drakenberg, T. & Stenflo, J. (1993). The effect of aspartate hydroxylation on calcium binding to epidermal growth factor-like modules in coagulation factors IX and X. *J. Biol. Chem.* **268**, 23339-23344.

Superti-Furga, A., Raghunath, M. & Willems, P. J. (1992). Deficiencies of fibrillin and decorin in fibroblast cultures of a patient with neonatal Marfan syndrome. *J. Med. Genet.* **29**, 875-878.

Thurmond, F. A. & Trotter, J. A. (1996). Morphology and biomechanics of the microfibrillar network of sea cucumber dermis. *J. Exp. Biol.* **199**, 1817-1828.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (2000). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296**, 1065-1078.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. Backbone dynamics of an apo cbEGF domain pair: Implications for the role of calcium in human fibrillin-1. Manuscript in preparation.

Wess, T. J., Purslow, P. P., Sherratt, M. J., Ashworth, J., Shuttleworth, C. A. & Kielty, C. M. (1998). Calcium determines the supramolecular organisation of fibrillin-rich microfibrils. *J. Cell Biol.* **141**, 829-837.

Whiteman, P., Downing, A. K., Smallridge, R., Winship, P. R. & Handford, P. A. (1998a). A Gly to Ser change causes defective folding *in vitro* of calcium binding epidermal growth factor-like domains from factor IX and fibrillin-1. *J. Biol. Chem.* **273**, 7807-7813.

Whiteman, P., Downing, A. K. & Handford, P. A. (1998b). NMR analysis of cbEGF domains gives new insights into the structural consequences of a P1148A substitution in fibrillin-1. *Protein Eng.* **11**, 957-959.

Yuan, X., Downing, A. K., Knott, V., & Handford, P. A. (1997). Solution structure of the transforming growth factor β-binding protein-like module, a domain associated with matrix fibrils. *EMBO J.* **16**, 6659-6666.

Yuan, X., Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (1998). Effects of proline *cis-trans* isomerisation on TB domain secondary structure. *Protein Sci.* **7**, 2127-2135.

Yuan, X., Lack, J., Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K.. Effects of the N2144S mutation on backbone dynamics of a TB-cbEGF domain pair from human fibrillin-1. Manuscript in preparation.

Chapter 5

Determination of the Solution Structure of cbEGF12-13

The cbEGF12-13 domain pair from human fibrillin-1 includes residues D1070-I1154 of the intact molecule (numbering according to Pereira *et al.*, 1993). In both chapters 5 and 6, these residues are referred to as D4-I88.

5.1 NMR Data Acquisition and Spectral Assignment of cbEGF12-13

5.1.1 NMR Sample Preparation

All samples for NMR data acquisition were adjusted to pH 6.5 and contained 20mM $CaCl_2$, with samples for heteronuclear experiments also containing 4.55mM Tris•HCl. At 20mM $CaCl_2$, in the absence of additional salt, both calcium binding sites of the pair were considered saturated based on previous calcium binding studies (Chapter 4; Smallridge *et al.*, 1999). Furthermore, 2D HSQC spectra (Kay *et al.*, 1992) recorded on a ^{15}N-labelled sample containing 10, 12, 14, 17 and 20mM $CaCl_2$ indicated that saturation was achieved with 20mM $CaCl_2$. Samples for homonuclear NMR experiments contained approximately 3mM or 3.6mM cbEGF12-13 in 90% H_2O/10% ^2H_2O or 99.996% ^2H_2O, respectively. The ^{15}N-labelled sample used to obtain both 3D data and HMQC-J data for use in structure calculations contained

3.8mM cbEGF12-13 in 90% H_2O/10% ^2H_2O. Samples for acquiring NH-exchange data contained ~1.6mM ^{15}N-labelled cbEGF12-13 in 99.996% ^2H_2O for ^2H_2O exchange experiments, and ~1mM ^{15}N-labelled cbEGF12-13 in 90% H_2O/10% ^2H_2O for H_2O exchange experiments. The concentrations of samples were estimated from $O.D._{280}$ measurements using ε_{280} = 3280$M^{-1}cm^{-1}$. The cbEGF12-13 pair has an isoelectric point of 4.27 and a molecular weight of 9646.86 units.

5.1.2 Acquisition and Processing of NMR Data for Spectral Assignment and Structure Calculations

Initial assignments were made using 3D gradient-enhanced ^{15}N-separated NOESY-HSQC spectra (Kay *et al.*, 1989; Kay *et al.*, 1992), recorded at 15 and 33°C on a home-built/GE Omega spectrometer operating at 600MHz. The spectrometer was fitted with a triple resonance probe with self-shielded pulsed field gradients. Two different temperatures were used to enable spectral overlap to be resolved, and ultimately 33°C data were used to derive restraints for structure calculations. 3D NOESY spectra were recorded over approximately 3 days, with acquisition times of 102ms in the direct ^1H (F_3) dimension, 10ms in the ^{15}N (F_2) dimension and within a range of 21.2-25.6ms in the indirect ^1H (F_1) dimension. All NOESY spectra were recorded with a mixing time of 150ms, and linear prediction was used to double the F_2 acquisition time to 20ms (Zhu & Bax, 1990). Gradient enhanced ^{15}N-separated TOCSY-HSQC spectra (Kay *et al.*, 1992) were also recorded on the same spectrometer at 15 and 33°C to enable identification of intraresidue cross-peaks. These spectra were recorded over approximately three days with the same acquisition times as for the NOESY, except for a 16ms acquisition time in the indirect ^1H (F_1) dimension. All TOCSY spectra were recorded with τ_m=46ms.

Assignments were confirmed by comparison of homonuclear 2D NOESY spectra (Jeener *et al.*, 1979; Macura *et al.*, 1981) with mixing times of 150ms, 2D

HOHAHA spectra (Aue *et al.*, 1976; Brown *et al.*, 1988) with τ_m=46ms, and 2D COSY spectra (Braunschweiler & Ernst, 1983; Bax & Davis, 1985). These spectra were recorded at 28 and 33°C, at field strengths of 600 and 750MHz. Spectral overlap due to sequence and structure repetition in the cbEGF12-13 domain pair frequently hindered the assignment process, hence NMR data were recorded at different temperatures and proton frequencies to help improve resolution. A relayed 2D NOESY experiment (Wagner, 1984) was also utilised for the elucidation of NH-NH sequential connectivities (see below).

^{1}H-^{15}N HMQC-J spectra (Kay & Bax, 1990) were recorded at 33°C at field strengths of 500 and 750MHz to allow derivation of $^{3}J_{HN\text{-}H\alpha}$ coupling constants. These spectra were recorded with acquisition times of 82ms in the direct ^{1}H (F_2) dimension and 261 and 254ms in the ^{15}N (F_1) dimension, respectively. Data recorded at 500MHz were subsequently used for HMQC-J analysis, and these data were processed using two different window functions in the ^{15}N (F_1) dimension, i.e., 1Hz exponential line narrowing and a sinebell function with 0° phase shift. These window functions were selected since they produced well-resolved splittings whilst maintaining adequate signal-to-noise. J-coupling data were subsequently extracted as described in section 2.3.2.2.

NH-exchange data in $^{2}H_2O$ were obtained by recording a series of ^{1}H-^{15}N HSQC spectra (Kay *et al.*, 1992) at 33°C after a fully protonated cbEGF12-13 sample had been dissolved in a $^{2}H_2O$ solution (Barsukov & Lian, 1993). NH-exchange data in H_2O at 33°C were obtained using the method of Böckmann & Guittet (1997; see section 2.3.2.3), using a 100ms time-scale for NH-exchange. This value was chosen by comparing data obtained using three different time-scales for exchange (10, 50 and 100ms) with secondary structure predicted from both homology modelling (section 5.1.4) and from initial structures calculated in the absence of these data. Hydrogen bonding restraints derived using these methods were only incorporated where

consistent with both NOE data and initial structures calculated in the absence of these data.

NOE intensity calibrations and all restraints for structure calculations were obtained from NMR spectra acquired at 33°C and pH 6.5. Felix 2.3 (Biosym, Inc.) was used to process all NMR data.

5.1.3 Spectral Assignment

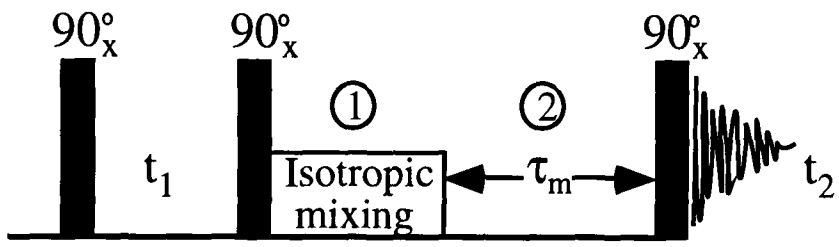
Several methods, in addition to recording NMR data at different temperatures/proton frequencies, were employed in this project to aid spectral assignment of the cbEGF12-13 domain pair. Initially attempts were made to sequentially assign residues of cbEGF12-13 by comparing strips taken from within the amide region of 3D ^{15}N-edited NOESY-HSQC spectra (Kay *et al.*, 1989; Kay *et al.*, 1992) using XEASY (Bartels *et al.*, 1995). This program allows identification of strong connectivities between strips, indicating that the corresponding residues may be sequentially positioned in a peptide chain. A problem with this feature, however, arises because identification of strong NH-NH connectivities between strips can imply either that the corresponding residues are sequentially linked or that they are closely positioned in three-dimensional space, e.g., on different strands of a β-sheet (Table 2.2). Since, by homology to cbEGF32-33 from human fibrillin-1 (Downing *et al.*, 1996), the cbEGF12-13 domain pair was predicted to be mainly composed of β-structure, using XEASY confused the initial assignment process.

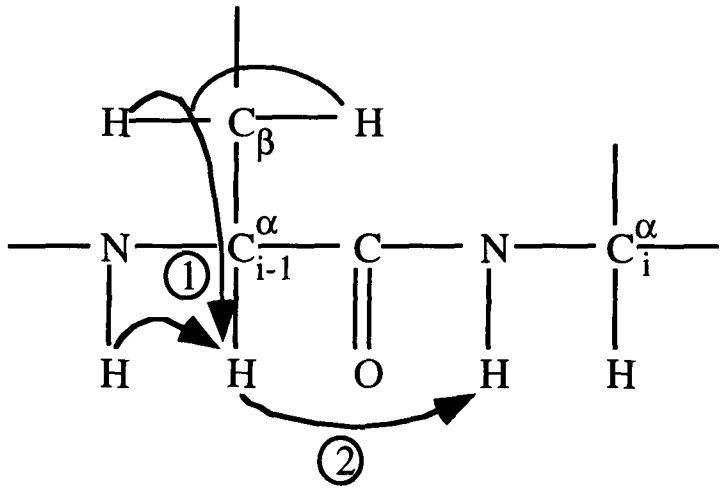
The structure of the cbEGF12-13 domain pair was subsequently modelled based on the coordinates of cbEGF32-33 (Downing *et al.*, 1996; see section 5.1.4), and this model, combined with the use of paper plots of NMR spectra, allowed main-chain directed assignments to be made. Knowledge of the predicted β-sheet structure of

cbEGF12-13 enabled identification of strong $d_{\alpha\alpha(i,j)}$, $d_{NN(i,j)}$ and $d_{\alpha N(i,j)}$ inter-strand connectivities (Table 2.2), allowing initial misassignment to be corrected.

A relayed 2D NOESY spectrum (Wagner, 1984) was also recorded to clarify $d_{NN(i,i+1)}$ and $d_{\beta N(i,i+1)}$ connectivities for residues in β-sheet structures, since these connectivities are often weak or not observed. In collaboration with Dr. Jonathan Boyd, a relayed NOESY pulse sequence was designed based on the method of Cavanagh & Rance (1990), with the aim of increasing the intensities of these NOE cross-peaks (Figure 5.1).

Figure 5.1 : The Pulse Sequence and Magnetisation Transfer Pathway of a Relayed NOESY Experiment





In this pulse sequence, the resonance positions of the NH_{i-1} and βH_{i-1} are frequency labelled during t_1. The DIPSI isotropic mixing period then transfers this magnetisation onto αH_{i-1}, before the NOE mixing period transfers it to NH_i (Figure 5.1). The result is intense sequential βH_{i-1}-NH_i and NH_{i-1}-NH_i cross-peaks, which result from indirect magnetisation transfer via αH_{i-1} and only appear on one side of the

diagonal. Since the cross-peak intensities are no longer directly dependent on the distance between the protons, these data should be used only for assignment purposes (Cavanagh *et al.*, 1996). A relayed 2D NOESY spectrum obtained of cbEGF12-13 did demonstrate increased cross-peak intensities for the connectivities discussed, however, it provided no additional data so its use in spectral assignment in this case was limited.

Once the majority of cross-peaks had been assigned using paper plots, the cross-peak labellings for both a 2D NOESY spectrum recorded on an unlabelled sample in 99.996% $^{2}H_2O$ and a 3D ^{15}N-edited NOESY-HSQC spectrum recorded on a ^{15}N-labelled sample in 90% H_2O/10% $^{2}H_2O$ were input into the program NMRView (Johnson & Blevins, 1994). These spectra were both recorded with a mixing time of 150ms at 33°C and pH 6.5. NMRView was used to assist with further NOE assignment, based on proton chemical shift values (see Appendix V) and quantitation of cross-peak intensities.

A further aid to spectral assignment was a 2D NOESY spectrum of the calcium saturated G1127S mutant cbEGF12-13 domain pair (see Chapter 7). The N- and C-termini of wild-type cbEGF12-13 were initially misassigned, due to spin system pattern similarity between these regions. Assignments for the regions D4-R9 and N83-I88 were interchanged upon inspection of mutant domain pair data.

Once spectral assignment was completed, it was observed that a number of peaks remained unidentified. For example, the 3D ^{15}N-edited NOESY-HSQC spectrum at 33°C contained approximately twenty strips more than were expected, although these "remainder" strips contained only weak resonances. In total less than ~5% of all NOE peaks remained unassigned, and four possible reasons were identified to account for the presence of these extra peaks,

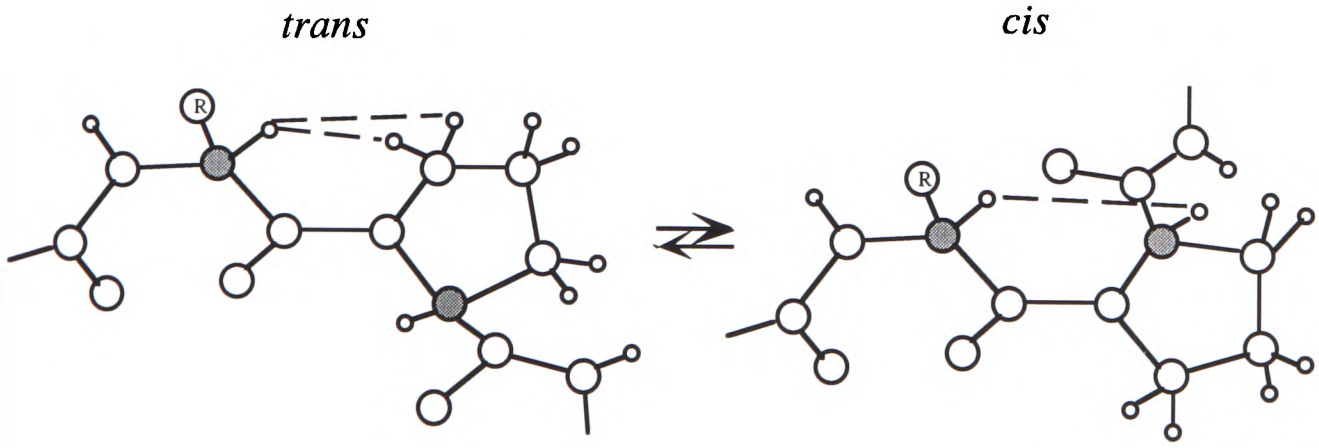
1) The region comprising residues G38-K43, inclusive, located around the top of the minor β-sheet of cbEGF12, has been shown to be relatively mobile based on heteronuclear ^{1}H-^{15}N NOE relaxation data recorded at 33°C at both 500 and 750MHz (see Chapter 6). The hypothesis that this region is flexible is supported by the

observation of a lower than average number of experimental restraints for several residues in this region (section 5.3). It was thought that this intrinsic flexibility could be responsible for a proportion of the weak peaks that remain unassigned, however, since this flexibility is on a fast time-scale in relation to the NMR experimental time-scale, this seems unlikely.

2) For a dipeptide, X-Pro, a *cis*-peptide bond is identified by strong NOE cross-peaks between αH_i and αH_{i+1} and between NH_i and αH_{i+1}, whereas a *trans*-peptide bond is identified by strong NOE cross-peaks between αH_i and $\delta CH_{2(i+1)}$ and between NH_i and $\delta CH_{2(i+1)}$ (Figure 5.2; Wüthrich, 1986). During assignment of cbEGF12-13, five of the six proline residues were assigned as *trans*, with P75 being *cis*. It was observed, however, that P76 appeared to exist in two conformations, i.e., a major *trans*- and minor *cis*-conformation. P76 is preceded by P75, therefore no strips were available for identifying connectivities from NH_i, but, despite overlap, a 2D NOESY spectrum highlighted the presence of intense αH_i and $\delta CH_{2(i+1)}$ connectivities and also a weak αH_i and αH_{i+1} connectivity (Figure 5.3).

Figure 5.2 : The Peptide Bond of a X-Pro Dipeptide Can Adopt *Cis* or *Trans* Conformations in Proteins

Shading is used to highlight Cα atoms, and dashed lines show the αH_i and $\delta CH_{2(i+1)}$ connectivities seen for *trans* peptide bonds and the αH_i and αH_{i+1} connectivity seen for *cis* peptide bonds. This diagram is modified from Wüthrich, 1986.



The theory that P76 exists in major and minor conformations is further supported by the behaviour of its neighbouring residue, G77. In 2D ^{1}H-^{15}N HSQC

Figure 5.3 : 2D NOESY Spectrum of cbEGF12-13 in ^2H_2O at 33°C Highlighting the Major *Trans* and Minor *Cis* Form of Proline 76

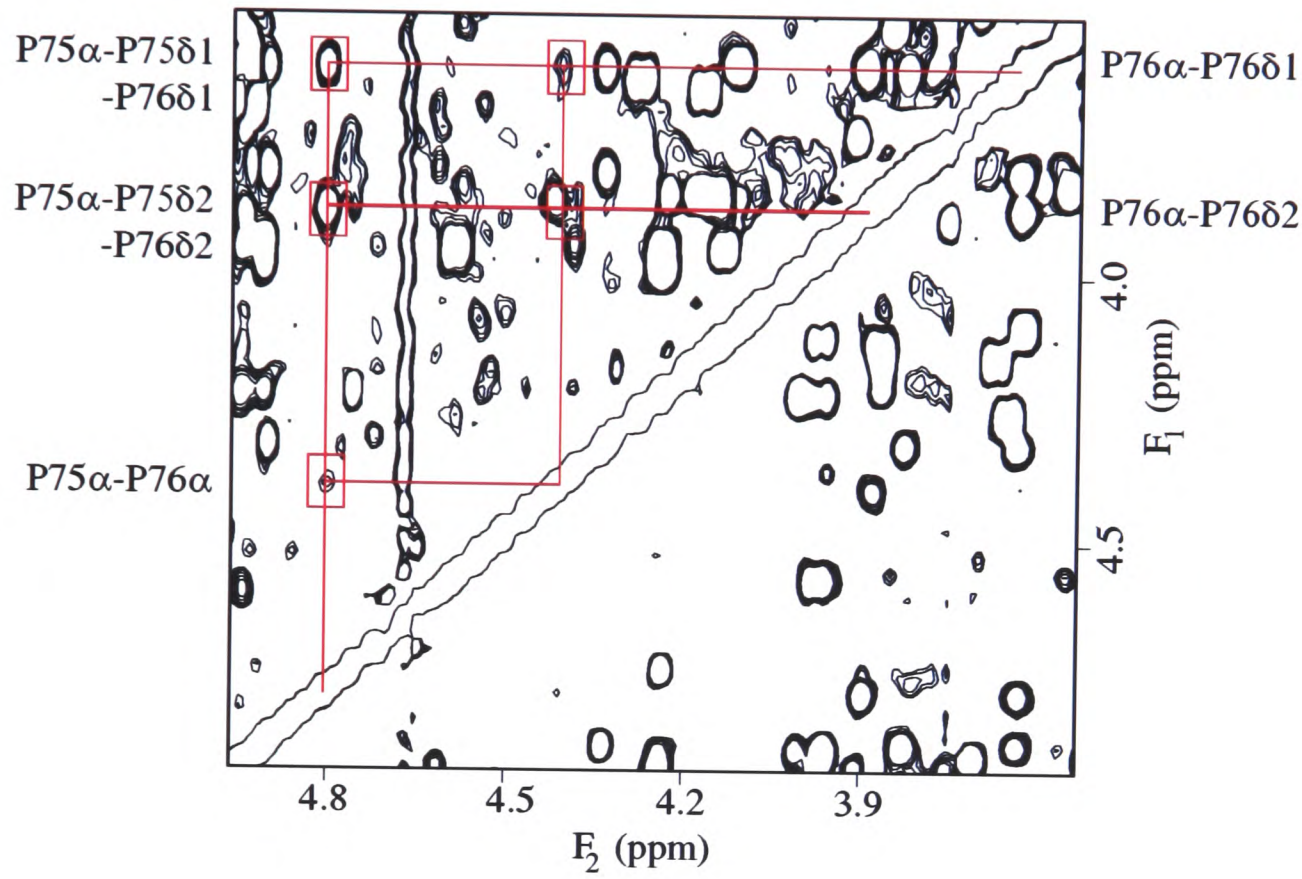
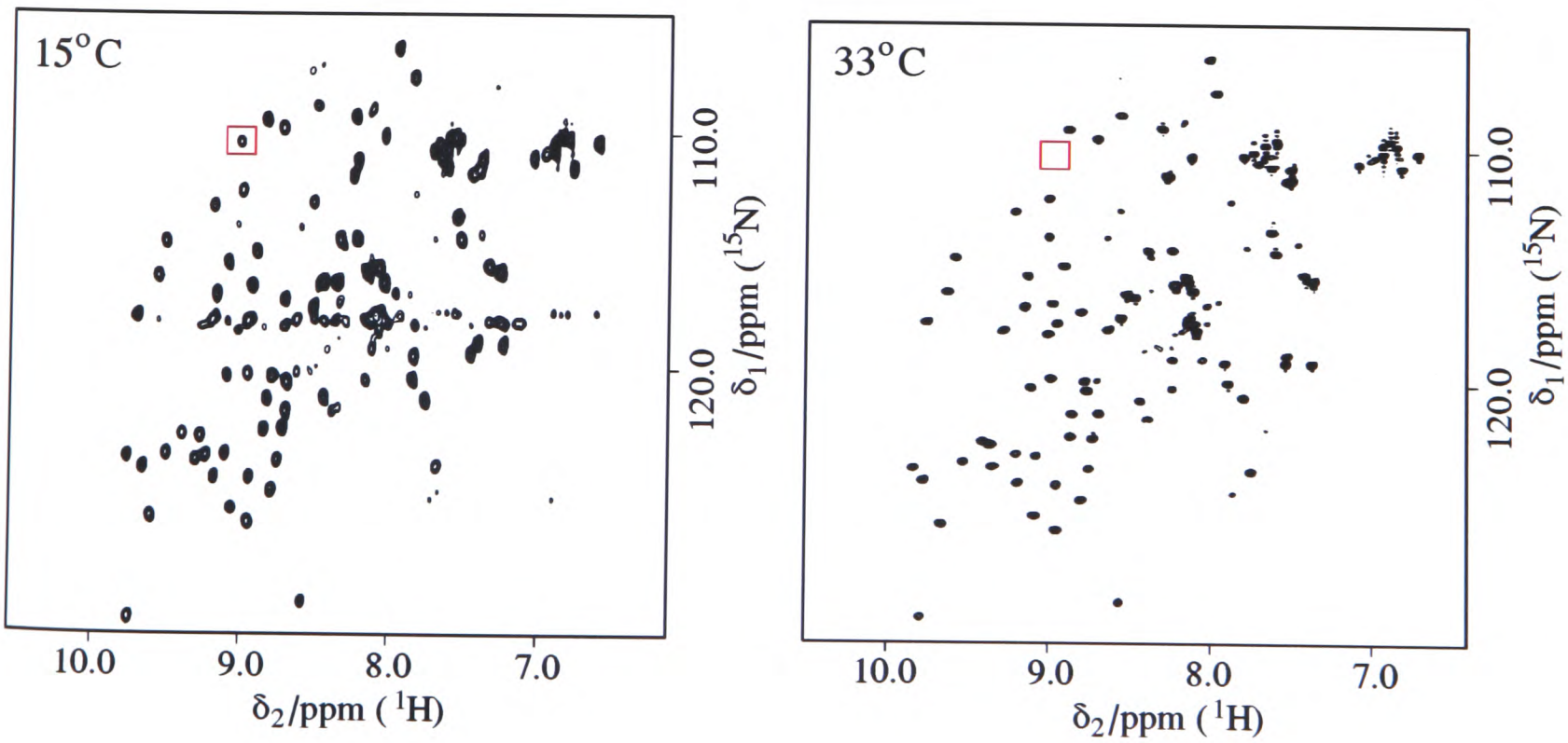


Figure 5.4 : 2D ^1H-^{15}N HSQC Spectra of the cbEGF12-13 Domain Pair at 15°C and 33°C

Red boxes highlight the position of the G77 resonance in both spectra.



spectra, the resonance for G77 cannot be observed at 33°C, whereas at 15°C this resonance can be seen clearly (Figure 5.4). This implies that at 33°C, G77 is flexible on the millisecond time-scale, whereas at 15°C it is stabilised in the major P76 *trans* conformation of cbEGF12-13. Attempts were made to assign the minor *cis*-P76 form of cbEGF12-13 to obtain restraints for use in separate structure calculations, a method used successfully for the TB6 domain from human fibrillin-1 (Yuan *et al.*, 1998). The extra peaks observed in this case were, however, too weak to be assigned with confidence.

3) A proportion of the extra peaks, i.e., those which correspond to distances only slightly greater than 5Å, may result from spin diffusion as a result of the 150ms NOESY mixing times.

4) High sample concentrations of cbEGF12-13 were used in this project to acquire 2D and 3D NOESY spectra with good signal-to-noise ratios for use in spectral assignment and structure calculations, i.e., 3.6mM and 3.8mM, respectively. It is possible that these high protein concentrations allowed multiple conformations of the cbEGF12-13 construct to be observed, particularly since this sample was produced by *in vitro* refolding and a single HPLC peak has previously been shown to contain multiple construct conformations (Yuan *et al.*, 1998).

The final ^{1}H-^{15}N spectral assignments for the cbEGF12-13 domain pair are shown in Figure 5.5 on a 2D HSQC spectrum, and an example of sequentially assigned strips extracted from a 3D ^{15}N-edited NOESY-HSQC spectrum is shown in Figure 5.6.

5.1.4 Modelling the Structure of the cbEGF12-13 Domain Pair

The structure of the cbEGF12-13 domain pair was modelled using the coordinates of the cbEGF32-33 domain pair from human fibrillin-1 (Downing *et al.*, 1996) and Modeller4 (Sali & Blundell, 1993). The sequences of cbEGF12-13 and cbEGF32-33 were aligned using the program ClustalX (Figure 5.7; Thompson *et al.*,

Figure 5.5 : 2D ^{1}H-^{15}N HSQC Spectrum of cbEGF12-13 in H_2O Recorded at pH 6.5, 33°C

NH resonances are labelled according to residue type and sequence number for main chain nitrogens. The NH_2 side chain cross-peaks are joined by horizontal lines. Positive and negative contours are shown with solid and dashed lines, respectively.

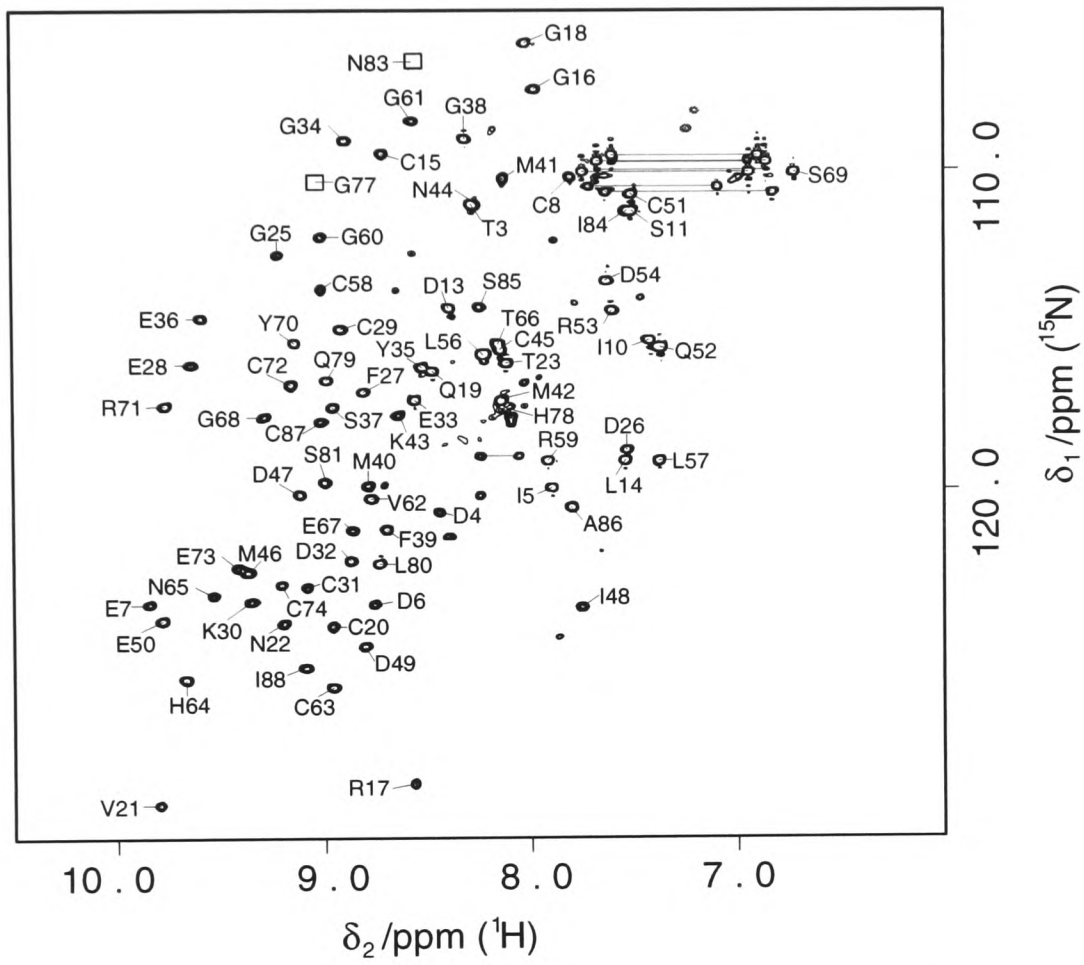


Figure 5.6 : Strips Extracted From a 3D ^{15}N-NOESY-HSQC Spectrum of cbEGF12-13 For Sequential Residues D13-C20

Intraresidue and sequential peaks for these residues are connected by horizontal lines.

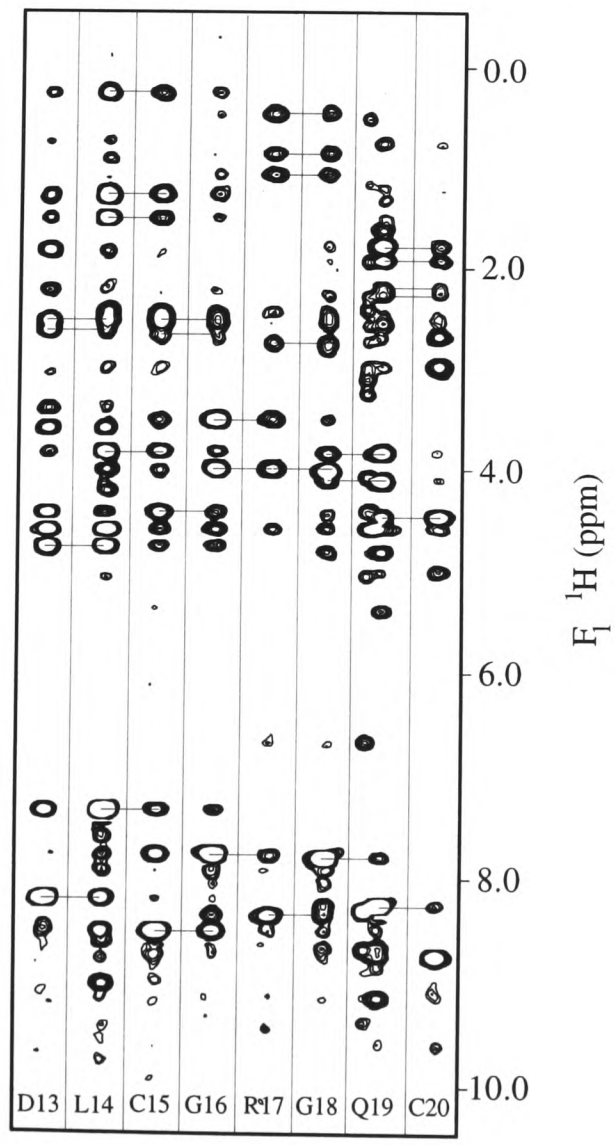


Figure 5.7 : ClustalX Alignment of the Sequences of cbEGF12-13 and cbEGF32-33 from Human Fibrillin-1

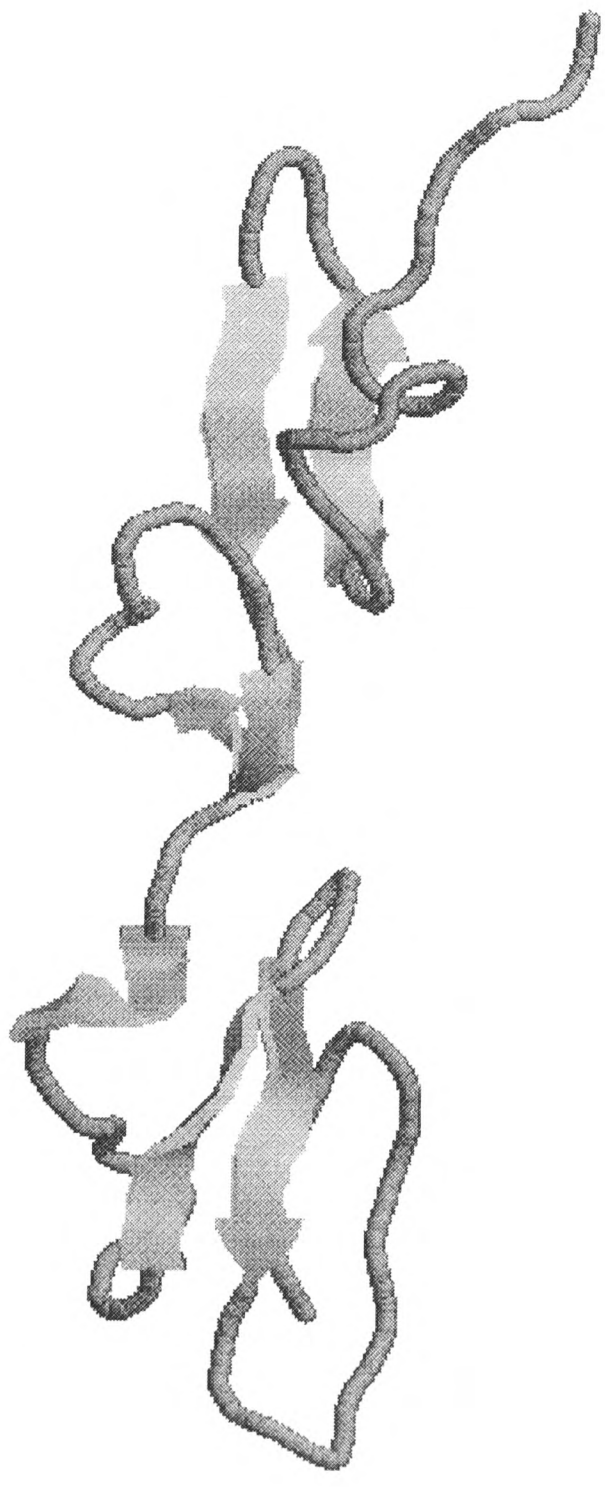
"*" indicates a fully conserved residue, ":" represents a strongly conserved residue and "." represents conservation within a weakly conserved group (for full details see Thompson *et al.*, 1997).

```
cbEGF12 : DIDECRISPDLCGRGQCVNTPGDFECKCDEGYESGFMMMKNCM
cbEGF32 : DMDECK-EPDVCKHGQCINTDGSYRCECPFGYI---LAGNECV
          *:***: .**:* :***:** *.:.*:*  **    :  ::*:

cbEGF13 : DIDECQR-DPLLCRGGVCHNTEGSYRCECPPGHQLSPNISACI
cbEGF32 : DTDECSVGNP--CGNGTCKNVIGGFECTCEEGFEPGPMMT-CE
          * ***.  :*  * .*.*:*. *.:.* *  *.: .* :: *
```

Figure 5.8 : A Model of the Structure of the cbEGF12-13 Domain Pair from Human Fibrillin-1

This model was based on the coordinates of the cbEGF32-33 domain pair from human fibrillin-1 (Downing *et al.*, 1996), and was calculated using Modeller4 (Sǎli & Blundell, 1993).
Calcium atoms were omitted.



1997), and this sequence alignment, together with the PDB file for cbEGF32-33, was used by Modeller4 to calculate 100 models. The best models were selected both on the basis of structural convergence, as visualised by structural overlays in MOLMOL (Koradi *et al.*, 1996; 1998), and from a variety of Modeller4 outputs, including minimum number of violated restraints. These selection criteria resulted in seven models being selected, and these were further analysed by PROCHECK (Laskowski *et al.*, 1993). A single model was selected on the basis of secondary structural content, with the final model containing six β-strands and an α-helix (Figure 5.8). This model did not contain protons and was not energy minimised, as neither were deemed necessary for its intended applications. The model proved invaluable to the progress of main-chain directed assignments, and was also useful when analysing violated restraints during initial rounds of structure calculations.

5.2 Conversion of NMR Data to Structural Restraints

5.2.1 NOE Data

The intensities of cross-peaks in 2D and 3D NOESY spectra labelled in NMRView were calibrated into four distance restraint categories, 2.8Å, 3.5Å, 5.0Å and 6.0Å, for use in structure calculations. The 2.8Å/3.5Å boundary was calculated as the average intensity of well-resolved NOEs for interproton interactions in regions of secondary structure (Table 2.2), i.e., $d_{\alpha\alpha(i,j)}$ and $d_{\alpha N(i,i+1)}$ NOEs for 2D and 3D data, respectively. Secondary structure was also exploited to calculate the 3.5Å/5.0Å boundary for 3D data, i.e., $d_{NN(i,j)}$ connectivities were used, whilst calibration of this boundary for homonuclear 2D data in ^2H_2O relied on knowledge of chemically fixed distances. For these 2D data, averages were taken of NOE cross-peak intensities for well-resolved δH*-αH and εH*-αH aromatic ring proton connectivities. A 6.0Å class

was introduced so that, if necessary, all distance classes of violated restraints could be "relaxed" by one distance group. Distance restraints thought to be correctly assigned from spectral analysis, but violated in structure calculations, were relaxed by one distance group to compensate for spin diffusion effects and motional averaging. These restraints remained relaxed only if they were confidently assigned and could not be reincorporated into the structures, without violation, in their unrelaxed form during the refinement process.

A total of 1892 restraints were derived from the 2D and 3D NOESY spectra (Appendix IV), including 504 intraresidue ($|i-j|=0$), 388 sequential ($|i-j|=1$), 211 short-range ($|i-j| \leq 4$), 378 long range ($|i-j| > 4$), and 411 ambiguous interproton distance restraints. The relatively high proportion of ambiguous restraints was a result of spectral overlap observed in the cbEGF12-13 NMR spectra. Of the short-range, long-range and ambiguous restraints, only 6% remained relaxed at the end of the structure determination process, and these were evenly distributed across the restraint classes.

5.2.2 HMQC-J Data

A ^{1}H-^{15}N HMQC-J spectrum, recorded at 33°C, pH 6.5 and 500MHz, used to derive $^{3}J_{HN-H\alpha}$ coupling constants is shown in Figure 5.9. Backbone ϕ torsion angles were restrained for 26 residues of cbEGF12-13, with six residues having small $^{3}J_{HN-H\alpha}$ coupling constants (<5.0Hz) and twenty having large $^{3}J_{HN-H\alpha}$ values (>8.0Hz) (Appendix IV). These data were incorporated into structure calculations as ϕ restraints with range $-60° \pm 30°$ and $-120° \pm 30°$, respectively. Several restraints in the latter class were violated during the structure calculation process, however, and were subsequently adjusted to $-100° \pm 30°$ or $-140° \pm 30°$ using $^{3}J_{HN-H\alpha}$ values and the Karplus curve (Figure 2.23). As the cbEGF domain fold contains no hydrophobic core (Rao *et al.*,

Figure 5.9 : 2D ^{1}H-^{15}N HMQC-J Spectrum of cbEGF12-13 in H_2O Recorded at pH 6.5, 33°C and 500MHz

Data shown were processed using a sinebell window function with a 0° phase shift.
Peaks used to derive phi angle data, that were incorporated into structure calculations, are labelled.

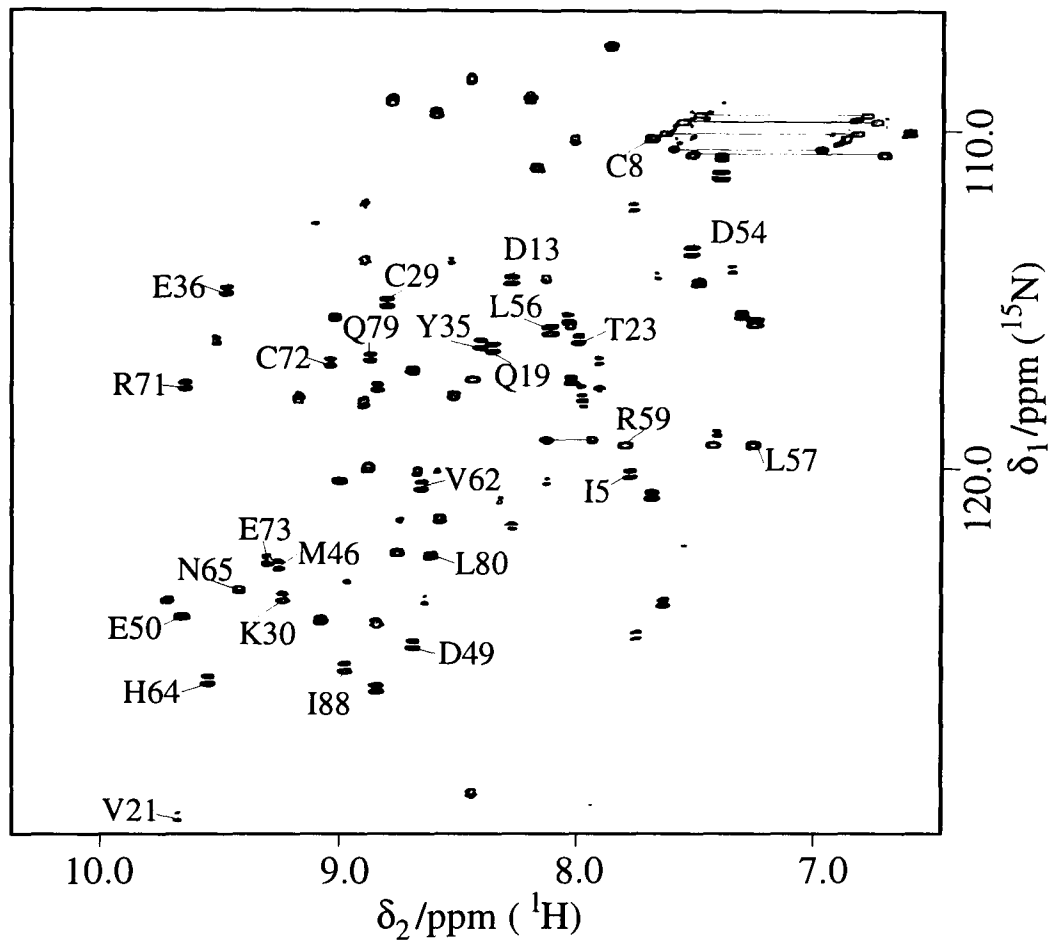
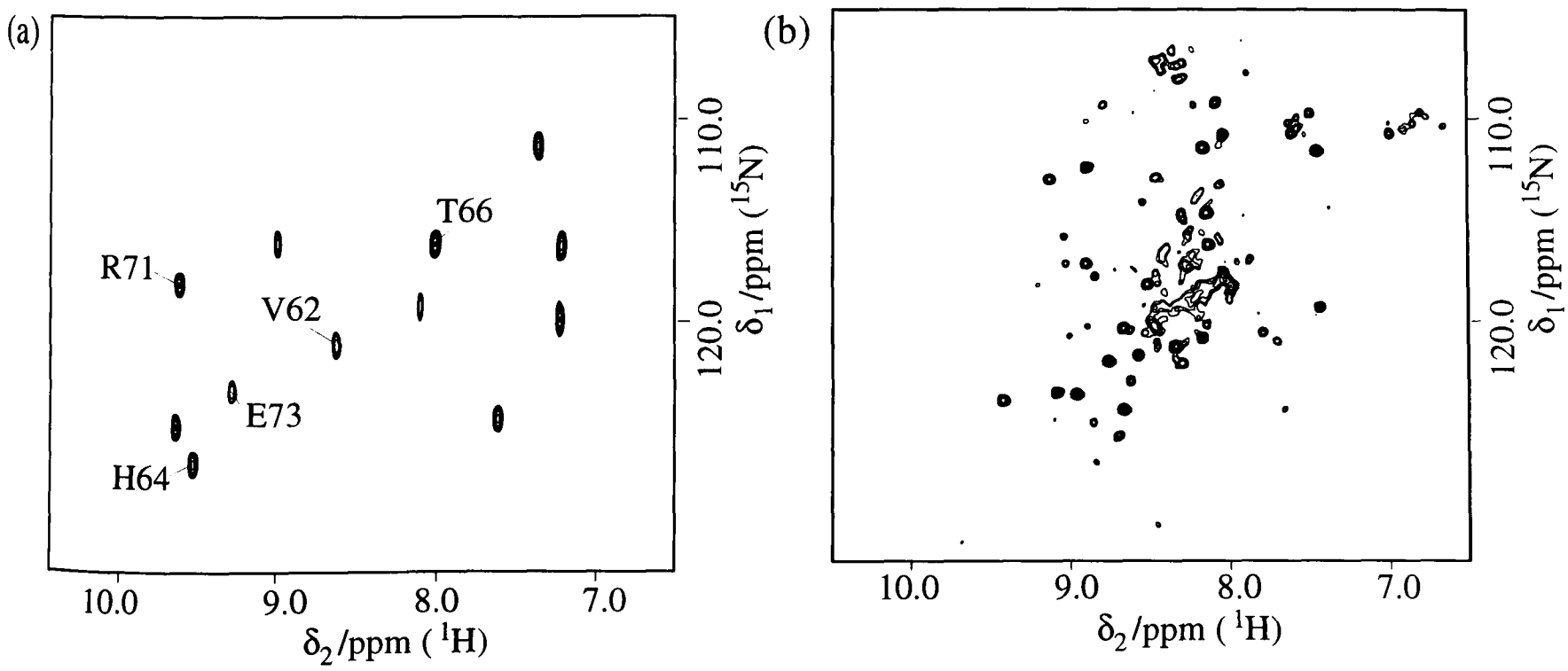


Figure 5.10 : 2D ^{1}H-^{15}N HSQC Spectra of cbEGF12-13 Recorded for NH-Exchange Analysis

The data in (a) shows NH resonances remaining 1.5 hours after a fully protonated sample was dissolved in $^{2}H_2O$. In (b), NH resonances that exchange with H_2O on the 100ms time-scale are shown. The resonances labelled in (a) and missing in (b) were incorporated into structure calculations as hydrogen bonds, once initial calculations had justified their inclusion.



1995; Downing *et al.*, 2000), backbone torsion angles are particularly important for defining its solution structure. However, ϕ angle restraints were only included in the structure calculation process where consistent with initial structures calculated in their absence.

5.2.3 NH-Exchange Data

2D ^{1}H-^{15}N HSQC spectra used to obtain NH-exchange data in both $^{2}H_2O$ and H_2O are shown in Figure 5.10. These spectra were used to identify "slowly" exchanging backbone amide protons, and those that could be assigned to regions of regular secondary structure were restrained to form HN-CO hydrogen bonds, using two distance restraints, $d_{O\text{-}N}$ = 3.3Å and $d_{O\text{-}HN}$ = 2.3Å, to prevent bond distortion (Appendix IV).

Several "fake" hydrogen bonds were also used, where NOE data supported their incorporation, during the middle of the structure calculation process to facilitate identification of incorrectly assigned NOEs. In final structures, 24 experimentally-derived distance restraints were used for 12 backbone hydrogen bonds in β-sheet structures. These restraints were incorporated into structure calculations only where consistent with both NOE data and initial structures calculated in their absence (Appendix IV).

5.2.4 Calcium Binding Restraints

In the cbEGF12-13 domain pair, six NOE restraints were used to ligate each calcium atom to oxygen atoms of the consensus calcium binding residues, as defined both by the factor IX crystal structure (Rao *et al.*, 1995) and by mutagenesis studies (Handford *et al.*, 1991). In the factor IX crystal structure, Ca-O distances

observed ranged from 2.24-2.53Å, and a hydrogen bond was observed between the carboxylate side chain of the second consensus calcium-binding residue and the amide proton of the following cysteine, forming a ASX turn (Rao *et al.*, 1995). Subsequently, to maintain correct geometry in these structure calculations, the calcium-oxygen distances were restrained to lie in the range 2.2-2.6Å, and side-chain-backbone hydrogen bonding restraints were included between D6-C8 and D49-C51, using two restraints for each bond as previously described (Appendix IV).

5.3 Structure Calculation and Refinement

To calculate the structure of the cbEGF12-13 domain pair, a modified version of XPLOR (Brünger, 1992) was used with the v4.01 topology and v4.02 parameter files (Appendix I). The modification allowed the incorporation of floating chirality protocols (kindly provided by M. Nilges, EMBL, Heidelberg). The SA structure calculation and refinement methods have been previously discussed in detail in Chapter 2, and protocols used for these calculations are shown in Appendix I. Table 5.1 summarises the parameters that were used in the SA structure calculation protocol (sa.inp) for calculation of cbEGF12-13 domain pair structure.

Sum-averaging (section 2.3.2.1.2) was used to incorporate ambiguous interproton distance restraints into the structure calculation and refinement processes (Nilges, 1995; 1997). Disulphide bonds were initially incorporated into structure calculations as NOE restraints (section 2.3.3.3.2 and Appendix I), and were only included as covalent bonds in the refinement process.

Floating chirality was used (section 2.3.2.1.3 and Appendix I and V), in which non-stereospecifically assigned valine and leucine methyl carbons, methylene group protons and glycine alpha protons could swap positions over the course of initial calculations. In the sa.inp protocol, swapping occurs five times in the high temperature conformational search phase, i.e., every 1300 steps, 20 times during slow cooling, 38

times during fast cooling, and once before the final Powell minimisation. During the cooling phases, swapping occurs every time the heat bath temperature is reduced by 50K.

After more than 100 rounds of structure calculations using sa.inp, 100 satisfactory structures were produced, and these were refined by three additional cycles of simulated annealing using "refine.inp" (Appendix I). Ultimately, a total of 1958 restraints were used to determine the structure of cbEGF12-13, i.e., an average of 22 restraints per residue. These include 504 intraresidue ($|i-j|=0$), 388 sequential ($|i-j|=1$), 211 short-range ($|i-j| \leq 4$), 378 long range ($|i-j| > 4$) and 411 ambiguous interproton distance restraints; 24 distance restraints for 12 backbone hydrogen bonds; 26 backbone torsion angle ϕ restraints; and 16 restraints to ligate calcium atoms.

Table 5.1 : Parameters Used in the SA Structure Calculation Protocol for Calculating cbEGF12-13 Domain Pair Structure

Stage	**High Temperature Conformational Search**[a]	**Cool1 (slow)**	**Cool2 (fast)**
Temperature (K)[b]	2000	2000→1000	1000→100
Energy constants[c]			
k_{NOE} (kcal mol^{-1} Å^{-2}) [d]	10/20/30/40/50	50	50
k_{ambi} (kcal mol^{-1} Å^{-2})	1.0	1→50	50
k_{hbond} (kcal mol^{-1} Å^{-2})	1.0	1→50	50
k_{rep} (kcal mol^{-1} Å^{-4})	0.02/0.04/0.06/0.08/0.1[e]	0.003→3.8	4.0
s_{rep} [f]	1.2	0.9→0.75	0.75
$k_{dihedral}$ (kcal mol^{-1} rad^{-2})	5	200	200
No. of simulation steps	6500	2500	1000

[a] The high temperature conformational search is omitted in refinement protocols.
[b] Temperature is maintained by coupling to a heat bath, with a coupling constant of 10 ps^{-1} (Berendsen *et al.*, 1984).
[c] Energy constants for bonds, angles and improper torsions are kept constant throughout the SA protocol, at 1000 kcal mol^{-1} Å^{-2}, 500 kcal mol^{-1} rad^{-2} and 500 kcal mol^{-1} rad^{-2}, respectively.
[d] A soft square-well potential with linear behaviour is used for NOE energy terms with upper limit violations greater than 1Å, with an asymptote slope of 2 (Nilges *et al.*, 1988).
[e] A reduced atom representation is used for non-bonded interactions, in which only two atoms, with van der Waals' radii of 2.25Å, are evaluated per amino acid (Folmer *et al.*, 1995).
[f] s_{rep} is the van der Waals' radius scale factor from Lennard-Jones parameters.

Final structures (25) were selected from three refinement rounds based on agreement with experimental data, with no distance violations greater than 0.3Å, no ϕ angle violations, i.e., final values for restrained ϕ angles remained within the $\pm 30^{\circ}$ boundary, and F_{NOE} less than $102 kJmol^{-1}$. An average coordinate structure was calculated and energy minimised based on all residues (Brünger, 1992), and the consensus secondary structure was derived using the program PROCHECK_NMR (Laskowski *et al.*, 1996). A summary of the NMR structural statistics for the cbEGF12-13 domain pair, in terms of agreement with experimental data and idealised covalent geometry, is shown in Table 5.2, and Figure 5.11 shows the number of unambiguous NOE restraints per residue in relation to secondary structure (see below).

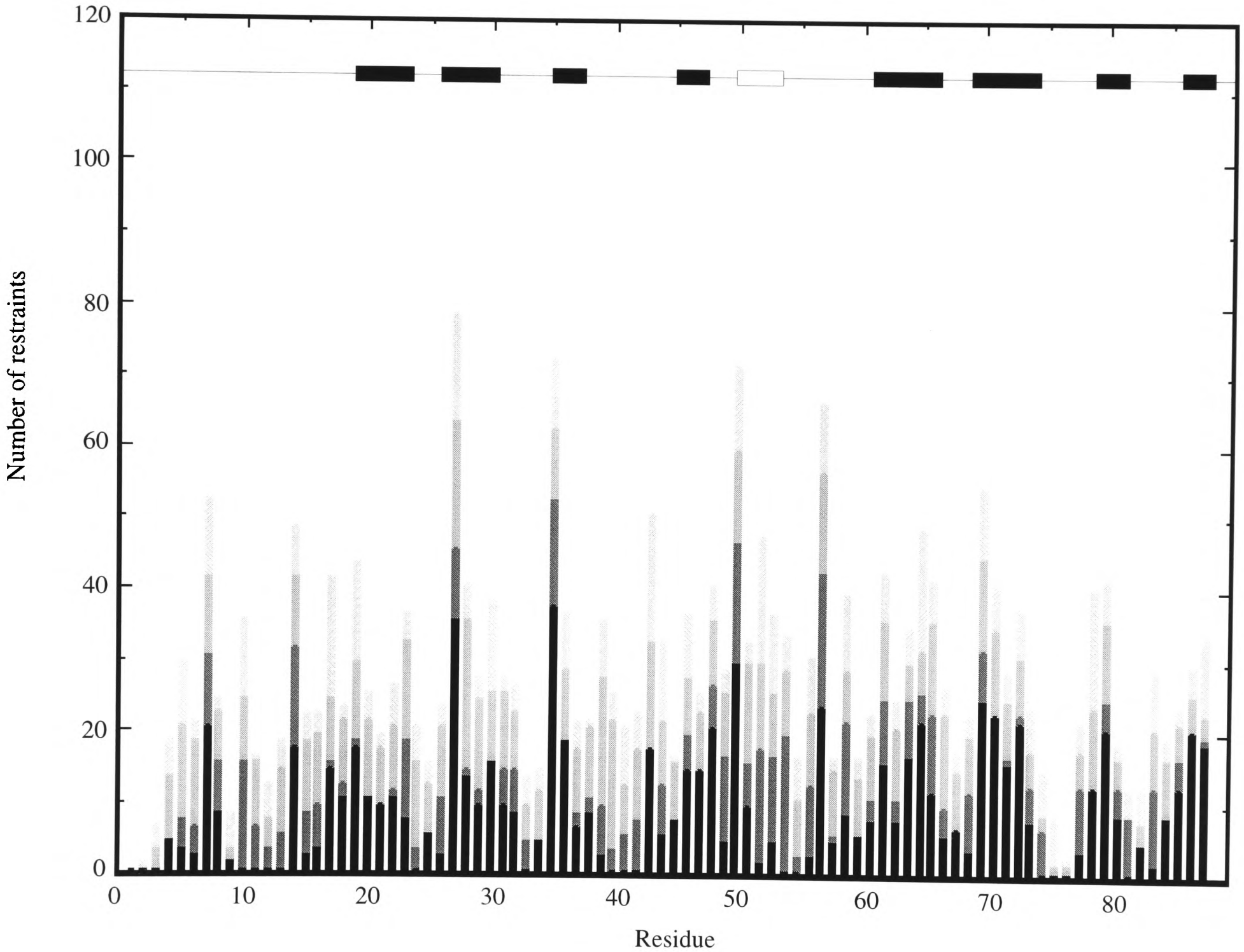
Table 5.2 : Statistics for the cbEGF12-13 Domain Pair Structure

<SA> relates to the family of 25 structures, whereas (SA) refers to the average minimised structure. The rmsd to (SA) for all heavy atoms was calculated using Insight II (version 98, MSI Inc.).

Statistic Category	<SA>	(SA)
Rmsd from experimental distance restraints (Å)		
All (1932)	0.013±0.001	0.013
Interproton distances		
Intraresidue (504)	0.009±0.002	0.001
Sequential (388)	0.011±0.002	0.001
Inter-residue short-range (|i-j|≤4) (211)	0.015±0.003	0.020
Inter-residue long-range (|i-j|>4) (378)	0.013±0.002	0.017
Ambiguous (411)	0.015±0.002	0.018
H-bonds (24)	0.014±0.004	0.013
Calcium (16)	0.012±0.005	0.003
Rmsd from experimental ϕ restraints (deg) (26)	0.177±0.093	0.203
Deviations from idealised covalent geometry		
Bonds (Å)	0.002±0.0001	0.002
Angles (deg)	0.412±0.014	0.345
Impropers (deg)	0.227±0.016	0.156
F_{NOE} (kJ mol^{-1})	64.1±13.8	28.7
$F_{dihedral}$ (kJ mol^{-1})	0.261±0.277	0.274
F_{repel} (kJ mol^{-1})	59.7±8.3	29.6
Rmsd for (C^{α}, C, N) residues 1-88 to (SA) (Å)	1.53±0.48	
Rmsd for all heavy atoms to (SA) (Å)	1.76±0.47	

Figure 5.11 : Number of Unambiguous NOE Restraints Per Residue for the cbEGF12-13 Domain Pair

The number of unambiguous NOE restraints per residue are shown, stacked from bottom to top with progressively lighter shading, for intra-residue, sequential, short-range ($2<|i-j|\leq 4$) and long-range ($|i-j|>4$) restraints. Secondary structure is shown at the top of the diagram (see section 5.4), with black rectangles representing β-strands and a white rectangle representing the α-helix.



5.4 The Structure of the cbEGF12-13 Domain Pair

5.4.1 Secondary Structure and Interdomain Packing

The 25 final NMR structures are shown overlaid on the average structure in Figure 5.12, and the consensus secondary structure is shown in Figure 5.13. As expected from homology to cbEGF32-33 (Downing *et al.*, 1996), the structure of the cbEGF12-13 domain pair is a near-linear, rod-like arrangement of two domains, with each domain comprising a major and minor two stranded anti-parallel β-sheet. The minor β-sheet of cbEGF13 depicted in Figure 5.13 is, however, non-ideal, which is most likely the result of,

(a) the small number of residues in this C-terminal region.

(b) the absence of the C-terminal cbEGF14 domain, which would be likely to stabilise the fold of this region through interdomain packing interactions (see below).

(c) the presence of a proline in this region (P82), which may distort β-sheet structure (see section 5.4.3).

(d) the minor *cis*-P76 conformation, which may have caused a small number of additional, incorrect distance restraints to be incorporated into this region.

In addition to β-structure, the cbEGF13 domain contains an α-helical region, which is present in cbEGF12 in 32% of final structures. The structure agrees well with the calculated model (Figure 5.8), highlighting the value of homology modelling, where relevant data are available, to the structure determination process.

Each domain of the cbEGF12-13 domain pair has no appreciable hydrophobic core, and is primarily stabilised by three disulphide bonds in each domain and by calcium ligation. The near-linear orientation of the two domains is maintained by calcium binding by the C-terminal domain and by an interdomain hydrophobic packing

Figure 5.12 : Stereoview Superposition of the 25 Final Structures of the cbEGF12-13 Domain Pair

The 25 structures (black) are shown overlaid on the average structure (red), and the overlay was performed using residues in regions of secondary structure, i.e., residues 19-23, 26-30, 35-37, 45-47, 50-53, 61-66, 69-74, 79-81 and 86-88.
This figure was created using Insight II (version 98, MSI Inc.).
Calcium atoms are not shown.

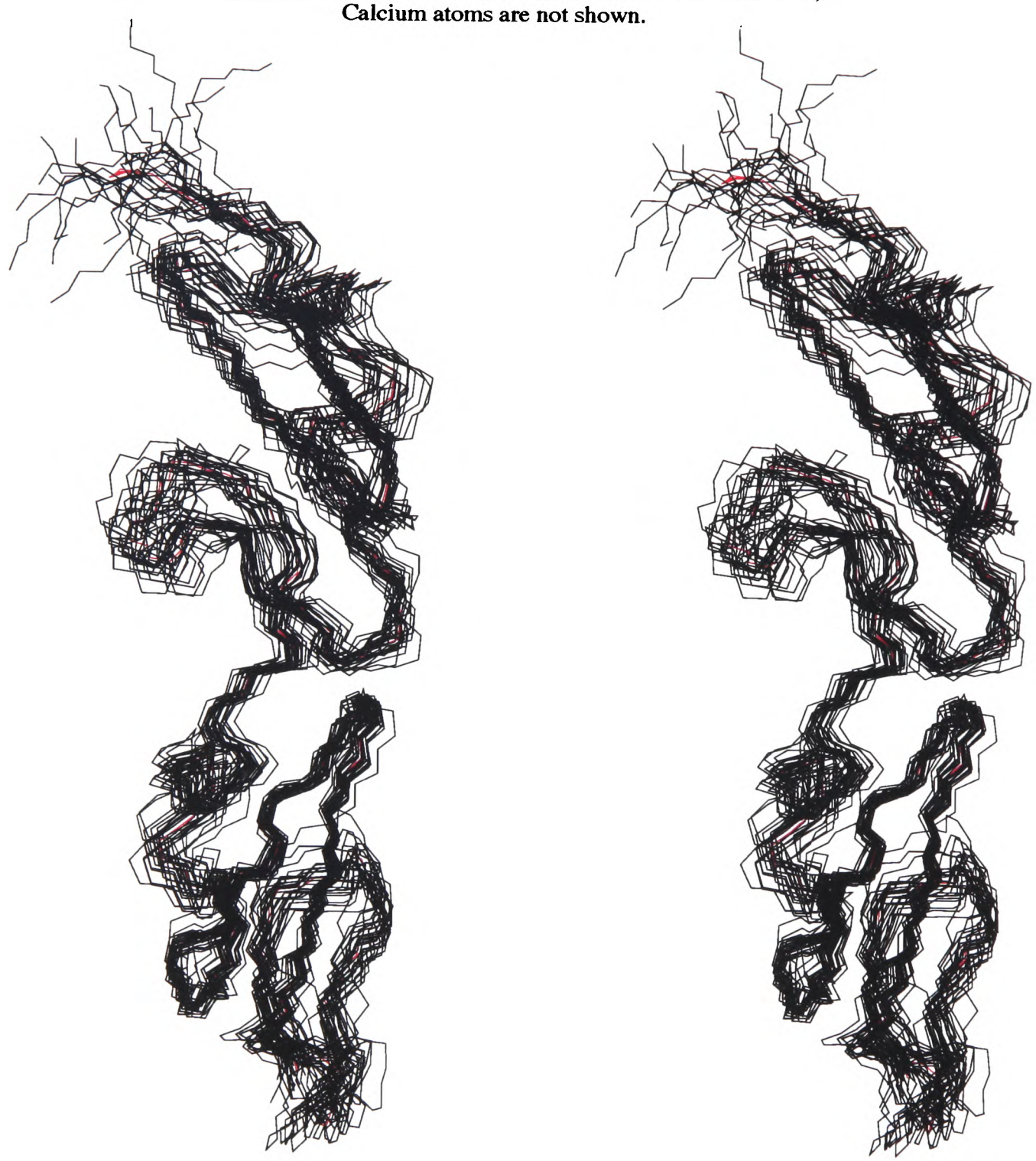
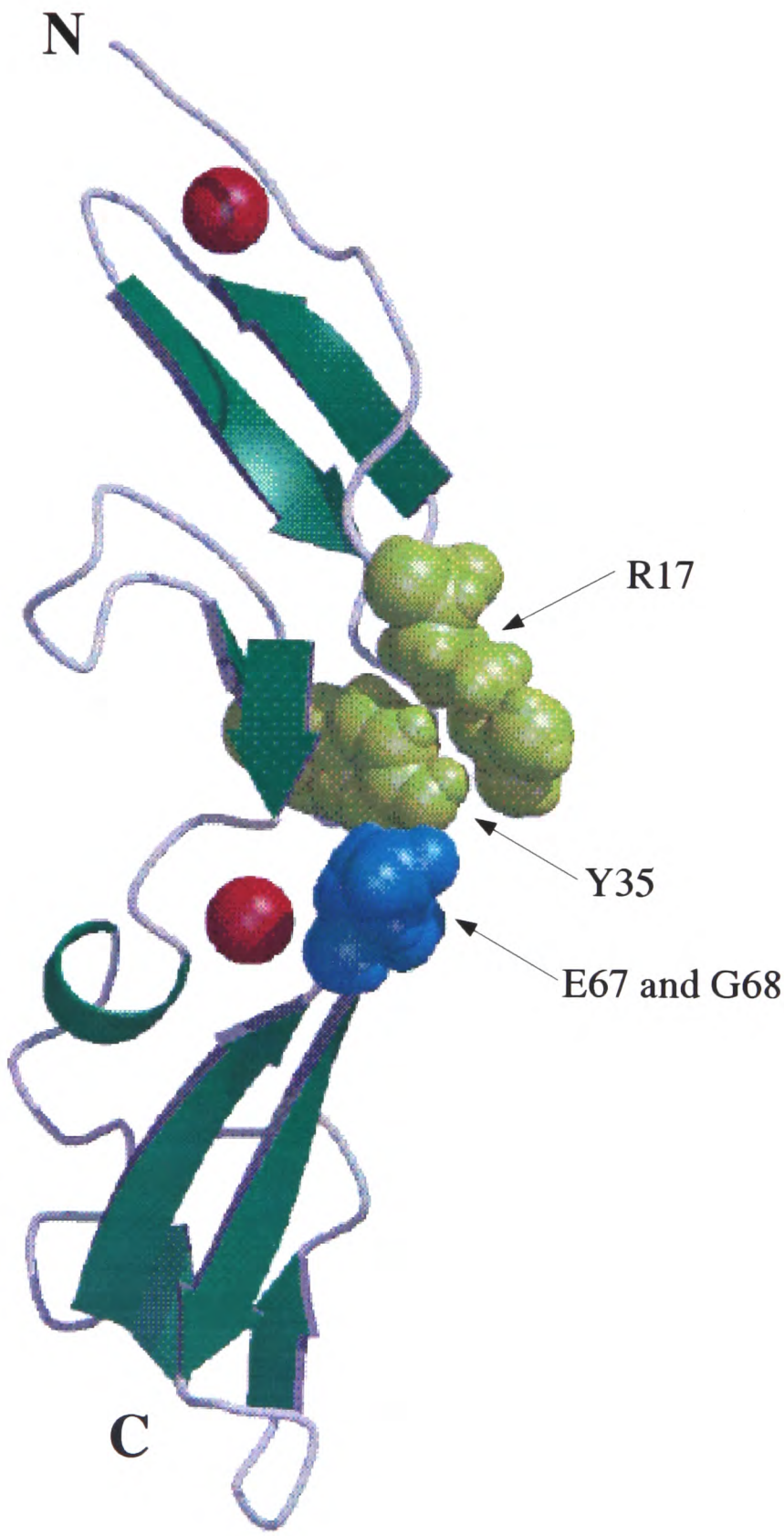


Figure 5.13 : Schematic Illustration of the Secondary Structure and Interdomain Packing Interactions of the cbEGF12-13 Domain Pair from Human Fibrillin-1

R17 and Y35 are shown in yellow, and E67 and G68 are shown in cyan. β-strands are depicted by green arrows, and calcium atoms are drawn as red spheres. This figure was rendered (Merritt & Murphy, 1994) from MOLSCRIPT (Kraulis, 1991) input.

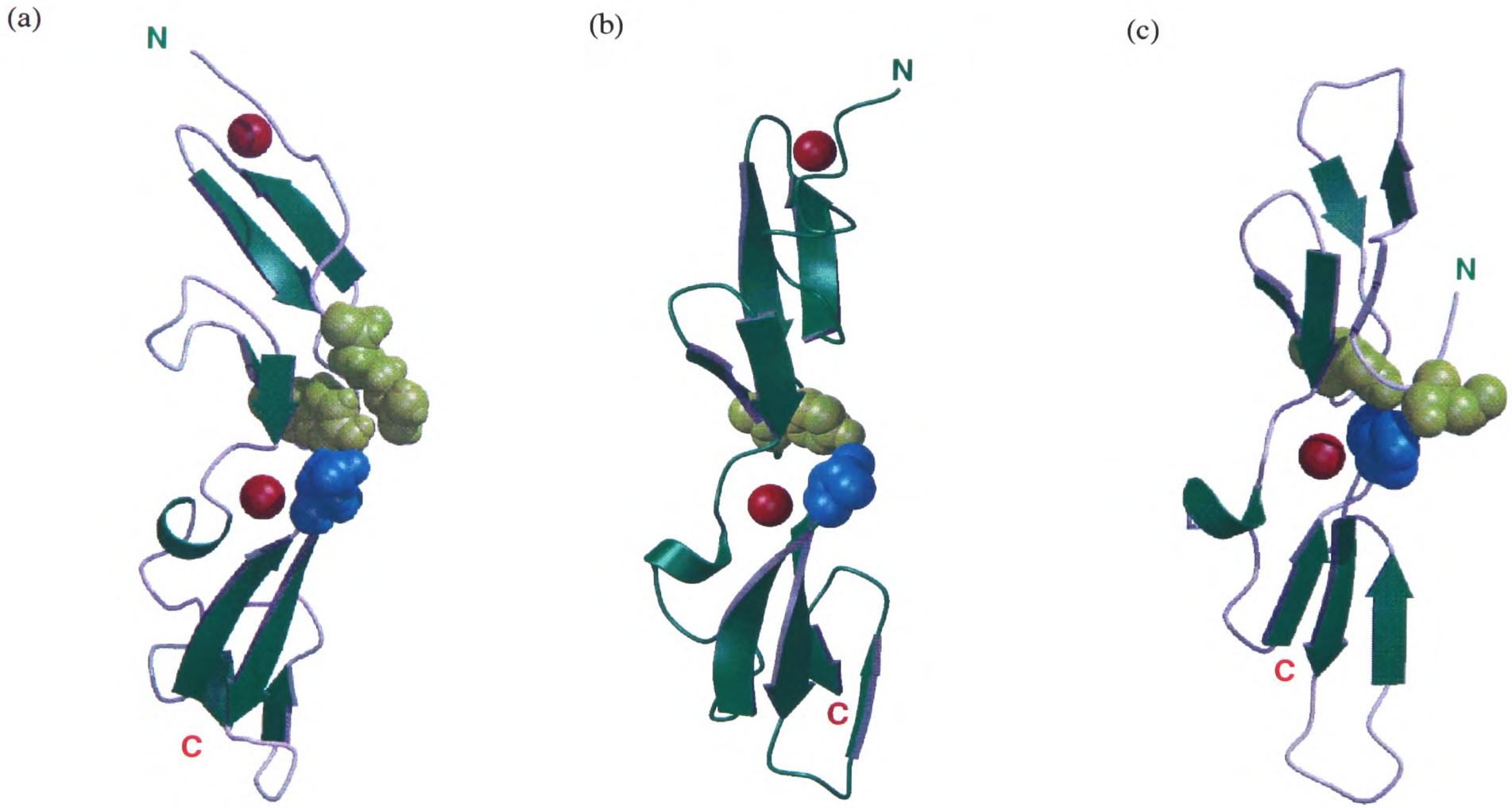


interaction, both previously observed in the cbEGF32-33 domain pair (Figure 1.12; Downing *et al.*, 1996). In cbEGF12-13, the conserved aromatic residue, Y35 (Y1101), at the open end of the minor β-sheet of cbEGF12 packs against the top of the major β-sheet of cbEGF13, and the main packing interaction involves the side chains of Y35 (Y1101) with the conserved G68 (G1134) and the methylene groups of E67 (E1133) (Figure 5.13). In this case, however, the interaction also appears to involve the methylene groups of R17 (R1083), which do not form interdomain contacts but pack against the side chain of Y35. R17 is located in a turn before the major β-sheet of cbEGF12, and, in cbEGF32, there is one less residue in this region between cysteines 1 and 3, so the residue corresponding to R17, H2139, does not interact with the conserved tyrosine in the cbEGF32-33 domain pair. The participation of R17 in interdomain packing in cbEGF12-13 may partially explain the improved stability of cbEGF13 in the cbEGF12-13 domain pair, as indicated by increased calcium binding affinity. The involvement of R17 could improve shielding of the calcium ion from solvent and/or provide a better defined binding site in cbEGF13.

Interdomain packing interactions of Class I EGF/cbEGF-cbEGF pairs of known structure are shown in Figure 5.14, i.e., cbEGF12-13 (this study), cbEGF32-33 from human fibrillin-1 (Downing *et al.*, 1996), and EGF5-cbEGF6 from human thrombomodulin (Fuentes-Prior *et al.*, 2000). The latter structure was solved as part of a larger section of thrombomodulin (EGF4-EGF5-cbEGF6) complexed with human α-thrombin (Fuentes-Prior *et al.*, 2000). The packing interactions shown in Figure 5.14 do not adhere only to the basic, conserved interaction originally observed for cbEGF32-33 (Downing *et al.*, 1996), as highlighted by the participation of R17 in packing in cbEGF12-13, and N391 in the interdomain interaction of EGF5-cbEGF6. EGF5 from thrombomodulin has an altered secondary structure and disulphide bonding pattern (1-2, 3-4 and 5-6) to that usually seen for EGF domains (section 1.3). It also has a non-conserved EGF core (see below), so that the tyrosine involved in interdomain packing

Figure 5.14 : Schematic Comparison of Interdomain Packing Interactions for Class I EGF/cbEGF-cbEGF Domain Pairs of Known Structure

Structures shown are (a) cbEGF12-13 from human fibrillin-1 (this study), (b) cbEGF32-33 from human fibrillin-1 (Downing *et al.*, 1996), and (c) EGF5-cbEGF6 from human thrombomodulin as part of EGF4-EGF5-cbEGF6 complexed with human α-thrombin (Fuentes-Prior *et al.*, 2000). Calcium atoms are shown as red spheres. In (a), the packing interaction comprises R1083 (R17) and Y1101 (Y35) of cbEGF12 (yellow) and E1133 (E67) and G1134 (G68) of cbEGF13 (cyan). In (b), this interaction comprises Y2157 of cbEGF32 (yellow) and I2185 and G2186 of cbEGF33 (cyan). In (c), the corresponding interaction comprises N391 and Y413 of EGF5 (yellow), and P441 and G442 of cbEGF6 (cyan). It should be noted that both the disulphide bonding pattern and secondary structure of EGF5 from human thrombomodulin differ from the consensus patterns usually seen for cbEGF domains and that the EGF core is not conserved. This figure was rendered (Merritt & Murphy, 1994) from MOLSCRIPT (Kraulis, 1991) input.



is located only 15, rather than 16, residues from the third cysteine (Fuentes-Prior *et al.*, 2000). The altered structure of EGF5 allows additional interdomain contacts to be formed, e.g., N391-P441, and, analogous to the situation seen for cbEGF12-13, these additional contacts have been proposed to explain the markedly low dissociation constant seen for cbEGF6 (~2μM; Light *et al.*, 1999) by improving protection of bound calcium from bulk solvent (Fuentes-Prior *et al.*, 2000). These variations highlight the plasticity that can exist for Class I pair interdomain packing interactions whilst still ensuring maintenance of the characteristic linear, rod-like appearance (Figure 5.14), and suggest that variations in interdomain packing interactions may be a major factor determining the calcium binding affinity exhibited by these domains.

5.4.2 Tilt and Twist Angles

The program mod2 (Bork *et al.*, 1996) was used to compute tilt and twist angles for the cbEGF12-13 domain pair family of structures, and, for each structure, the residues D4-C45 and D47-C87 were used to define the boundaries of cbEGF12 and cbEGF13, respectively. The domains were then superimposed on the coordinates of Cα atoms in the EGF core, where the core is defined as the constant number of residues (16) observed for the majority of cbEGF domains between the third cysteine and the conserved aromatic residue between the fifth and sixth cysteine (Downing *et al.*, 1996). This EGF core comprises the major β-sheet structure of cbEGF domains, and using this region, the tilt and twist angles for the cbEGF12-13 domain pair were calculated as 30° ± 15° and 152° ± 13°, respectively.

Corresponding tilt and twist angles of 18°±6° and 159°±6° and 132° and 47° have been previously calculated for the cbEGF32-33 domain pair and the two domains in the factor IX crystals, respectively (Downing *et al.*, 1996). A 110° tilt angle has also

been observed for the EGF pair from factor IXa (Brandstetter *et al.*, 1995). The angles for the fibrillin cbEGF pairs differ substantially from those seen for factor IX and IXa, whereas the angles observed for the fibrillin cbEGF pairs are comparable. The latter observation, combined with structural data for EGF5-cbEGF6 from human thrombomodulin (Figure 5.14c; Fuentes-Prior *et al.*, 2000), supports the hypothesis that Class I EGF-cbEGF/cbEGF-cbEGF pairs, with one residue in their interdomain linker, maintain a near-linear, rod-like structure.

The variation in interdomain tilt angle (± 15°) observed across the family of cbEGF12-13 structures could be the result of a hinge-like motion between the domains, or could be due to imperfections in the structure calculation process. The former is unlikely, however, as dynamics data indicate a rigid, interdomain linker (Chapter 6). The larger average tilt angle seen for the cbEGF12-13 domain pair when compared with cbEGF32-33, could result from the participation of R17 in interdomain packing. From Figure 5.13, it appears that R17 causes cbEGF12 to tilt relative to cbEGF13 purely due to the size and location of the R17 side chain. The single residue insertion between cysteines 1 and 3 in cbEGF12, compared with cbEGF32, may sterically prevent the formation of a structure as linear as that seen for the cbEGF32-33 domain pair. The variation could, however, be due to a lack of long-range distance information for cbEGF12-13. Until recently, structure determination by NMR relied only on short-range distance information, which complicates definition of interdomain orientation if only a small interface exists between domains. Residual dipolar coupling (RDC) data can now be obtained, which provide long-range structural information that can significantly improve the accuracy of protein structure determination by NMR (Tjandra & Bax, 1997; Clore *et al.*, 1998). RDC data have been obtained for the cbEGF12-13 domain pair, and future studies will involve incorporation of these data into the structural refinement process, verifying whether or not the larger tilt angle seen for the cbEGF12-13 domain pair is real.

5.4.3 Structural Convergence

Table 5.3 shows the root-mean-square deviation (rmsd) values for superposition of the final family of 25 NMR structures on the unminimised average structure for backbone (Cα, C, N) atoms. A variety of criteria were used to select residues for superposition of the structures, including selection on the basis of secondary structure and average backbone dynamics. From Table 5.3, it is clear that the best structural convergence is obtained from overlays on the basis of secondary structure, and Figure 5.12 shows the structures overlaid using this criterion. The significant difference observed between rmsd values obtained when overlaying residues spanning the domain pair and those localised to a single domain reflects the relatively large variation in tilt and twist angles seen for this domain pair.

Table 5.3 : Root-Mean-Square Deviations (Å) of the 25 Final NMR Structures to the Unminimised Average Structure for Backbone (Cα, C, N) Atoms

Overlay	All residues	Dynamics basis[1]	Secondary structure[2]	Order parameters[3]
cbEGF12-13	1.53 ± 0.48	1.12 ± 0.41	0.96 ± 0.44	1.09 ± 0.41
cbEGF12	1.17 ± 0.22	0.68 ± 0.21	0.37 ± 0.09	0.36 ± 0.06
cbEGF13	0.45 ± 0.13	0.32 ± 0.07	0.31 ± 0.06	0.32 ± 0.06

[1] Structures overlaid based on average backbone dynamics, i.e., average T_2 relaxation times (127.1 ± 34.6ms; Chapter 6). This method was used for superposition of the cbEGF32-33 family of structures (Downing *et al.*, 1996).
[2] Structures overlaid on regions of secondary structure.
[3] Structures overlaid using residues with ϕ angle order parameters ≥0.98.

Each domain of the cbEGF12-13 domain pair is very well defined. The backbone (Cα, C, N) rmsd values to the unminimised average coordinates, based on regions of secondary structure, are 0.37 ± 0.09Å and 0.31 ± 0.06Å for cbEGF12 and

cbEGF13, respectively. The tertiary folds of the two domains are very similar, with both comprising a major and a minor double-stranded β-sheet (Figure 5.15).

The most notable differences include the presence of an α-helix between cysteines 1 and 2 in cbEGF13, which is present in only 32% of cbEGF12 structures. This region contains an identically placed proline in both domains, and it is likely that the α-helix in cbEGF12 would be stabilised by the presence of a N-terminal cbEGF domain. The length and/or conformation of some loop regions between cysteines vary between these domains, and these differences are likely to relate to the specificity of cbEGF domain interactions (Downing *et al.*, 1996). The loop region between cysteines 5 and 6 is one residue shorter in cbEGF13 and contains three proline residues (P75,P76 and P82). This region contains no prolines in cbEGF12, and the tip of this extended loop, consisting of residues G38-N44 (GFMMMKN), is unstructured and may be involved in protein-protein interactions (section 5.5). Other differences between domains 12 and 13 include the orientation of the minor β-sheet (Figure 5.15), which is most likely a result of the interdomain packing interactions. This "β-sheet" in cbEGF13 is distorted by a proline residue (P82), as is the case for cbEGF33 in cbEGF32-33 (Downing *et al.*, 1996).

The known structures of cbEGF domains were overlaid on cbEGF12 and cbEGF13 based, where possible, on the backbone coordinates of the EGF core, and the backbone rmsd values for the structures are shown in Table 5.4. The results are analogous to previous atomic coordinate comparisons of cbEGF domains which have highlighted the structural variability of these domains (1.0-2.3Å), most likely a result of the absence of a hydrophobic core (Downing *et al.*, 1996; 2000). From Table 5.4, it can be seen that domain 12, in general, overlays better with the structures of other cbEGF domains than domain 13. This could be a consequence of the change in conformation of cbEGF13 caused by interdomain packing (Figures 5.13 and 5.15)

Figure 5.15 : Schematic Comparison of the Structures of cbEGF12 and cbEGF13 from Human Fibrillin-1

The loop connecting the first two cysteines is helical in cbEGF13 and in 32% of cbEGF12 structures. The conformations of the major and minor β-strands differ slightly due to inter-domain packing interactions. This figure was rendered (Merritt & Murphy, 1994) from MOLSCRIPT (Kraulis, 1991) input.

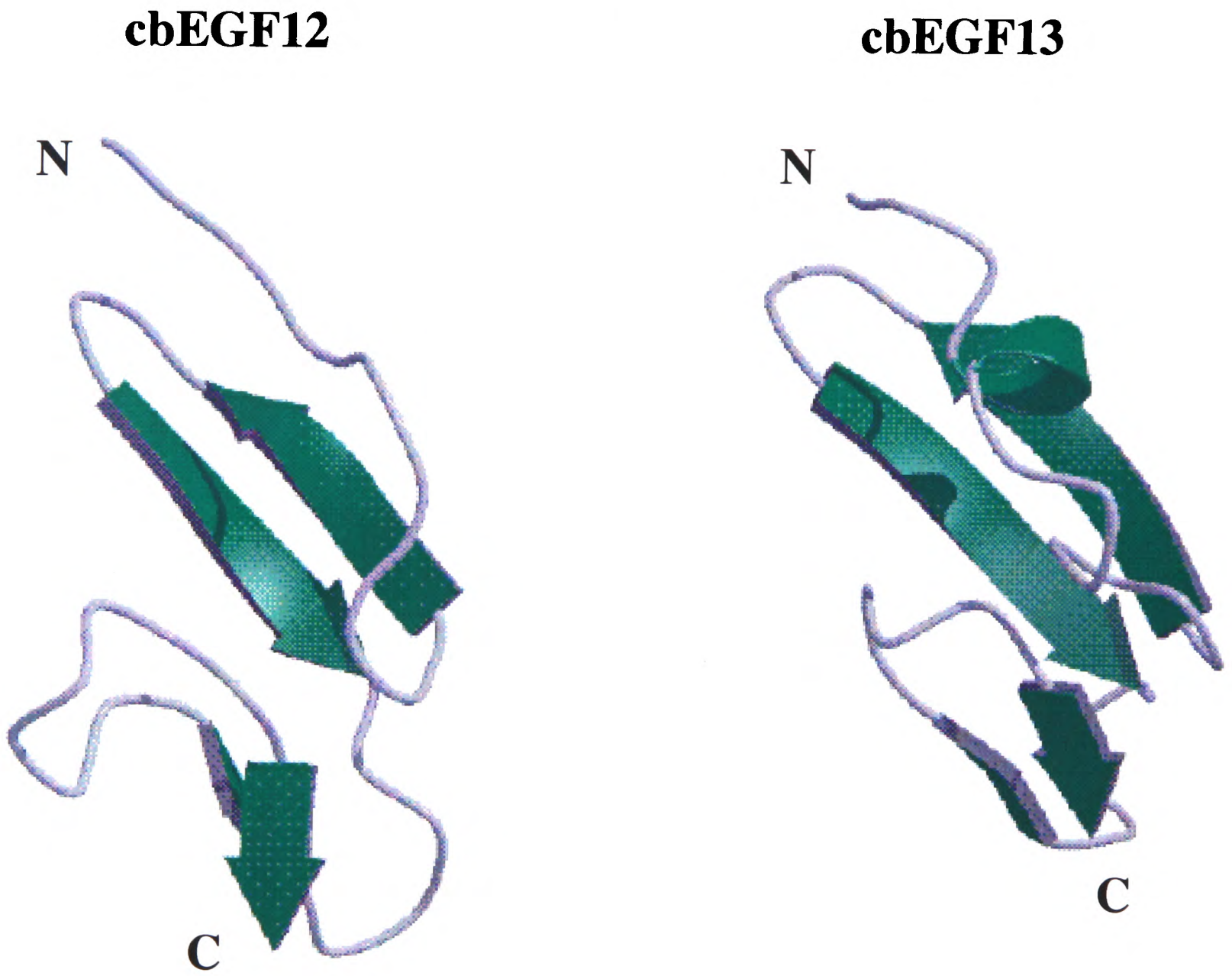


Table 5.4 : Atomic Coordinate Comparison of cbEGF Structures: Backbone (Cα, C, N) R.M.S. Deviations for Conserved Residues

Residues were overlaid based on the EGF core region, except for cbEGF6 from thrombomodulin (Fuentes-Prior *et al.*, 2000) where the EGF core is not conserved (see text). For this domain, the overlay was performed on residues 437-452, i.e., using 16 residues starting with cysteine 3. Structures were overlaid using Insight II (version 98, MSI Inc.). For details of previous atomic coordinate comparisons of cbEGF structures the reader is referred to Downing *et al.*, 1996 and 2000.

cbEGF structure	Method	Ca^{2+}	Residues used for overlay	Rmsd (Å) to cbEGF12	Rmsd (Å) to cbEGF13	Ref.
Isolated cbEGF domains						
cbEGF1, human factor VII	NMR	-	61-76	1.2	1.8	Muranyi *et al.*, 1998.
cbEGF1, human factor VII	NMR	+	61-76	1.3	1.7	Kao *et al.*, 1999.
cbEGF1, human factor IX	NMR	-	62-77	1.3	1.9	Baron *et al.*, 1992.
cbEGF1, human factor IX	X-ray (1.5Å)	+	62-77	1.1	2.0	Rao *et al.*, 1995.
cbEGF1, human factor X	NMR	-	61-76	1.4	2.1	Ullner *et al.*, 1992.
cbEGF1, human factor X	NMR	+	61-76	1.8	2.2	Selander-Sunnerhagen *et al.*, 1992.
cbEGF1, human C1r	NMR	-	148-163	1.7	2.1	Bersch *et al.*, 1998.
cbEGF domain pairs						
cbEGF12-13, human fibrillin-1						
cbEGF12,	NMR	+	20-35	-	1.9	This study.
cbEGF13,	NMR	+	63-78	1.9	-	This study.
cbEGF32-33, human fibrillin-1						
cbEGF32	NMR	+	2142-2157	2.1	2.3	Downing *et al.*, 1996.
cbEGF33	NMR	+	2181-2196	1.3	1.9	Downing *et al.*, 1996.
cbEGF domains as part of larger structures						
cbEGF1 in porcine factor IXa	X-ray (3.0Å)	-	62-63, 65-77	1.6	2.3	Brandstetter *et al.*, 1995.
cbEGF1 from human factor X in a Gla-cbEGF pair	NMR	+	61-62 64-76	2.0	2.5	Sunnerhagen *et al.*, 1996.
cbEGF domains in complexes						
cbEGF1 in human factor VIIa: tissue factor complex	X-ray (2.0Å)	+	61-76	1.1	1.9	Banner *et al.*, 1996.
cbEGF6 in human thrombomodulin: thrombin complex	X-ray (2.3Å)	+	437-452	2.7	2.9	Fuentes-Prior *et al.*, 2000.

and/or could be related to the increased calcium binding affinity observed for this domain, which could make it more compact. The cbEGF6 domain from thrombomodulin (Fuentes-Prior *et al.*, 2000) is most structurally different to the cbEGF12 and 13 domains, not surprising, since the EGF core is not conserved in this domain. Lack of conservation of an aromatic residue at the usual position in cbEGF6 from thrombomodulin may be because this domain is the last EGF domain in a run of six, hence the usual aromatic residue may no longer be required for interdomain interactions. The next most dissimilar domains are cbEGF1 from factor X in a Gla-cbEGF construct (Sunnerhagen *et al.*, 1996) and cbEGF32 from fibrillin-1 in cbEGF32-33 (Downing *et al.*, 1996). The larger rmsd values found for these domains when overlaid with cbEGF12 and 13 are most likely a result of the different packing interactions observed between domains of these pairs.

5.5 Biological Significance of the cbEGF12-13 Structure

5.5.1 Implications for Connective Tissue Microfibrils

The structure of the cbEGF12-13 domain pair further validates the proposal of Downing *et al.* (1996), i.e., that Class I EGF-cbEGF/cbEGF-cbEGF domain pairs, with one residue in their interdomain linker, adopt a rigid, rod-like structure. The cbEGF12-13 structure, however, appears less linear than the cbEGF32-33 structure (Downing *et al.*, 1996), although this difference remains to be verified using RDC data. Each domain of cbEGF12-13 spans ~2.7nm, comparable to the span observed for cbEGF32-33, hence the structure and dimensions of cbEGF12-13 further confirm the most recent model for incorporation of fibrillin-1 monomers into the 10-12nm connective tissue microfibrils (Figure 1.4).

The inter-cysteine loop lengths for cbEGF domains of known structure are shown in Table 5.5. The relatively long stretch of residues between cysteines 5 and 6

of cbEGF12 (13 residues) means that several side chains in this region are highly solvent accessible. The extended loop, which joins strands of the minor β-sheet, consists of residues G38-N44 (GFMMMKN). This region is unstructured (Figure 5.13) and has reduced ^{1}H-^{15}N heteronuclear NOE values, indicative of increased flexibility (Chapter 6). This region may therefore be important for protein-protein interactions. It could be directly involved in microfibril assembly, i.e., directly interacting with other fibrillin-1 monomers, or it may be involved in the incorporation of non-homologous proteins into the 10-12nm microfibrils. No MFS-causing mutations have been identified in this region to date, so the importance of this flexible loop remains to be verified. The cbEGF6 domain from thrombomodulin is the only other cbEGF of known structure with a 13 residue loop connecting cysteines 5 and 6, and a residue in this region (D461) has been shown to be involved in complex formation with thrombin (Fuentes-Prior *et al.*, 2000). Similarly, the cbEGF domain from C1r has an unusually large loop region connecting cysteines 1 and 2. This region does not possess a unique structure, and has therefore been proposed to play a role in domain-domain interactions in C1r or in protein-protein interactions within the C1 complex (Bersch *et al.*, 1998).

Table 5.5 : Inter-Cysteine Loop Lengths for cbEGF Domains of Known Structure

cbEGF domain	**1-2**	**2-3**	**3-4**	**4-5**	**5-6**
cbEGF1, factor VII	4	5	8	1	8
cbEGF1, factor IX	4	5	8	1	8
cbEGF1, factor X	4	5	8	1	8
cbEGF1, Clr	14	3	8	1	12
cbEGF12, fibrillin-1	6	4	8	1	13
cbEGF13, fibrillin-1	6	4	8	1	12
cbEGF32, fibrillin-1	5	4	8	1	10
cbEGF33, fibrillin-1	5	4	8	1	11
cbEGF6, thrombomodulin	5	3	8	1	13

Domain 13 of the cbEGF12-13 domain pair contains a potential N-linked glycosylation site (^{83}NIS) in the loop connecting strands of the minor β-sheet. Whether

this site is N-glycosylated *in vivo*, and the potential importance of this site to the structure and function of the microfibrils, remains to determined. O-fucosylation of cbEGF1 from factor VII has been shown to cause a two-fold enhancement of calcium binding affinity in the absence of dramatic structural alterations (Kao *et al.*, 1999). It remains to be seen whether glycosylation of cbEGF13 could further enhance the calcium binding affinity of this already high affinity site.

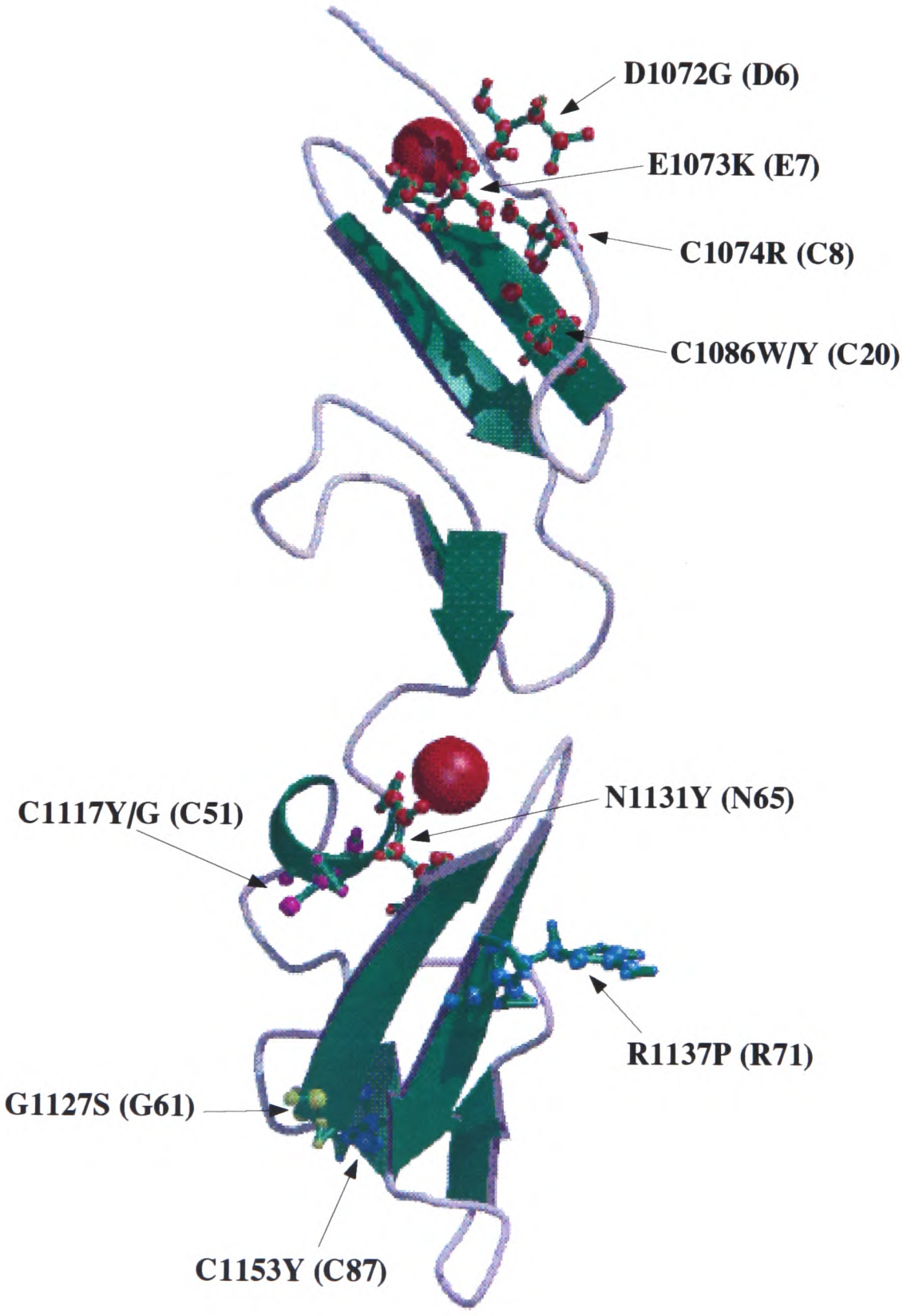
5.5.2 Implications for the Marfan Syndrome

Known missense mutations in cbEGF12-13 that cause MFS are highlighted on the structure of the domain pair in Figure 5.16, and these mutations can be classified into three groups depending on the residue affected (Downing *et al.*, 1996; Table 1.3). Mutations affecting cysteine residues are likely to alter disulphide bond formation, thereby disrupting the correct fold, and mutations affecting residues in the calcium binding consensus sequence are likely to reduce calcium binding affinity, leading to destabilisation of the interdomain linkage. The two remaining missense mutations in this domain pair, G1127S and R1137P, which have been shown by NMR studies to impair folding in cbEGF13 (Whiteman *et al.*, 1998; Wu *et al.*, 1995), are likely to cause secondary structure disruption. Mutation of R1137 to a proline could greatly distort the major β-sheet of cbEGF13, as could exchanging G1127 for a less flexible serine residue at the start of this β-structure (Chapter 7).

Despite these insights, mapping MFS mutations onto the structure of the cbEGF12-13 pair provides little clarification as to why these mutations produce such varying phenotypes. Why does a N1131Y mutation in cbEGF13 cause nMFS, whereas a mutation affecting the same calcium ligand in cbEGF14 (N1173K) causes incomplete MFS? Why does the E1073K mutation in cbEGF12 cause nMFS, whereas the corresponding mutation (E1200G) in cbEGF15 produces a classical phenotype? With a view to addressing these questions, the positions of mutations in cbEGF11-15 were

Figure 5.16 : Schematic Diagram Highlighting the Location of MFS-Causing Mutations in the cbEGF12-13 Domain Pair from Human Fibrillin-1

Mutations that cause nMFS, MFS and a mild variant of MFS are shown in red, blue and yellow, respectively. Mutation of the residue depicted in purple can cause MFS (C1117Y) or nMFS (C1117G). This figure was rendered (Merritt & Murphy, 1994) from MOLSCRIPT (Kraulis, 1991) input.



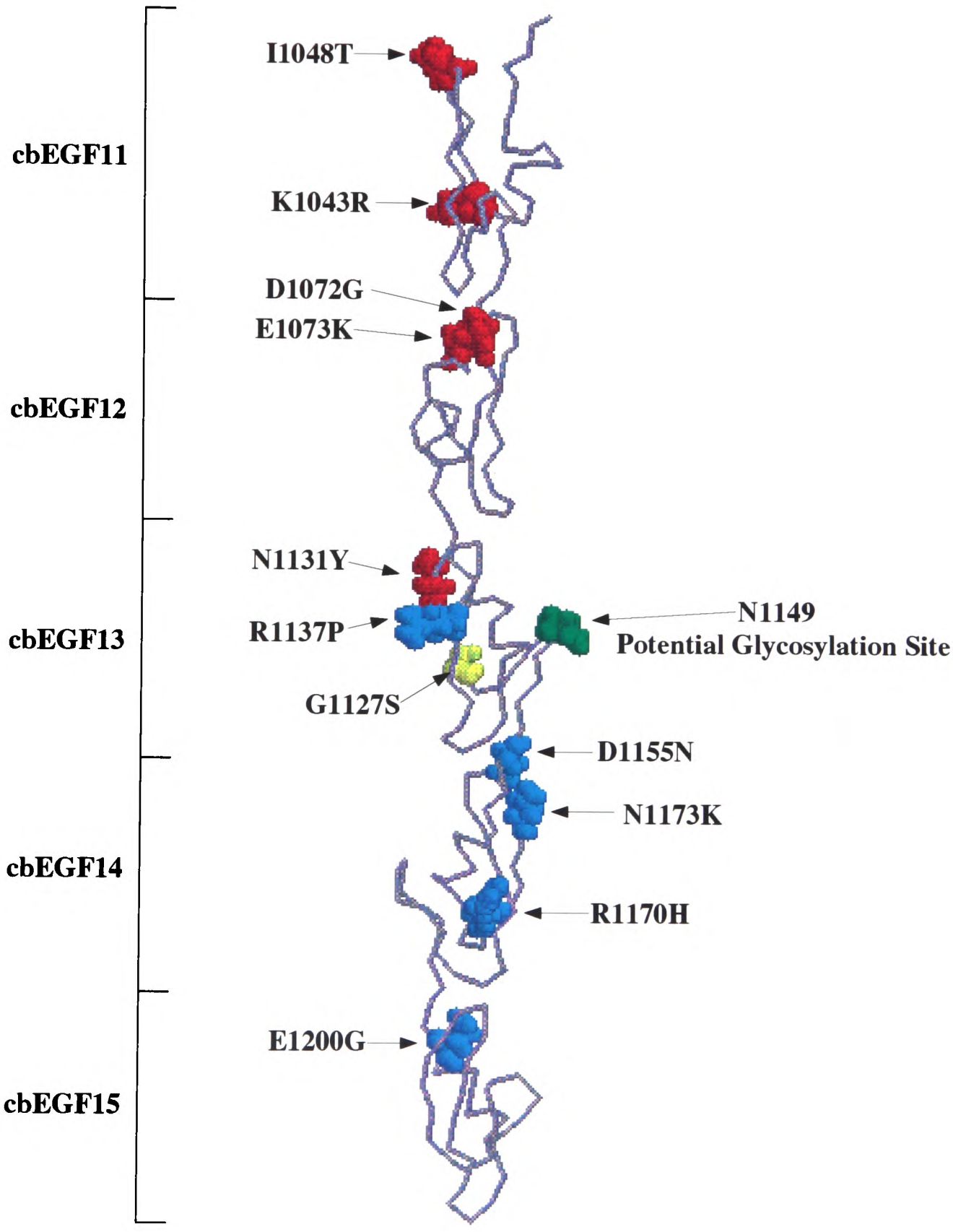
overlaid with the corresponding residues in a model of cbEGF32-36 (Downing *et al.*, 1996), and the result is shown in Figure 5.17. The cbEGF32-36 model was used since tilt and twist angles are comparable between the cbEGF12-13 and cbEGF32-33 domain pairs (section 5.4.2). A rigid, rod-like structure is predicted for the cbEGF11-15 region of fibrillin-1, and, if the potential glycosylation site in cbEGF13 is used to define the outside face, it can be seen that all nMFS-causing mutations in this region appear to cluster on the inside face of cbEGF11-13. The extended, flexible loop region between cysteines 5 and 6 of cbEGF12 also projects into solution from this face, so this region could therefore be a candidate region for involvement in protein-protein interactions. Cysteine mutations were not considered in this analysis, since they are more likely to affect protein folding than protein-protein interactions.

K1043, I1048, D1072, E1073, N1131 and the flexible loop of cbEGF12 may be directly involved in protein-protein interactions, either with other fibrillin-1 monomers or with non-homologous components of the 10-12nm microfibrils. The importance of the neonatal region of fibrillin-1 for microfibril assembly has been highlighted by immunofluorescence studies of neonatal MFS fibrils, which showed them to be short, fragmented and frayed relative to wild-type microfibrils (Superti-Furga *et al.*, 1992; Godfrey *et al.*, 1995). The specific mutation(s) giving rise to this morphology has, however, not been identified. Aspartate, glutamate and asparagine residues have recently been shown to be important for the interaction of EGF5-cbEGF6 from thrombomodulin with thrombin (Fuentes-Prior *et al.*, 2000), and, in relation to cbEGF11, the back face of the major β-sheet of the factor VIIa cbEGF has been shown to be important for complex formation with tissue factor (Banner *et al.*, 1996).

It is possible that the slightly kinked structure of cbEGF12-13, which remains to be verified, could be important for orientating residues, and the cbEGF12 loop, into the correct position for protein binding, and if this region is involved in protein-protein interactions, mutation of only one of these residues could directly disrupt complex formation. Neonatal MFS mutations disrupting calcium binding consensus residues

Figure 5.17 : Model of cbEGF11-15 from Human Fibrillin-1 Highlighting MFS-Causing Mutations in this Region

The cbEGF11-15 model was created from a model of cbEGF32-36 (Downing *et al.*, 1996), using Insight II (version 98, MSI Inc.) to mutate relevant residues. Calcium atoms are omitted. Residues causing nMFS, MFS and a mild variant of MFS are shown in red, cyan and yellow, respectively. A potential N-linked glycosylation site in cbEGF13 is highlighted in green. Only non-cysteine mutations were considered here, since cysteine mutations are likely to affect protein folding rather than protein-protein interactions.



could also have an indirect effect on protein binding. Destabilisation of cbEGF domain linkage, caused by mutations which reduce calcium binding affinity, could lead to increased interdomain flexibility, distorting the presentation of a binding site. The reduction in calcium binding affinity could also lead to increased susceptibility of this region to proteolysis, as shown recently for the E1073K mutation (Reinhardt *et al.*, 2000). The cbEGF domains in the neonatal region of fibrillin-1 may have increased calcium binding affinities (Chapter 4) to ensure provision of the rigidity and stability required to maintain correct presentation of this potential site.

Insights gained from the structure of the cbEGF12-13 domain pair and the cbEGF11-15 model have provided a focus for future functional studies, which may be utilised to verify the potential importance of the neonatal region to the structure and function of the 10-12nm connective tissue microfibrils. Gene targeting experiments in mice have previously shown that fibrillin-1 monomers missing five domains, i.e., cbEGF8, cbEGF9, hybrid2, cbEGF10 and TB3, are still capable of forming microfibrils (Pereira *et al.*, 1997), hence this method could be targeted to cbEGF11-13 in the future to verify its importance in microfibril assembly.

References

Aue, W. P., Bartholdi, E. & Ernst, R. R. (1976). Two-dimensional spectroscopy: application to nuclear magnetic resonance. *J. Chem. Phys.* **64,** 2229-2246.

Banner, D. W., D'Arcy, A., Chene, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y. & Kirchhofer, D. (1996). The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature.* **380**, 41-46.

Baron, M., Norman, D. G., Harvey, T. S., Handford, P. A., Mayhew, M., Tse, A. G. D., Brownlee, G. G. & Campbell, I. D. (1992). The three-dimensional structure of the first EGF-like module of human factor IX: comparison with EGF and TGF-α. *Prot. Sci.* **1**, 81-90.

Barsukov, I. L. & Lian, L.-Y. (1993). Structure determination from NMR data I. Analysis of NMR data. In NMR of macromolecules. G. C. K. Roberts, ed., Oxford University Press Inc., New York, pp. 315-357.

Bartels, C., Xia, T.-H., Billeter, M., Güntert P. & Wüthrich K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR.* **5**, 1-10.

Bax, A. & Davis, D. G. (1985). MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **65**, 355-360.

Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A. & Haak, J. R. (1984). Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690.

Bersch, B., Hernandez, J.-F., Marion, D. & Arlaud, G. J. (1998). Solution structure of the epidermal growth factor (EGF)-like module of human complement protease C1r, an atypical member of the EGF family. *Biochemistry.* **37**, 1204-1214.

Böckmann, A. & Guittet, E. (1997). Determination of fast proton exchange rates of biomolecules by NMR using water selective diffusion experiments. *FEBS Lett.* **418**, 127-130.

Bork, P., Downing, A. K., Kieffer, B. & Campbell, I. D. (1996) Structure and distribution of modules in extracellular proteins. *Quart. Rev. Biophys.* **29**, 119-167.

Brandstetter, H., Bauer, M., Huber, R., Lollar, P. & Bode, W. (1995). X-ray structure of clotting factor IXa: Active site and module structure related to Xase activity and hemophilia B. *Proc. Natl. Acad. Sci. USA.* **92**, 9796-9800.

Braunschweiler, L. & Ernst, R. R. (1983). Coherence transfer by isotropic mixing - application to proton correlation spectroscopy. *J. Magn. Reson.* **53**, 521-528.

Brown, S. C., Weber, P. L. & Mueller, L. (1988). Toward complete ^{1}H NMR spectra in proteins. *J. Magn. Reson.* **77**, 166-169.

Brünger, A. T. (1992). X-PLOR (version 3.1): A system for X-ray crystallography and NMR. Yale University, New Heaven, CT, USA.

Cavanagh, J. & Rance, M. (1990). A combined two-dimensional relayed-NOESY TOCSY experiment. *J. Magn. Reson.* **87**, 408-414.

Cavanagh, J., Fairbrother, W. J., Palmer III, A. G. & Skelton, N. J. (1996). Protein NMR spectroscopy. Academic Press, Inc., San Diego.

Clore, G. M., Gronenborn, A. M. & Tjandra, N. (1998). Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.* **131**, 159-162.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Downing, A. K., Campbell, I. D. & Handford, P. A. (2000). Chapter 4 : Calcium binding EGF-like domains. In Calcium Homeostasis. Carafoli, Krebs, eds. *Topics Biol. Inorg. Chem.*, volume 3.

Folmer, R. H. A., Nilges, M., Konings, R. N. H. & Hilbers, C. W. (1995). Solution structure of the single-stranded-DNA binding-protein of the filamentous pseudomonas phage Pf3 - similarity to other proteins binding to single-stranded nucleic-acids. *EMBO J.* **14**, 4132-4142.

Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D. R & Bode, W. (2000). Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. *Nature*. **404**, 518-525.

Godfrey, M., Raghunath, M., Cisler, J., Bevins, C. L., de Paepe, A., Di Rocco, M., Gregoritch, J., Imaizumi, K., Kaplan, P., Kuroki, Y., Silberbach, M., Superti-Furga, A., Van Thienen, M. N., Vetter, U. & Steinmann, B. (1995). Abnormal morphology of fibrillin microfibrils in fibroblast cultures from patients with neonatal Marfan syndrome. *Am. J. Pathol.* **146**, 1414-1421.

Handford, P. A., Mayhew, M., Baron, M., Winship, P. R., Campbell, I. D. & Brownlee, G. G. (1991). Key residues involved in calcium binding motifs in EGF-like domains. *Nature*. **351**, 164-167.

Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys*. **71**, 4546-4553.

Johnson, B. A. & Blevins, R. A. (1994). NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*. **4**, 603-614.

Kao, Y.-H., Lee, G. F., Wang, Y., Starovasnik, M. A., Kelley, R. F., Spellman, M. W. & Lerner, L. (1999). The effect of O-fucosylation on the first EGF-like domain from human blood coagulation Factor VII. *Biochemistry*. **38**, 7097-7110.

Kay, L. E., Marion, D. & Bax, A. (1989). Practical aspects of 3D heteronuclear NMR of proteins. *J. Magn. Reson*. **83**, 72-84.

Kay, L. E. & Bax, A. (1990). New methods for the measurement of HN-CαH coupling constants in ^{15}N-labeled proteins. *J. Magn. Reson*. **86**, 110-126.

Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient-enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc*. **114**, 10663-10665.

Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL : a program for display and analysis of macromolecular structures. *J. Mol. Graphics*. **14**, 51-55.

Koradi, R. & Billeter, M. (1998). MOLMOL : a program for display and analysis of macromolecular structures. *PDB Newsletter*. **84**, 5-7.

Kraulis, P. J. (1991). MOLSCRIPT : A program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst*. **24**, 946-950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK : a program to check the stereochemical quality of protein structures. *J. Appl. Cryst*. **26**, 283-291.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR : Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486.

Light, D. R., Glaser, C. B., Betts, M., Blasko, E., Campbell, E., Clarke, J. H., McCaman, M., McLean, K., Nagashima, M., Parkinson, J. F., Rumennik, G., Young, T. & Morser, J. (1999). The interaction of thrombomodulin with Ca^{2+}. *Eur. J. Biochem*. **262**, 522-533.

Macura, S., Huang, Y., Suter, D. & Ernst, R. R. (1981). Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.* **43**, 259-281.

Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D, version 2.0 : A program for photorealistic molecular graphics. *Acta Cryst.* **50**, 869-873.

Muranyi, A., Finn, B. E., Gippert, G. P., Forsen, S., Stenflo, J. & Drakenberg, T. (1998). Solution structure of the N-terminal EGF-like domain from human factor VII. *Biochemistry.* **37**, 10605-10615.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* **229**, 317-324.

Nilges, M. (1995). Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J. Mol. Biol.* **245**, 645-660.

Nilges, M. (1997). Ambiguous distance data in the calculation of NMR structures. *Folding & Design.* **2**, S53-S57.

Pereira, L., D'Alessio, M., Ramirez, F., Lynch, J. R., Sykes, B., Pangilinan, T. & Bonadio, J. (1993). Genomic organisation of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Hum. Mol. Genet.* **2**, 961-968.

Pereira, L., Andrikopoulos, K., Tian, J., Lee, S. Y., Keene, D. R., Ono, R., Reinhardt, D. P., Sakai, L. Y., Biery, N. J., Bunton, T., Dietz, H. C. & Ramirez, F. (1997). Targetting of the gene encoding fibrillin-1 recapitulates the vascular aspect of Marfan syndrome. *Nat. Genet.* **17**, 218-222.

Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G. & Stuart, D. (1995). The structure of a Ca^{2+}-binding epidermal growth factor-like domain: Its role in protein-protein interactions. *Cell.* **82**, 131-141.

Reinhardt, D. P., Ono, R. N., Notbohm, H., Muller, P. K., Bachinger, H. P. & Sakai, L. Y. (2000). Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis: A potential disease-causing mechanism in Marfan syndrome. *J. Mol. Biol.* **275**, 12339-12345.

Sali A. & Blundell T. L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815.

Selander-Sunnerhagen, M., Ullner, M., Persson, E., Teleman, O., Stenflo, J. & Drakenberg, T. (1992). How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the NH_2-terminal EGF-like domain in coagulation factor X. *J. Biol. Chem.* **267**, 19642-19649.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol.* **286**, 661-668.

Sunnerhagen, M., Olah, G. A., Stenflo, J., Forsen, S., Drakenberg, T. & Trewhella, J. (1996). The relative orientation of Gla and EGF domains in coagulation- factor-X is

altered by Ca^{2+} binding to the first EGF domain - a combined NMR small-angle X-ray-scattering study. *Biochemistry.* **35**, 11547-11559.

Superti-Furga, A., Raghunath, M. & Willems, P. J. (1992). Deficiencies of fibrillin and decorin in fibroblast cultures of a patient with neonatal Marfan syndrome. *J. Med. Genet.* **29**, 875-878.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The ClustalX windows interface : flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **25**, 4876-4882.

Tjandra, N. & Bax, A. (1997). Direct measurement of distances and angles in biomolecules by NMR in a dilute crystalline medium. *Science.* **278**, 1111-1114.

Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T. & Teleman, O. (1992). Three-dimensional structure of the Apo form of the N-terminal EGF-like module of blood coagulation factor X as determined by NMR spectroscopy and simulated folding. *Biochemistry.* **31**, 5974-5983

Wagner, G. (1984). Two-dimensional relayed coherence transfer-NOE spectroscopy. *J. Magn. Reson.* **57**, 497-505.

Whiteman, P., Downing, A. K., Smallridge, R., Winship, P. R. & Handford, P. A. (1998). A Gly to Ser change causes defective folding *in vitro* of calcium binding epidermal growth factor-like domains from factor IX and fibrillin-1. *J. Biol. Chem.* **273**, 7807-7813.

Wu, Y. -S., Bevilacqua, V. L. H. & Berg, J. M. (1995). Fibrillin domain folding and calcium binding : significance to Marfan Syndrome. *Chem. Biol.* **2**, 91-97.

Wüthrich, K. (1986). NMR of proteins and nucleic acids. John Wiley & Sons, Inc., New York.

Yuan, X., Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (1998). Effects of proline *cis-trans* isomerization on TB domain secondary structure. *Protein Sci.* **7**, 2127-2135.

Zhu, G. & Bax, A. (1990). Improved linear prediction for truncated signals of known phase. *J. Magn. Reson.* **90**, 405-410.

Chapter 6

Backbone Dynamics of cbEGF12-13

6.1 Acquisition and Processing of NMR Data for Relaxation Studies

6.1.1 Sample Preparation

The sample used to obtain NMR data for ^{15}N backbone dynamics analysis contained ~1mM ^{15}N-labelled cbEGF12-13 in 4.55mM Tris•HCl, 20mM $CaCl_2$ at pH 6.5. To ensure data were not affected by sample aggregation, measurements were also performed on another sample which contained a higher concentration of cbEGF12-13 (1.6mM). Calcium binding sites in the cbEGF12-13 pair were deemed saturated in both samples based on previous calcium binding studies (Chapter 4; Smallridge *et al.*, 1999).

6.1.2 Data Acquisition

All NMR spectra were recorded at 35°C on GE-Omega/home-built spectrometers using pulsed field gradients for coherence selection and water suppression (Kay *et al.*, 1992a). The temperature and pH were selected to allow comparison with previous results obtained for the cbEGF32-33 pair from human fibrillin-1 (Downing *et al.*, 1996; Werner *et al.*, 2000a). ^{15}N-T_1, ^{15}N-T_2 and ^{1}H-^{15}N heteronuclear NOE data were measured at 500 and 750MHz using previously described methods (Kay *et al.*, 1989; Farrow *et al.*, 1994), and at least each series of T_1, T_2 and the pair of heteronuclear ^{1}H-^{15}N-NOE experiments were acquired in a single session. In

T_1 and T_2 experiments, acquisition times were 102.4 and 110.5ms in the ^1H (F_2) and the ^{15}N (F_1) dimension, respectively. At 500 and 750MHz, 12 and 8 scans were recorded for each acquisition with recycle delays of 1.2s and 2s, respectively, to ensure sufficient recovery of ^1H magnetisation. ^{15}N-T_2 data were measured using a spin-echo pulse sequence with a CPMG delay of 325μs. Dipolar and CSA cross-correlations were removed by application of proton 180° pulses every 5ms for T_1 experiments and in the middle of the basic CPMG block for T_2 experiments (Boyd *et al.*, 1990; Kay *et al.*, 1992b). Inaccuracies in T_2 measurements associated with sample heating were reduced by inclusion of a train of ^{15}N refocusing pulses and delays after the pulse sequence, making the total number of ^{15}N refocusing pulses the same for each T_2 experiment (Wang & Bax, 1993). The effects of heating were assessed by comparing NH chemical shift changes in T_2 spectra recorded with the minimum and maximum number of ^{15}N refocusing pulses. In this case, no difference was observed, so the external temperature control did not need to be adjusted at either field strength to compensate for heating effects. Each experiment was preceded by ~15 minutes of dummy scans to allow peak intensities to be reproduced reliably, and each series of experiments consisted of 11 to 12 auto-correlation spectra with increasing ^{15}N relaxation time delays, including one repeat experiment (see below). Relaxation delays used to collect T_1 data were 20.0, 40.1 (measured twice), 80.2, 120.3, 180.4, 260.6, 360.8, 481.0, 701.5 and 1002.1ms at 500MHz, and 20.0, 40.1 (measured twice), 60.1, 120.2, 200.4, 300.6, 501.0, 701.4, 1001.9 and 1502.9ms at 750MHz. The first data point was, however, excluded from T_1 fits for all residues, as in all cases it did not fit the decay curve. Delays used to collect T_2 data were 6, 12 (measured twice), 18, 36, 48, 66, 90, 120, 150, 210 and 300ms at both 500 and 750MHz. Delays were selected to approximately sample all intensity decays, with the largest delay being at least 1.5 times the average relaxation time constant.

Pairs of ^1H-detected ^1H-^{15}N heteronuclear NOE experiments with and without NOE were recorded at 500 and 750MHz with acquisition times of 102.4 and 81.9ms in

F_2 (^1H) and 110.4 and 79.7ms in F_1 (^{15}N), respectively. In experiments with NOE, ^1H saturation was effected by means of a train of 120° flip-angle pulses at 10ms intervals for 3 and 4.5s at 500 and 750MHz, respectively. 128 scans and 96 scans were recorded at these field strengths, respectively, for each t_1 transient.

6.1.3 Data Processing

All spectra were processed using Felix 2.3 (Biosym, Inc.) with mild resolution enhancement in both dimensions to optimise resolution whilst maintaining good signal-to-noise. All spectra recorded in one series were processed identically. Overlapped resonances were excluded from the analysis.

6.2 Analysis and Results of ^{15}N Relaxation Data

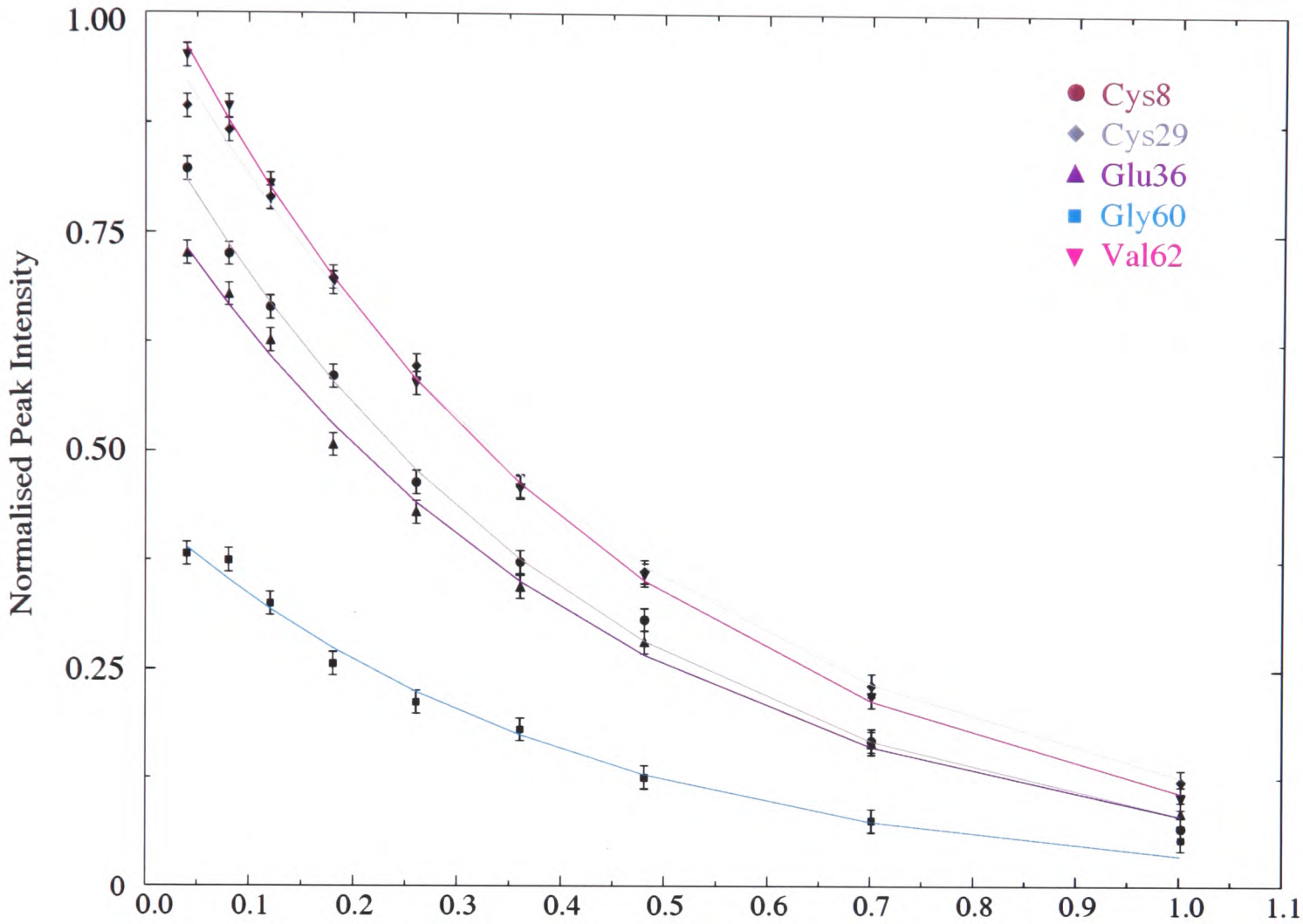
6.2.1 ^{15}N-T_1, ^{15}N-T_2 and ^1H-^{15}N heteronuclear NOE

Non-overlapping peaks in all spectra were picked and assigned, using assignments determined in Chapter 5 (Figure 5.5), and the intensities of these peaks were extracted. T_1 and T_2 relaxation time constants were derived from two-parameter exponential fits to resonance intensities, and fits to representative T_1 and T_2 intensity decays are shown in Figure 6.1. Errors were determined from the standard deviations of differences in the peak intensities in the two spectra that were recorded with the same relaxation delay (Palmer *et al.*, 1991). Several residues had to be excluded from the analysis at this point, since use of neither a two- nor a three-parameter model for fitting the resonance intensities was statistically justified. These fits may have been inadequate due to incorrect estimation of noise contributions using comparison of two spectra, which can lead to unreasonably high χ^2 values and subsequent rejection of some fits, especially for high signal-to-noise data (Werner *et al.*, 2000b). The heteronuclear NOE

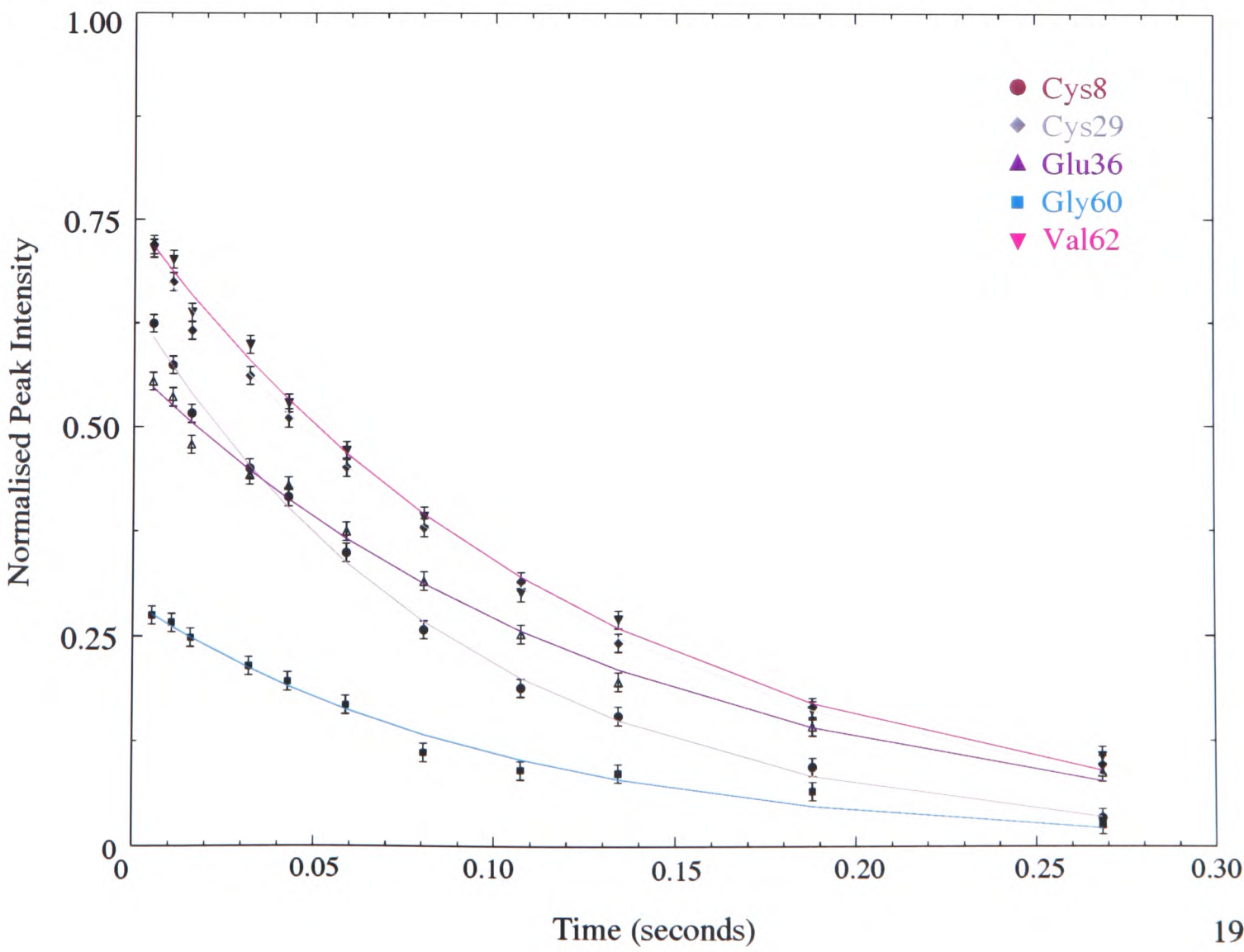
Figure 6.1 : Intensity Decays of (a) T_1 and (b) T_2 Relaxation Measurements for a Representative Set of Residues from cbEGF12-13 at pH 6.5, 35°C and 500MHz

The curves represent the best fit to the data using mono-exponential decays.

(a)



(b)



effect was calculated as the ratio of resonance intensities in spectra recorded with and without NOE. Errors were estimated from the signal-to-noise ratio of each spectrum.

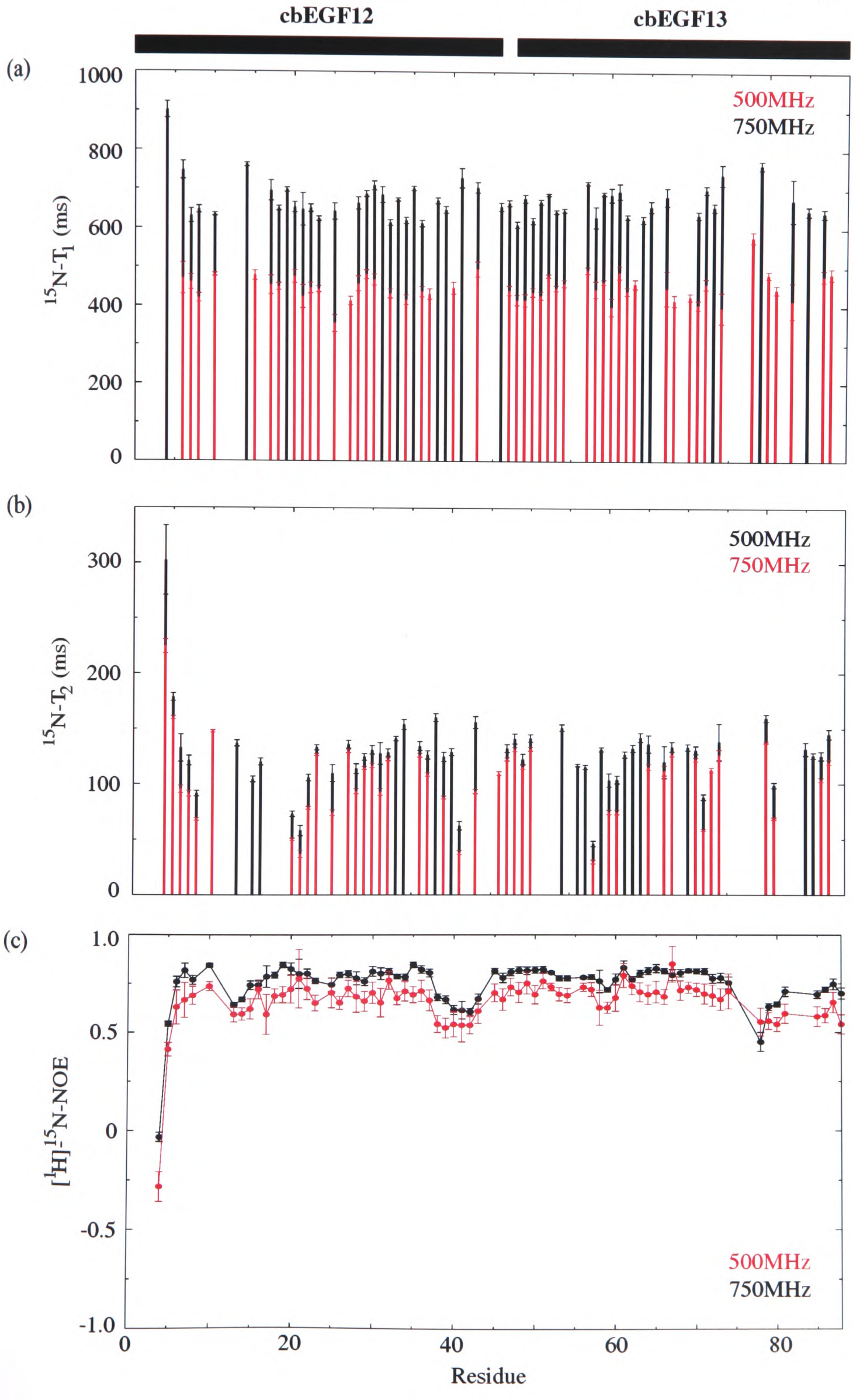
Experimental T_1, T_2 and NOE values at 500 and 750MHz, i.e., 11.7 and 17.6T, are plotted versus residue number in Figure 6.2. At 500MHz, the average T_1 value is 449ms, increasing to 674ms at 750MHz (Figure 6.2a). The average T_1 value for cbEGF12-13 at 500MHz is almost identical to that observed for the cbEGF32-33 domain pair, but a pronounced difference in the average T_1 values is observed for these pairs at 750MHz, i.e., 674ms and 619ms for cbEGF12-13 and cbEGF32-33, respectively (Werner *et al.*, 2000a). Although the residue sets used to calculate these averages may differ slightly, the difference in T_1 values observed at 750MHz may be indicative of increased anisotropy of the cbEGF12-13 pair.

For cbEGF32-33, a striking difference in T_2 values was observed for each domain of the pair, since cbEGF33 showed only small variations in T_2, whereas many residues in cbEGF32 had reduced T_2 values (Werner *et al.*, 2000a). For the cbEGF12-13 pair, this difference is not observed, and, in contrast, several residues in the C-terminal cbEGF13 domain have markedly low T_2 values (Figure 6.2b). At 500MHz, the average T_2 values are 127ms for cbEGF12-13, 129ms for cbEGF12 and 125ms for cbEGF13. At 750MHz, the T_2 average value for the pair is 104ms, with values of 99 and 104ms for the 12 and 13 domains, respectively. The lack of a prominent difference in T_2 values across the cbEGF12-13 pair, in contrast to results for cbEGF32-33, may relate to the different calcium binding affinities observed for these domains (Chapter 4; see below).

The average ^{1}H-^{15}N NOE values for residues in regions of secondary structure, are 0.69 and 0.77 at 500 and 750MHz, respectively. These values are comparable to the corresponding values obtained for cbEGF32-33 (0.64 and 0.75, respectively; Werner *et al.*, 2000a). The observed field dependence is consistent with a molecule in the slow tumbling limit (Figure 6.2c). Significantly reduced NOEs are observed for residues at the N-terminus of cbEGF12, before the start of the major β-sheet of this domain; for

Figure 6.2 : Relaxation Parameters (a) T_1, (b) T_2, and (c) NOE at 500 and 750MHz for cbEGF12-13 as a Function of Residue Number

The boundaries of cbEGF12 and cbEGF13 are indicated at the top of the figure.
For full details of these experimental values and the associated experimental uncertainties see Appendix VI.

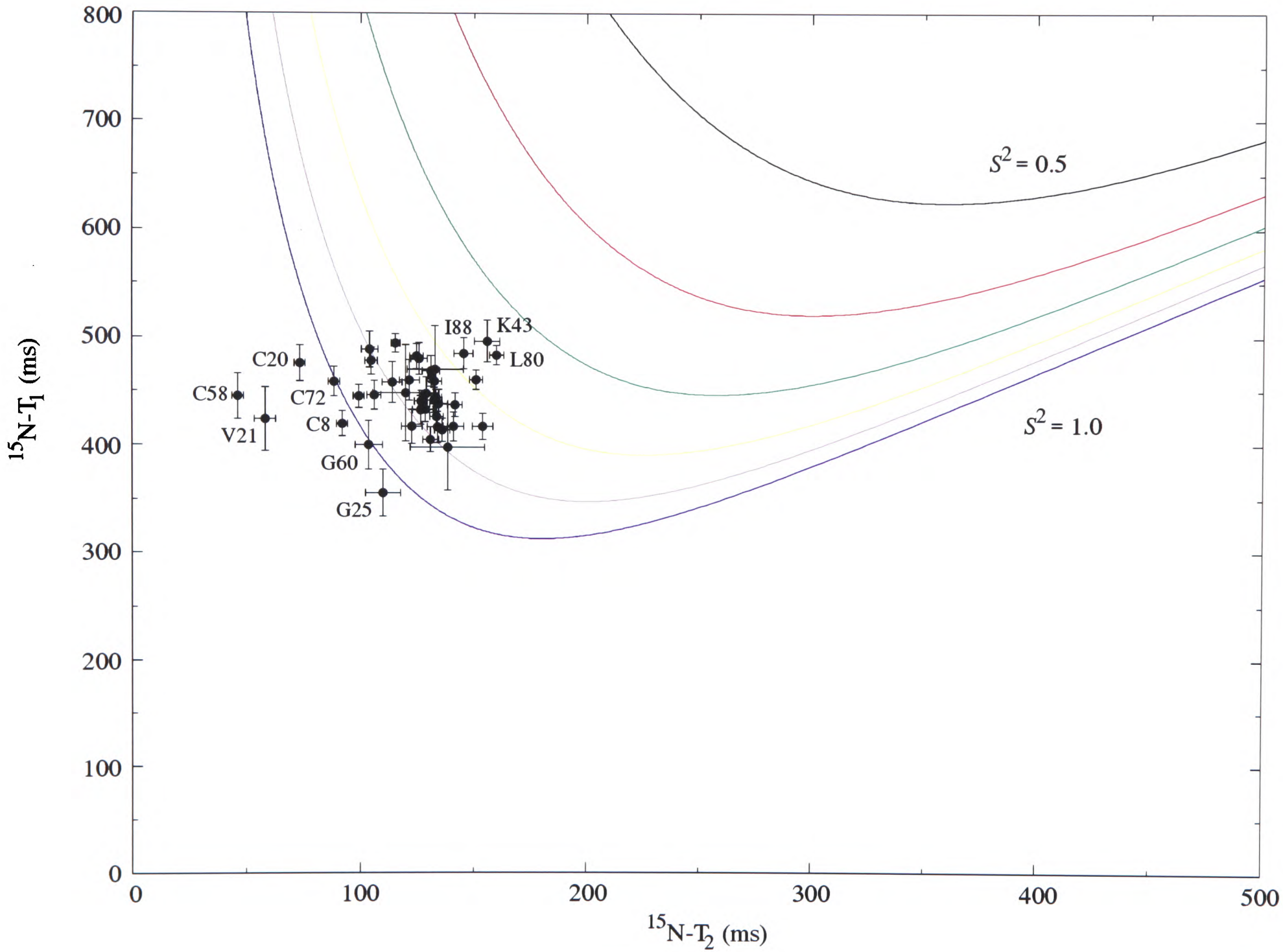


G38-K43 of cbEGF12, located in the turn joining the strands of the minor β-sheet; and for residues in the minor β-sheet (Q79-I88) of cbEGF13. It is interesting that the former and latter of these observations were noted for the cbEGF32-33 domain pair (Werner *et al.*, 2000a), indicating that they may be characteristics of all Class I cbEGF domain pairs. A significantly reduced NOE value was not seen for the single linker residue (M46) of the cbEGF12-13 domain pair, indicating, in agreement with cbEGF32-33 data (Werner *et al.*, 2000a), that Class I cbEGF domain pairs possess a rigid interdomain linker.

A plot of T_1 versus T_2 for each residue, overlaid on parametric curves of T_1 and T_2 as a function of order parameter, S^2, was used to provide a qualitative assessment of relaxation data for the cbEGF12-13 domain pair (Figure 6.3; Werner *et al.*, 2000b). It can be seen that residues cluster together in a relatively small region of this plot indicating that their relaxation properties can be described by diffusion of the molecule as a whole. Also, there are no systematic differences between residues from cbEGF12 and cbEGF13, suggesting that these domains tumble as a single unit. Residues with T_1 and T_2 values shifted to the left, i.e., small T_2 values, which lie outside the theoretical curves, are affected by motions on the μs to ms time-scale (R_{ex}), and Figure 6.3 indicates that residues C8, C20, V21, G25, C58, G60 and C72 are affected by such motions. Residues with T_1 and T_2 values shifted to the upper right region are likely to be affected by motions on a time-scale smaller than the overall tumbling time of the molecule, and residues located here are expected to have significantly reduced NOE values. Figure 6.3 indicates K43, L80 and I88 are likely to be affected by these fast motions, in agreement with the NOE results discussed above.

Figure 6.3 : Backbone ^{15}N Relaxation Time Constants for cbEGF12-13 Measured at 500MHz

Outlying residues are labelled by residue name. The continuous lines are calculated T_1 and T_2 values as a function of the order parameter, S^2, using the Lipari & Szabo model (1982a; 1982b). Curves with values of S^2 equalling 0.5 (black), 0.6 (red), 0.7 (green), 0.8 (yellow), 0.9 (brown) and 1.0 (blue) are shown, with the lowest and highest values labelled accordingly.



6.2.2 Motion on the Micro- to Millisecond Time-Scale

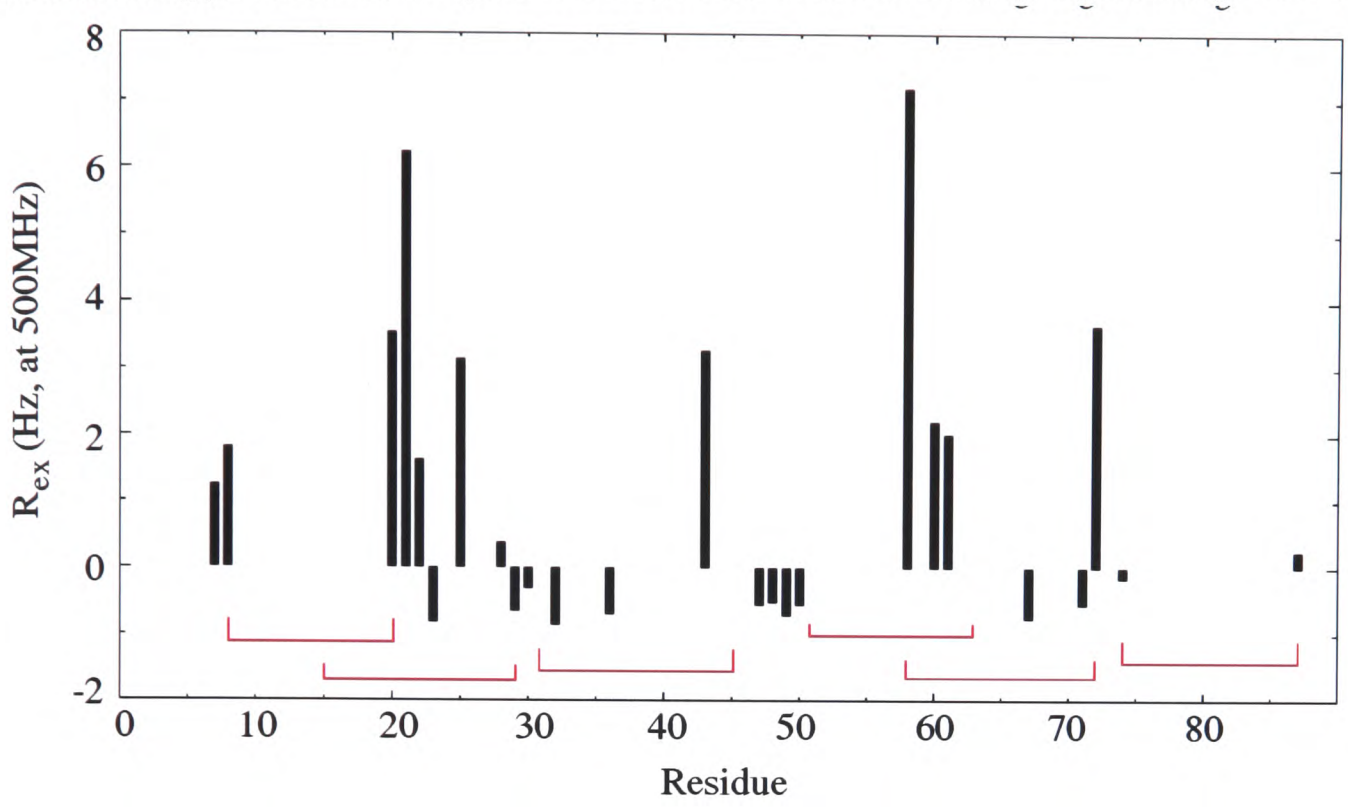
To allow a more quantitative analysis of motion on the µs to ms time-scale, R_{ex} terms were determined from the field dependence of R_2-R_1/2 (Phan *et al.*, 1996). As described in section 2.4.6, provided that the spectral density at the proton frequency is small compared to the spectral density at zero frequency, it has been shown that R_2-R_1/2 is a linear function of the square of the spectrometer field strength (Phan *et al.*, 1996). The slope depends on the CSA value, generally assumed to be -170ppm (Tjandra *et al.*, 1996a), and on the R_{ex} terms.

R_{ex} contributions to 25 residues of the cbEGF12-13 pair, 13 in cbEGF12 and 12 in cbEGF13, could be estimated using this analysis, and the results are shown in Figure 6.4. Of the residues that could be analysed, the results indicate significant R_{ex} terms, $> \pm 1$, for seven residues of cbEGF12 and four of cbEGF13, including all residues predicted to have R_{ex} terms from Figure 6.3. The observation that cbEGF13 of this domain pair contains four residues with $R_{ex} > \pm 2$, i.e., C58, G60, G61 and C72, contrasts the results obtained for cbEGF32-33, where the C-terminal domain contained only a single residue with $R_{ex} > \pm 2$ (Werner *et al.*, 2000a). C58 and C72 form the 2-4 disulphide bond of cbEGF13, suggesting that disulphide bond isomerisation may play a role in the slow dynamics of this domain. Such slow motions have been associated with BPTI (Otting *et al.*, 1993; Loria *et al.*, 1999). This isomerisation could potentially affect the dynamics of G60 and G61 due to the close proximity of C58 to these residues. Conversely, the presence of two glycines, the most flexible amino acid, near C58 could be responsible for the µs to ms motions observed for the 2-4 disulphide bond of cbEGF13. In cbEGF12, cbEGF32 and cbEGF33, only a single glycine is located at this position (see section 7.1). G60 and G61 are located in a turn before, and at the

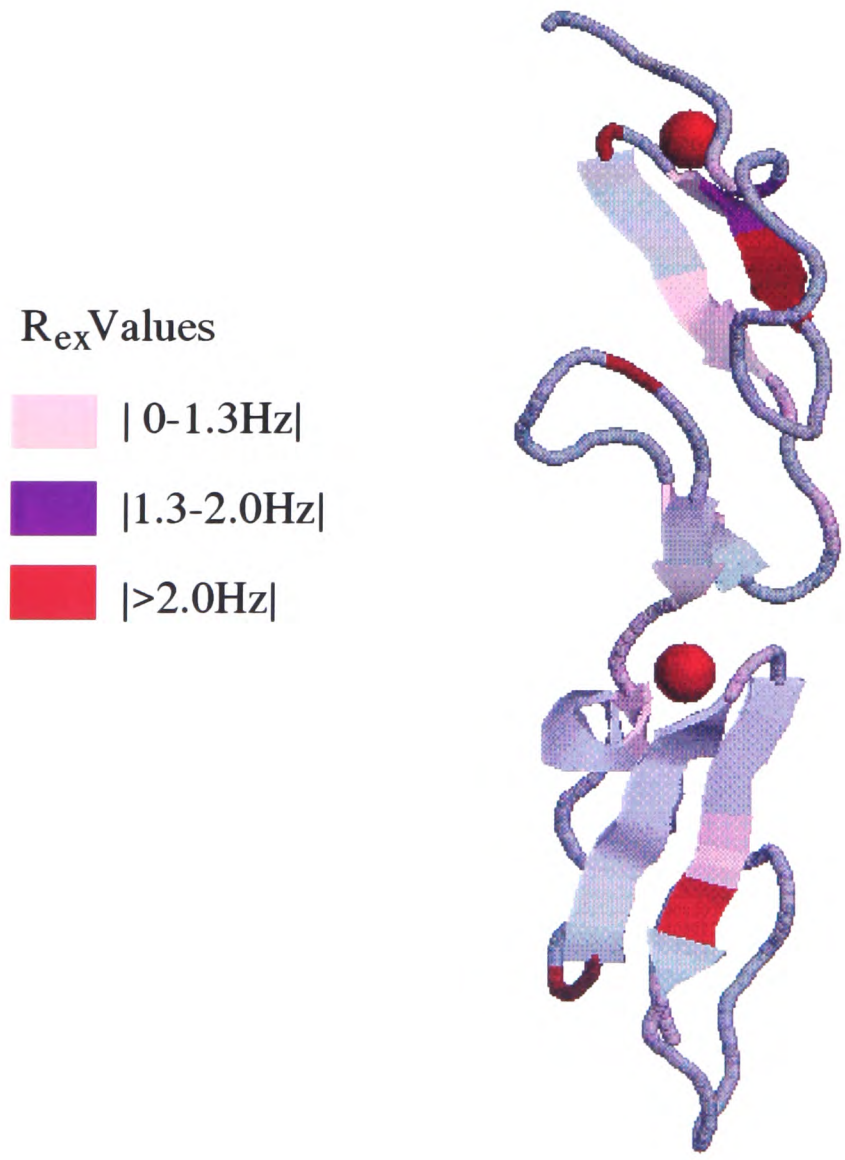
Figure 6.4 : R_{ex} Contributions (in Hz, at 500MHz) from Analysis of the Field Dependence of T_1 and T_2 Relaxation Time Constants of Calcium Bound cbEGF12-13; (a) as a Function of Residue and (b) Mapped onto the Average NMR Solution Structure

In (a), the positions of disulphide bonds are indicated in red and in (b), calcium atoms are represented by red spheres. R_{ex} contributions are usually positive, therefore the negative values indicated in (a) may have no meaning and may purely reflect the errors in the fitting procedure. An exception to this case occurs when the exchange regime changes between the two fields.

(a)



(b)



edge of, the major β-sheet of cbEGF13, respectively, and G61 corresponds to the G1127S mutation which causes a mild variant of MFS (Chapter 7).

Cysteines 8 and 20 of cbEGF12, which form the 1-3 disulphide bond, also demonstrate significant exchange contributions, whilst other cysteine residues that could be analysed by this method, i.e., C29, C74 and C87, are not significantly affected. The C8-C20 disulphide bond of cbEGF12 connects the N-terminus to the major β-sheet, and the corresponding disulphide bond in cbEGF32 has also been shown to be affected by R_{ex} terms (Werner *et al.*, 2000a). Several residues close to C8 and C20 of cbEGF12 are affected by R_{ex} terms in the range ~1-6Hz, i.e., E7, V21, N22, an effect more dramatically observed in cbEGF32 (see below; Werner *et al.*, 2000a). The final two residues in cbEGF12-13 with significant R_{ex} terms are in loop regions, i.e., G25, in the turn joining strands of the major β-sheet of cbEGF12, and K43, in the loop joining strands of the minor β-sheet of cbEGF12.

The calcium binding sites in cbEGF12-13 comprise residues D4, I5, D6, E7, N22 and T23 in cbEGF12, and D47, I48, D49, E50, N65 and T66 in cbEGF13. Domain 13 binds calcium at least 53 times more strongly than domain 12, and calcium saturation of both sites was achieved in this study using results of previous calcium binding studies (Chapter 4; Smallridge *et al.*, 2000), and by monitoring calcium-dependent changes in 2D HSQC spectra (section 5.1.1). Exchange contributions could be determined for the ligands E7, N22, T23, D47, I48, D49 and D50, which all have R_{ex} values less than ±2Hz. Values are slightly larger for cbEGF12 ligands, although this could result from their proximity to the C8-C20 disulphide bond. The majority of R_{ex} terms for calcium ligands in the cbEGF32-33 pair similarly had values less than ±2Hz (Werner *et al.*, 2000a), and it seems that in cbEGF12-13, consistent with the situation observed in cbEGF32-33, a substantial part of cbEGF12, in particular the N-terminal region, may be affected by μs to ms motion even when this site is fully

occupied by calcium. This is also consistent with changes in dynamics observed upon ion binding for calmodulin (Barbato *et al.*, 1992; Tjandra *et al.*, 1995; Zhang *et al.*, 1995) and calbindin D_{9K} (Akke *et al.*, 1993). In addition, in the N-terminal domain of chicken skeletal troponin C, increased disorder in the ion-free state has been shown to correlate with decreased affinity (Gagne *et al.*, 1998).

The fact that several residues of the major β-sheet of cbEGF12 have significant R_{ex} terms, an effect not seen for cbEGF13, suggests that the lower calcium binding affinity observed for this domain relative to cbEGF13 (~1.6mM versus < 30μM) may be a result of slow dynamic processes that compromise the formation of a well-defined binding site. It is noteworthy that cbEGF32 in cbEGF32-33, which has a calcium binding affinity more than five times weaker than cbEGF12 (~9mM; Knott *et al.*, 1996), has R_{ex} terms with an upper limit of 21Hz (Werner *et al.*, 2000a) compared to an upper limit of 6Hz for cbEGF12. The lack of significant exchange contributions, where data were available, for residues in the minor β-sheet of cbEGF12 and the major β-sheet of cbEGF13 could indicate that cbEGF13 is able to form a well-defined calcium binding site, explaining the relatively high calcium binding affinity observed for this domain. Exchange contributions for C58, G60, G61 and C72 are unlikely to affect calcium binding by cbEGF13, since these residues are not located close to this site.

6.2.3 Estimation of the Rotational Diffusion Tensor

The isotropic correlation times of cbEGF12-13, and the individual domains, were estimated at 500 and 750MHz using T_1/T_2 ratios of residues that did not exhibit significant exchange effects or have significantly reduced NOE values. The results are shown in Table 6.1, and, at each field strength, the estimated correlation times are the same within experimental error. The correlation times do, however, slightly differ when comparing data from different field strengths. Variation in isotropic

correlation time at different field strengths has been previously noted for the HIV-1 protease (Tjandra *et al.*, 1996b), although the cause of this variation is still unclear. The 1.6mM sample was used to confirm that non-specific aggregation effects were not significant.

Table 6.1 : Isotropic Correlation Times for cbEGF12-13, cbEGF12 and cbEGF13 Determined from T_1/T_2 Ratios at Two Different Field Strengths

	cbEGF12-13 (ns)	**cbEGF12 (ns)**	**cbEGF13 (ns)**
500MHz	5.8 ± 0.1	6.0 ± 0.1	5.7 ± 0.2
750MHz	5.4 ± 0.1	5.4 ± 0.2	5.5 ± 0.1

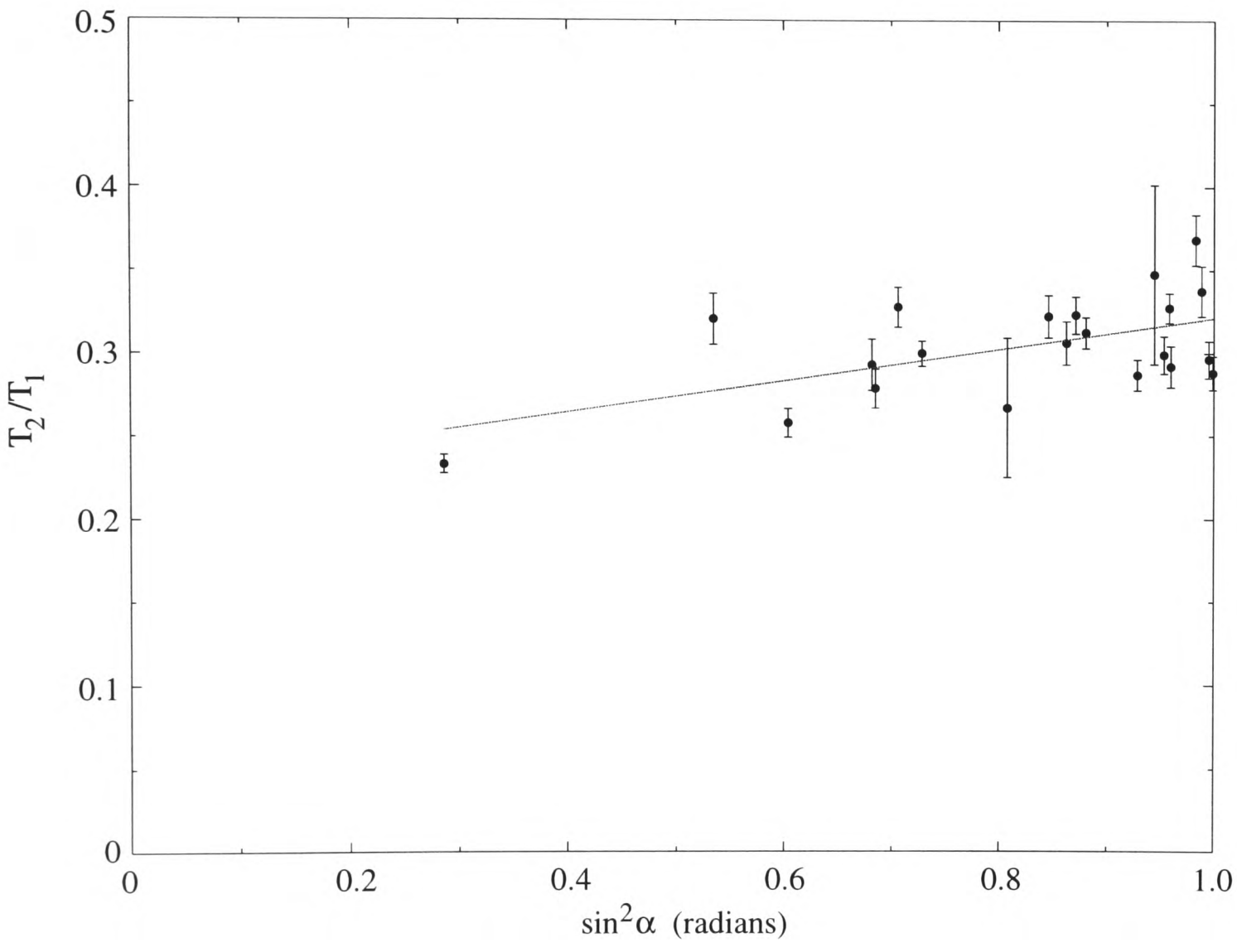
From the elongated structure observed for the cbEGF12-13 domain pair (Chapter 5), and from homology to cbEGF32-33, it was decided that quantitative analysis of cbEGF12-13 relaxation data should include anisotropic diffusion. For the analysis, the diffusion tensor was obtained by a global least-squares fit of the spectral density functions of the T_1/T_2 ratios to the experimental values (Palmer *et al.*, 1991; Mandel *et al.*, 1995; Boyd & Redfield, 1998). Three models were tested : a sphere ($D_x=D_y=D_z$; $D=(D_x+D_y+D_z)/3$), a symmetric top ($D_{\|}=D_z$ and $D_{\perp}=D_x+D_y$), and a fully asymmetric tensor $D_xD_yD_z$ (Woessner, 1962). At higher field strengths, due to the quadratic field dependence of CSA, the angle between the CSA tensor and the dipolar vector can become significant (Boyd & Redfield, 1998). As this can complicate relaxation data analysis for anisotropic molecules, 500MHz data were used for this analysis.

Relaxation studies of cbEGF32-33 suggested that this molecule was best represented by a symmetric top, prolate, model (Werner *et al.*, 2000a). For this model, T_2/T_1 ratios are expected to depend on the angle, α, that the NH bond vector makes with the main axis of the diffusion tensor (Figure 6.5). For cbEGF32-33, whilst T_2/T_1 ratios depended on α, the specific topology of this pair resulted in many more NH bond

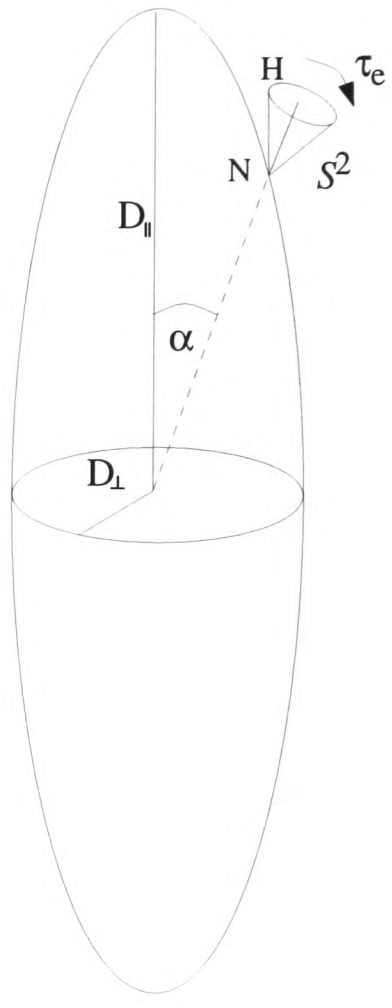
Figure 6.5 : (a) Rotational Diffusion Anisotropy of cbEGF12-13 and (b) Schematic Diagram of a Prolate, Symmetric Top Model with Principal Axes $D_{\parallel}$ and $D_{\perp}$

In (a), the diffusion tensor for the average structure was calculated and the angle, α, between the NH bond vector and the largest principal component of the diffusion tensor (see (b)) was determined. The linear least-squares fit to the data is shown for illustration, and the slope is indicative of a prolate molecule. In (b), the order parameter, S^2, and the internal correlation time , τ_e, are visualised in a diffusion-in-a-cone interpretation.

(a)



(b)



vectors with α = ~90° rather than with α = 0° (Werner et al., 2000a). This was also found to be the case for cbEGF12-13 (Figure 6.5a).

Local fast and slow dynamics, as indicated by decreased NOE values and significant R_{ex} terms, respectively, affect several residues of the cbEGF12-13 pair, yielding T_2/T_1 ratios smaller or larger than would be predicted by anisotropic diffusion alone. A 21 residue subset was therefore selected from the 500MHz data, with NOE values > 0.65 (lower boundary of ^{15}N-NOE average) and R_{ex} terms < ±1, for estimation of the rotational diffusion tensor. The representative 21 residues were from the central region of the domain pair, i.e., between and including T23 and C74. The T_1/T_2 ratios of these residues were used to determine the size and orientation of the diffusion tensor using a global least-squares fit to these data (Palmer *et al.*, 1991; Mandel *et al.*, 1995). For cbEGF32-33, it was found that fitting relaxation data to both the average structure and to all 22 members of the final NMR structural ensemble yielded comparable results (Werner *et al.*, 2000a). Since cbEGF12-13 is homologous to cbEGF32-33, and has a well-defined structure (Chapter 5), estimation of the rotational diffusion tensor on the basis of a single average structure was considered justified. The results of fitting the 21 residue subset to the isotropic, symmetric top and fully asymmetric models are shown in Table 6.2. An 18 residue subset, in which three residues with relatively large χ^2 values were excluded, was also tested (Table 6.2).

From Table 6.2, it can be seen that for both residue subsets the best fit to the data was obtained using a prolate, symmetric top model. For the 21 and 18 residue subsets, comparison of isotropic and prolate models indicates an improvement with only a 0.005 and 0.001% probability, respectively, of occurring by chance. A fully asymmetric model was also tested, and comparison of prolate and asymmetric models for the 21 and 18 residue sets showed that any improvement in χ^2 had a 100% and a 14% probability, respectively, of occurring by chance. This reflects the near-degeneracy of the small axes of the diffusion tensor for cbEGF12-13. Inclusion of an

angle of 0 to +40° between the dipolar and CSA tensor had no statistically significant effect on the results (Boyd & Redfield, 1998).

Table 6.2: Diffusion Parameters for cbEGF12-13 at pH 6.5 and T=35°C

Parameters were obtained by fitting experimental T_1/T_2 values at 500MHz for various residue subsets to the average structure of cbEGF12-13.

Residues	$(6D)^{-1}$ (ns)	$^*2D_{zz}/(D_{xx}+D_{yy})$	$^*D_{xx}/D_{yy}$	θ (deg)	φ (deg)	ψ (deg)	χ^2_{total}	$Q_1^{\#}$	$Q_2^{\#}$
21 Residue Subset									
Isotropic	5.82	-	-	-	-	-	193.5	-	-
Axially symmetric									
Oblate¶	5.84	0.68	-	88	14	-	160.0	-	-
Prolate¶	5.81	1.87	-	2	-35	-	82.1	4.6×10^{-5}	-
Asymmetric	5.81	1.90	1.03	2	-72	142	82.1	-	1.0
18 Residue Subset									
Isotropic	5.81	-	-	-	-	-	141.1	-	-
Axially symmetric									
Oblate¶	5.95	0.71	-	95	-166	-	89.1	-	-
Prolate¶	5.91	1.86	-	1	-15	-	38.5	1.1×10^{-5}	-
Asymmetric	5.95	1.87	1.16	5	-99	8	29.4	-	0.14

* Values in 10^7Hz.

¶ Fits were performed assuming a prolate ($2D_{zz}/(D_{xx}+D_{yy}) > 1.0$) or an oblate ($2D_{zz}/(D_{xx}+D_{yy}) < 1.0$) diffusion tensor.

Q_1 was calculated for comparison of isotropic and axially symmetric (prolate) models, whereas Q_2 was calculated to compare axially symmetric (prolate) and asymmetric models (see section 2.4.5).

Despite the slight improvement in χ^2 on moving from the 21 to the 18 residue subset, the prolate, symmetric top models yield comparable values of τ_m, $D_{\parallel/\perp}$, θ, and φ for both sets. Data from the 21 residue set were therefore employed in further relaxation analyses, since this improved the sampling of cbEGF12-13. The good alignment of the diffusion tensor with respect to the inertia tensor (θ = 2 and 1° for the 21 and 18 residue set, respectively), and the dramatic improvement of χ^2 using the prolate, symmetric top instead of the isotropic model, confirms the extended structure of cbEGF12-13 seen in Chapter 5.

6.2.4 Model-Free Analysis of Internal Dynamics

After determination of the diffusion tensor for cbEGF12-13, T_1, T_2 and NOE data were interpreted in terms of internal dynamics, using an extended version of the model-free approach described in section 2.4.4 (Lipari & Szabo, 1982a; 1982b). Fast motions on two time-scales were included in the limit that the faster internal correlation time could be neglected (Clore *et al.*, 1990). Within this time-scale, faster motions could be characterised by S^2_f, whereas slower motions were characterised by S^2_s and τ_e. This analysis was implemented using the program Modelfree4 (Palmer *et al.*, 1991; Mandel *et al.*, 1995; program kindly provided by A. G. Palmer, Columbia University). Various dynamic models of increasing complexity (Table 2.4) were chosen for each residue using χ^2 values and F-statistics (Mandel *et al.*, 1995; Werner *et al.*, 2000b; section 2.4.5). Models were compared using the F-test, followed by calculation of the probability, Q, that this F could have been obtained by chance (Press *et al.*, 1990; section 2.4.5). To assign a model to each residue, the diffusion tensor for the average structure was fixed to the values obtained for the 21 residue subset described above. The nomenclature and selection process of the five models of internal dynamics considered here initially followed Mandel *et al.* (1995). After model selection, the prolate, symmetric top diffusion tensor and the parameters for internal dynamics were optimised using all available data for the average cbEGF12-13 structure. The final statistics were based on 500 Monte Carlo simulations.

A problem arose when all parameters were freely simulated, however, since, instead of remaining close to their starting values (Table 6.2), τ_m and $D_{\parallel/\perp}$ drifted to approximately 5.7ns and 2.4, respectively. This drift could not be prevented by the exclusion of residues with high χ^2 values or by model adjustment, and Figure 6.5a highlights a potential reason for this drift. It is clear that not all residues can be satisfied

by a single $D_{\parallel/\perp}$, and, since only one residue has a $\sin^2\alpha$ value less than 0.53 radians, the residue set used to plot this graph can greatly affect the slope produced and therefore the $D_{\parallel/\perp}$ value obtained. As the rotational diffusion tensor obtained for a prolate model using 21 residues from cbEGF12-13 was considered accurate based on comparison of two representative residue sets (Table 6.2), it was concluded that the best solution was to keep the rotational diffusion tensor fixed during 500 Monte Carlo simulations. The 21 residues used to determine this tensor were also fixed to model 1, whilst all remaining residues were simulated. The results of model-free analysis of cbEGF12-13, using a prolate, symmetric top model for the molecular tumbling, are shown in Table 6.3 and Figure 6.6.

The average order parameter for all residues in secondary structure, where data were available, is $\langle S^2 \rangle = 0.83$ (Figure 6.6a). This is consistent with the corresponding $\langle S^2 \rangle$ value obtained for cbEGF32-33 (0.79) (Werner *et al.*, 2000a), and suggests that for both these domain pairs fast motions, in general, only have small amplitudes. This is also confirmed by the high ^1H-^{15}N NOE ratios for both pairs (Figure 6.2c; Werner *et al.*, 2000a).

Two residues of cbEGF12-13 have S^2 values of 1.0 from model-free analysis (Figure 6.2a), i.e., G25, located in a turn at the top of the major β-sheet of cbEGF12, and G60, located in a turn before the major β-sheet of cbEGF13. Since glycines are the most flexible naturally occurring amino acid, and both of these residues were observed to have significant exchange contributions from the field dependence of R_2-$R_1/2$, S^2 values of 1.0 seem highly unlikely. Reanalysis of spectral data showed that resonances arising from these residues are weak and broadened, which could have greatly hindered fitting of R_{ex} terms using Modelfree4. It is therefore likely that these residues have exchange contributions and S^2 values < 1.0, a conclusion supported by the lower than average T_2 values observed for these residues at 750MHz (Figure 6.2b).

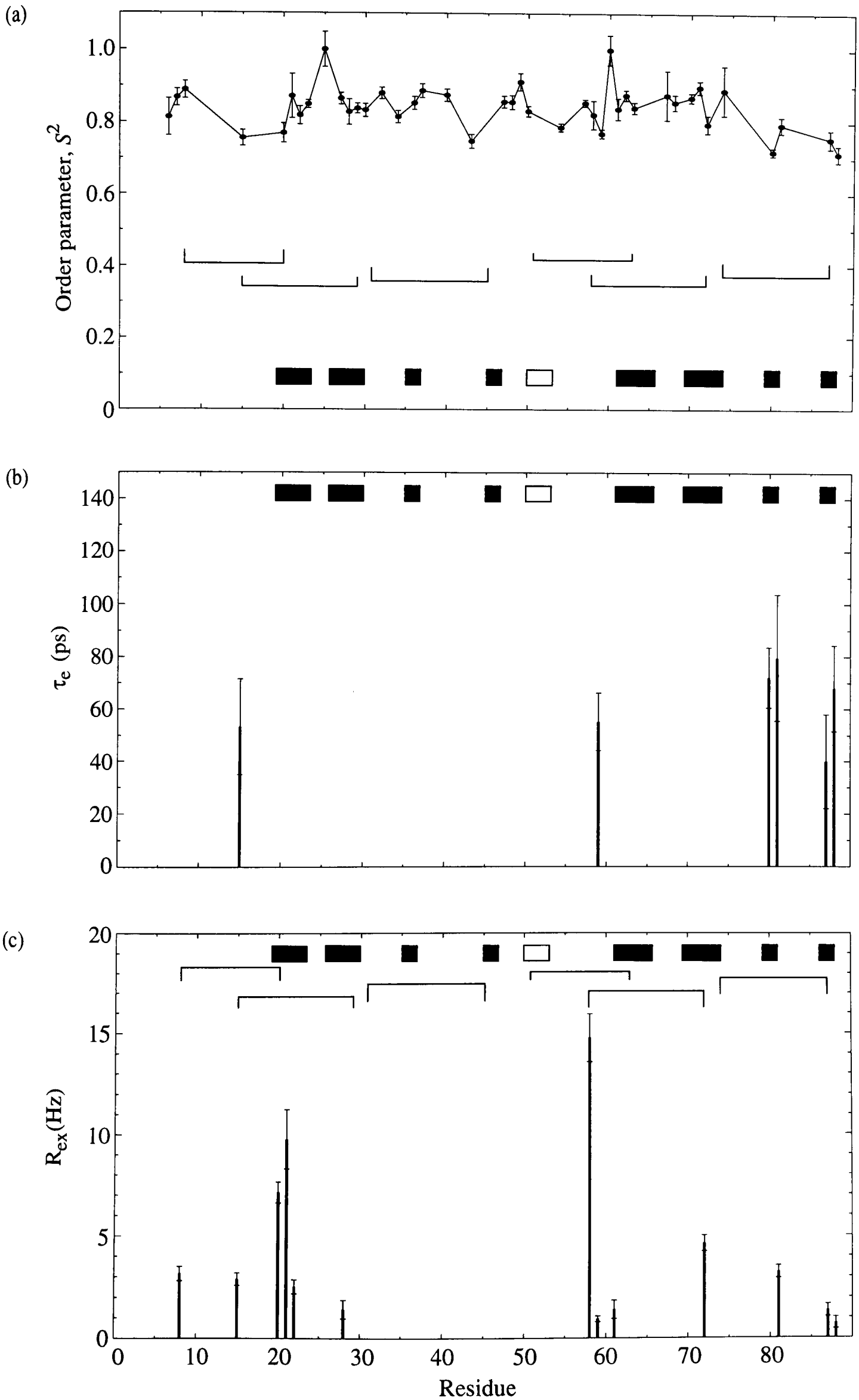
Table 6.3 : Lipari-Szabo Motional Parameters for Calcium Bound cbEGF12-13 at pH 6.5 and 35°C

M represents the model of internal dynamics selected for each residue (see Table 2.4), and model 5 was not required for this analysis. Residues marked with * form the 21 residue subset used to determine the rotational diffusion tensor, and these residues were kept fixed at model 1 during the final simulation. The diffusion tensor was also fixed to the values determined in Table 6.2 for the final simulations (see text).

Residue	M	$\langle S^2 \rangle \pm \langle \Delta S^2 \rangle$	$\langle \tau_e \rangle \pm \langle \Delta \tau_e \rangle$	$\langle R_{ex} \rangle \pm \langle \Delta R_{ex} \rangle$	$\langle \chi^2 \rangle$	$\langle \chi^2_{95\%} \rangle$
D6	1	0.81 ± 0.05			5.0	6.6
E7	1	0.87 ± 0.02			1.6	6.0
C8	3	0.89 ± 0.02		3.2 ± 0.4	2.9	4.1
C15	4	0.76 ± 0.02	53 ± 18	2.9 ± 0.3	0.0	0.0
C20	3	0.77 ± 0.03		7.1 ± 0.5	0.3	4.0
V21	3	0.87 ± 0.06		9.8 ± 1.5	0.0	5.0
N22	3	0.82 ± 0.03		2.5 ± 0.3	0.5	3.7
T23*	1	0.85 ± 0.01			9.9	6.2
G25	1	1.00 ± 0.05			4.0	7.3
F27*	1	0.87 ± 0.02			10.9	5.4
E28	3	0.83 ± 0.04		1.4 ± 0.5	1.3	3.6
C29*	1	0.84 ± 0.01			11.4	5.9
K30*	1	0.83 ± 0.02			2.5	6.0
D32*	1	0.88 ± 0.02			5.6	5.9
G34*	1	0.82 ± 0.02			11.6	7.0
E36*	1	0.85 ± 0.02			1.0	5.9
S37*	1	0.89 ± 0.02			8.0	5.3
M40	1	0.88 ± 0.02			17.4	5.9
K43	1	0.75 ± 0.02			6.1	5.9
D47*	1	0.86 ± 0.02			3.6	6.0
I48*	1	0.86 ± 0.02			2.2	5.9
D49*	1	0.91 ± 0.02			0.2	5.4
E50*	1	0.83 ± 0.01			3.6	6.0
D54*	1	0.79 ± 0.01			5.0	5.1
L57*	1	0.85 ± 0.01			5.7	5.8
C58	3	0.82 ± 0.04		14.7 ± 1.2	2.3	3.8
R59	4	0.77 ± 0.01	55 ± 11	0.9 ± 0.1	0.0	0.0
G60	1	1.00 ± 0.04			5.8	7.0
G61	3	0.84 ± 0.03		1.4 ± 0.4	0.0	3.7
V62*	1	0.88 ± 0.01			9.7	6.5
C63*	1	0.84 ± 0.02			12.3	6.0
E67*	1	0.88 ± 0.07			1.4	7.3
G68*	1	0.86 ± 0.02			13.4	6.3
Y70*	1	0.87 ± 0.01			2.1	6.4
R71*	1	0.90 ± 0.02			3.3	6.2
C72	3	0.80 ± 0.02		4.6 ± 0.4	1.8	4.2
C74*	1	0.89 ± 0.07			0.7	6.0
L80	2	0.72 ± 0.01	72 ± 12		0.3	3.6
S81	4	0.79 ± 0.02	80 ± 24	3.2 ± 0.3	0.0	0.0
C87	4	0.76 ± 0.03	40 ± 18	1.3 ± 0.3	0.0	0.0
I88	4	0.71 ± 0.02	68 ± 16	0.7 ± 0.3	0.0	0.0

Figure 6.6 : Model-Free Analysis of cbEGF12-13 Using a Symmetric Top Diffusion Model for Molecular Tumbling

Generalised order parameters (a), internal correlation times (b), and exchange contributions (c) are plotted versus residue number. Positions of β-strands are shown as black rectangles and the α-helix is shown as a white rectangle. Lines are drawn joining cysteine residues engaged in disulphide bonds.



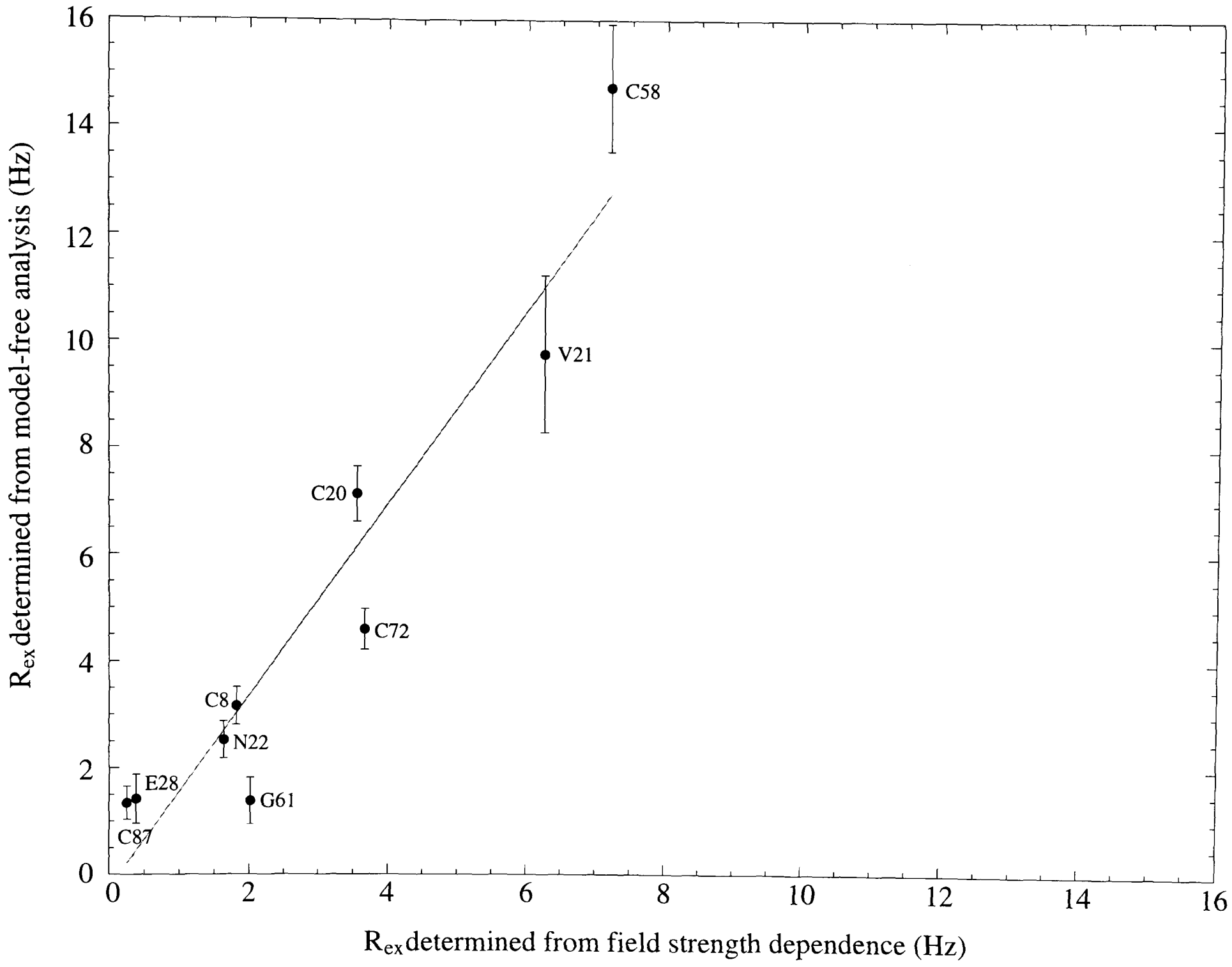
Order parameters could be determined for two residues involved in interdomain packing, E67 and G68, which have values of 0.88 and 0.86, respectively. These data, combined with the fact that the single interdomain linker residue, M46, and the aromatic residue involved in interdomain packing, Y35, have above average ^{15}N-NOE values (Figure 6.2c), indicate a well-defined molecule interface. This is consistent with results obtained for cbEGF32-33 (Werner *et al.*, 2000a), and with the similar isotropic correlation times obtained for the individual domains of cbEGF12-13 (Table 6.1). The NOE value for R17 is, however, below average, which is perhaps not surprising since this residue is not conserved in packing interactions in all Class I cbEGF domain pairs.

A significant number of residues require exchange terms in the fits (Table 6.3 and Figure 6.6c), and the exchange contributions fitted by model-free analysis are similar to those calculated from the field dependence of R_2-R_1/2 (section 6.2.2). For nine residues, R_{ex} values determined by these two methods could be directly compared, and the results of this comparison are shown in Figure 6.7. Although values obtained from model-free analysis are generally larger than those determined from the field dependence, there is an approximately linear relationship between these values. A difference in R_{ex} values determined by different methods was also observed for cbEGF32-33 (Werner *et al.*, 2000a), and this variation is likely to arise from difficulties in accurately estimating the relaxation rates of residues with broad resonances.

In contrast to cbEGF32-33 (Werner *et al.*, 2000a), residues requiring R_{ex} terms in cbEGF12-13 are not mainly localised to the N-terminal domain of this pair (Figure 6.6c). R_{ex} terms were required for six and seven residues of cbEGF12 and cbEGF13, respectively. In agreement with field dependence data, residues involved in and around the 1-3 disulphide bond of cbEGF12 have significant exchange contributions. Additional information was also obtained from model-free analysis for C15 of cbEGF12, which has a R_{ex} term of 2.9Hz, as well as a significant τ_e term. This cysteine residue forms the 2-4 disulphide bond with C29, hence, in agreement with the situation

Figure 6.7 : Comparison of Exchange Contributions Determined by Model-Free Analysis and the Field Dependence of R_2-R_1/2

A direct agreement between the two data sets would produce the dashed line, and the linear least-squares fit to the data is shown as a solid line.



observed for cbEGF32 (Werner *et al.*, 2000a), this bond may also be affected by motions on the μs to ms time-scale. C29, however, has no significant exchange contributions.

In cbEGF13, additional R_{ex} information was obtained for R59 and residues of the minor β-sheet of cbEGF13, i.e., S81, C87 and I88, with these residues also requiring τ_e terms. The fact that R59 requires an exchange term is not surprising, since it is located adjacent to C58 and G60-G61, which were all shown to be significantly affected by μs to ms time-scale motions from field dependence analysis. Whether these motions are a result of isomerisation of the C58-C72 disulphide bond or the intrinsic flexibility of glycine residues, or both, is unclear. The minor β-sheet of cbEGF13 is a highly dynamic structure, comprising residues manifesting reduced NOE values, as well as significant R_{ex} and τ_e terms. These motions may partially explain the non-ideal structure of this β-sheet (Chapter 5). The dynamics of this region are, however, unlikely to affect calcium binding by cbEGF13, since this β-like structure is not involved in formation of the calcium binding site.

6.3 Discussion of Results : Biological Significance

6.3.1 Correlation of Backbone Flexibility and Calcium Affinity

Relaxation data presented above confirm an extended structure for the calcium-bound cbEGF12-13 domain pair (Chapter 5). The diffusion tensor is a symmetric top with an axial ratio of 1.87:1 that aligns well with the inertia tensor. The increased size of this diffusion tensor, compared to that observed for cbEGF32-33 (1.55:1; Werner *et al.*, 2000a), implies that cbEGF12-13 is a more anisotropic

molecule. For cbEGF32-33, it was noted that the size of the diffusion tensor was considerably smaller than expected from the structure of this pair (Downing *et al.*, 1996), and this was attributed to the dynamic N- and C-terminal halves of cbEGF32 and cbEGF33, respectively (Werner *et al.*, 2000a). The increased anisotropy observed for cbEGF12-13 could therefore be due to increased rigidity of the N-terminal half of cbEGF12. The calcium affinity of cbEGF12 in the 12-13 pair (Chapter 4; Smallridge *et al.*, 1999) is more than five times stronger than that observed for cbEGF32 in cbEGF32-33 (Knott *et al.*, 1996), making it likely that the N-terminal half of cbEGF12 is more rigid. In support of this argument, comparison of dynamics data for the N-terminal halves of cbEGF12 (this study) and cbEGF32 (Werner *et al.*, 2000a) indicate that residues from this region of cbEGF12 have larger S^2 values and smaller R_{ex} terms than the corresponding residues of cbEGF32.

The high affinity of cbEGF13 could also improve the stability of the interdomain interface, aiding production of a more extended structure (Chapter 4; Smallridge *et al.*, 1999). The calcium affinity of cbEGF13 in the 12-13 pair is more than 11 times stronger than that observed for cbEGF33 in cbEGF32-33 (Knott *et al.*, 1996). A qualitative comparison of the number of unambiguous, interdomain NOEs observed for cbEGF12-13 and cbEGF32-33 (Downing *et al.*, 1996) highlighted that interdomain packing is stabilised by a similar number of restraints in both pairs, i.e., 37 and 35, respectively. In cbEGF12-13, however, an extra residue, R17, appears to stabilise the packing interaction without forming interdomain contacts (Chapter 5). Despite the flexibility of R17 indicated by a lower than average NOE value (Figure 6.2c), the methylene groups of this residue form hydrophobic contacts with Y35. This interaction could improve shielding of the calcium ion from solvent, provide a better defined binding site in cbEGF13 and/or stabilise interdomain packing interactions. Residues involved in interdomain packing and the single linker residue have been shown to be well ordered from S^2 and NOE values (section 6.2.4), confirming a well-defined domain interface for this pair, as well as the relatively small interdomain tilt

angle observed for the final structural ensemble (Chapter 5). Increased stability of the interdomain interface of cbEGF12-13 compared to cbEGF32-33 (Downing *et al.*, 1996; Werner *et al.*, 2000a), as a result of additional packing interactions and increased calcium binding affinity, could therefore also partially explain the observed difference in diffusion tensor magnitude.

The calcium binding affinities of cbEGF12 and cbEGF13 in a tandemly linked pair differ by a factor of more than 53, and dynamics data presented above highlight a potential explanation for this difference. Residues in the N-terminal half of the first domain of cbEGF12-13, including residues around the low affinity calcium binding site of cbEGF12, appear, in agreement with results for cbEGF32-33 (Werner *et al.*, 2000a), to be affected by μs to ms fluctuations. This suggests that the calcium binding site in cbEGF12 is not maintained in the correct conformation to bind calcium, which would be expected to reduce the affinity of this site due to unfavourable entropic and enthalpic effects. In contrast, only residues in the C-terminal half of cbEGF13, localised far from the high affinity calcium binding site, are affected by μs to ms motions. The observation, from studies of cbEGF12-13 (this study) and cbEGF32-33 (Werner *et al.*, 2000a), that motions on the μs to ms time-scale only appear to affect the N-terminal binding site of cbEGF domain pairs may have important implications for both the role of dynamics in determining calcium binding affinity of cbEGF domains and the role of calcium in relation to the molecular organisation of tandem cbEGF domains. This is likely to be relevant to many functional environments, since tandem cbEGF domains are found in a wide variety of proteins (Chapter 1; Downing *et al.*, 2000).

A correlation between linker flexibility and calcium binding affinity has been proposed for cbEGF domains (Chapter 4; Kettle *et al.*, 1999; Smallridge *et al.*, 1999; Werner *et al.*, 2000a; Yuan *et al.*, manuscript in preparation), and results obtained here further support this proposition. From available data for fibrillin-1 (Chapter 4; Figure

4.6), cbEGF domains located C-terminally in the Class I cbEGF pairs have significantly higher affinities than when either isolated or linked to a N-terminal, non-homologous domain by a flexible linker, e.g., TB6-cbEGF32 (Kettle *et al.*, 1999; Yuan *et al.*, manuscript in preparation), and the two cbEGF pairs investigated from human fibrillin-1, i.e., cbEGF12-13 and cbEGF32-33 (Werner *et al.*, 2000a), have both been confirmed to have rigid, interdomain linkers. In addition, comparison of data for cbEGF12-13 and cbEGF32-33 (Werner *et al.*, 2000a) has also highlighted that variations in C-terminal calcium binding affinity in Class I cbEGF domain pairs may be a consequence of variations in the conserved packing interaction. Establishing a correlation between changes in calcium affinity and both the rigidity of the interdomain linker and the rigidity/composition of the packing interaction will be of value in understanding the properties of other constructs.

6.3.2 Implications for Connective Tissue Microfibrils and MFS

Comparison of the slow dynamic behaviour of cbEGF12 and cbEGF13 suggests that the presence of a preceding domain can greatly influence the dynamics of these domains. In the absence of a N-terminally-linked domain, the N-terminal region of cbEGF12 exhibits conformational exchange on the μs to ms time-scale, an effect not observed for the N-terminal region of cbEGF13. These results are in agreement with those for cbEGF32-33 (Werner *et al.*, 2000a). It also appears, however, that the C-terminal half of the cbEGF13, far from the high affinity site, exhibits conformational exchange on this time-scale, an effect not observed for cbEGF33 (Werner *et al.*, 2000a). The cbEGF13 and cbEGF33 domains, however, both present a systematic decrease in heteronuclear NOE values towards their C-terminus, indicative of fast motions, which suggests that this region of cbEGF domain pairs may be sensitive to calcium binding and/or pairwise domain interactions. Hence, in the presence of calcium, the most ordered region of cbEGF domain pairs investigated to date (this

study; Werner *et al.*, 2000a) is localised to the interdomain interface, with the central region of a pair, i.e., the minor β-sheet of the N-terminal domain and the major β-sheet of the C-terminal domain, forming a common dynamic unit. The fact that cbEGF12-13 and cbEGF32-33 have dynamic N- and C-terminal regions (this study; Werner *et al.*, 2000a) may explain why cbEGF domains are usually present in proteins as multiple tandem repeats (Chapter 1; Downing *et al.*, 2000).

Although covalent linkage of a C-terminal cbEGF domain has been shown to have no effect on the calcium binding affinity of the preceding cbEGF domain in human fibrillin-1 (Chapter 4; Smallridge *et al.*, 1999), linkage of a C-terminal cbEGF may stabilise the C-terminal region of the preceding domain through interdomain packing interactions. Proteolysis studies of cbEGF12-13, cbEGF12-14 and the corresponding G1127S mutant constructs recently highlighted changes in cbEGF13 indicative of an influence of cbEGF14, hence, a C- to N-structural inter-dependence may exist for cbEGF domains of fibrillin-1 (section 7.5.1; Whiteman *et al.*, manuscript submitted). This is consistent with the observation of increased flexibility across the full length of a cbEGF domain pair upon removal of calcium (Werner *et al.*, manuscript in preparation), and suggests that the C-terminal half of cbEGF13 may be less severely affected by ps to ms motions when covalently linked to cbEGF14.

The dynamic behaviour of cbEGF12-13 and cbEGF32-33 (Werner *et al.*, 2000a) highlights the importance of calcium in determining the global organisation of these domains in fibrillin-1 and subsequently the 10-12nm connective tissue microfibrils (Chapter 1; Figure 1.3). A MFS-causing mutation in a cbEGF domain that affects a calcium ligand, or reduces calcium binding affinity indirectly, could destabilise an interdomain interface, decreasing anisotropy and increasing interdomain flexibility of a localised region of fibrillin-1. Mutation of a residue that participates in interdomain packing could also have this effect. If the defective fibrillin-1 monomers were subsequently incorporated into the 10-12nm microfibrils, the microfibrils would appear less extended and rigid, in agreement with electron micrographs of microfibrils in the

absence of calcium (Figure 1.3). This change could result in production of a MFS phenotype due to increased susceptibility of the microfibrils to proteolysis (Reinhardt *et al.*, 2000).

A MFS phenotype could alternatively result from the loss of a protein binding site. For example, consider the calcium ligand mutations D1072G, E1073K and N1131Y in cbEGF12-13, which produce a nMFS phenotype and form part of the potential protein binding site proposed in section 5.5.2. These mutations could, as a result of defective calcium binding, decrease the anisotropy and increase the intrinsic flexibility of the neonatal region of fibrillin-1. A more compressed, flexible structure could distort a potential binding interface, which could affect the microfibril assembly process and/or interactions with non-homologous components of the microfibrils by making protein binding entropically and enthalpically unfavourable. This could partially explain the increased calcium binding affinity of the neonatal region of fibrillin-1 compared with other regions thus far investigated (Chapter 4; Smallridge *et al.*, 1999).

Dynamics data highlighted the presence of a flexible loop region joining the strands of the minor β-sheet of cbEGF12, confirming the unstructured appearance of this region in the structure of the cbEGF12-13 domain pair (Chapter 5). No MFS-causing mutations have been identified in this region to date, so the proposed participation of this region in protein-protein interactions (sections 5.5.1 and 5.5.2) remains to be verified.

To summarise, backbone dynamics studies of the calcium saturated cbEGF12-13 pair, combined with corresponding results for cbEGF32-33 (Werner *et al.*, 2000a), have highlighted a potential dynamics signature for Class I cbEGF domain pairs. This signature includes μs to ms fluctuations for residues in the N-terminal half of the N-terminal domain, including at least the first disulphide bond, and fast (ps to ns) motions for residues of the minor β-sheet of the C-terminal domain. It includes a rigid interdomain interface and linker residue, with the central region of the domain pair

forming a common dynamic unit. This signature may explain why cbEGF domains are usually arranged as multiple tandem repeats. The studies have further confirmed that calcium stabilises cbEGF domain pair structure and modulates backbone dynamics over a wide time-scale (Werner *et al.*, 2000a). These studies have also suggested that calcium affinity may be a factor in determining the anisotropy of a Class I cbEGF pair, with increased affinity resulting in the production of a more extended structure. MFS-causing mutations along the length of fibrillin-1 that result in defective calcium binding by cbEGF domains and/or modify interdomain packing interactions are therefore likely to produce a less extended, more flexible structure for a region of fibrillin-1, which could increase proteolytic susceptibility and/or distort potential protein binding sites. The severity of the disease phenotype produced will depend on both the nature of the fibrillin-1 defect and its location within the fibrillin-1 monomer and the 10-12nm connective tissue microfibrils.

References

Akke, M., Skelton, N. J., Kordel, J., Palmer, A. G. & Chazin, W. J. (1993). Effects of ion-binding on the backbone dynamics of calbindin-D_9k determined by ^{15}N NMR relaxation. *Biochemistry.* **32**, 9832-9844.

Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W. & Bax, A. (1992). Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected 2-dimensional NMR-spectroscopy - the central helix is flexible. *Biochemistry.* **31**, 5269-5278.

Boyd, J., Hommel, U. & Campbell, I. D. (1990). Influence of cross-correlation between dipolar and anisotropic chemical-shift relaxation mechanisms upon longitudinal relaxation rates of ^{15}N in macromolecules. *Chem. Phys. Lett.* **175**, 477-482.

Boyd, J. & Redfield, C. (1998). Defining the orientation of the ^{15}N shielding tensor using ^{15}N NMR relaxation data for a protein in solution. *J. Am. Chem. Soc.* **120**, 9692-9693.

Clore, G. M., Szabo, A., Bax A., Kay, L. E., Driscoll, P. C. & Gronenborn, A. M. (1990). Deviations from the simple two parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Amer. Chem. Soc.* **112**, 4989-4991.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Downing, A. K., Campbell, I. D. & Handford, P. A. (2000). Chapter 4 : Calcium binding EGF-like domains. In Calcium Homeostasis. Carafoli, Krebs, eds. *Topics Biol. Inorg. Chem.*, volume 3.

Farrow, N. A., Zhang, O. W., Formankay, J. D. & Kay, L. E. (1994). A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical-exchange rates of systems in slow equilibrium. *J. Biomol. NMR.* **4**, 727-734.

Gagne, S. M., Tsuda, S., Spyracopoulos, L., Kay, L. E. & Sykes, B. D. (1998). Backbone and methyl dynamics of the regulatory domain of troponin C: Anisotropic rotational diffusion and contribution of conformational entropy to calcium affinity. *J. Mol. Biol.* **278**, 667-686.

Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry.* **28**, 8972-8979.

Kay, L. E., Keifer, P. & Saarinen, T. (1992a). Pure absorption gradient-enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665.

Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A. & Torchia, D. A. (1992b). Pulse sequences for removal of the effects of cross-correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T_1 and T_2 values in proteins. *J. Magn. Reson.* **97**, 359-375.

Kettle, S., Yuan, X., Grundy, G., Knott, V., Downing, A. K. & Handford, P. A. (1999). Defective calcium binding to fibrillin-1: consequence of an N2144S change for fibrillin-1 structure and function. *J. Mol. Biol.* **285**, 1277-1287.

Knott, V., Downing, A. K., Cardy, C. M. & Handford, P. (1996). Calcium binding properties of an EGF-like domain pair from human fibrillin-1. *J. Mol. Biol.* **255**, 22-27.

Lipari, G. & Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559.

Lipari, G. & Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559-4570.

Loria, J. P., Rance, M. & Palmer, A. G. (1999). A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 2331-2332.

Mandel, A. M., Akke, M. & Palmer, A. G. (1995). Backbone dynamics of *Escherichia-coli* ribonuclease HI - correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144-163.

Otting, G., Liepinsh, E. & Wuthrich, K. (1993). Disulfide bond isomerization in BPTI and BPTI(G36S) - an NMR-study of correlated mobility in proteins. *Biochemistry.* **32**, 3571-3582.

Palmer, A. G., Rance, M. & Wright, P. E. (1991). Intramolecular motions of a zinc finger DNA-binding domain from xfin characterized by proton-detected natural abundance ^{12}C heteronuclear NMR-spectroscopy. *J. Am. Chem. Soc.* **113**, 4371-4380.

Phan, I. Q. H., Boyd, J. & Campbell, I. D. (1996). Dynamics studies of a fibronectin type I module pair at three frequencies : Anisotropic modelling and direct determination of conformational exchange. *J. Biomol. NMR.* **8**, 369-378.

Press, W. H., Flannery, B. P., Teukosky, S. A. & Vetterling, W. T. (1990). In Numerical Recipes. Second edition, Cambridge University Press, New York.

Reinhardt, D. P., Ono, R. N., Notbohm, H., Muller, P. K., Bachinger, H. P. & Sakai, L. Y. (2000). Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis: A potential disease-causing mechanism in Marfan syndrome. *J. Mol. Biol.* **275**, 12339-12345.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol.* **286**, 661-668.

Tjandra, N., Kuboniwa, H., Ren, H. & Bax, A. (1995). Rotational-dynamics of calcium-free calmodulin studied by ^{15}N-NMR relaxation measurements. *Eur. J. Biochem.* **230**, 1014-1024.

Tjandra, N., Szabo, A. & Bax, A. (1996a). Protein backbone dynamics and ^{15}N chemical-shift anisotropy from quantitative measurement of relaxation interference effects. *J. Am. Chem. Soc.* **118**, 6986-6991.

Tjandra, N., Wingfield, P., Stahl, S. & Bax, A. (1996b). Anisotropic rotational diffusion of perdeuterated HIV protease from ^{15}N NMR relaxation measurements at two magnetic fields. *J. Biomol. NMR.* **8**, 273-284.

Wang, A. C. & Bax, A. (1993). Minimizing the effects of radiofrequency heating in multidimensional NMR experiments. *J. Biomol. NMR.* **3**, 715-720.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (2000a). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296**, 1065-1078.

Werner, J. M., Campbell, I. D. & Downing, A. K. (2000b). Shape and dynamics of calcium binding proteins investigated by nitrogen-15 relaxation. In Calcium binding protein protocols, H. J. Vogel, ed., Meth. Mol. Biol. Humana Press, in press.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. Backbone dynamics of an apo cbEGF domain pair: Implications for the role of calcium in human fibrillin-1. Manuscript in preparation.

Whiteman, P., Smallridge, R. S., Knott, V., Cordle, J. J., Downing, A. K. & Handford, P. A.. A glycine to serine amino acid change in calcium-binding epidermal growth factor-like domain 13 of human fibrillin-1 causes short range conformational effects. Manuscript submitted.

Woessner, D. E. (1962). Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion. *J. Chem. Phys.* **37**, 647-654.

Yuan, X., Lack, J., Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K.. Effects of the N2144S mutation on backbone dynamics of a TB-cbEGF domain pair from human fibrillin-1. Manuscript in preparation.

Zhang, M. J., Tanaka, T., Farrow, N. A., Kay, L. E. & Ikura, M. (1995). Multidimensional NMR-studies of the structure and dynamics of calcium-free calmodulin. *Prot. Eng.* **8**, 30.

Chapter 7

NMR Studies of the cbEGF12-13 G1127S Mutant Domain Pair

7.1 Introduction to the G1127S Mutation

A number of diseases are caused by missense mutations to cbEGF domains, highlighting the biological importance of this domain type (section 1.3), and the G1127S missense mutation in cbEGF13 of human fibrillin-1 was the focus of studies for this dissertation (numbering according to Pereira *et al.*, 1993). This mutation converts a conserved glycine residue to a serine residue and produces a "partial" MFS phenotype. It results in a mild inherited weakness of the elastic tissue, which predisposes to ascending aortic aneurysm and dissection in later life (Franke *et al.*, 1995). Interestingly, a corresponding missense mutation has been identified in human Factor IX, G60S, where it is associated with mild haemophilia B (Figure 7.1; Denton *et al.*, 1988; Whiteman *et al.*, 1998).

From structural knowledge of cbEGF domains, it was predicted that the G1127S mutation would result in defective folding (Whiteman *et al.*, 1998). This conserved glycine is located at the edge of the two stranded antiparallel β-sheet of cbEGF13, the main secondary structural feature of this type of domain (Rao *et al.*, 1995; Downing *et al.*, 1996). When this Gly to Ser change was introduced into cbEGF13 by molecular modelling, the serine side chain clashed sterically with the

consensus disulphide bond between the fifth and sixth cysteine residues (C1140 and C1150). Also, the G1127S mutation is located centrally between the second and third cysteine residues of cbEGF13 (C1124 and C1129), hence it was speculated that the G1127S mutation could interfere with the folding pathway associated with formation of disulphide bonds (Whiteman *et al.*, 1998).

Figure 7.1 : Sequence Alignment of cbEGF Domains from Factor IX and Fibrillin-1

The G1127S mutation of fibrillin-1 and the G60S mutation of Factor IX are underlined. Conserved cysteines are highlighted in bold, as are the glycines conserved between the second and third cysteine residues in many cbEGF domains.

```
Factor IX : DGDQCESNP  CLNGGSCKDDINSYECWCPFGFEGKN      C
cbEGF32   : DMDEC KEPDVCKH GQCINTDGSYRCECPFGYILAG   NEC
cbEGF33   : DTDECSVGNP CGN GTCKNVIGGFECTCEEGFEPGP  MMTC
cbEGF12   : DIDECRISPDLCGR GQCVNTPGDFECKCDEGYESGFMMMKNC
cbEGF13   : DIDECQRDPLLCR GGVCHNTEGSYRCECPPGHQLSP NISAC
cbEGF14   : DINECELSAHLCPN GRCVNLIGKYQCACNPGYHSTP DRLFC
```

Of the twenty naturally occurring amino acids, glycine is the most flexible. It can adopt a wider range of conformations than any other residue, allowing unusual main chain conformations in proteins (Shulz & Schirmer, 1990). It can therefore play an important structural role in proteins. In cbEGF13, the flexibility given by G1127 could be essential for disulphide bond formation, as a glycine is conserved between the second and third cysteine residues in many cbEGF domains, perhaps highlighting its importance (Figure 7.1; Handford *et al.*, 1991). The G1127S mutation could drastically reduce flexibility, decreasing the number of possible conformations of this region. The G1127S mutation was therefore predicted to result in domain misfolding by affecting disulphide bond formation or solely by affecting protein flexibility.

The positions of the G1127S mutation in the structure of the cbEGF12-13 domain pair, the cbEGF13 single domain and the cbEGF13-14 domain pair are shown in Figure 7.2. It has been shown previously that the G1127S mutation does indeed result in defective folding *in vitro* when introduced into the single cbEGF domain 13. When present in the covalently-linked cbEGF13-14 domain pair, the C-terminal cbEGF14 domain adopts the native fold and retains calcium binding properties, despite the fact that the adjacent domain 13 is again misfolded. These results demonstrated the localised effects of this mutation (Whiteman *et al.*, 1998). In this study, since it has been previously shown that the presence of a N-terminal cbEGF domain can have a significant effect on the properties of a C-terminal cbEGF domain from human fibrillin-1 (Smallridge *et al.*, 1999; Werner *et al.*, 2000), the consequences of the G1127S mutation in the cbEGF12-13 domain pair were investigated. Two-dimensional homonuclear NMR methods were used to qualitatively assess the calcium binding and structural properties of this mutant domain pair. This study was performed to further the understanding of the pathogenic effects of the G1127S mutation in fibrillin-1, and to improve understanding of the genotype/phenotype relationship of MFS.

7.2 Sample Preparation

The wild-type and G1127S mutant cbEGF12-13 domain pair protein samples were prepared as described in Chapter 3 (Whiteman *et al.*, 1998, and manuscript submitted). 11.6mg of wild-type and 11.2mg of G1127S mutant cbEGF12-13 were dissolved in 600µl 90% H_2O/10% ^2H_2O, 5mM Tris•HCl at pH 6.5, to yield final protein concentrations of 2.00mM and 1.93mM, respectively. 550µl of these samples were subsequently used for NMR analyses. NaCl was not utilised in these investigations, since high ionic strength may compromise spectral quality.

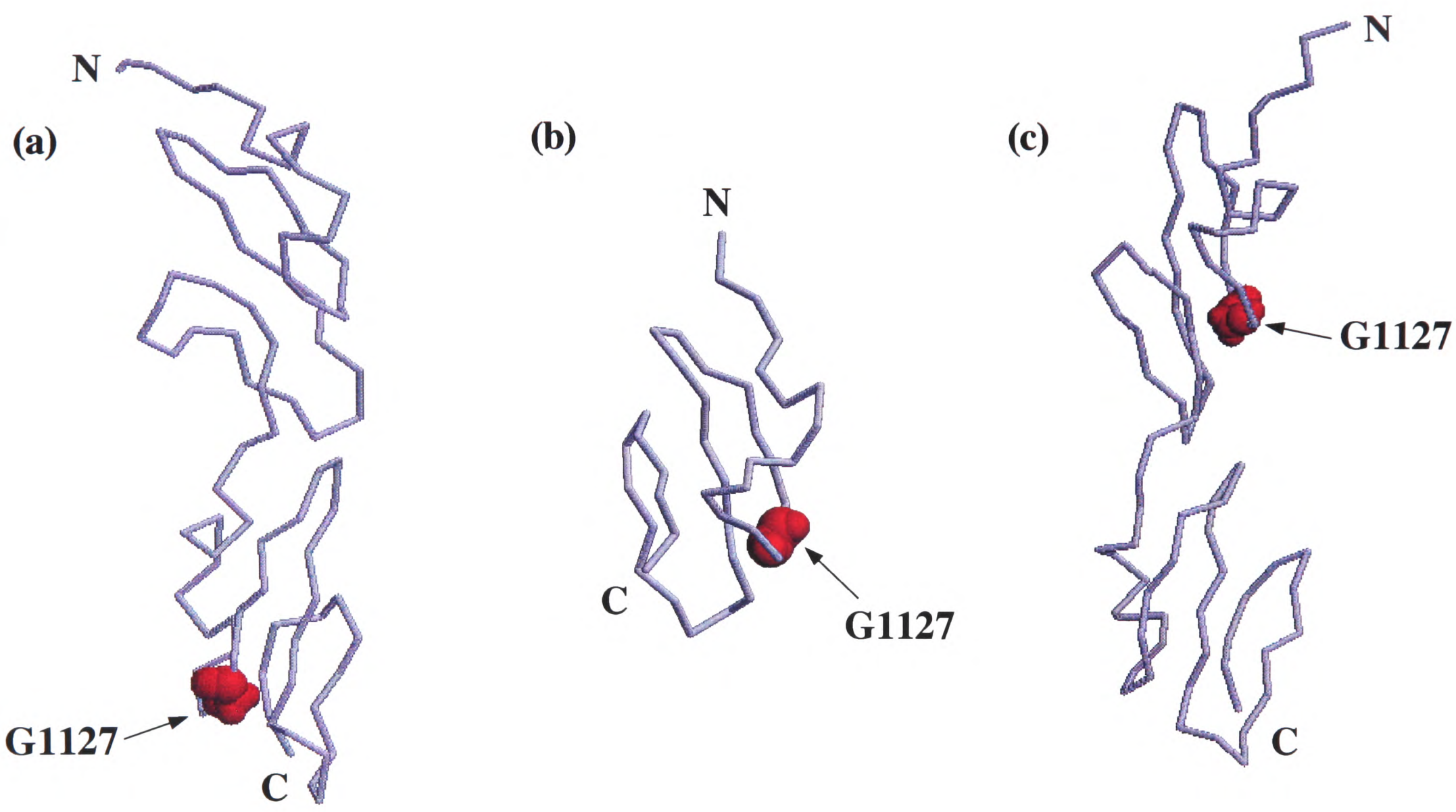


Figure 7.2 : Comparison of the Position of the G1127S Missence Mutation in (a) the cbEGF12-13 Domain Pair, (b) the cbEGF13 Domain and (c) the cbEGF13-14 Domain Pair from Human Fibrillin-1

The average NMR solution structure of the cbEGF12-13 domain pair is shown in (a) (see Chapter 5). The structures in (b) and (c) were modelled on the coordinates of cbEGF32-33 (Downing *et al.*, 1996). The cbEGF13 model was created using Modeller4 (Šali & Blundell, 1993) and the cbEGF13-14 model was created using Insight v2.3 (Biosym, Inc.). Calcium atoms have been omitted to increase structural clarity.

7.3 Data Acquisition

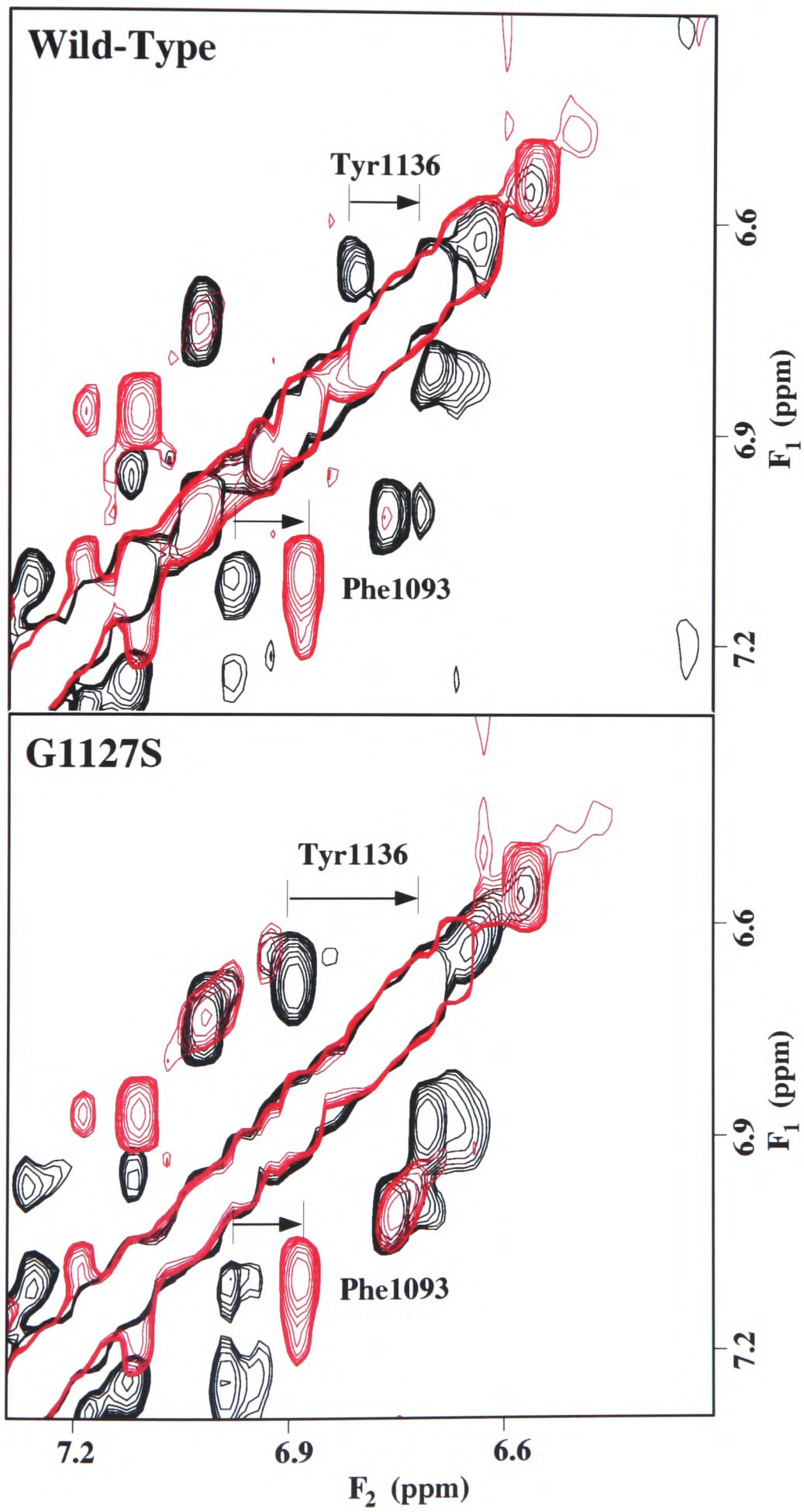
Two-dimensional NOESY spectra were acquired for both samples at 0mM $CaCl_2$ and 12.5mM $CaCl_2$ to qualitatively assess the effect of the G1127S mutation on both the calcium binding and structural properties of the cbEGF12-13 domain pair. At 12.5mM $CaCl_2$ the two calcium binding sites in the wild-type domain pair are saturated based on previous calcium binding studies of this domain pair (Chapter 4; Smallridge *et al.*, 1999). All NMR spectra were recorded on a home-built/GE Omega spectrometer at 600MHz, with a mixing time of 150ms at T = 33°C. 1024 complex points were acquired in F_2 and F_1 for each experiment with a spectral width of 8000Hz in each dimension. Data were processed using Felix 2.3 (Biosym, Inc.). Spectra were referenced with respect to the H^2HO resonance and were zero-filled to 8K in the F_2 dimension to yield a digital resolution of 0.98Hz/pt (Whiteman *et al.*, manuscript submitted).

7.4 Results

The aromatic regions of the 2D NOESY spectra acquired at 0mM and 12.5mM $CaCl_2$ are overlaid for both the wild-type and G1127S mutant cbEGF12-13 domain pair in Figure 7.3. Chemical shift changes of the Hδ* resonances of the consensus, aromatic, calcium binding residues for the N- and C-terminal domains of each pair, i.e., Phe1093 and Tyr1136, respectively, were used to qualitatively assess calcium binding. The overlay of the wild-type NOESY spectra at zero and saturating calcium show the expected, native binding properties previously observed (Chapter 4; Smallridge *et al.*, 1999). Comparison of the wild-type spectral overlay with that of the G1127S mutant reveals that both domain 12 and 13 retain calcium binding properties in the G1127S mutant. Domain 12 appears to retain native calcium binding properties in the

Figure 7.3 : An Overlay of the Aromatic Region of 2D NOESY Spectra of the Wild-Type cbEGF12-13 Domain Pair and the G1127S Mutant cbEGF12-13 Domain Pair at Zero Calcium (black) and Saturating Calcium (red)

The calcium-dependent migration of the Phe1093 and Tyr1136 cross-peaks is highlighted.



mutant pair, since the Phe1093 peak movement is identical to that seen in the wild-type spectra. The peak movement seen for Tyr1136 is slightly altered, however, with a larger chemical shift change associated with binding by the mutant domain (Figure 7.3; Whiteman *et al.*, manuscript submitted).

In addition to qualitative assessment of calcium binding by the cbEGF12-13 G1127S mutant pair, the structural consequences of this mutation were also assessed. A qualitative comparison of the calcium saturated 2D NOESY spectra indicated that, unlike the case for the cbEGF13-14 wild-type and mutant spectra (Whiteman *et al.*, 1998), a significant number of peaks corresponding to domain 13 are unaffected or only mildly affected by the presence of the G to S mutation (Figure 7.4). Using these 2D NOESY spectra, combined with a knowledge of the complete spectral assignments for the cbEGF12-13 wild-type domain pair (Chapter 5), it was possible to gain insight into the degree of disruption of the native fold caused by the G1127S mutation in the cbEGF12-13 construct (Figure 7.5).

7.5 Discussion

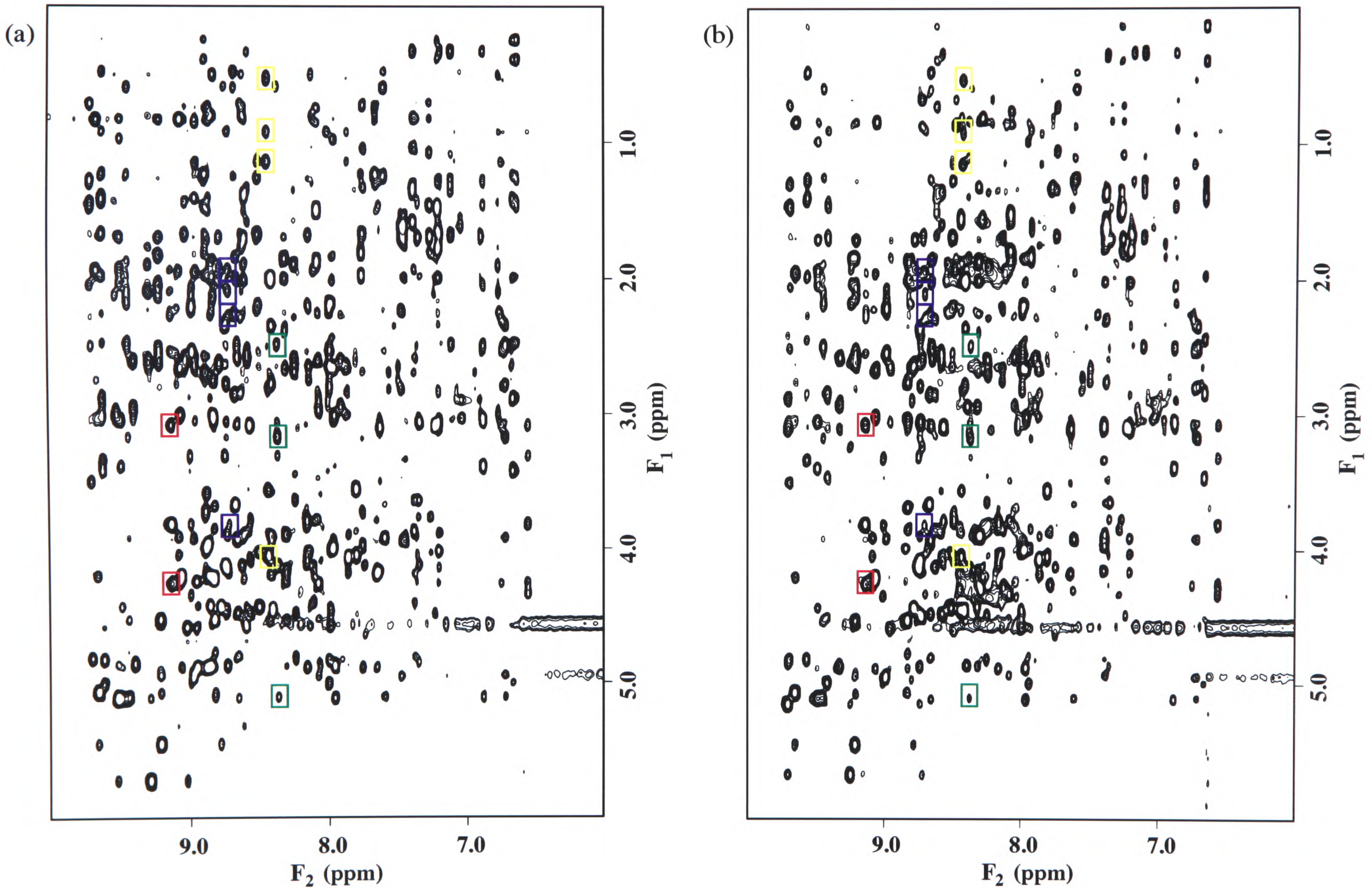
7.5.1 Chaperoning of Folding by Tandem Linkage of cbEGF Domains

The demonstration that domain 13 in the G1127S mutant cbEGF12-13 pair continues to bind calcium is unlike the situation previously seen for both the mutant cbEGF13 single domain and the mutant cbEGF13-14 domain pair. As calcium binding by cbEGF domains is used to probe correct refolding, the fact that cbEGF13 continues to bind calcium in the mutant cbEGF12-13 pair suggests that this domain preserves a "native-like" fold.

The hypothesis that cbEGF13 preserves a "native-like" fold in the cbEGF12-13 G1127S mutant pair is further supported by a qualitative comparison of the calcium saturated 2D NOESY spectra for the wild-type and G1127S mutant cbEGF12-13

Figure 7.4 : Two-Dimensional NOESY Spectra of (a) the Wild-Type cbEGF12-13 Domain Pair and (b) the G1127S Mutant cbEGF12-13 Domain Pair from Human Fibrillin-1

Resonances resulting from R1083, Y1101, E1133 and G1134, i.e., residues involved in interdomain packing in the cbEGF12-13 pair, are highlighted in both spectra by yellow, green, blue and red boxes, respectively (see section 7.5.1).





Colour Code

residues undisrupted in the mutant

residues whose cross-peaks are slightly shifted in the mutant

residues whose cross-peaks are significantly shifted or broadened in the mutant

residues whose cross-peak positions in the mutant are unclear due to spectral overlap in homonuclear spectra

Figure 7.5 : The Structure of the cbEGF12-13 Domain Pair from Human Fibrillin-1 Indicating the Degree of Disruption of the Native Fold in the Presence of the G1127S Mutation

The mutated glycine is highlighted in red, and calcium atoms have been omitted for structural clarity.

domain pairs (Figures 7.4 and 7.5). These results further demonstrate that the effects of the G1127S mutation are localised to domain 13 in the mutant 12-13 pair, and that even within domain 13 there are several residues that remain unaffected by the presence of the mutation. Measurement of the calcium binding affinities of the two sites in the cbEGF12-13 wild-type pair has previously demonstrated that covalent linkage of a N-terminal cbEGF domain can have a stabilising effect on the adjacent C-terminal cbEGF domain (Smallridge *et al.*, 1999), and the data obtained here suggest that the presence of a N-terminal cbEGF domain can also significantly influence the folding of a C-terminal cbEGF domain. The fact that the G1127S mutation has a less severe effect on folding when preceded by cbEGF12 suggests a chaperoning role for cbEGF12 in the cbEGF12-13 mutant pair.

How this chaperoning effect is mediated, however, is unclear. As previously discussed in section 1.3, cbEGF domains lack any appreciable hydrophobic core, and are stabilised by a combination of three disulphide bonds and calcium binding to the N-terminus (Rao *et al.*, 1995; Downing *et al.*, 2000). From the results obtained here it is clear that the calcium binding properties of the cbEGF13 domain in the mutant cbEGF12-13 domain pair are altered, with resonances of several of the calcium binding residues, i.e., E1116, N1131 and T1132, slightly shifted in the presence of the mutation. From the 2D NOESY spectral comparisons, it can also be seen that the resonances of all the six cysteine residues of cbEGF13 are affected by the presence of the G1127S mutation. The packing interactions between domains 12 and 13 in the cbEGF12-13 pair remain, however, unaffected by the mutation (Figures 7.4 and 7.5). Resonances of R1083, Y1101, E1133 and G1134 are identically located within the wild-type and mutant spectra (Figure 7.4), suggesting a role for these packing interactions in stabilising the folding of the mutant domain 13.

The exact differences in terms of calcium binding, structural and/or dynamic properties between the "native-like" fold of cbEGF13 in the cbEGF12-13 G1127S mutant pair and the true, native fold of cbEGF13 in the wild-type pair remain to be

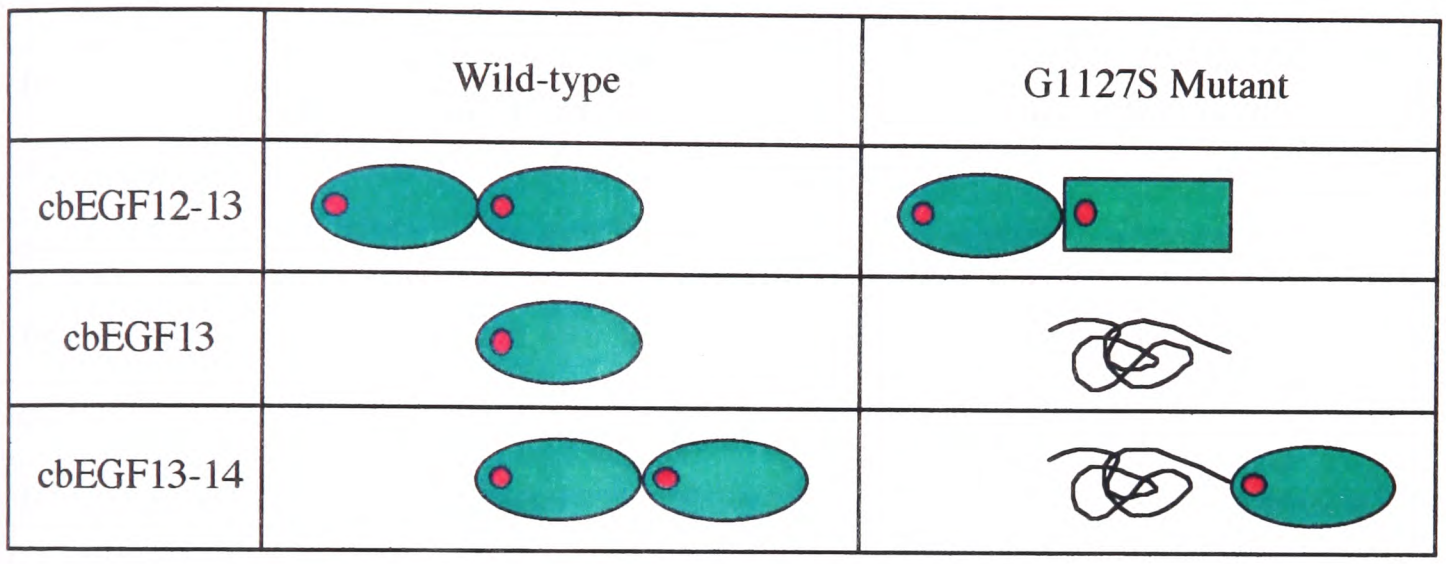
quantitatively determined and are currently under investigation. Overlap in the homonuclear 2D NOESY spectra of the wild-type and mutant cbEGF12-13 domain pair hindered a full analysis of the structural consequences of this mutation (Figures 7.4 and 7.5), hence a ^{15}N-cbEGF12-13 G1127S sample has subsequently been expressed to enable more detailed analysis. Furthermore, in the calcium saturated 2D NOESY spectrum of the cbEGF12-13 G1127S mutant (Figure 7.4b), a central cluster of broadened peaks can be observed, perhaps indicating an increase in the intrinsic flexibility of the mutant construct. This initially seems counterintuitive, as a glycine to serine change would be expected to decrease the intrinsic flexibility of the mutant construct, however, the calcium binding studies indicated that the calcium binding properties of the wild-type and mutant cbEGF12-13 pairs differ (Figure 7.3). If the mutant pair binds with reduced affinity, this could decrease the stability of the interdomain interface, decreasing the anisotropy and reducing the rigidity of this structure. The ^{15}N-cbEGF12-13 G1127S sample will therefore also be used in future dynamics studies, the results of which will be compared to the results of dynamics studies on the wild-type cbEGF12-13 domain pair (Chapter 6), to confirm whether or not changes in intrinsic mobility could be a factor in the production of a disease phenotype. A schematic summary of the NMR results obtained for the cbEGF12-13 domain pair, the cbEGF13 single domain and cbEGF13-14 domain pair is shown in Figure 7.6.

Protease digestion studies have recently been performed on both the wild-type and G1127S mutant cbEGF12-13 domain pairs, and the wild-type and mutant cbEGF12-14 triple constructs by Ms J. Cordle under the supervision of Dr. Pat Whiteman (Laboratory of Dr. Penny A. Handford, Biochemistry Department, University of Oxford; Whiteman *et al.*, manuscript submitted). The results of these studies were consistent with the NMR data described above, in that all the identified cleavage sites, including sites in the mutant domain, showed calcium-dependent protection from proteolysis. In addition, consistent with the NMR data for the

cbEGF12-13 and cbEGF13-14 domain pairs, the location of the cleavage sites in the flanking domains 12 and 14 were the same in the mutant and wild-type constructs, which is indicative of extremely localised changes.

Figure 7.6 : Schematic Diagram to Summarise the Consequences of the G1127S Mutation in cbEGF13 of Human Fibrillin-1

The results of studies on the cbEGF12-13 domain pair (this study), the cbEGF13 single domain and the cbEGF13-14 domain pair (Whiteman *et al.*, 1998) are summarised. Calcium atoms are depicted as red circles. The rectangle used to represent the state of cbEGF13 in the cbEGF12-13 G1127S mutant pair highlights the difference between the "native-like" and native fold.



Comparison of wild-type and G1127S mutant protease digestion products indicated the presence of extra cleavage sites in the mutant cbEGF13 domain, i.e., one extra site in the G1127S mutant cbEGF12-13 domain pair and three extra sites in the G1127S mutant cbEGF12-14 triple. The digestion site common to both the domain pair and triple construct was at G1134 (Whiteman *et al.*, manuscript submitted), i.e., a residue near the top of the major β-hairpin of cbEGF13, involved in the packing interactions between the domains of the cbEGF12-13 pair. As previously mentioned, the cross-peak positions for resonances arising from this residue are identical in the calcium saturated wild-type and G1127S mutant cbEGF12-13 2D NOESY spectra (Figure 7.4). The two extra cleavage sites located in the mutant cbEGF12-14 triple construct were at G1126, adjacent to the mutation site, and C1138, diagonally opposite G1127 across the major β-sheet of cbEGF13 (Whiteman *et al.*, manuscript submitted). From the NMR data, the resonances arising from G1126 are, as expected, significantly

shifted in the 2D NOESY spectrum of the cbEGF12-13 mutant pair, whereas the C1138 resonances are only slightly shifted.

The finding that further changes are apparent in the triple construct compared with the domain pair suggests an influence of the C-terminal domain 14 on the preceding domain, with the result that additional cleavage sites become available in the mutant domain. Modulation of the calcium binding affinity of a N-terminal domain by the addition of a C-terminal domain has been reported for multi-domain EGF fragments from protein S (Stenberg *et al.*, 1997a; 1997b; 1999), an effect not observed thus far for human fibrillin-1 (Chapter 4; Smallridge *et al.*, 1999). The results of Whiteman *et al.* (manuscript submitted), however, indicate that a C- to N-structural inter-dependence may exist for cbEGF domains of fibrillin-1. This is consistent with the observation of increased flexibility across the full length of a cbEGF domain pair upon removal of calcium (Werner *et al.*, manuscript in preparation). Additional structural and dynamic NMR analyses of the cbEGF12-14 construct are required, and will be performed in the future, to confirm whether or not this effect exists.

7.5.2 Implications for the Pathogenic Effect of the G1127S Mutation

Since the effect of the G1127S mutation is confined to domain 13 and this domain retains a "native-like" fold when covalently-linked to cbEGF12, mutant fibrillin-1 monomers are likely to be secreted by the cell with the pathogenic effect of this mutation being exerted either on or after incorporation into the 10-12nm connective tissue microfibrils. This is consistent with analysis of a fibroblast cell line obtained from a patient with the G1127S mutation, which indicated normal synthesis of total fibrillin-1 but reduced deposition in the matrix (Franke *et al.*, 1995).

From this study, the G1127S mutation is likely to produce a "partial" MFS phenotype in one, or more, of three ways,

1) The presence of the G1127S mutation may slightly disrupt the fold of the cbEGF13 domain in human fibrillin-1.

2) The presence of the mutation may alter the dynamic properties of the cbEGF13 domain.

3) The presence of the mutation may alter the calcium binding properties of the cbEGF13 domain either directly, and/or indirectly as a consequence of the structural and/or dynamic changes. The alteration of calcium binding properties is not insignificant, since defective calcium binding at only a single site within human fibrillin-1 is linked to the MFS phenotype.

The mutation-induced changes, localised to cbEGF13 of human fibrillin-1, could affect the microfibril assembly process, disrupting binding sites involved in this process and/or resulting in the loss of calcium ions required for mediating these interactions. They could also disrupt other protein-protein interactions that occur during or after microfibril assembly. The alterations could also cause a structural weakness of the 10-12nm connective tissue microfibrils once the mutant fibrillin-1 monomers have been incorporated. The changes in the calcium binding, structural and/or dynamics properties of cbEGF13 could result in an inherent structural weakness and/or decrease the rigidity/stability of the connective tissues microfibrils containing the defective protein, affecting the native biomechanical properties of fibrillin-1.

The disruption of the properties of cbEGF13 could render mutant fibrillin-1 more susceptible to proteolysis *in vivo*. The calcium-dependent protection from proteolytic degradation demonstrated by fibrillin-1, and other molecules containing tandem of cbEGF domains, has led to the hypothesis that missense mutations can result in increased proteolysis *in vivo* (Reinhardt *et al.*, 2000). This degradation could occur during secretion, in the initial stages of microfibril assembly in the extracellular matrix, or after the mutant fibrillin-1 monomers have been incorporated into the 10-12nm microfibrils.

Further NMR analyses are currently in progress with the aim of quantitatively characterising the calcium binding, structural and dynamic properties of the G1127S cbEGF13 mutant domain when located both C-terminally in the cbEGF12-13 domain pair and in the centre of the cbEGF12-14 triple construct, i.e., in a more "native" setting. These future studies will allow clarification of the exact consequences of the G1127S on the cbEGF13 domain, giving greater insight into the pathogenic mechanism of this mutation, and will also enhance understanding of the chaperoning role of cbEGF12 in folding of the G1127S mutant cbEGF12-13 domain pair.

References

Denton, P. H., Fowlkes, D. M., Lord, S. T. & Reisner, H. M. (1988). Hemophilia-B Durham: A mutation in the first EGF-like domain of factor IX that is characterised by polymerase chain reaction. *Blood.* **72**, 1407-1411.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Downing, A. K., Campbell, I. D. & Handford, P. A. (2000). Chapter 4 : Calcium binding EGF-like domains. In Calcium Homeostasis. Carafoli, Krebs, eds. *Topics Biol. Inorg. Chem.*, volume 3.

Franke, U., Berg, M. A., Tynan, K., Brenn, T., Liu, W., Aoyama, T., Gasner, C. Miller, D. C. & Furthmayr, H. (1995). A Gly1127Ser mutation in an EGF-like domain of the fibrillin-1 gene is a risk factor for ascending aortic aneurysm and dissection. *Am. J. Hum. Genet.* **56**, 1287-1296.

Handford, P. A., Mayhew, M., Baron, M., Winship, P. R., Campbell, I. D. & Brownlee, G. G. (1991). Key residues involved in calcium binding motifs in EGF-like domains. *Nature.* **351**, 164-167.

Pereira, L., D'Alessio, M., Ramirez, F., Lynch, J. R., Sykes, B., Pangilinan, T. & Bonadio, J. (1993). Genomic organisation of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Hum. Mol. Genet.* **2**, 961-968.

Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G. & Stuart, D. (1995). The structure of a Ca^{2+}-binding epidermal growth factor-like domain: Its role in protein-protein interactions. *Cell.* **82**, 131-141.

Reinhardt, D. P., Ono, R. N., Notbohm, H., Muller, P. K., Bachinger, H. P. & Sakai, L. Y. (2000). Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis: A potential disease-causing mechanism in Marfan syndrome. *J. Mol. Biol.* **275**, 12339-12345.

Sali A. & Blundell T. L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815.

Shulz, G. E. & Schirmer, R. H. (1990). Principles of protein structure. Ninth edition. Springer-Verlag, New York.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol.* **286**, 661-668.

Stenberg, Y., Linse, S., Drakenberg, T. & Stenflo J. (1997a). The high affinity calcium-binding sites in the epidermal growth factor module region of vitamin K-dependent protein S. *J. Biol. Chem.* **272**, 23255-23260.

Stenberg, Y., Julenius, K., Dahlqvist, I., Drakenberg, T. & Stenflo, J. (1997b). Calcium-binding properties of the third and fourth epidermal-growth-factor-like modules in vitamin-K-dependent protein S. *Eur. J. Biochem.* **248**, 163-170.

Stenberg, Y., Muranyi, A., Steen, C., Thulin, E., Drakenberg, T. & Stenflo J. (1999). EGF-like module pair 3-4 in vitamin K-dependent protein S: Modulation of calcium affinity of module 4 by module 3, and interaction with factor X. *J. Mol. Biol.* **293**, 653-665.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (2000). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296**, 1065-1078.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. Backbone dynamics of an apo cbEGF domain pair: Implications for the role of calcium in human fibrillin-1. Manuscript in preparation.

Whiteman, P., Downing, A. K., Smallridge, R., Winship, P. R. & Handford, P. A. (1998). A Gly to Ser change causes defective folding *in vitro* of calcium binding epidermal growth factor-like domains from factor IX and fibrillin-1. *J. Biol. Chem.* **273**, 7807-7813.

Whiteman, P., Smallridge, R. S., Knott, V., Cordle, J. J., Downing, A. K. & Handford, P. A.. A glycine to serine amino acid change in calcium-binding epidermal growth factor-like domain 13 of human fibrillin-1 causes short range conformational effects. Manuscript submitted.

Chapter 8

Summary of Conclusions

8.1 The Calcium Binding Properties of cbEGF12-13

Calcium binding studies of cbEGF12-13 from human fibrillin-1, using a combination of 2D NMR and fluorescence techniques, allowed determination of K_d values of ~1.6mM and <30μM for the N- and C-terminal sites, respectively. Domain 13 of this pair contains the highest affinity site thus far investigated from human fibrillin-1, and comparison of these data with previous results showed that covalent linkage of a N-terminal domain can enhance the affinity of a C-terminal cbEGF domain, with the magnitude of the effect depending on domain type. Conversely, for regions of fibrillin-1 studied to date, C-terminal linkage of a cbEGF domain has no effect on the calcium binding affinity of a N-terminal cbEGF domain (Figure 4.6).

These studies highlighted that the "neonatal" region of human fibrillin-1 appears to have an increased calcium binding affinity compared to other regions previously investigated. This increased affinity implies that sites in this region are more likely to be saturated, allowing maintenance of the predicted rigid, rod-like structure. This stable, extended structure may be important for fibrillin-1 assembly into the 10-12nm connective tissue microfibrils, interactions with other components of the microfibrils and/or the biomechanical properties of the microfibrillar network, potentially explaining why mutations to calcium ligands in this region can produce the severe, nMFS phenotype.

8.2 The Structure of the cbEGF12-13 Domain Pair

The NMR solution structure of the cbEGF12-13 domain pair is a near-linear, rod-like arrangement of two domains, with each domain comprising a major and minor two stranded anti-parallel β-sheet. In addition to β-structure, cbEGF13 contains an α-helical region, which is present in cbEGF12 in 32% of final structures (Figure 5.13). A comparison of the atomic coordinates of cbEGF12 and cbEGF13 with the coordinates of other cbEGF domains of known structure yielded backbone rmsd values for conserved residues of only 1.1-2.5Å (Table 5.4), highlighting the structural similarity of cbEGF domains. The tilt and twist angles for cbEGF12-13 were calculated as 30° ± 15° and 152° ± 13°, respectively, and these results further support the hypothesis that Class I EGF/cbEGF-cbEGF domain pairs, characterised by a single interdomain linker residue, possess a rigid, rod-like structure (Figure 5.14; Downing *et al.*, 1996). The dimensions of cbEGF12-13 also support the most recent model for the incorporation of fibrillin-1 monomers into the 10-12nm microfibrils (Figure 1.4).

The structure of cbEGF12-13 is stabilised both by calcium binding by the C-terminal domain and by interdomain hydrophobic packing interactions, both previously observed for cbEGF32-33 from fibrillin-1 (Downing *et al.*, 1996). In cbEGF12-13, however, an extra residue, R1083, is involved in the packing interaction without forming interdomain contacts. This residue may be responsible for the increased calcium binding affinity of cbEGF13 compared to cbEGF33 (Smallridge *et al.*, 1999; Knott *et al.*, 1996), as R1083 may improve shielding of the calcium ion from solvent and/or provide a better defined binding site in cbEGF13 (Figure 5.14). This is similar to the situation seen for the Class I EGF5-cbEGF6 pair from human thrombomodulin, where additional interdomain contacts have been proposed to explain the high affinity of cbEGF6 (Figure 5.14; Light *et al.*, 1999; Fuentes-Prior *et al.*, 2000). These results

suggest that interdomain contacts play a major role in determining the calcium binding affinity of cbEGF domains.

A model of cbEGF11-15 from the neonatal region of human fibrillin-1, created using structural data from cbEGF12-13 and a model of cbEGF32-36 (Downing *et al.*, 1996), has highlighted a potential protein binding interface. A rigid, rod-like structure is predicted for cbEGF11-15 of fibrillin-1, and, if the potential glycosylation site in cbEGF13 is used to define the outside face, it can be seen that all nMFS-causing mutations in this region appear to cluster on the inside face of cbEGF11-13 (Figure 5.17). An extended, flexible, unstructured loop region (G38-K43), which joins the strands of the minor β-sheet of cbEGF12, also projects into solution from this face, making it a candidate region for involvement in protein-protein interactions. This site may be important for fibrillin-1 assembly into the 10-12nm microfibrils and/or its interactions with other components of the microfibrils, and may potentially explain why many mutations in this region produce the devastating nMFS phenotype.

8.3 Dynamics of the cbEGF12-13 Domain Pair

NMR relaxation data for calcium-bound cbEGF12-13 confirmed an extended structure for this pair, where the diffusion tensor is a symmetric top with an axial ratio of 1.87:1 that aligns well with the inertia tensor. The increased size of this diffusion tensor, compared to that observed for cbEGF32-33 (1.55:1; Werner *et al.*, 2000), is most likely related to the increased calcium binding affinity observed for sites of cbEGF12-13 (Smallridge *et al.*, 1999). The more than five fold higher affinity of cbEGF12 compared to cbEGF32 (Knott *et al.*, 1996) appears to increase the rigidity of the N-terminal half of cbEGF12, which may explain the increased anisotropy of this pair. In addition, the greater than 11-fold increase in affinity of cbEGF13 compared to cbEGF33 (Knott *et al.*, 1996) may also improve the stability of the interdomain interface, aiding production of a more extended structure.

Residues in the N-terminal half of the cbEGF12 appear, in agreement with results for cbEGF32-33 (Werner *et al.*, 2000), to be affected by μs to ms fluctuations, suggesting that the calcium binding site in cbEGF12 is not maintained in the correct conformation to bind calcium. In contrast, only residues in the C-terminal half of cbEGF13, localised far from the high affinity calcium binding site, are affected by μs to ms motions. This may explain the lower affinity generally seen for isolated cbEGF domains or those located N-terminally in a pair. In addition, the μs to ms fluctuations seen for cbEGF32 (Werner *et al.*, 2000) were larger than those seen for cbEGF12, which may explain both the five fold increase in calcium affinity of cbEGF12 compared to cbEGF32 (Smallridge *et al.*, 1999; Knott *et al.*, 1996) and the variation in anisotropy observed for cbEGF12-13 and cbEGF32-33 (this study; Werner *et al.*, 2000).

Results obtained here support the proposed correlation between linker flexibility and calcium binding affinity for cbEGF domains (Kettle *et al.*, 1999; Smallridge *et al.*, 1999; Werner *et al.*, 2000; Yuan *et al.*, manuscript in preparation). The interdomain linker residue between cbEGF12 and 13 has been shown to be rigid, which correlates with the high affinity seen for cbEGF13 (Smallridge *et al.*, 1999). In addition, comparison of data for cbEGF12-13 and cbEGF32-33 (Werner *et al.*, 2000) has highlighted that variations in C-terminal calcium binding affinity in Class I cbEGF domain pairs may be a consequence of variations in the conserved packing interaction.

Dynamics data highlighted the presence of a flexible loop region, G38-K43, joining strands of the minor β-sheet of cbEGF12. This confirmed the unstructured appearance of this region of cbEGF12-13 (Figure 5.13), and suggested that this loop may participate in protein-protein interactions.

Comparison of the slow dynamic behaviour of cbEGF12 and cbEGF13 suggested that the presence of a preceding and subsequent domain can greatly influence the dynamics of these domains. In the absence of N- and C-terminally-linked domains, the N-terminal and C-terminal halves of cbEGF12 and cbEGF13, respectively, exhibit

conformational exchange on the μs to ms time-scale, with cbEGF13 also presenting a systematic decrease in heteronuclear NOE values towards its C-terminus.

Backbone dynamics studies of calcium saturated cbEGF12-13, combined with corresponding results for cbEGF32-33 (Werner *et al.*, 2000), have highlighted a potential dynamics signature for Class I cbEGF domain pairs. This includes μs to ms fluctuations for residues in the N-terminal half of the N-terminal domain, including at least the first disulphide bond, and fast (ps to ns) motions for residues of the minor β-sheet of the C-terminal domain. The fact that these domain pairs have dynamic N- and C-terminal regions may explain why cbEGF domains are usually present in proteins as multiple tandem repeats (section 1.3; Downing *et al.*, 2000). The signature also includes a rigid interdomain interface and linker residue, with the central region of the domain pair, i.e., the minor β-sheet of the N-terminal domain and the major β-sheet of the C-terminal domain, forming a common dynamic unit. The studies have further confirmed that calcium stabilises cbEGF domain pair structure and modulates backbone dynamics over a wide time-scale (Werner *et al.*, 2000), and have also suggested that calcium affinity may be factor in determining the anisotropy of a Class I cbEGF pair, with increased affinity resulting in the production of a more extended structure. Finally, these studies have provided insight into how mutations that result in defective calcium binding by cbEGF domains and/or modify interdomain packing interactions may result in production of a MFS phenotype.

8.4 The G1127S Mutation

The G1127S mutation of human fibrillin-1, positioned at the edge of the major β-sheet of cbEGF13, produces a mild variant of the MFS phenotype, predisposing to ascending aortic aneurysm and dissection in later life (Franke *et al.*, 1995). NMR

studies of a cbEGF12-13 G1127S mutant construct showed that domain 13 continues to bind calcium and also maintains some wild-type structural properties upon *in vitro* refolding, suggesting that this domain preserves a "native-like" fold despite the presence of this mutation. This contrasts the situation previously observed for both the mutant cbEGF13 single domain and the mutant cbEGF13-14 domain pair, where, in both cases, cbEGF13 was misfolded (Whiteman *et al.*, 1998). Calcium binding studies of wild-type cbEGF12-13 demonstrated that covalent linkage of a N-terminal cbEGF domain can stabilise an adjacent C-terminal cbEGF domain (Smallridge *et al.*, 1999), and data obtained here suggest that the presence of a N-terminal cbEGF domain can also influence folding of a C-terminal cbEGF domain. The G1127S mutation has a less severe effect on folding when cbEGF13 is preceded by cbEGF12 suggesting a chaperoning role for cbEGF12 in the cbEGF12-13 mutant pair, and this chaperoning effect is most likely mediated through interdomain packing interactions which are not disrupted by the presence of this mutation.

These studies also demonstrated that the effects of the G1127S mutation are localised, since cbEGF12 and cbEGF14 exhibit native calcium binding and structural properties in G1127S cbEGF12-13 and cbEGF13-14 mutant pairs, respectively. This mutation is therefore likely to produce a "partial" MFS phenotype solely by alteration of the properties of cbEGF13, i.e., structural, dynamic and/or calcium binding properties. Any of these changes could result in an inherent structural weakness and /or decrease the rigidity/stability of the connective tissues microfibrils containing the defective protein, affecting the native biomechanical properties of fibrillin-1. The disruption could also render mutant fibrillin-1 more susceptible to proteolysis *in vivo* (Reinhardt *et al.*, 2000).

Further NMR analyses are currently in progress to quantitatively characterise the properties of the G1127S cbEGF13 mutant domain when located both C-terminally in cbEGF12-13 and centrally in the cbEGF12-14 triple construct, i.e., a more "native" setting. These future studies will clarify the exact consequences of the G1127S

mutation on cbEGF13, giving greater insight into the pathogenic mechanism of this mutation, and enhancing understanding of the chaperoning role of cbEGF12 in folding of the G1127S mutant cbEGF12-13 domain pair.

References

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. & Handford, P. A. (1996). Solution structure of a pair of cbEGF-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell.* **85**, 597-605.

Downing, A. K., Campbell, I. D. & Handford, P. A. (2000). Chapter 4 : Calcium binding EGF-like domains. In Calcium Homeostasis. Carafoli, Krebs, eds. *Topics Biol. Inorg. Chem.*, volume 3.

Franke, U., Berg, M. A., Tynan, K., Brenn, T., Liu, W., Aoyama, T., Gasner, C. Miller, D. C. & Furthmayr, H. (1995). A Gly1127Ser mutation in an EGF-like domain of the fibrillin-1 gene is a risk factor for ascending aortic aneurysm and dissection. *Am. J. Hum. Genet.* **56**, 1287-1296.

Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D. R & Bode, W. (2000). Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. *Nature.* **404**, 518-525.

Kettle, S., Yuan, X., Grundy, G., Knott, V., Downing, A. K. & Handford, P. A. (1999). Defective calcium binding to fibrillin-1: consequence of an N2144S change for fibrillin-1 structure and function. *J. Mol. Biol.* **285**, 1277-1287.

Knott, V., Downing, A. K., Cardy, C. M. & Handford, P. (1996). Calcium binding properties of an EGF-like domain pair from human fibrillin-1. *J. Mol. Biol.* **255**, 22-27.

Light, D. R., Glaser, C. B., Betts, M., Blasko, E., Campbell, E., Clarke, J. H., McCaman, M., McLean, K., Nagashima, M., Parkinson, J. F., Rumennik, G., Young, T. & Morser, J. (1999). The interaction of thrombomodulin with Ca^{2+}. *Eur. J. Biochem.* **262**, 522-533.

Reinhardt, D. P., Ono, R. N., Notbohm, H., Muller, P. K., Bachinger, H. P. & Sakai, L. Y. (2000). Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis: A potential disease-causing mechanism in Marfan syndrome. *J. Mol. Biol.* **275**, 12339-12345.

Smallridge, R. S., Whiteman, P., Doering, K., Handford, P. A. & Downing, A. K. (1999). EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. *J. Mol. Biol.* **286**, 661-668.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K. (2000). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296**, 1065-1078.

Whiteman, P., Downing, A. K., Smallridge, R., Winship, P. R. & Handford, P. A. (1998). A Gly to Ser change causes defective folding *in vitro* of calcium binding

epidermal growth factor-like domains from factor IX and fibrillin-1. *J. Biol. Chem.* **273**, 7807-7813.

Yuan, X., Lack, J., Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D. & Downing, A. K.. Effects of the N2144S mutation on backbone dynamics of a TB-cbEGF domain pair from human fibrillin-1. Manuscript in preparation.

Appendix I

Geometric Energy Function Parameters and NMR Structure Calculation Protocols for Proteins

1. Geometric Energy Function Parameters

topallhdg.pro (version 4.01, M. Nilges):
ftp://ftp.NMR.EMBL-Heidelberg.DE/pub/local/nilges/parameters/

parallhdg.pro (version 4.02, M. Nilges):
ftp://ftp.NMR.EMBL-Heidelberg.DE/pub/local/nilges/parameters/

2. NMR Structure Calculation Protocols

****

generate_fib1213.inp
Used to generate the cbEGF12-13 structure file. This was used to generate fib1213.psf, which contains only disulphide patches (section 2.3.3.3.2), for use in sa.inp. A modified version of generate_fib1213.inp was used to generate fib1213_SS.psf, a structure file containing disulphide bonds, for use in refine.inp.

generate_fib1213_template.inp:
Used to generate an extended template for cbEGF12-13 with ideal geometry for each amino acid residue.

sa.inp:
Structure calculation protocol. This protcol employs several other X-PLOR scripts, which are also shown in this appendix.

refine.inp:
Structure refinement protocol.

****

generate_fib1213.inp:

```
 remarks  file  nmr/generate_fib1213.inp
 remarks  Generates a structure file for a protein
 remarks  using the SA parameter and topology files.

 topology
   @topallhdg.pro
 end                    {* read topology file *}

 segment                                {* generate protein          *}
  name="    "                           {* this name has to match the  *}
                                        {* four characters in columns 73*}
                                        {* through 76 in the coordinate*}
                                        {* file, in XPLOR this name is *}
                                        {* name is referred to as segid *}
chain
```

```
    @toph19.pep                {* defines standard protein peptide bonds and termini   *}
sequence                                    {*ten residues per line, 3 letter code, caps   *}
      SER ALA THR ASP ILE ASP GLU CYS ARG ILE
      SER PRO ASP LEU CYS GLY ARG GLY GLN CYS
      VAL ASN THR PRO GLY ASP PHE GLU CYS LYS
      CYS ASP GLU GLY TYR GLU SER GLY PHE MET
      MET MET LYS ASN CYS MET ASP ILE ASP GLU
      CYS GLN ARG ASP PRO LEU LEU CYS ARG GLY
      GLY VAL CYS HIS ASN THR GLU GLY SER TYR
      ARG CYS GLU CYS PRO PRO GLY HIS GLN LEU
      SER PRO ASN ILE SER ALA CYS ILE
    end
 end
end
segment                                     {* add in calcium atoms *}
      name = CA1
      chain
       coor @cal1.pdb end
end
segment
      name = CA2
      chain
       coor @cal2.pdb end
end

!---------------------------------------------------------------------
! apply a special disulfide patch that only deletes the hydrogen
! but does not add a bond
patch DISN reference=1=(resid 8 )    reference=2=(resid 20) end
patch DISN reference=1=(resid 15 )  reference=2=(resid 29) end
patch DISN reference=1=(resid 31 )  reference=2=(resid 45) end
patch DISN reference=1=(resid 51 )  reference=2=(resid 63) end
patch DISN reference=1=(resid 58 )  reference=2=(resid 72) end
patch DISN reference=1=(resid 74 )  reference=2=(resid 87) end

patch CIPP                                  {* edit in a cis-proline for P75 *}
   reference=-=( resid 74 )
    reference=+=( resid 75 )
end

write structure output=fib1213.psf end

stop
```

generate_fib1213_template.inp:

```
remarks  file  nmr/generate_fib1213_template.inp
remarks  Generates a "template" coordinate set.  This produces
remarks  an arbitrary extended conformation with ideal geometry.
remarks  author: Axel T. Brunger

{====>}
structure @fib1213.psf end                  {* read structure file *}

parameter
{====>}
  @parallhdg.pro                            {* read parameter file *}
end
topology
```

```
  presidue NDIS                          {* break disulphide bonds *}
    delete bond 1SG 2SG
    delete angle 1CB 1SG 2SG
    delete angle 1SG 2SG 2CB
  end
end

{====>}
 {* if your protein contains S-S bridges appropriately modify and      *}
 {* then uncomment the following lines.                                *}
patch ndis reference=1=( residue 8    ) reference=2=( residue 20 ) end
patch ndis reference=1=( residue 15 ) reference=2=( residue 29 ) end
patch ndis reference=1=( residue 31 ) reference=2=( residue 45 ) end
patch ndis reference=1=( residue 51 ) reference=2=( residue 63 ) end
patch ndis reference=1=( residue 58 ) reference=2=( residue 72 ) end
patch ndis reference=1=( residue 74 ) reference=2=( residue 87 ) end

vector ident (x) ( all )              {* set up the initial Cartesian space coordinates*}
vector do (x=x/10.) ( all )
vector do (y=random(0.5) ) ( all )
vector do (z=random(0.5) ) ( all )

vector do (fbeta=50) (all)           {* friction coefficient, in 1/ps *}
vector do (mass=100) (all)           {* heavy masses, in amu       *}

parameter
  nbonds
     cutnb=5.5 rcon=20. nbxmod=-2 repel=0.9  wmin=0.1 tolerance=1.
     rexp=2 irexp=2 inhibit=0.25
  end
end
```

```
{* nbonds                                                             *}
{* applies to both electrostatic and van der Waals energy calculations.  *}
{* It sets up global parameters for the nonbonded interaction list     *}
{* generation and determines the form of subsequent nonbonded          *}
{* energy calculations                                                 *}

{* cutnb specifies the nonbonded interaction cutoff for the nonbonded  *}
{*       list generation (default: 8.5 Å)                              *}


{* r_con specifies the energy constant  for the repel function         *}
{*       (default: 100.0).                                             *}

{* nbxmod=-2 excludes nonbonded interactions between bonded atoms      *}

{* repel specifies k_rep : if > 0, this option turns on the repel function *}
{* and turns off the electrostatic energy.  k_rep specifies the factor *}
{*       by which to multiply the van der Waals radius  (default: 0).  *}

{* w_min specifies the threshold distance for close contact warnings,  *}
{* i.e., a warning is issued when a pair of atoms gets closer than this *}
{* distance unless the nonbonded interaction is excluded by the        *}
{* nbxmod option (default: 1.5 Å).                                     *}

{* tolerance specifies the distance that any atom is allowed to move   *}
{* before the nonbonded list gets updated.                             *}

{* r_exp specifies the exponent r_exp for the repel function           *}
{* i_rep specifies the exponent i_rexp for the repel function          *}
```

```
{* inhibit specifies the distance  (Eq. 4.7) between two atoms below        *}
{* which the van der Waals potential (Eq. 4.8) is truncated (default:      *}
{*       0.25 Å).                                                          *}
flags exclude * include bond angle vdw end

minimize powell nstep=50  nprint=10 end

flags include impr end

minimize powell nstep=50 nprint=10 end

dynamics  verlet                    {* molecular dynamics simulation     *}
  nstep=50  timestep=0.001 iasvel=maxwell  firsttemp= 300.
  tcoupling = true  tbath = 300.   nprint=50  iprfrq=0
end

{* timestep assigns the time step  value for the finite-difference       *}
{* integration in psec (default: 0.001 psec)                             *}

{* iasvel determines the                                                 *}
{* mode of initial velocity assignments (default: zero velocities).      *}

{* tcoupling is a logical flag that turns on Berendsen's temperature-    *}
{* coupling method. This flag works in conjunction with tbath            *}

{* tbath specifies the temperature of the heatbath or the target         *}
{* temperature of Berendsen's temperature-coupling method.               *}

{* iprfrq prints energy statistics every iprfrq steps if not equal to 0  *}

parameter
  nbonds
     rcon=2. nbxmod=-3 repel=0.75
  end
end

minimize powell nstep=100 nprint=25 end

dynamics  verlet
  nstep=500  timestep=0.005 iasvel=maxwell  firsttemp= 300.
  tcoupling = true  tbath = 300.   nprint=100  iprfrq=0
end

flags exclude vdw elec end
vector do (mass=1.) ( name h* )
hbuild selection=( name h* ) phistep=360 end

{* phistep specifies the step size in degrees for the dihedral angle     *}
{* search in cases where the placement of the hydrogens is not unique    *}
{* (default:10).                                                         *}

flags include vdw elec end

minimize powell nstep=500 nprint=50 end
                                                    {* write coordinates *}
remarks produced by nmr/generate_fib1213_template.inp
write coordinates output=fib1213_template.pdb end

stop
```

sa.inp:

```
remarks sa.inp
remarks Author: Michael Nilges, EMBL
evaluate ($ini_count = 0)
evaluate ($end_count = 100)

evaluate ($iniseed = 776688)          {* random seed value*}

evaluate ($init_t = 2000 )           {* initial simulated annealing temperature*}
evaluate ($high_steps = 6500)          {* total number of steps at high temp. *}
evaluate ($cool1_steps = 2500)         {* total number of steps during cooling.*}
evaluate ($cool2_steps = 1000)

evaluate ($fileroot = "fib1213_swap_")

structure   @@fib1213.psf end                {* read structure file (patched disulphides)*}
parameter  @@parallhdg.pro end                {* read parameter file                  *}

parameter  nbfix S  S  462  13.6  462  13.6 end {* reduce sulphur atom size so       *}
                                                {* cysteines in patched disulphide   *}
                                               {* bonds can get close to one another *}

parameter                                       {* set bond and angle energy constants*}
  bond (name sg) (name sg) 0.0 TOKEN           {* for disulphide bonds between      *}
  angle (all) (name sg) (name sg)  0.0 TOKEN {* cysteine sulphur atoms to zero     *}
end

noe                              {* parameters for NOE effective energy term. *}

     reset                                      {* erases current NOE database *}
     nrestraints = 5000          {* allocate space for NOEs                    *}
     ceiling 100                 {* maximum energy constant that is allowed.  *}
     set echo off message off end
     class     dist       @newfib12132d.tbl   {* read unambiguous distance restraints  *}
                          @newfib12133d.tbl
                          @newfib12133dsidechains.tbl
                          @fib1213_hbonds.tbl                {* read H-bonds restraints* }
                          @fib1213_cal.tbl           {* read calcium binding restraints* }
     class     ambi       @fib12132d_ambig.tbl {* read ambiguous distance restraints * }
                          @fib12132d_extraambig.tbl
                          @fib12133d_ambig.tbl
                          @fib12133d_extraambig.tbl
!    set echo on message on end

   averaging    * sum
   potential    * soft                {* the square-well function is used for distances   * }
                                      {* within a specified "switching" region, whereas * }
                                      {* outside this region a "soft" asymptote is used  * }
   scale        * 1.0                {* scales the restraint class (default: 1)          * }
   sqconstant   * 1.0                {* is an additional scale factor                    * }
   sqexponent   * 2          {* an exponent for the square-well or soft-square function * }
   soexponent   * 1                    {* an exponent for the soft-square function      * }
   rswitch      * 1.0                 {* the switch parameter between the square-well* }
                                      {* and the asymptote                              * }
   sqoffset     * 0.0              {* applies a negative offset value to all upper bounds* }
   asymptote    * 2.0        {* the slope c of the asymptote of the soft-square function * }
end

@@fib1213_dihe.tbl              {* read dihedral (phi) angle restraints * }
```

```
noe                {* patch in S-S bonds using distance restraints *}
  assign (resid   8  and  name SG) (resid  20 and  name SG)  2.02 0.1 0.1
  assign (resid   15 and  name SG) (resid  29 and  name SG)  2.02 0.1 0.1
  assign (resid   31 and  name SG) (resid  45 and  name SG)  2.02 0.1 0.1
  assign (resid   51 and  name SG) (resid  63 and  name SG)  2.02 0.1 0.1
  assign (resid   58 and  name SG) (resid  72 and  name SG)  2.02 0.1 0.1
  assign (resid   74 and  name SG) (resid  87 and  name SG)  2.02 0.1 0.1

end

flags exclude * include bonds angle impr vdw noe cdih end

@setup_swap.xplor {* defines symmetrical groups for use in floating *}
                  {* chirality protocol (swap15v.xplor)              *}

!set echo on message on end

vector ident (store2) (store1)          {* copy store 1 to store 2 *}
parameter
   improper (store2) (store2) (all) (all) 0.0 TOKEN TOKEN
   improper (all) (all) (store2) (store2) 0.0 TOKEN TOKEN
end                                {* set impropers to zero to allow floating chirality *}
set echo false message false end

vector do (fbeta=10) (all)      {* friction coefficient for MD heatbath, in 1/ps *}
vector do (mass=100) (all)      {* uniform heavy masses to speed molecular dynamics.*}

evaluate ($kcdih = 5)                   {* scale dihedral angle force constant *}
restraints dihedral
    scale=$kcdih
end

evaluate ($count = $ini_count)
while ($count < $end_count ) loop main

   evaluate ($count=$count+1)
   evaluate ($nreassign = 0)
   coor @@fib1213_template.pdb              {* read template file *}
   evaluate ($seed = $count*$iniseed)
   set seed $seed end
   @sa_1_randomchain.xplor          {* randomise phi and psi angles *}

   evaluate ($cpu1 = $cpu)

   evaluate ($final1_t = 1000)    { K }
   evaluate ($final2_t = 100)     { K }
   evaluate ($tempstep = 50)      { K }

   evaluate ($ncycle = ($init_t-$final1_t)/$tempstep)
   evaluate ($nstep = int($cool1_steps/$ncycle))

   @@sa_1_initial_values.xplor
                        {* initial values for dynamics simulation and minimisations *}
   parameter
     angle (store2) (all) (store2) $ini_flt TOKEN
     angle (all)      (all) (store2) $ini_flt TOKEN
   end

   parameter nbonds
      atom cutnb 12 tolerance 3.5 repel=1.2 wmin 0.5
       rexp=2 irexp=2 rcon=1. nbxmod 4
```

```
  end end

  @@sa_l_reduced.xplor {* defines store1, where residues are all reduced to  *}
                       {* two atoms (Gly to one) to reduce calculation time *}
  constraints
    interaction (all) (not store1) weights * 1 angl 1.0 impr 1.0 vdw 0.0 elec 0 end
    interaction (store1) (store1)  weights * 1 angl 1.0 impr 1.0 vdw 0.1 elec 0 end
  end       {* set vdw force constant to zero for anything outside store 1 and to 0.1 for *}
          {* reduced residues in store 1. Allows anything outside store 1 to move freely *}

{* 1 ========================================== initial minimization  *}
  restraints dihedral  scale=5.  end
  noe potential * soft scale * 1.0 asymptote * 2.0  end
  minimize powell nstep=50 drop=10.  nprint=25 end

{* 2 ================================  high temperature dynamics *}

  @sa_l_hightemp.xplor

{* 3 ============================================== cooling 1 *}

  @sa_l_cool1.xplor

{* 4 ============================================== cooling 2 *}

  @sa_l_cool2.xplor

{* 5 ========================================== final minimization *}

  evaluate ($swap = 1.001)
  flags exclude * include noe end
  @swap15v.xplor      {* systematically swaps randomly selected methylene protons *}
                      {* and isopropyl groups and evaluates the noe energy term   *}
  flags include bond angl impr vdw noe cdih end

  minimize powell nstep=200 drop=10.0 nprint=25 end

{* 6 ================================ write out the final structure(s) *}

evaluate ($filename=$fileroot+ encode($count)+ ".pdb")
  print threshold=0.4 noe
  evaluate ($rms_noe=$result)
  evaluate ($violations_noe=$violations)
  print threshold=3. cdih
  evaluate ($rms_cdih=$result)
  evaluate ($violations_cdih=$violations)
  print thres=0.05 bonds
  evaluate ($rms_bonds=$result)
  print thres=5. angles
  evaluate ($rms_angles=$result)
  print thres=5. impropers
  evaluate ($rms_impropers=$result)
  remarks initial random number seed: $seed
    remarks ============================================================
  remarks          overall,bonds,angles,improper,vdw,noe,cdih
  remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
    remarks ============================================================
  remarks          bonds,angles,impropers,noe,cdih
  remarks rms-dev.: $rms_bonds,$rms_angles,$rms_impropers,$rms_noe,$rms_cdih
    remarks ============================================================
  remarks           noe,  cdih
```

```
  remarks violations.: $violations_noe, $violations_cdih
    remarks  ============================================================
  write coordinates sele= (all) output =$filename end
  evaluate ($fname="fib1213_swap_"+encode($count)+".prt")
  set print = $fname end
  noe print thresh = 0.1 end {* write out .prt files, which contain information about   *}
                             {* distance restraints violated in each structure          *}
  close $fname end

end loop main

stop
```

X-PLOR scripts used by sa.inp

1. setup_swap.xplor:

```
vector do (q = 0) (all)
vector do (store1 = 0) (all)
vector do (store1 = store1 + 1)
        ((resn leu and name cd*) or (resn val and name cg*))

set echo on message on end
for $id in id (
  (resn gly and name ha1)
  or (resn ile and name hg11)
  or (resn leu and name hb1)
  or (resn phe and name hb1)
  or (resn pro and (name hb1 or name hg1 or name hd1))
  or (resn met and (name hb1 or name hg1))
  or (resn trp and name hb1)
  or (resn cys and name hb1)
  or (resn ser and name hb1)
  or (resn asn and name hb1)
  or (resn gln and (name hb1 or name hg1))
  or (resn tyr and name hb1)
  or (resn his and name hb1)
  or (resn asp and name hb1)
  or (resn asp and name od1)
  or (resn glu and (name hb1 or name hg1))
  or (resn lys and (name hb1 or name hg1 or name hd1 or name he1))
  or (resn arg and (name hb1 or name hg1 or name hd1)))

loop meth
  coor sele= (bondedto(bondedto(id $id)) and name h*) end
  if ($select = 2) then
    vector do (store1 = store1 + 1) (bondedto(bondedto(id $id)) and name h*)
  end if
end loop meth
```

2. sa_l_randomchain.xplor:

```
REMARK randomchain.inp
REMARK randomize a structure by random rotations around the N-Cα and Cα-C bonds.
REMARK input:  a coordinate set obtained with generate_template.inp or any
REMARK         other pdb file with good geometry
REMARK output: the coordinates randomized around phi and psi. chi_i are not
REMARK         affected.
REMARK the input file can be used also to generate the randomized coordinates
REMARK "on the fly" as part of the sa.inp protocol. the lines setting the seed,
```

```
REMARK reading the structure and the template coordinates at the beginning,
REMARK and writing out the randomized coordinates and the stop command at the end,
REMARK should then be deleted.
REMARK this file should be read after the "coor swap" and "coor copy" statements
REMARK at the beginning of the main loop in sa.inp.
REMARK Author: M. Nilges, EMBL, Oct. 1993

set message off echo off end

! The vector store1 is set to 1 for all atoms that have already been
! rotated.
vector do (store1=0) (all)
vector do (store3=0) (all)

! loop through all ca atoms.
for $id in id (name ca) loop pept

  vector show element (resid) (id $id)
        {* shows selected elements and stores the last element in $RESULT *}
  evaluate ($res = $result)
  vector show element (segid) (id $id)
        {* segid: segment or chain identifier ( string) *}
  evaluate ($seg = $result)
  vector do (store3=1) (byres (id $id))

  evaluate ($phi = random()*360-180)
  evaluate ($psi = random()*360-180)

! all atoms up to the present ca are held fixed for phi rotation
  vector do (store1=1)
    ((store3 and (name n or name hn or name ht* or name ca)))

! rotate phi
  coor rotate sele= (attribute store1=0)
     center (head (store3 and name ca))
     axis   (head (store3 and name ca)
            tail (store3 and name n)) $phi
  end

! fix the sidechain of the current residue for psi rotation
  vector do (store1=1)
    ((store3 and not (name c or name o or name ot#)))

! rotate psi
  coor rotate sele= (attribute store1=0 and segid $seg)
     center (head (store3 and name c) )
     axis   (head (store3 and name c)
            tail (store3 and name ca)) $psi
  end

! fix all of current residue
  vector do (store1=1) (store3)
  vector do (store3=0) (byres (id $id))

end loop pept

set message on echo on end
```

3. sa_1_initial_values.xplor

```
  evaluate ($ini_rad  = 0.9)      evaluate ($fin_rad  = 0.75)
```

```
evaluate ($ini_con=  0.003)  evaluate ($fin_con=  4.0)
evaluate ($ini_flt = 25)     evaluate ($fin_flt = 500)
evaluate ($ini_ang = 1.0)    evaluate ($fin_ang = 1.0)
evaluate ($ini_imp = 1.0)    evaluate ($fin_imp = 1.0)
evaluate ($ini_amb = 1.0)    evaluate ($fin_amb = 50.0)

evaluate ($bath  = $init_t)
evaluate ($k_vdw = $ini_con)
evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
evaluate ($radius=    $ini_rad)
evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))
evaluate ($k_flt = $ini_flt)
evaluate ($flt_fac = ($fin_flt/$ini_flt)^(1/$ncycle))
evaluate ($k_ang = $ini_ang)
evaluate ($ang_fac = ($fin_ang/$ini_ang)^(1/$ncycle))
evaluate ($k_imp = $ini_imp)
evaluate ($imp_fac = ($fin_imp/$ini_imp)^(1/$ncycle))
evaluate ($k_amb = $ini_amb)
evaluate ($amb_fac = ($fin_amb/$ini_amb)^(1/$ncycle))
```

4. sa_l_reduced.xplor

```
vector ident (store1)
((resn Gly  and name CA) or
 (resn Ala  and (name CA or name CB  )) or
 (resn Val  and (name CA or name CG1 )) or
 (resn Ile  and (name CA or name CG1 )) or
 (resn Leu  and (name CA or name CD1 )) or
 (resn Phe  and (name CA or name CD1 )) or
 (resn Pro  and (name CA or name CG  )) or
 (resn Met  and (name CA or name SD  )) or
 (resn Trp  and (name CA or name CE2 )) or
 (resn Cys  and (name CA or name SG  )) or
 (resn Ser  and (name CA or name CB  )) or
 (resn Thr  and (name CA or name CG2 )) or
 (resn Asn  and (name CA or name CG  )) or
 (resn Gln  and (name CA or name CD  )) or
 (resn Tyr  and (name CA or name CD1 )) or
 (resn His  and (name CA or name CD2 )) or
 (resn Asp  and (name CA or name CG  )) or
 (resn Glu  and (name CA or name CD  )) or
 (resn Lys  and (name CA or name CD  )) or
 (resn Arg  and (name CA or name CD  )))
```

5. sa_l_hightemp.xplor

```
evaluate ($nstep1=int($high_steps / 5.0 ) )

   vector do (vx = maxwell($init_t)) (all)
   vector do (vy = maxwell($init_t)) (all)
   vector do (vz = maxwell($init_t)) (all)


   for $knoe in (10 20 30 40 50) loop anne

noe scale dist $knoe scale ambi $ini_amb scale hbon $ini_amb end

    flags exclude * include noe end
   evaluate ($swap = 1.001)
   @swap15v.xplor    {* systematically swaps randomly selected methylene protons *}
                     {* and isopropyl groups and evaluates the noe energy term   *}
```

```
 flags include bond angl impr vdw noe cdih end

evaluate ($kvdw = $knoe * 0.002)
constraints
   interaction (all) (not store1) weights * 1 angl 1.0 impr 1.0 vdw   0.0 elec 0 end
   interaction (store1) (store1)  weights * 1 angl 1.0 impr 1.0 vdw $kvdw elec 0 end
end


dynamics  verlet
   nstep=$nstep1   timestep=0.005   iasvel=current   firstt=$init_t
    tcoupling=true  tbath=$init_t  nprint=50  iprfrq=0
   ntrfrq=9999
end

evaluate ($critical=$temp/$init_t)
if ($critical > 1.2) then
   evaluate ($nreassign = $nreassign + 1)
   mini powell nstep 100 end
   vector do (vx=maxwell($init_t)) ( all )
   vector do (vy=maxwell($init_t)) ( all )
   vector do (vz=maxwell($init_t)) ( all )
end if
end loop anne
```

6. sa_l_cool1.xplor

```
vector do (vx=maxwell($init_t)) ( all )
vector do (vy=maxwell($init_t)) ( all )
vector do (vz=maxwell($init_t)) ( all )

{*4 ========================================== cool the system *}
evaluate ($swap = 0.05)

restraints dihedral   scale=200.   end

parameter                        {* parameters for the repulsive energy term *}
 nbonds
  repel=1.                           {* initial value for repel - modified later *}
   rexp=2 irexp=2 rcon=1.
   nbxmod=3
   wmin=0.01
    cutnb=4.5 ctonnb=2.99 ctofnb=3.
   tolerance=0.5
end
end

evaluate ($ncycle = ($init_t-$final1_t)/$tempstep)
evaluate ($nstep = int($cool1_steps/$ncycle))


evaluate ($i_cool = 0)
while ($i_cool < $ncycle) loop cool
   evaluate ($i_cool=$i_cool+1)

   evaluate ($bath  = $bath  - $tempstep)
   evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
    evaluate ($radius=max($fin_rad,$radius*$radfact))
   evaluate ($k_flt = $k_flt*$flt_fac)
    evaluate ($k_ang = $k_ang*$ang_fac)
```

```
evaluate ($k_imp = $k_imp*$imp_fac)
evaluate ($k_amb = $k_amb*$amb_fac)


flags exclude * include noe end
@swap15v.xplor
flags include bond angl impr vdw noe cdih end

parameter
  bond  (name sg) (name sg) $k_flt TOKEN
  angle (all) (name sg) (name sg)  $k_flt TOKEN
  angle (store2) (all) (store2) $k_flt TOKEN
  angle (all)    (all) (store2) $k_flt TOKEN
  nbonds repel=$radius end
end

constraints
  interaction  (all) (all) weights * 1 vdw $k_vdw end
end
 noe scale ambi $k_amb scale hbon $k_amb end

dynamics  verlet
    nstep=$nstep time=0.005 iasvel=current firstt=$bath
     tcoup=true tbath=$bath nprint=$nstep iprfrq=0
    ntrfrq=99999
end

evaluate ($critical=$temp/$bath)
if ($critical > 1.2) then
   evaluate ($nreassign = $nreassign + 1)
   mini powell nstep 100 end
   vector do (vx=maxwell($bath)) ( all )
   vector do (vy=maxwell($bath)) ( all )
   vector do (vz=maxwell($bath)) ( all )
   end if

end loop cool
```

7. sa_l_cool2.xplor

```
constraints interaction (all) (all) weights * 1. vdw 4. end end

  evaluate ($ncycle = ($bath-$final2_t)/$tempstep)
  evaluate ($nstep = int($cool2_steps/$ncycle))


  evaluate ($i_cool = 0)
  while ($i_cool < $ncycle) loop cool
    evaluate ($i_cool=$i_cool+1)
    evaluate ($bath = $bath - $tempstep)

    flags exclude * include noe end
    @swap15v.xplor
    flags include bond angl impr vdw noe cdih end


    dynamics  verlet
       nstep=$nstep time=0.005 iasvel=current firstt=$bath
        tcoup=true tbath=$bath nprint=$nstep iprfrq=0
      ntrfrq=9999
    end
```

```
{====>}                                        {* abort condition *}
    evaluate ($critical=$temp/$bath)
    if ($critical > 1.2) then
      evaluate ($nreassign = $nreassign + 1)
      mini powell nstep 100 end
      vector do (vx=maxwell($bath)) ( all )
      vector do (vy=maxwell($bath)) ( all )
      vector do (vz=maxwell($bath)) ( all )
    end if

  end loop cool
```

8. swap15v.xplor

```
REMARK Swaps, systematically, randomly selected methylene protons
REMARK and accepts the conformation with lower noe energy.
REMARK Can be used in conjunction with floating assignment.
REMARK The floating groups have to be identified in store2
REMARK (with setup_swap.xplor and vector ident (store2) (store1)).
REMARK Uses the new "coor flip" and "coor vflip" commands.
REMARK   Author: Michael Nilges, EMBL.

set echo on message on end

evaluate ($reject = 0)
evaluate ($tries = 0)
evaluate ($newnoe=0)
energy end
evaluate ($oldnoe=$noe)

if ($swap > 0) then
  vector do (q = q * ran()) (attr q > -0.5)
end if

for $id in id (store2 and name h*1 and attr q < $swap and attr q > -0.5) loop meth
  vector show element (resid) (id $id)
  eval ($resid = $result)
  vector show element (name) (id $id)
  eval ($name1 = $result)
  eval ($name2 = $result - "1" + "2")
  coor flip sele= (resid $resid and (name $name1 or name $name2)) end

  energy end
  evaluate ($newnoe=$noe)
  evaluate ($tries = $tries + 1)

  if ($newnoe >= $oldnoe) then
    coor flip sele= (resid $resid and (name $name1 or name $name2)) end
    evaluate ($reject = $reject+1)
  else
    evaluate ($oldnoe = $newnoe)
    coor vflip sele= (resid $resid and (name $name1 or name $name2)) end
  end if
end loop meth

for $id in id (store2 and ((resn val and name cg1) or (resn leu and name cd1))
            and attr q < $swap and attr q > -0.5) loop meth
  coor copy end

  vector show element (resid) (id $id)
```

```
eval ($resid = $result)
vector show element (name) (id $id)
eval ($name1 = $result)
eval ($name2 = $name1 - "1" + "2")

coor flip sele= (resid $resid and (name $name1 or name $name2)) end

for $hycount in (1 2 3) loop myle
   eval ($hname1 = "H" + $name1 - "C" + encode($hycount))
   eval ($hname2 = "H" + $name2 - "C" + encode($hycount))
    coor flip sele= (resid $resid and (name $hname1 or name $hname2)) end
end loop myle

energy end
evaluate ($newnoe=$noe)
evaluate ($tries = $tries + 1)

if ($newnoe >= $oldnoe) then
  coor swap end
  evaluate ($reject = $reject+1)
else
  evaluate ($oldnoe = $newnoe)
  coor vflip sele= (resid $resid and (name $name1 or name $name2)) end
  for $hycount in (1 2 3) loop myle
    eval ($hname1 = "H" + $name1 - "C" + encode($hycount))
    eval ($hname2 = "H" + $name2 - "C" + encode($hycount))
     coor vflip sele= (resid $resid and (name $hname1 or name $hname2)) end
  end loop myle
 end if
end loop meth

if ($tries > 0) then evaluate ($reject = 100 * $reject / $tries) end if

vector do (q=1) (attr q < $swap and attr q > -0.5)

display  swaps tried: $tries   swaps rejected: $reject percent

set echo on message on end
```

refine.inp:

```
remarks file nmr/refine.inp  -- Simulated annealing refinement
remarks                          for NMR structure determination.
remarks  Authors: Michael Nilges, John Kuszewski, and Axel T. Brunger

{====>}
evaluate ($init_t = 1500)          {* initial annealing temperature, in K *}
{====>}
evaluate ($cool_steps = 4000 )     {* total number of steps during cooling*}

parameter                              {* read the parameter file       *}
{====>}
   @parallhdg.pro
end

{====>} structure @fib1213_SS.psf end  {* read structure file            *}
                                       {* (contains disulphide bonds) *}

set echo = off end
noe
{====>}
```

```
 nres=4000                          {* estimate greater than the actual number of NOEs. *}
     set echo on message on end
     class     dist       @newfib12132d.tbl   {* read unambiguous distance restraints *}
                          @newfib12133d.tbl
                          @newfib12133dsidechains.tbl
                          @fib1213_hbonds.tbl                {* read H-bonds restraints *}
                          @fib1213_cal.tbl            {* read calcium binding restraints *}
     class     ambi       @fib12132d_ambig.tbl {* read ambiguous distance restraints *}
                          @fib12132d_extraambig.tbl
                          @fib12133d_ambig.tbl
                          @fib12133d_extraambig.tbl
      set echo on message on end
end
set echo = on end

{====>}
@fib1213_dihe.tbl                    {* read dihedral (phi) angle restraints*}

vector do (fbeta=10) (all)        {* friction coefficient for MD heatbath, in 1/ps *}
vector do (mass=100) (all)        {* heavy masses to speed molecular dynamics * }

noe                          {* parameters for NOE effective energy term *}
   ceiling=1000
   averaging      * sum
   potential      * square
   sqconstant     * 1.
   sqexponent     * 2
   scale          * 50.          {* constant NOE scale throughout the protocol *}
end
parameter                      {* parameters for the repulsive energy term *}
   nbonds
        repel=0.5                 {* initial value for repel--modified later *}
        rexp=2 irexp=2 rcon=1.
        nbxmod=3
        wmin=0.01
        cutnb=4.5 ctonnb=2.99 ctofnb=3.
        tolerance=0.5
      end
   end

restraints dihedral
      scale=200.
end

{====>}
evaluate ($end_count= 100)          {* loop through a family of 100 structures.*}

evaluate ($count = 0)
while ($count < $end_count ) loop main

   evaluate ($count=$count+1)

{====>}                             {* filename(s) for embedded coordinates.*}
   evaluate ($filename="fib1213_swap_"+encode($count)+".pdb")

   coor @@$filename

   flags exclude * include bond angl impr vdw noe cdih end

   vector do (vx=maxwell($init_t)) ( all )
   vector do (vy=maxwell($init_t)) ( all )
```

```
vector do (vz=maxwell($init_t)) ( all )

evaluate ($final_t = 100)     { K }
evaluate ($tempstep = 50)     { K }

evaluate ($ncycle = ($init_t-$final_t)/$tempstep)
evaluate ($nstep = int($cool_steps/$ncycle))

evaluate ($ini_rad  = 0.9)        evaluate ($fin_rad  = 0.75)
evaluate ($ini_con=  0.003)       evaluate ($fin_con=  4.0)

evaluate ($bath  = $init_t)
evaluate ($k_vdw = $ini_con)
evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
evaluate ($radius=    $ini_rad)
evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))

evaluate ($i_cool = 0)
while ($i_cool < $ncycle) loop cool
  evaluate ($i_cool=$i_cool+1)

  evaluate ($bath  = $bath  - $tempstep)
  evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
  evaluate ($radius=max($fin_rad,$radius*$radfact))

  parameter  nbonds repel=$radius   end end
  constraints interaction (all) (all) weights * 1. vdw $k_vdw end end

  dynamics  verlet
     nstep=$nstep time=0.005 iasvel=current firstt=$bath
     tcoup=true tbath=$bath nprint=$nstep iprfrq=0
  end
{====>}                                  {* abort condition.*}
    evaluate ($critical=$temp/$bath)
    if ($critical >  10. ) then
       display  ****&&&& rerun job with smaller timestep (i.e., 0.003)
      exit cool
    end if
  end loop cool

{* ========================================== final  minimisation.*}
  minimize powell nstep= 200 nprint=25 end

{* ================================== write out the final structure(s).*}
  evaluate ($fname="fib1213_refine"+encode($count)+".prt")
  set print = $fname end
  noe print thresh = 0.1 end
  close $fname end
  print threshold=0.1 noe
  evaluate ($rms_noe=$result)
  evaluate ($violations_noe=$violations)
  print threshold=3. cdih
  evaluate ($rms_cdih=$result)
  evaluate ($violations_cdih=$violations)
  print thres=0.05 bonds
  evaluate ($rms_bonds=$result)
  print thres=5. angles
  evaluate ($rms_angles=$result)
  print thres=5. impropers
  evaluate ($rms_impropers=$result)
```

```
    remarks  ============================================================
  remarks           overall,bonds,angles,improper,vdw,noe,cdih
  remarks energies: $ener, $bond, $angl, $impr, $vdw, $noe, $cdih
    remarks  ============================================================
  remarks           bonds,angles,impropers,noe,cdih
  remarks rms-d: $rms_bonds,$rms_angles,$rms_impropers,$rms_noe,$rms_cdih
    remarks  ============================================================
  remarks             noe,  cdih
  remarks violations.: $violations_noe, $violations_cdih
    remarks  ============================================================

{====>}                        {* name(s) of the family of final structures.*}
   evaluate ($filename="fib1213_refine_"+encode($count)+".pdb")

   write coordinates output =$filename end

end loop main

stop
```

Appendix II

Minimal Media Labelling

For 1 litre;

6.0g Na_2HPO_4
3.0g KH_2PO_4
0.5g NaCl
1.0g $^{15}NH_4Cl$ (0.7g sufficed)
1ml 1M $MgSO_4$ (filter sterilised)
1ml 0.1M $CaCl_2$ (filter sterilised)
1ml 1M thiamine (filter sterilised)
10ml of *E.coli* trace elements (filter sterilised) -see below*
25ml 20% Glucose (filter sterilised)
1.5g of Yeast Nitrogen Base without ammonium chloride and amino acids (filter sterilised)

**E. coli* Trace Elements;
5g EDTA were dissolved in 800ml H_2O and the pH adjusted to 7.0 with HCl. The following were then added, returning the pH to 7.0 after every addition,

0.5g $FeCl_3$
0.05g $ZnCl_2$
0.01g $CuCl_2$
0.01g $CoCl_2.6H_2O$
0.01g H_3BO_3

Appendix III

Amino Acid Analysis

In the calcium titrations, two methods were used to calculate the total concentration of protein present. The first method involved the measurement of the $O.D._{280}$ of the sample. This method has several flaws, as imperfect cuvettes or tiny solid particles in the solution could greatly influence the $O.D._{280}$ obtained. Therefore a second, more accurate, method of determining protein concentration was employed, i.e., amino acid analysis.

In each of the titrations performed, 10-20µl of the sample were reserved for amino acid analysis (a.a.analysis) prior to any calcium additions. Also, during and/or at the end of the titrations, 10-20µl were reserved for a.a.analysis to check that no loss of protein, or contamination by protein, had occurred. An exact volume between 1-20µl was submitted and made up to 100µl with water. This was divided into three runs of 20µl (A, B, and C). Each 20µl sample solution was analysed against an internal standard of Nor-Leucine. The actual results of a.a.analysis on the cbEGF12-13 sample can be seen in Table A1. For each run, the final column of the table shows the concentration of each amino acid present in a known volume of sample submitted, and to calculate the protein concentration the number of times each amino acid appears in the peptide's sequence must be taken into account. Only certain amino acids are reliable for calculation of protein concentration, e.g., Arg, Ala, Val, Leu and Phe. Calculations of protein concentration from analysis results of cbEGF12-13 are shown in Tables A2 and A3.

The a.a.analysis method is more accurate than measuring the $O.D._{280}$, as it has many safety nets:

1. The internal Nor-Leu standard highlights any odd results that may arise as a result of the analysis itself.
2. Running the sample in three separate runs increases result confidence.
3. Protein contamination can be detected by this method. If contamination had occurred, large inconsistencies would have arisen whilst calculating [protein].

Therefore, for maximum accuracy, all $[Ca^{2+}]_{free}$ values in this study were computed using protein concentration values calculated from a.a.analysis results.

Table A1 : Results of Amino Acid Analysis on the cbEGF12-13 Domain Pair

The table shows the results of one 20µl run, and the final column (PMOL correc. INT STD) shows the concentration of each amino acid present in the 20µl of sample submitted. Amino acid analysis was kindly performed by Dr. Tony Willis.

Sample ID: RS20 (initiated 7/21/98 2:11am) BASELINE CORRECTED

Turntable Position: 4 C
Data Start : 4.00 min
Data Duration : 16.00 min
Peak Ht Threshold : 100 uAU

Sampling Interval: 1.0 sec
Samples In Run : 30
Operator ID : MRC IMMUNO
Int. Std. Amt : 2000 pmol

Calibration File : STDH
Reference Time : 10.00 min
Reference Offset 1: 0.02 min
Reference Offset 2: 0.00 min

(initiated 7/21/98 1:25pm)
ISTD Peak ID : N-LEU

Integration Interval: 0.0 to 15.0 min

PEAK ID	RET. TIME min	CAL. TIME min	PEAK HEIGHT uAU	PMOL BY HEIGHT	PMOL correc. INT STD
Aspartic Acid	1.47	1.42	16927	125.41	1149.27
Glutamic Acid	1.93	1.90	12218	107.91	988.84
	3.48	-----	105	----------	----------
	3.78	-----	105	----------	----------
Serine	3.05	3.03	6295	74.36	681.38
Glycine	3.72	3.72	10151	111.73	1023.90
Histidine	4.00	4.00	1589	19.41	169.63
	4.50	-----	665	----------	----------
Arginine	4.70	4.72	6398	49.66	455.07
	4.93	-----	168	----------	----------
Threonine	5.18	5.20	3468	31.75	290.98
Alanine	5.58	5.59	2969	26.14	239.57
Proline	5.90	5.93	6362	53.03	485.96
	6.15	-----	5985	----------	----------
	6.95	-----	112	----------	----------
	7.37	-----	4419	----------	----------
	7.62	-----	528	----------	----------
	7.95	-----	838	----------	----------
	8.00	-----	337181	----------	----------
	8.58	-----	140	----------	----------
	8.78	-----	2362	----------	----------
Tyrosine	8.92	8.90	1331	9.71	88.96
	9.32	-----	150	----------	----------
	9.57	-----	303	----------	----------
Valine	9.88	9.88	1867	14.25	130.56
	9.98	-----	1483	----------	----------
Methionine	10.38	10.40	265	2.30	21.07
	10.93	-----	278	----------	----------
Cysteine	11.38	11.45	266	1.26	11.50
	11.78	-----	1899	----------	----------
Isoleucine	12.05	12.05	6734	42.98	393.88
Leucine	12.25	12.25	6124	41.65	381.64
N-LEU	12.62	12.62	31233	218.25	2000.00
	12.87	-----	269	----------	----------
Phenylalanine	13.10	13.10	2848	19.47	178.40
	13.40	-----	4315	----------	----------
	13.93	-----	4761	----------	----------
Lysine	14.10	14.10	5066	20.70	189.67
	14.42	-----	13845	----------	----------
PCP	14.87	14.78	332	2.96	27.09
	15.43	-----	527	----------	

Table A2 : Calculation of [cbEGF12-13] Present in the NMR Calcium Titration

The results given in column two of the table below were for 1.25µl of sample submitted.

Amino Acid	Average of two runs (pmols present in 1.25µl sample)	No. of particular amino acid type present in cbEGF12-13	Correction for multiplicity of a.a's (pmols in 1.25µl sample)
Leucine	605.76	4	151.44
Valine	343.11	2	171.555
Arginine	731.865	5	146.373
Alanine	368.12	2	184.06

Average no. of pmols present in 1.25µl = 163.357pmols

$$[\text{protein}] = 163.357 * \frac{550\mu l}{1.25\mu l} * \frac{1000ml}{0.55ml} = 1.30685 \times 10^{8} \text{pmols}$$

$$\Rightarrow [\text{cbEGF12-13}] = 0.131\text{mM}$$

Table A3 : Calculation of [cbEGF12-13] Present in One of the Fluorescence Calcium Titrations

The results given in column two of the table below were for 20µl of sample submitted.

Amino Acid	Average of three runs (pmols present in 20µl sample)	No. of particular amino acid type present in cbEGF12-13	Correction for multiplicity of a.a's (pmols in 20µl sample)
Leucine	381.64	4	95.41
Valine	130.56	2	65.28
Arginine	455.07	5	91.014
Alanine	239.57	2	119.785
Phenylalanine	178.40	2	89.2

Average no. of pmols present in 20µl = 92.1378pmols

$$[\text{protein}] = 92.1378 * \frac{2000\mu l}{20\mu l} * \frac{1000ml}{2ml} = 4.606890 \times 10^{6} \text{pmols}$$

$$\Rightarrow [\text{cbEGF12-13}] = 4.6\mu\text{M}$$

Appendix IV

Experimental Restraints Used for Structure Calculations of cbEGF12-13

(1) Unamgiuous Interproton Distance Restraints

```
assign ( resid   2  and name   ha  ) ( resid   2  and name   hb#  ) 0.0 0.0 3.5
assign ( resid   3  and name   ha  ) ( resid   3  and name   hb   ) 0.0 0.0 3.5
assign ( resid   3  and name   ha  ) ( resid   3  and name   hg2# ) 0.0 0.0 2.8
assign ( resid   3  and name   hb  ) ( resid   3  and name   hg2# ) 0.0 0.0 2.8
assign ( resid   4  and name   ha  ) ( resid   4  and name   hb1  ) 0.0 0.0 3.5
assign ( resid   4  and name   hn  ) ( resid   3  and name   ha   ) 0.0 0.0 3.5
assign ( resid   4  and name   hn  ) ( resid   3  and name   hb   ) 0.0 0.0 5.0
assign ( resid   4  and name   hn  ) ( resid   3  and name   hg2# ) 0.0 0.0 5.0
assign ( resid   4  and name   hn  ) ( resid   4  and name   ha   ) 0.0 0.0 2.8
assign ( resid   4  and name   hn  ) ( resid   4  and name   hb1  ) 0.0 0.0 5.0
assign ( resid   4  and name   hn  ) ( resid   4  and name   hb2  ) 0.0 0.0 5.0
assign ( resid   4  and name   hn  ) ( resid   5  and name   hn   ) 0.0 0.0 5.0
assign ( resid   4  and name   hn  ) ( resid   5  and name   hg2# ) 0.0 0.0 6.0
assign ( resid   4  and name   hb2 ) ( resid   4  and name   hb1  ) 0.0 0.0 2.8
assign ( resid   5  and name   ha  ) ( resid   5  and name   hd1# ) 0.0 0.0 3.5
assign ( resid   5  and name   ha  ) ( resid   5  and name   hg1# ) 0.0 0.0 3.5
assign ( resid   5  and name   ha  ) ( resid   5  and name   hg2# ) 0.0 0.0 3.5
assign ( resid   5  and name   hb  ) ( resid   5  and name   hg1# ) 0.0 0.0 2.8
assign ( resid   5  and name   hb  ) ( resid   5  and name   hg2# ) 0.0 0.0 2.8
assign ( resid   5  and name   hn  ) ( resid   4  and name   ha   ) 0.0 0.0 2.8
assign ( resid   5  and name   hn  ) ( resid   4  and name   hn   ) 0.0 0.0 6.0
assign ( resid   5  and name   hn  ) ( resid   4  and name   hb1  ) 0.0 0.0 5.0
assign ( resid   5  and name   hn  ) ( resid   4  and name   hb2  ) 0.0 0.0 5.0
assign ( resid   5  and name   hn  ) ( resid   5  and name   ha   ) 0.0 0.0 3.5
assign ( resid   5  and name   hn  ) ( resid   5  and name   hd1# ) 0.0 0.0 3.5
assign ( resid   5  and name   hn  ) ( resid   5  and name   hg1# ) 0.0 0.0 5.0
assign ( resid   5  and name   hn  ) ( resid   6  and name   hn   ) 0.0 0.0 5.0
assign ( resid   5  and name   hn  ) ( resid   7  and name   hn   ) 0.0 0.0 6.0
assign ( resid   5  and name   hn  ) ( resid  23  and name   hn   ) 0.0 0.0 6.0
assign ( resid   5  and name   hn  ) ( resid  25  and name   hn   ) 0.0 0.0 5.0
assign ( resid   5  and name   hg1# ) ( resid   5  and name   hg2# ) 0.0 0.0 2.8
assign ( resid   6  and name   ha  ) ( resid   5  and name   hg2# ) 0.0 0.0 5.0
assign ( resid   6  and name   hn  ) ( resid   5  and name   ha   ) 0.0 0.0 3.5
assign ( resid   6  and name   hn  ) ( resid   5  and name   hn   ) 0.0 0.0 5.0
assign ( resid   6  and name   hn  ) ( resid   5  and name   hd1# ) 0.0 0.0 5.0
assign ( resid   6  and name   hn  ) ( resid   5  and name   hg1# ) 0.0 0.0 6.0
assign ( resid   6  and name   hn  ) ( resid   5  and name   hg2# ) 0.0 0.0 5.0
assign ( resid   6  and name   hn  ) ( resid   6  and name   ha   ) 0.0 0.0 6.0
assign ( resid   6  and name   hn  ) ( resid   6  and name   hb1  ) 0.0 0.0 6.0
assign ( resid   6  and name   hn  ) ( resid   6  and name   hb2  ) 0.0 0.0 5.0
assign ( resid   6  and name   hn  ) ( resid   7  and name   hn   ) 0.0 0.0 5.0
assign ( resid   7  and name   ha  ) ( resid   7  and name   hb1  ) 0.0 0.0 3.5
assign ( resid   7  and name   ha  ) ( resid   7  and name   hb2  ) 0.0 0.0 3.5
assign ( resid   7  and name   ha  ) ( resid   7  and name   hg1  ) 0.0 0.0 3.5
assign ( resid   7  and name   ha  ) ( resid   7  and name   hg2  ) 0.0 0.0 5.0
```

```
assign ( resid   7  and name   ha  ) ( resid  10  and name   hb  ) 0.0 0.0 5.0
assign ( resid   7  and name   ha  ) ( resid  10  and name  hg12  ) 0.0 0.0 5.0
assign ( resid   7  and name   ha  ) ( resid  10  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid   7  and name   ha  ) ( resid  14  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid   7  and name   ha  ) ( resid  14  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   5  and name   ha  ) 0.0 0.0 6.0
assign ( resid   7  and name   hn  ) ( resid   6  and name   ha  ) 0.0 0.0 3.5
assign ( resid   7  and name   hn  ) ( resid   6  and name   hn  ) 0.0 0.0 6.0
assign ( resid   7  and name   hn  ) ( resid   6  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   6  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   7  and name   ha  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   7  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   7  and name  hg1  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   7  and name  hg2  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid   8  and name   hn  ) 0.0 0.0 5.0
assign ( resid   7  and name   hn  ) ( resid  10  and name  hg12  ) 0.0 0.0 6.0
assign ( resid   7  and name   hn  ) ( resid  27  and name  hb2  ) 0.0 0.0 6.0
assign ( resid   7  and name   hn  ) ( resid  27  and name  hd#  ) 0.0 0.0 6.0
assign ( resid   7  and name  hb1  ) ( resid  10  and name  hg11  ) 0.0 0.0 5.0
assign ( resid   7  and name  hb1  ) ( resid  14  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid   7  and name  hg1  ) ( resid   5  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid   7  and name  hg1  ) ( resid   7  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   7  and name  hg1  ) ( resid   7  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   7  and name  hg2  ) ( resid   5  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid   7  and name  hg2  ) ( resid   7  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   7  and name  hg2  ) ( resid  14  and name  hd1#  ) 0.0 0.0 6.0
assign ( resid   8  and name   ha  ) ( resid   8  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   8  and name   ha  ) ( resid  14  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid   8  and name   ha  ) ( resid  14  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid   8  and name   ha  ) ( resid  20  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   6  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   6  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   6  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   6  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   7  and name   ha  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   7  and name   hn  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   7  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   7  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   8  and name   hn  ) ( resid   7  and name  hg1  ) 0.0 0.0 6.0
assign ( resid   8  and name   hn  ) ( resid   8  and name   ha  ) 0.0 0.0 3.5
assign ( resid   8  and name   hn  ) ( resid   9  and name  hb2  ) 0.0 0.0 6.0
assign ( resid   8  and name   hn  ) ( resid  10  and name   hb  ) 0.0 0.0 6.0
assign ( resid   8  and name   hn  ) ( resid  27  and name  hd#  ) 0.0 0.0 6.0
assign ( resid   8  and name  hb2  ) ( resid  20  and name  hb1  ) 0.0 0.0 5.0
assign ( resid   9  and name   ha  ) ( resid   9  and name  hb2  ) 0.0 0.0 3.5
assign ( resid   9  and name  hb2  ) ( resid   9  and name  hb1  ) 0.0 0.0 3.5
assign ( resid   9  and name  hb2  ) ( resid   9  and name  hg#  ) 0.0 0.0 3.5
assign ( resid   9  and name  hd#  ) ( resid   9  and name  hb2  ) 0.0 0.0 5.0
assign ( resid   9  and name  hd#  ) ( resid   9  and name  hg#  ) 0.0 0.0 2.8
assign ( resid  10  and name   ha  ) ( resid  10  and name   hb  ) 0.0 0.0 3.5
assign ( resid  10  and name   ha  ) ( resid  10  and name  hg11  ) 0.0 0.0 5.0
assign ( resid  10  and name   ha  ) ( resid  10  and name  hg12  ) 0.0 0.0 5.0
assign ( resid  10  and name   ha  ) ( resid  10  and name  hg2#  ) 0.0 0.0 2.8
assign ( resid  10  and name   hn  ) ( resid   6  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  10  and name   hn  ) ( resid   6  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  10  and name   hn  ) ( resid   7  and name   ha  ) 0.0 0.0 3.5
assign ( resid  10  and name   hn  ) ( resid   8  and name   ha  ) 0.0 0.0 5.0
assign ( resid  10  and name   hn  ) ( resid   8  and name   hn  ) 0.0 0.0 5.0
assign ( resid  10  and name   hn  ) ( resid   8  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  10  and name   hn  ) ( resid   9  and name   ha  ) 0.0 0.0 3.5
assign ( resid  10  and name   hn  ) ( resid  10  and name   ha  ) 0.0 0.0 3.5
assign ( resid  10  and name   hn  ) ( resid  10  and name   hb  ) 0.0 0.0 5.0
assign ( resid  10  and name   hn  ) ( resid  10  and name  hg11  ) 0.0 0.0 3.5
assign ( resid  10  and name   hn  ) ( resid  10  and name  hg12  ) 0.0 0.0 5.0
assign ( resid  10  and name   hn  ) ( resid  10  and name  hg2#  ) 0.0 0.0 5.0
```

```
assign ( resid  10  and name   hn  ) ( resid  11  and name   ha  ) 0.0 0.0 6.0
assign ( resid  10  and name   hn  ) ( resid  11  and name   hb# ) 0.0 0.0 5.0
assign ( resid  10  and name  hd1# ) ( resid  10  and name  hg12 ) 0.0 0.0 2.8
assign ( resid  10  and name  hg12 ) ( resid  10  and name  hg2# ) 0.0 0.0 5.0
assign ( resid  11  and name   ha  ) ( resid  12  and name  hd1  ) 0.0 0.0 2.8
assign ( resid  11  and name   hn  ) ( resid   8  and name   hn  ) 0.0 0.0 6.0
assign ( resid  11  and name   hn  ) ( resid  10  and name   ha  ) 0.0 0.0 3.5
assign ( resid  11  and name   hn  ) ( resid  10  and name   hb  ) 0.0 0.0 3.5
assign ( resid  11  and name   hn  ) ( resid  10  and name  hg11 ) 0.0 0.0 6.0
assign ( resid  11  and name   hn  ) ( resid  10  and name  hg12 ) 0.0 0.0 5.0
assign ( resid  11  and name   hn  ) ( resid  10  and name  hg2# ) 0.0 0.0 5.0
assign ( resid  11  and name   hn  ) ( resid  11  and name   ha  ) 0.0 0.0 3.5
assign ( resid  11  and name  hb#  ) ( resid  10  and name  hg2# ) 0.0 0.0 5.0
assign ( resid  11  and name  hb#  ) ( resid  14  and name   hg  ) 0.0 0.0 3.5
assign ( resid  11  and name  hb#  ) ( resid  14  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  11  and name  hb#  ) ( resid  14  and name  hd1# ) 0.0 0.0 3.5
assign ( resid  11  and name  hb#  ) ( resid  14  and name  hd2# ) 0.0 0.0 5.0
assign ( resid  12  and name   ha  ) ( resid  12  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  12  and name   ha  ) ( resid  12  and name  hd2  ) 0.0 0.0 5.0
assign ( resid  12  and name   ha  ) ( resid  14  and name  hd1# ) 0.0 0.0 5.0
assign ( resid  12  and name  hd1  ) ( resid  12  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  12  and name  hd2  ) ( resid  12  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  12  and name  hd2  ) ( resid  12  and name  hd1  ) 0.0 0.0 2.8
assign ( resid  13  and name   ha  ) ( resid  13  and name  hb2  ) 0.0 0.0 2.8
assign ( resid  13  and name   hn  ) ( resid  10  and name  hg2# ) 0.0 0.0 6.0
assign ( resid  13  and name   hn  ) ( resid  12  and name   ha  ) 0.0 0.0 3.5
assign ( resid  13  and name   hn  ) ( resid  12  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  13  and name   hn  ) ( resid  12  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  13  and name   hn  ) ( resid  13  and name   ha  ) 0.0 0.0 3.5
assign ( resid  13  and name   hn  ) ( resid  13  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  13  and name   hn  ) ( resid  13  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  13  and name   hn  ) ( resid  14  and name   ha  ) 0.0 0.0 5.0
assign ( resid  13  and name   hn  ) ( resid  14  and name   hn  ) 0.0 0.0 3.5
assign ( resid  13  and name   hn  ) ( resid  14  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  13  and name   hn  ) ( resid  15  and name   hn  ) 0.0 0.0 6.0
assign ( resid  14  and name   ha  ) ( resid  10  and name   hb  ) 0.0 0.0 6.0
assign ( resid  14  and name   ha  ) ( resid  11  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  14  and name   ha  ) ( resid  14  and name  hd1# ) 0.0 0.0 3.5
assign ( resid  14  and name   ha  ) ( resid  14  and name  hd2# ) 0.0 0.0 5.0
assign ( resid  14  and name   hg  ) ( resid  14  and name  hd1# ) 0.0 0.0 3.5
assign ( resid  14  and name   hn  ) ( resid  10  and name  hg2# ) 0.0 0.0 6.0
assign ( resid  14  and name   hn  ) ( resid  12  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  14  and name   hn  ) ( resid  12  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  14  and name   hn  ) ( resid  13  and name   ha  ) 0.0 0.0 3.5
assign ( resid  14  and name   hn  ) ( resid  13  and name   hn  ) 0.0 0.0 3.5
assign ( resid  14  and name   hn  ) ( resid  13  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  14  and name   hn  ) ( resid  14  and name   ha  ) 0.0 0.0 2.8
assign ( resid  14  and name   hn  ) ( resid  14  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  14  and name   hn  ) ( resid  15  and name   hn  ) 0.0 0.0 5.0
assign ( resid  14  and name   hn  ) ( resid  16  and name   hn  ) 0.0 0.0 5.0
assign ( resid  14  and name   hn  ) ( resid  20  and name   hn  ) 0.0 0.0 5.0
assign ( resid  14  and name  hb2  ) ( resid  14  and name  hd1# ) 0.0 0.0 3.5
assign ( resid  14  and name  hb2  ) ( resid  14  and name  hd2# ) 0.0 0.0 3.5
assign ( resid  15  and name   ha  ) ( resid  15  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  15  and name   ha  ) ( resid  15  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  15  and name   hn  ) ( resid  13  and name   ha  ) 0.0 0.0 5.0
assign ( resid  15  and name   hn  ) ( resid  13  and name   hn  ) 0.0 0.0 6.0
assign ( resid  15  and name   hn  ) ( resid  14  and name   ha  ) 0.0 0.0 3.5
assign ( resid  15  and name   hn  ) ( resid  14  and name   hn  ) 0.0 0.0 5.0
assign ( resid  15  and name   hn  ) ( resid  14  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  15  and name   hn  ) ( resid  15  and name   ha  ) 0.0 0.0 3.5
assign ( resid  15  and name   hn  ) ( resid  15  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  15  and name   hn  ) ( resid  16  and name   hn  ) 0.0 0.0 3.5
```

```
assign ( resid  15  and name  hn  ) ( resid  16  and name  ha1  ) 0.0 0.0 6.0
assign ( resid  15  and name  hn  ) ( resid  16  and name  ha2  ) 0.0 0.0 5.0
assign ( resid  15  and name  hn  ) ( resid  20  and name  hn  ) 0.0 0.0 5.0
assign ( resid  15  and name  hn  ) ( resid  44  and name  ha  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  13  and name  ha  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  14  and name  ha  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  14  and name  hn  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  14  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  15  and name  ha  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  15  and name  hn  ) 0.0 0.0 3.5
assign ( resid  16  and name  hn  ) ( resid  15  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  16  and name  ha1  ) 0.0 0.0 2.8
assign ( resid  16  and name  hn  ) ( resid  16  and name  ha2  ) 0.0 0.0 2.8
assign ( resid  16  and name  hn  ) ( resid  17  and name  hn  ) 0.0 0.0 5.0
assign ( resid  16  and name  hn  ) ( resid  17  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  16  and name  hn  ) ( resid  17  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  16  and name  hn  ) ( resid  44  and name  hn  ) 0.0 0.0 6.0
assign ( resid  16  and name  hn  ) ( resid  45  and name  hn  ) 0.0 0.0 5.0
assign ( resid  16  and name  ha2  ) ( resid  16  and name  ha1  ) 0.0 0.0 2.8
assign ( resid  17  and name  ha  ) ( resid  17  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  17  and name  ha  ) ( resid  17  and name  hd1  ) 0.0 0.0 3.5
assign ( resid  17  and name  ha  ) ( resid  17  and name  hd2  ) 0.0 0.0 3.5
assign ( resid  17  and name  ha  ) ( resid  17  and name  hg#  ) 0.0 0.0 2.8
assign ( resid  17  and name  hn  ) ( resid  15  and name  hn  ) 0.0 0.0 5.0
assign ( resid  17  and name  hn  ) ( resid  16  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  17  and name  hn  ) ( resid  17  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  17  and name  hn  ) ( resid  17  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  17  and name  hn  ) ( resid  17  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  17  and name  hn  ) ( resid  17  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  17  and name  hn  ) ( resid  35  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  17  and name  hb2  ) ( resid  17  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  17  and name  hb2  ) ( resid  17  and name  hg#  ) 0.0 0.0 3.5
assign ( resid  17  and name  hd1  ) ( resid  17  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  17  and name  hd1  ) ( resid  17  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  17  and name  hd1  ) ( resid  17  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  17  and name  hd2  ) ( resid  17  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  17  and name  hd2  ) ( resid  17  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  17  and name  hd2  ) ( resid  17  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  17  and name  hg#  ) ( resid  17  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  18  and name  hn  ) ( resid  15  and name  hn  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  16  and name  ha1  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  17  and name  hn  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  17  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  17  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  17  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  17  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  18  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  18  and name  hn  ) ( resid  18  and name  ha2  ) 0.0 0.0 3.5
assign ( resid  18  and name  hn  ) ( resid  19  and name  hn  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  30  and name  hn  ) 0.0 0.0 6.0
assign ( resid  18  and name  hn  ) ( resid  31  and name  ha  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  32  and name  hn  ) 0.0 0.0 5.0
assign ( resid  18  and name  hn  ) ( resid  45  and name  hn  ) 0.0 0.0 5.0
assign ( resid  19  and name  ha  ) ( resid  19  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  19  and name  ha  ) ( resid  19  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  19  and name  hn  ) ( resid  15  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  19  and name  hn  ) ( resid  18  and name  hn  ) 0.0 0.0 5.0
assign ( resid  19  and name  hn  ) ( resid  18  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  19  and name  hn  ) ( resid  18  and name  ha2  ) 0.0 0.0 3.5
assign ( resid  19  and name  hn  ) ( resid  19  and name  ha  ) 0.0 0.0 3.5
assign ( resid  19  and name  hn  ) ( resid  19  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  19  and name  hn  ) ( resid  19  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  19  and name  hn  ) ( resid  19  and name  hg2  ) 0.0 0.0 5.0
```

```
assign ( resid  19  and name   hn  ) ( resid  20  and name   ha  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  20  and name   hn  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  29  and name   ha  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  29  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  30  and name   hn  ) 0.0 0.0 3.5
assign ( resid  19  and name   hn  ) ( resid  30  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  30  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  30  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  30  and name  hg2  ) 0.0 0.0 6.0
assign ( resid  19  and name   hn  ) ( resid  31  and name   ha  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  31  and name   hn  ) 0.0 0.0 5.0
assign ( resid  19  and name   hn  ) ( resid  31  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  19  and name   hn  ) ( resid  32  and name   hn  ) 0.0 0.0 6.0
assign ( resid  19  and name  hg1  ) ( resid  30  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  19  and name  hg2  ) ( resid  19  and name  hg1  ) 0.0 0.0 3.5
assign ( resid  19  and name  he21 ) ( resid  19  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  19  and name  he21 ) ( resid  19  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  19  and name  he21 ) ( resid  30  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  19  and name  he22 ) ( resid  19  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  19  and name  he22 ) ( resid  30  and name  hg1  ) 0.0 0.0 6.0
assign ( resid  19  and name  ne21 ) ( resid  19  and name  he21 ) 0.0 0.0 2.8
assign ( resid  19  and name  ne21 ) ( resid  19  and name  he22 ) 0.0 0.0 2.8
assign ( resid  19  and name  ne22 ) ( resid  19  and name  he21 ) 0.0 0.0 2.8
assign ( resid  19  and name  ne22 ) ( resid  19  and name  he22 ) 0.0 0.0 2.8
assign ( resid  20  and name   ha  ) ( resid  20  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  20  and name   ha  ) ( resid  20  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  20  and name   ha  ) ( resid  29  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid   8  and name   ha  ) 0.0 0.0 6.0
assign ( resid  20  and name   hn  ) ( resid  19  and name   ha  ) 0.0 0.0 3.5
assign ( resid  20  and name   hn  ) ( resid  19  and name   hn  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid  19  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid  19  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid  19  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid  20  and name   ha  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid  20  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  20  and name   hn  ) ( resid  21  and name   hn  ) 0.0 0.0 5.0
assign ( resid  20  and name  hb1  ) ( resid  14  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  21  and name   ha  ) ( resid   8  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  21  and name   ha  ) ( resid  21  and name   hb  ) 0.0 0.0 3.5
assign ( resid  21  and name   hn  ) ( resid  20  and name   ha  ) 0.0 0.0 3.5
assign ( resid  21  and name   hn  ) ( resid  20  and name   hn  ) 0.0 0.0 6.0
assign ( resid  21  and name   hn  ) ( resid  20  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  21  and name   hn  ) ( resid  21  and name   ha  ) 0.0 0.0 5.0
assign ( resid  21  and name   hn  ) ( resid  22  and name   hn  ) 0.0 0.0 6.0
assign ( resid  21  and name   hn  ) ( resid  27  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  21  and name   hn  ) ( resid  28  and name   hn  ) 0.0 0.0 5.0
assign ( resid  21  and name   hn  ) ( resid  28  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  21  and name   hn  ) ( resid  29  and name   ha  ) 0.0 0.0 5.0
assign ( resid  21  and name   hn  ) ( resid  30  and name   hn  ) 0.0 0.0 5.0
assign ( resid  22  and name   ha  ) ( resid  23  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  22  and name   hn  ) ( resid   8  and name   hn  ) 0.0 0.0 6.0
assign ( resid  22  and name   hn  ) ( resid  21  and name   ha  ) 0.0 0.0 2.8
assign ( resid  22  and name   hn  ) ( resid  21  and name   hb  ) 0.0 0.0 5.0
assign ( resid  22  and name   hn  ) ( resid  21  and name   hn  ) 0.0 0.0 5.0
assign ( resid  22  and name   hn  ) ( resid  22  and name   ha  ) 0.0 0.0 5.0
assign ( resid  22  and name   hn  ) ( resid  22  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  22  and name   hn  ) ( resid  23  and name   hn  ) 0.0 0.0 5.0
assign ( resid  22  and name   hn  ) ( resid  27  and name   ha  ) 0.0 0.0 5.0
assign ( resid  22  and name  hd21 ) ( resid   6  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  22  and name  hd21 ) ( resid   7  and name   hn  ) 0.0 0.0 6.0
assign ( resid  22  and name  hd21 ) ( resid   9  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  22  and name  hd21 ) ( resid  23  and name   hn  ) 0.0 0.0 5.0
assign ( resid  22  and name  hd22 ) ( resid   6  and name  hb1  ) 0.0 0.0 5.0
```

```
assign ( resid  22 and name  hd22 ) ( resid   7 and name    hn ) 0.0 0.0 6.0
assign ( resid  22 and name  nd21 ) ( resid  22 and name  hd21 ) 0.0 0.0 2.8
assign ( resid  22 and name  nd21 ) ( resid  22 and name  hd22 ) 0.0 0.0 2.8
assign ( resid  22 and name  nd22 ) ( resid  22 and name  hd21 ) 0.0 0.0 2.8
assign ( resid  22 and name  nd22 ) ( resid  22 and name  hd22 ) 0.0 0.0 2.8
assign ( resid  23 and name   ha ) ( resid  23 and name  hg2# ) 0.0 0.0 2.8
assign ( resid  23 and name   ha ) ( resid  24 and name   hb2 ) 0.0 0.0 6.0
assign ( resid  23 and name   ha ) ( resid  24 and name   hd1 ) 0.0 0.0 2.8
assign ( resid  23 and name   ha ) ( resid  24 and name   hd2 ) 0.0 0.0 2.8
assign ( resid  23 and name   ha ) ( resid  24 and name   hg1 ) 0.0 0.0 5.0
assign ( resid  23 and name   ha ) ( resid  24 and name   hg2 ) 0.0 0.0 5.0
assign ( resid  23 and name   hb ) ( resid  23 and name  hg2# ) 0.0 0.0 2.8
assign ( resid  23 and name   hn ) ( resid  22 and name    hn ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  22 and name   hb1 ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  23 and name    hb ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  23 and name  hg2# ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  24 and name   hd1 ) 0.0 0.0 6.0
assign ( resid  23 and name   hn ) ( resid  24 and name   hd2 ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  26 and name    hn ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  27 and name    ha ) 0.0 0.0 3.5
assign ( resid  23 and name   hn ) ( resid  27 and name    hn ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  27 and name   hb1 ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  27 and name   hb2 ) 0.0 0.0 5.0
assign ( resid  23 and name   hn ) ( resid  27 and name   hd# ) 0.0 0.0 6.0
assign ( resid  23 and name   hn ) ( resid  28 and name    hn ) 0.0 0.0 5.0
assign ( resid  24 and name  hb2 ) ( resid  24 and name   hg2 ) 0.0 0.0 2.8
assign ( resid  24 and name  hd1 ) ( resid  23 and name  hg2# ) 0.0 0.0 5.0
assign ( resid  24 and name  hd1 ) ( resid  24 and name   hb1 ) 0.0 0.0 3.5
assign ( resid  24 and name  hd2 ) ( resid  23 and name  hg2# ) 0.0 0.0 3.5
assign ( resid  24 and name  hd2 ) ( resid  24 and name   hb2 ) 0.0 0.0 5.0
assign ( resid  24 and name  hd2 ) ( resid  24 and name   hd1 ) 0.0 0.0 2.8
assign ( resid  24 and name  hd2 ) ( resid  24 and name   hg2 ) 0.0 0.0 3.5
assign ( resid  24 and name  hd2 ) ( resid  24 and name   hd1 ) 0.0 0.0 2.8
assign ( resid  24 and name  hd2 ) ( resid  24 and name   hg2 ) 0.0 0.0 3.5
assign ( resid  25 and name   hn ) ( resid   4 and name    ha ) 0.0 0.0 2.8
assign ( resid  25 and name   hn ) ( resid   4 and name   hb1 ) 0.0 0.0 5.0
assign ( resid  25 and name   hn ) ( resid   4 and name   hb2 ) 0.0 0.0 5.0
assign ( resid  25 and name   hn ) ( resid   5 and name    hn ) 0.0 0.0 6.0
assign ( resid  25 and name   hn ) ( resid  24 and name   hb2 ) 0.0 0.0 5.0
assign ( resid  25 and name   hn ) ( resid  24 and name   hd1 ) 0.0 0.0 6.0
assign ( resid  25 and name   hn ) ( resid  24 and name   hg2 ) 0.0 0.0 6.0
assign ( resid  25 and name   hn ) ( resid  25 and name   ha1 ) 0.0 0.0 3.5
assign ( resid  25 and name   hn ) ( resid  25 and name   ha2 ) 0.0 0.0 5.0
assign ( resid  25 and name   hn ) ( resid  26 and name    hn ) 0.0 0.0 5.0
assign ( resid  25 and name  ha2 ) ( resid  25 and name   ha1 ) 0.0 0.0 2.8
assign ( resid  26 and name   ha ) ( resid  26 and name   hb1 ) 0.0 0.0 2.8
assign ( resid  26 and name   hn ) ( resid   4 and name    ha ) 0.0 0.0 3.5
assign ( resid  26 and name   hn ) ( resid   7 and name   hg1 ) 0.0 0.0 5.0
assign ( resid  26 and name   hn ) ( resid  22 and name  hd22 ) 0.0 0.0 6.0
assign ( resid  26 and name   hn ) ( resid  23 and name    hn ) 0.0 0.0 5.0
assign ( resid  26 and name   hn ) ( resid  23 and name  hg2# ) 0.0 0.0 5.0
assign ( resid  26 and name   hn ) ( resid  24 and name   hb1 ) 0.0 0.0 6.0
assign ( resid  26 and name   hn ) ( resid  24 and name   hb2 ) 0.0 0.0 6.0
assign ( resid  26 and name   hn ) ( resid  24 and name   hd1 ) 0.0 0.0 6.0
assign ( resid  26 and name   hn ) ( resid  25 and name    hn ) 0.0 0.0 3.5
assign ( resid  26 and name   hn ) ( resid  25 and name   ha1 ) 0.0 0.0 3.5
assign ( resid  26 and name   hn ) ( resid  25 and name   ha2 ) 0.0 0.0 3.5
assign ( resid  26 and name   hn ) ( resid  26 and name   hb2 ) 0.0 0.0 3.5
assign ( resid  26 and name   hn ) ( resid  27 and name    ha ) 0.0 0.0 5.0
assign ( resid  26 and name   hn ) ( resid  27 and name    hn ) 0.0 0.0 6.0
assign ( resid  26 and name   hn ) ( resid  27 and name   hb1 ) 0.0 0.0 6.0
assign ( resid  26 and name  hb2 ) ( resid  26 and name   hb1 ) 0.0 0.0 2.8
assign ( resid  27 and name   ha ) ( resid  22 and name    ha ) 0.0 0.0 3.5
assign ( resid  27 and name   ha ) ( resid  23 and name  hg2# ) 0.0 0.0 3.5
assign ( resid  27 and name   ha ) ( resid  27 and name   hb2 ) 0.0 0.0 2.8
```

```
assign ( resid  27  and name   ha  ) ( resid  28  and name   ha  ) 0.0 0.0 5.0
assign ( resid  27  and name   ha  ) ( resid  28  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid   7  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  27  and name   hn  ) ( resid  23  and name   hn  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  23  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  26  and name   hn  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  26  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  27  and name   ha  ) 0.0 0.0 3.5
assign ( resid  27  and name   hn  ) ( resid  27  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  27  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  27  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  27  and name  he#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  28  and name   ha  ) 0.0 0.0 6.0
assign ( resid  27  and name   hn  ) ( resid  28  and name   hn  ) 0.0 0.0 5.0
assign ( resid  27  and name   hn  ) ( resid  28  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  27  and name   hz  ) ( resid  14  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hz  ) ( resid  14  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hz  ) ( resid  27  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  27  and name   hz  ) ( resid  27  and name  he#  ) 0.0 0.0 2.8
assign ( resid  27  and name   hz  ) ( resid  43  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  27  and name   hz  ) ( resid  43  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  27  and name   hz  ) ( resid  43  and name  he#  ) 0.0 0.0 6.0
assign ( resid  27  and name   hz  ) ( resid  43  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  27  and name  hb1  ) ( resid   7  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name  hb1  ) ( resid   7  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  27  and name  hb1  ) ( resid  26  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name  hb2  ) ( resid   7  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name  hb2  ) ( resid   7  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  27  and name  hb2  ) ( resid  27  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  27  and name  hd#  ) ( resid   7  and name   ha  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid   7  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  27  and name  hd#  ) ( resid   7  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid   7  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  27  and name  hd#  ) ( resid   7  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid  14  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid  14  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid  22  and name   ha  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid  27  and name   ha  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid  27  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  27  and name  hd#  ) ( resid  27  and name  hb2  ) 0.0 0.0 2.8
assign ( resid  27  and name  hd#  ) ( resid  28  and name   ha  ) 0.0 0.0 5.0
assign ( resid  27  and name  hd#  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid   7  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid   7  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  14  and name   hg  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  14  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  14  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  14  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  27  and name  he#  ) ( resid  14  and name  hd2#  ) 0.0 0.0 3.5
assign ( resid  27  and name  he#  ) ( resid  27  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  27  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  27  and name  hd#  ) 0.0 0.0 2.8
assign ( resid  27  and name  he#  ) ( resid  28  and name   ha  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  43  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  27  and name  he#  ) ( resid  43  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  27  and name  he#  ) ( resid  43  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  28  and name   ha  ) ( resid  28  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  28  and name   ha  ) ( resid  28  and name  hg#  ) 0.0 0.0 3.5
assign ( resid  28  and name   ha  ) ( resid  43  and name  hd2  ) 0.0 0.0 5.0
assign ( resid  28  and name   ha  ) ( resid  43  and name  he1  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid   7  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  28  and name   hn  ) ( resid  21  and name   ha  ) 0.0 0.0 6.0
assign ( resid  28  and name   hn  ) ( resid  22  and name   ha  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid  23  and name   hn  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid  23  and name  hg2#  ) 0.0 0.0 5.0
```

```
assign ( resid  28  and name   hn  ) ( resid  26  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  28  and name   hn  ) ( resid  27  and name    ha  ) 0.0 0.0 3.5
assign ( resid  28  and name   hn  ) ( resid  27  and name    hn  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid  27  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid  27  and name   hd#  ) 0.0 0.0 3.5
assign ( resid  28  and name   hn  ) ( resid  27  and name   he#  ) 0.0 0.0 6.0
assign ( resid  28  and name   hn  ) ( resid  28  and name    ha  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid  29  and name    hn  ) 0.0 0.0 5.0
assign ( resid  28  and name   hn  ) ( resid  29  and name   hb#  ) 0.0 0.0 6.0
assign ( resid  28  and name  hb1  ) ( resid  23  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  28  and name  hb2  ) ( resid  23  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  28  and name  hb2  ) ( resid  28  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  28  and name  hg#  ) ( resid  23  and name  hg2#  ) 0.0 0.0 3.5
assign ( resid  28  and name  hg#  ) ( resid  28  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  29  and name   ha  ) ( resid  20  and name    ha  ) 0.0 0.0 2.8
assign ( resid  29  and name   ha  ) ( resid  28  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  29  and name   ha  ) ( resid  28  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  29  and name   ha  ) ( resid  29  and name   hb#  ) 0.0 0.0 2.8
assign ( resid  29  and name   hn  ) ( resid  20  and name    ha  ) 0.0 0.0 5.0
assign ( resid  29  and name   hn  ) ( resid  21  and name    hn  ) 0.0 0.0 5.0
assign ( resid  29  and name   hn  ) ( resid  27  and name   hd#  ) 0.0 0.0 6.0
assign ( resid  29  and name   hn  ) ( resid  27  and name   he#  ) 0.0 0.0 6.0
assign ( resid  29  and name   hn  ) ( resid  28  and name    ha  ) 0.0 0.0 3.5
assign ( resid  29  and name   hn  ) ( resid  28  and name    hn  ) 0.0 0.0 5.0
assign ( resid  29  and name   hn  ) ( resid  28  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  29  and name   hn  ) ( resid  28  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  29  and name   hn  ) ( resid  28  and name   hg#  ) 0.0 0.0 3.5
assign ( resid  29  and name   hn  ) ( resid  29  and name    ha  ) 0.0 0.0 3.5
assign ( resid  29  and name   hn  ) ( resid  29  and name   hb#  ) 0.0 0.0 3.5
assign ( resid  29  and name   hn  ) ( resid  30  and name    hn  ) 0.0 0.0 5.0
assign ( resid  29  and name   hn  ) ( resid  43  and name   hg#  ) 0.0 0.0 5.0
assign ( resid  30  and name   ha  ) ( resid  30  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  30  and name   ha  ) ( resid  30  and name   hg1  ) 0.0 0.0 5.0
assign ( resid  30  and name   ha  ) ( resid  30  and name   hg2  ) 0.0 0.0 2.8
assign ( resid  30  and name   hn  ) ( resid  18  and name   ha1  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  19  and name    hn  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  19  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  20  and name    ha  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  21  and name    hn  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  29  and name    ha  ) 0.0 0.0 3.5
assign ( resid  30  and name   hn  ) ( resid  29  and name    hn  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  29  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  30  and name    ha  ) 0.0 0.0 3.5
assign ( resid  30  and name   hn  ) ( resid  30  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  30  and name   hn  ) ( resid  30  and name   hd#  ) 0.0 0.0 6.0
assign ( resid  30  and name   hn  ) ( resid  30  and name   hg1  ) 0.0 0.0 6.0
assign ( resid  30  and name   hn  ) ( resid  30  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  30  and name   hn  ) ( resid  31  and name    hn  ) 0.0 0.0 6.0
assign ( resid  30  and name  hb2  ) ( resid  30  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  30  and name  hb2  ) ( resid  30  and name   hd#  ) 0.0 0.0 3.5
assign ( resid  30  and name  hb2  ) ( resid  30  and name   hg1  ) 0.0 0.0 3.5
assign ( resid  30  and name  hb2  ) ( resid  30  and name   hg2  ) 0.0 0.0 3.5
assign ( resid  30  and name  hd#  ) ( resid  30  and name   hg2  ) 0.0 0.0 2.8
assign ( resid  31  and name   ha  ) ( resid  18  and name   ha2  ) 0.0 0.0 3.5
assign ( resid  31  and name   ha  ) ( resid  35  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  31  and name   hn  ) ( resid  18  and name   ha1  ) 0.0 0.0 6.0
assign ( resid  31  and name   hn  ) ( resid  19  and name   hb2  ) 0.0 0.0 6.0
assign ( resid  31  and name   hn  ) ( resid  30  and name    hn  ) 0.0 0.0 5.0
assign ( resid  31  and name   hn  ) ( resid  30  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  31  and name   hn  ) ( resid  30  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  31  and name   hn  ) ( resid  30  and name   hg1  ) 0.0 0.0 6.0
assign ( resid  31  and name   hn  ) ( resid  30  and name   hg2  ) 0.0 0.0 6.0
assign ( resid  31  and name   hn  ) ( resid  31  and name    ha  ) 0.0 0.0 6.0
```

```
assign ( resid  31  and name  hn  ) ( resid  31  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  31  and name  hn  ) ( resid  32  and name  hn  ) 0.0 0.0 5.0
assign ( resid  31  and name  hn  ) ( resid  35  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  32  and name  ha  ) ( resid  32  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  32  and name  ha  ) ( resid  32  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  32  and name  hn  ) ( resid  17  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  17  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  32  and name  hn  ) ( resid  17  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  18  and name  hn  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  18  and name  ha1  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  18  and name  ha2  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  31  and name  ha  ) 0.0 0.0 3.5
assign ( resid  32  and name  hn  ) ( resid  31  and name  hn  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  31  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  31  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  32  and name  hn  ) ( resid  32  and name  ha  ) 0.0 0.0 3.5
assign ( resid  32  and name  hn  ) ( resid  33  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  32  and name  hn  ) ( resid  35  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  35  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  32  and name  hn  ) ( resid  35  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  32  and name  hb2  ) ( resid  32  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  33  and name  ha  ) ( resid  33  and name  hb#  ) 0.0 0.0 2.8
assign ( resid  33  and name  ha  ) ( resid  33  and name  hg#  ) 0.0 0.0 2.8
assign ( resid  33  and name  hn  ) ( resid  32  and name  ha  ) 0.0 0.0 3.5
assign ( resid  33  and name  hn  ) ( resid  32  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  33  and name  hn  ) ( resid  33  and name  ha  ) 0.0 0.0 3.5
assign ( resid  33  and name  hn  ) ( resid  33  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  33  and name  hn  ) ( resid  35  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  34  and name  hn  ) ( resid  33  and name  ha  ) 0.0 0.0 3.5
assign ( resid  34  and name  hn  ) ( resid  33  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  34  and name  hn  ) ( resid  34  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  34  and name  hn  ) ( resid  34  and name  ha2  ) 0.0 0.0 3.5
assign ( resid  34  and name  hn  ) ( resid  35  and name  ha  ) 0.0 0.0 6.0
assign ( resid  34  and name  hn  ) ( resid  35  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  34  and name  hn  ) ( resid  47  and name  hn  ) 0.0 0.0 6.0
assign ( resid  34  and name  ha2  ) ( resid  34  and name  ha1  ) 0.0 0.0 2.8
assign ( resid  34  and name  ha2  ) ( resid  47  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  34  and name  ha2  ) ( resid  48  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  35  and name  ha  ) ( resid  35  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  35  and name  ha  ) ( resid  35  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  35  and name  ha  ) ( resid  47  and name  ha  ) 0.0 0.0 3.5
assign ( resid  35  and name  hn  ) ( resid  17  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  35  and name  hn  ) ( resid  31  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  35  and name  hn  ) ( resid  33  and name  ha  ) 0.0 0.0 5.0
assign ( resid  35  and name  hn  ) ( resid  34  and name  hn  ) 0.0 0.0 3.5
assign ( resid  35  and name  hn  ) ( resid  34  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  35  and name  hn  ) ( resid  34  and name  ha2  ) 0.0 0.0 3.5
assign ( resid  35  and name  hn  ) ( resid  35  and name  ha  ) 0.0 0.0 3.5
assign ( resid  35  and name  hn  ) ( resid  35  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  35  and name  hn  ) ( resid  35  and name  hd#  ) 0.0 0.0 3.5
assign ( resid  35  and name  hn  ) ( resid  36  and name  hn  ) 0.0 0.0 5.0
assign ( resid  35  and name  hn  ) ( resid  46  and name  hn  ) 0.0 0.0 6.0
assign ( resid  35  and name  hn  ) ( resid  47  and name  ha  ) 0.0 0.0 5.0
assign ( resid  35  and name  hn  ) ( resid  47  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  35  and name  hb2  ) ( resid  31  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  35  and name  hb2  ) ( resid  33  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  35  and name  hb2  ) ( resid  35  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  35  and name  hd#  ) ( resid  17  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  35  and name  hd#  ) ( resid  17  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  35  and name  hd#  ) ( resid  17  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  35  and name  hd#  ) ( resid  31  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  35  and name  hd#  ) ( resid  32  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  35  and name  hd#  ) ( resid  33  and name  hb#  ) 0.0 0.0 5.0
```

```
assign ( resid  35  and name  hd# ) ( resid  35  and name   ha ) 0.0 0.0 3.5
assign ( resid  35  and name  hd# ) ( resid  35  and name  hb2 ) 0.0 0.0 3.5
assign ( resid  35  and name  hd# ) ( resid  46  and name   ha ) 0.0 0.0 3.5
assign ( resid  35  and name  hd# ) ( resid  47  and name   ha ) 0.0 0.0 3.5
assign ( resid  35  and name  hd# ) ( resid  68  and name  ha1 ) 0.0 0.0 5.0
assign ( resid  35  and name  hd# ) ( resid  68  and name  ha2 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  17  and name  hb1 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  17  and name  hb2 ) 0.0 0.0 3.5
assign ( resid  35  and name  he# ) ( resid  17  and name  hd# ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  17  and name  hd1 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  17  and name  hg# ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  32  and name  hb1 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  32  and name  hb2 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  35  and name   ha ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  35  and name  hb2 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  46  and name   ha ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  47  and name   ha ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  47  and name  hb1 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  67  and name   ha ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  67  and name  hb1 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  67  and name  hb2 ) 0.0 0.0 5.0
assign ( resid  35  and name  he# ) ( resid  67  and name  hg# ) 0.0 0.0 3.5
assign ( resid  35  and name  he# ) ( resid  68  and name  ha1 ) 0.0 0.0 3.5
assign ( resid  35  and name  he# ) ( resid  68  and name  ha2 ) 0.0 0.0 3.5
assign ( resid  36  and name   ha ) ( resid  36  and name  hb1 ) 0.0 0.0 2.8
assign ( resid  36  and name   ha ) ( resid  36  and name  hb2 ) 0.0 0.0 3.5
assign ( resid  36  and name   ha ) ( resid  36  and name  hg1 ) 0.0 0.0 5.0
assign ( resid  36  and name   ha ) ( resid  36  and name  hg2 ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  31  and name  hb2 ) 0.0 0.0 6.0
assign ( resid  36  and name   hn ) ( resid  35  and name   ha ) 0.0 0.0 3.5
assign ( resid  36  and name   hn ) ( resid  35  and name   hn ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  35  and name  hb2 ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  35  and name  hd# ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  36  and name   ha ) 0.0 0.0 3.5
assign ( resid  36  and name   hn ) ( resid  36  and name  hb1 ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  36  and name  hb2 ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  37  and name   hn ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  45  and name   ha ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  45  and name  hb# ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  46  and name   ha ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  46  and name   hn ) 0.0 0.0 3.5
assign ( resid  36  and name   hn ) ( resid  47  and name   ha ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  48  and name   ha ) 0.0 0.0 6.0
assign ( resid  36  and name   hn ) ( resid  48  and name   hn ) 0.0 0.0 5.0
assign ( resid  36  and name   hn ) ( resid  48  and name  hd1# ) 0.0 0.0 6.0
assign ( resid  36  and name  hb1 ) ( resid  48  and name  hd1# ) 0.0 0.0 5.0
assign ( resid  36  and name  hb2 ) ( resid  48  and name  hg1# ) 0.0 0.0 5.0
assign ( resid  36  and name  hg2 ) ( resid  36  and name  hg1 ) 0.0 0.0 2.8
assign ( resid  37  and name   hn ) ( resid  31  and name  hb1 ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  36  and name   ha ) 0.0 0.0 2.8
assign ( resid  37  and name   hn ) ( resid  36  and name   hn ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  36  and name  hb# ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  36  and name  hg# ) 0.0 0.0 3.5
assign ( resid  37  and name   hn ) ( resid  37  and name   ha ) 0.0 0.0 2.8
assign ( resid  37  and name   hn ) ( resid  37  and name  hb1 ) 0.0 0.0 3.5
assign ( resid  37  and name   hn ) ( resid  37  and name  hb2 ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  38  and name   hn ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  38  and name  ha2 ) 0.0 0.0 6.0
assign ( resid  37  and name   hn ) ( resid  39  and name  hb# ) 0.0 0.0 6.0
assign ( resid  37  and name   hn ) ( resid  43  and name  hd# ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  43  and name  he# ) 0.0 0.0 5.0
assign ( resid  37  and name   hn ) ( resid  46  and name   hn ) 0.0 0.0 5.0
assign ( resid  37  and name  hb2 ) ( resid  29  and name  hb# ) 0.0 0.0 5.0
```

```
assign ( resid  37  and name  hb2  ) ( resid  37  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  38  and name  hn  ) ( resid  37  and name  ha  ) 0.0 0.0 2.8
assign ( resid  38  and name  hn  ) ( resid  37  and name  hn  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  38  and name  ha2  ) 0.0 0.0 2.8
assign ( resid  38  and name  hn  ) ( resid  39  and name  hn  ) 0.0 0.0 3.5
assign ( resid  38  and name  hn  ) ( resid  39  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  39  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  38  and name  hn  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  43  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  43  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  45  and name  ha  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  46  and name  hn  ) 0.0 0.0 5.0
assign ( resid  38  and name  hn  ) ( resid  46  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  38  and name  hn  ) ( resid  46  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  39  and name  ha  ) ( resid  39  and name  hb#  ) 0.0 0.0 2.8
assign ( resid  39  and name  hn  ) ( resid  37  and name  ha  ) 0.0 0.0 5.0
assign ( resid  39  and name  hn  ) ( resid  38  and name  hn  ) 0.0 0.0 3.5
assign ( resid  39  and name  hn  ) ( resid  38  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  39  and name  hn  ) ( resid  38  and name  ha2  ) 0.0 0.0 3.5
assign ( resid  39  and name  hn  ) ( resid  39  and name  ha  ) 0.0 0.0 3.5
assign ( resid  39  and name  hn  ) ( resid  39  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  39  and name  hn  ) ( resid  39  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  39  and name  hn  ) ( resid  40  and name  hn  ) 0.0 0.0 2.8
assign ( resid  39  and name  hn  ) ( resid  40  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  39  and name  hn  ) ( resid  41  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  39  and name  hn  ) ( resid  41  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  39  and name  hn  ) ( resid  46  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  39  and name  hz  ) ( resid  40  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  39  and name  hd#  ) ( resid  39  and name  ha  ) 0.0 0.0 5.0
assign ( resid  39  and name  hd#  ) ( resid  39  and name  hb#  ) 0.0 0.0 2.8
assign ( resid  39  and name  hd#  ) ( resid  40  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  39  and name  hd#  ) ( resid  40  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  39  and name  hd#  ) ( resid  40  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  39  and name  he#  ) ( resid  39  and name  ha  ) 0.0 0.0 6.0
assign ( resid  39  and name  hd#  ) ( resid  40  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  39  and name  he#  ) ( resid  39  and name  ha  ) 0.0 0.0 6.0
assign ( resid  39  and name  he#  ) ( resid  39  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  39  and name  he#  ) ( resid  40  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  39  and name  he#  ) ( resid  40  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  39  and name  he#  ) ( resid  40  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  40  and name  ha  ) ( resid  40  and name  hg1  ) 0.0 0.0 3.5
assign ( resid  40  and name  hn  ) ( resid  39  and name  ha  ) 0.0 0.0 3.5
assign ( resid  40  and name  hn  ) ( resid  39  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  40  and name  hn  ) ( resid  39  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  40  and name  hn  ) ( resid  40  and name  ha  ) 0.0 0.0 2.8
assign ( resid  40  and name  hn  ) ( resid  40  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  40  and name  hn  ) ( resid  40  and name  hg#  ) 0.0 0.0 3.5
assign ( resid  40  and name  hn  ) ( resid  41  and name  hn  ) 0.0 0.0 3.5
assign ( resid  40  and name  hn  ) ( resid  41  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  40  and name  hn  ) ( resid  41  and name  hg1  ) 0.0 0.0 6.0
assign ( resid  41  and name  ha  ) ( resid  41  and name  hb#  ) 0.0 0.0 2.8
assign ( resid  41  and name  ha  ) ( resid  41  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  41  and name  ha  ) ( resid  41  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  38  and name  ha1  ) 0.0 0.0 6.0
assign ( resid  41  and name  hn  ) ( resid  39  and name  ha  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  39  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  41  and name  hn  ) ( resid  40  and name  hn  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  40  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  40  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  41  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  41  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  41  and name  hn  ) ( resid  41  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  41  and name  hg1  ) ( resid  41  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  41  and name  hg2  ) ( resid  41  and name  hg1  ) 0.0 0.0 2.8
assign ( resid  42  and name  ha  ) ( resid  42  and name  hb#  ) 0.0 0.0 2.8
```

```
assign ( resid  42  and name   ha  ) ( resid  42  and name  hg2  ) 0.0 0.0 3.5
assign ( resid  42  and name   hn  ) ( resid  38  and name  ha1  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  39  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  40  and name   hn  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  40  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  40  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  41  and name   ha  ) 0.0 0.0 3.5
assign ( resid  42  and name   hn  ) ( resid  42  and name   ha  ) 0.0 0.0 3.5
assign ( resid  42  and name   hn  ) ( resid  42  and name  hg2  ) 0.0 0.0 3.5
assign ( resid  42  and name   hn  ) ( resid  43  and name   hn  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  42  and name   hn  ) ( resid  43  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  42  and name   hn  ) ( resid  43  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  42  and name   hn  ) ( resid  44  and name   hn  ) 0.0 0.0 5.0
assign ( resid  42  and name  hg2  ) ( resid  42  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  43  and name   ha  ) ( resid  43  and name  hb#  ) 0.0 0.0 2.8
assign ( resid  43  and name   ha  ) ( resid  43  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  43  and name   ha  ) ( resid  43  and name  hd2  ) 0.0 0.0 3.5
assign ( resid  43  and name   ha  ) ( resid  43  and name  hg#  ) 0.0 0.0 3.5
assign ( resid  43  and name   hn  ) ( resid  27  and name  he#  ) 0.0 0.0 6.0
assign ( resid  43  and name   hn  ) ( resid  42  and name   ha  ) 0.0 0.0 3.5
assign ( resid  43  and name   hn  ) ( resid  42  and name   hn  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  42  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  42  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  42  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  43  and name   ha  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  43  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  43  and name   hn  ) ( resid  43  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  43  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  43  and name   hn  ) ( resid  44  and name   hn  ) 0.0 0.0 5.0
assign ( resid  43  and name  hb#  ) ( resid  43  and name  hg#  ) 0.0 0.0 3.5
assign ( resid  43  and name  hd2  ) ( resid  43  and name  hd1  ) 0.0 0.0 2.8
assign ( resid  43  and name  hd2  ) ( resid  43  and name  hg#  ) 0.0 0.0 2.8
assign ( resid  43  and name  he1  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  43  and name  hd2  ) ( resid  43  and name  hg#  ) 0.0 0.0 2.8
assign ( resid  43  and name  he1  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  43  and name  he1  ) ( resid  43  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  43  and name  he1  ) ( resid  43  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  43  and name  he2  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  43  and name  he2  ) ( resid  43  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  43  and name  he2  ) ( resid  43  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  43  and name  hg#  ) ( resid  43  and name  hd1  ) 0.0 0.0 2.8
assign ( resid  44  and name   ha  ) ( resid  44  and name  hb2  ) 0.0 0.0 2.8
assign ( resid  44  and name   hn  ) ( resid  37  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  44  and name   hn  ) ( resid  38  and name  ha2  ) 0.0 0.0 6.0
assign ( resid  44  and name   hn  ) ( resid  42  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  44  and name   hn  ) ( resid  43  and name   ha  ) 0.0 0.0 3.5
assign ( resid  44  and name   hn  ) ( resid  43  and name   hn  ) 0.0 0.0 3.5
assign ( resid  44  and name   hn  ) ( resid  43  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  44  and name   hn  ) ( resid  43  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  44  and name   hn  ) ( resid  44  and name   ha  ) 0.0 0.0 2.8
assign ( resid  44  and name   hn  ) ( resid  44  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  44  and name   hn  ) ( resid  45  and name   hn  ) 0.0 0.0 5.0
assign ( resid  44  and name  hb2  ) ( resid  44  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  44  and name  hd21 ) ( resid  39  and name   ha  ) 0.0 0.0 6.0
assign ( resid  44  and name  hd21 ) ( resid  43  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  44  and name  hd21 ) ( resid  44  and name   ha  ) 0.0 0.0 5.0
assign ( resid  44  and name  hd21 ) ( resid  44  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  44  and name  hd21 ) ( resid  45  and name   hn  ) 0.0 0.0 6.0
assign ( resid  44  and name  hd21 ) ( resid  46  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  44  and name  hd21 ) ( resid  46  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  44  and name  hd22 ) ( resid  44  and name   ha  ) 0.0 0.0 5.0
assign ( resid  44  and name  hd22 ) ( resid  46  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  44  and name  hd22 ) ( resid  46  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  44  and name  nd21 ) ( resid  44  and name  hd21 ) 0.0 0.0 2.8
```

```
assign ( resid  44  and name  nd21  ) ( resid  44  and name  hd22  ) 0.0 0.0 2.8
assign ( resid  44  and name  nd22  ) ( resid  44  and name  hd21  ) 0.0 0.0 2.8
assign ( resid  44  and name  nd22  ) ( resid  44  and name  hd22  ) 0.0 0.0 2.8
assign ( resid  45  and name    hn  ) ( resid  16  and name   ha2  ) 0.0 0.0 6.0
assign ( resid  45  and name    hn  ) ( resid  17  and name    hn  ) 0.0 0.0 5.0
assign ( resid  45  and name    hn  ) ( resid  44  and name    hn  ) 0.0 0.0 5.0
assign ( resid  45  and name    hn  ) ( resid  46  and name    hn  ) 0.0 0.0 5.0
assign ( resid  45  and name    hn  ) ( resid  46  and name   hg#  ) 0.0 0.0 5.0
assign ( resid  45  and name   hb2  ) ( resid  45  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  46  and name    ha  ) ( resid  46  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  46  and name    ha  ) ( resid  46  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  46  and name    ha  ) ( resid  46  and name   hg#  ) 0.0 0.0 3.5
assign ( resid  46  and name    hn  ) ( resid  35  and name    ha  ) 0.0 0.0 5.0
assign ( resid  46  and name    hn  ) ( resid  35  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  46  and name    hn  ) ( resid  35  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  46  and name    hn  ) ( resid  36  and name    hn  ) 0.0 0.0 3.5
assign ( resid  46  and name    hn  ) ( resid  44  and name    hn  ) 0.0 0.0 6.0
assign ( resid  46  and name    hn  ) ( resid  45  and name    ha  ) 0.0 0.0 2.8
assign ( resid  46  and name    hn  ) ( resid  45  and name    hn  ) 0.0 0.0 5.0
assign ( resid  46  and name    hn  ) ( resid  45  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  46  and name    hn  ) ( resid  46  and name   hb#  ) 0.0 0.0 3.5
assign ( resid  46  and name    hn  ) ( resid  46  and name   hg#  ) 0.0 0.0 5.0
assign ( resid  46  and name    hn  ) ( resid  47  and name    hn  ) 0.0 0.0 6.0
assign ( resid  46  and name   hb2  ) ( resid  46  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  46  and name   hg1  ) ( resid  46  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  46  and name   hg2  ) ( resid  46  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  46  and name   hg2  ) ( resid  46  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  47  and name    ha  ) ( resid  47  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  47  and name    ha  ) ( resid  47  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  47  and name    hn  ) ( resid  35  and name   hd#  ) 0.0 0.0 3.5
assign ( resid  47  and name    hn  ) ( resid  36  and name    hn  ) 0.0 0.0 6.0
assign ( resid  47  and name    hn  ) ( resid  46  and name    ha  ) 0.0 0.0 3.5
assign ( resid  47  and name    hn  ) ( resid  46  and name    hn  ) 0.0 0.0 5.0
assign ( resid  47  and name    hn  ) ( resid  47  and name    ha  ) 0.0 0.0 5.0
assign ( resid  47  and name    hn  ) ( resid  48  and name    hn  ) 0.0 0.0 5.0
assign ( resid  48  and name    ha  ) ( resid  36  and name   hg1  ) 0.0 0.0 5.0
assign ( resid  48  and name    ha  ) ( resid  36  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  48  and name    ha  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  48  and name    ha  ) ( resid  36  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  48  and name    ha  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  34  and name   ha1  ) 0.0 0.0 6.0
assign ( resid  48  and name    hn  ) ( resid  35  and name    ha  ) 0.0 0.0 3.5
assign ( resid  48  and name    hn  ) ( resid  35  and name    hn  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  35  and name   hb2  ) 0.0 0.0 6.0
assign ( resid  48  and name    hn  ) ( resid  35  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  36  and name    hn  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  36  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  36  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  47  and name    ha  ) 0.0 0.0 3.5
assign ( resid  48  and name    hn  ) ( resid  47  and name    hn  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  47  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  47  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  48  and name    ha  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  48  and name    hn  ) ( resid  49  and name    hn  ) 0.0 0.0 5.0
assign ( resid  48  and name  hg11  ) ( resid  48  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  48  and name  hg2#  ) ( resid  48  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  49  and name    hn  ) ( resid  48  and name    ha  ) 0.0 0.0 3.5
assign ( resid  49  and name    hn  ) ( resid  48  and name    hn  ) 0.0 0.0 5.0
assign ( resid  49  and name    hn  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  49  and name    hn  ) ( resid  49  and name    ha  ) 0.0 0.0 3.5
assign ( resid  49  and name    hn  ) ( resid  49  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  49  and name    hn  ) ( resid  49  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  49  and name    hn  ) ( resid  50  and name    hn  ) 0.0 0.0 5.0
assign ( resid  49  and name    hn  ) ( resid  52  and name   hb1  ) 0.0 0.0 6.0
```

```
assign ( resid  49  and name   hn  ) ( resid  52  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  49  and name   hn  ) ( resid  53  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  49  and name  hb2  ) ( resid  52  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  49  and name  hb2  ) ( resid  52  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  50  and name   ha  ) ( resid  48  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  50  and name   ha  ) ( resid  50  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  50  and name   ha  ) ( resid  50  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  50  and name   ha  ) ( resid  50  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  50  and name   ha  ) ( resid  53  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  50  and name   ha  ) ( resid  53  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  50  and name   ha  ) ( resid  53  and name   hg#  ) 0.0 0.0 5.0
assign ( resid  50  and name   ha  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  50  and name   ha  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  48  and name   ha  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  48  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  50  and name   hn  ) ( resid  49  and name   ha  ) 0.0 0.0 2.8
assign ( resid  50  and name   hn  ) ( resid  49  and name   hn  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  49  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  49  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  50  and name   ha  ) 0.0 0.0 3.5
assign ( resid  50  and name   hn  ) ( resid  50  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  50  and name   hn  ) ( resid  50  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  50  and name   hn  ) ( resid  50  and name   hg1  ) 0.0 0.0 3.5
assign ( resid  50  and name   hn  ) ( resid  50  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  51  and name   ha  ) 0.0 0.0 6.0
assign ( resid  50  and name   hn  ) ( resid  51  and name   hn  ) 0.0 0.0 3.5
assign ( resid  50  and name   hn  ) ( resid  51  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  52  and name   hn  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  53  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  53  and name   hb2  ) 0.0 0.0 6.0
assign ( resid  50  and name   hn  ) ( resid  57  and name  hd1#  ) 0.0 0.0 6.0
assign ( resid  50  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 6.0
assign ( resid  50  and name   hn  ) ( resid  65  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  65  and name  hd21  ) 0.0 0.0 5.0
assign ( resid  50  and name   hn  ) ( resid  65  and name  hd22  ) 0.0 0.0 5.0
assign ( resid  50  and name  hb1  ) ( resid  57  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  50  and name  hb1  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  50  and name  hb2  ) ( resid  50  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  50  and name  hb2  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  50  and name  hg1  ) ( resid  50  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  50  and name  hg2  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  50  and name  hg2  ) ( resid  50  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  50  and name  hg2  ) ( resid  50  and name   hg1  ) 0.0 0.0 2.8
assign ( resid  50  and name  hg2  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  51  and name   ha  ) ( resid  51  and name   hb#  ) 0.0 0.0 2.8
assign ( resid  51  and name   ha  ) ( resid  57  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  51  and name   ha  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  49  and name   hn  ) 0.0 0.0 6.0
assign ( resid  51  and name   hn  ) ( resid  49  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  50  and name   ha  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  50  and name   hn  ) 0.0 0.0 3.5
assign ( resid  51  and name   hn  ) ( resid  50  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  51  and name   hn  ) ( resid  50  and name   hg1  ) 0.0 0.0 6.0
assign ( resid  51  and name   hn  ) ( resid  50  and name   hg2  ) 0.0 0.0 6.0
assign ( resid  51  and name   hn  ) ( resid  51  and name   ha  ) 0.0 0.0 3.5
assign ( resid  51  and name   hn  ) ( resid  51  and name   hb#  ) 0.0 0.0 3.5
assign ( resid  51  and name   hn  ) ( resid  52  and name   ha  ) 0.0 0.0 6.0
assign ( resid  51  and name   hn  ) ( resid  52  and name   hn  ) 0.0 0.0 3.5
assign ( resid  51  and name   hn  ) ( resid  52  and name   hg1  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  53  and name   hn  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
```

```
assign ( resid  51  and name   hn  ) ( resid  65  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  65  and name  hd21  ) 0.0 0.0 5.0
assign ( resid  51  and name   hn  ) ( resid  65  and name  hd22  ) 0.0 0.0 5.0
assign ( resid  51  and name  hb#  ) ( resid  63  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  52  and name   ha  ) ( resid  52  and name   hg1  ) 0.0 0.0 5.0
assign ( resid  52  and name   ha  ) ( resid  52  and name   hg2  ) 0.0 0.0 3.5
assign ( resid  52  and name   hn  ) ( resid  48  and name  hd1#  ) 0.0 0.0 6.0
assign ( resid  52  and name   hn  ) ( resid  49  and name   ha  ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  49  and name   hn  ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  49  and name   hb2 ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  50  and name   ha  ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  50  and name   hn  ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  50  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  51  and name   ha  ) 0.0 0.0 3.5
assign ( resid  52  and name   hn  ) ( resid  51  and name   hn  ) 0.0 0.0 3.5
assign ( resid  52  and name   hn  ) ( resid  51  and name   hb# ) 0.0 0.0 3.5
assign ( resid  52  and name   hn  ) ( resid  52  and name   ha  ) 0.0 0.0 3.5
assign ( resid  52  and name   hn  ) ( resid  52  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  52  and name   hg1 ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  52  and name   hg2 ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  53  and name   hn  ) 0.0 0.0 3.5
assign ( resid  52  and name   hn  ) ( resid  53  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  52  and name   hn  ) ( resid  65  and name  hd21  ) 0.0 0.0 5.0
assign ( resid  52  and name  hb2  ) ( resid  52  and name   hb1 ) 0.0 0.0 2.8
assign ( resid  52  and name  hg1  ) ( resid  52  and name   hb1 ) 0.0 0.0 3.5
assign ( resid  52  and name  hg2  ) ( resid  52  and name   hb1 ) 0.0 0.0 3.5
assign ( resid  52  and name  hg2  ) ( resid  52  and name   hg1 ) 0.0 0.0 2.8
assign ( resid  52  and name  he21 ) ( resid  49  and name   hb2 ) 0.0 0.0 6.0
assign ( resid  52  and name  he21 ) ( resid  52  and name   ha  ) 0.0 0.0 6.0
assign ( resid  52  and name  he21 ) ( resid  52  and name   hg2 ) 0.0 0.0 5.0
assign ( resid  52  and name  he22 ) ( resid  49  and name   hb2 ) 0.0 0.0 6.0
assign ( resid  52  and name  he22 ) ( resid  52  and name   hg1 ) 0.0 0.0 6.0
assign ( resid  52  and name  he22 ) ( resid  52  and name   hg2 ) 0.0 0.0 6.0
assign ( resid  52  and name  ne21 ) ( resid  52  and name  he21  ) 0.0 0.0 2.8
assign ( resid  52  and name  ne21 ) ( resid  52  and name  he22  ) 0.0 0.0 2.8
assign ( resid  52  and name  ne22 ) ( resid  52  and name  he21  ) 0.0 0.0 2.8
assign ( resid  52  and name  ne22 ) ( resid  52  and name  he22  ) 0.0 0.0 2.8
assign ( resid  53  and name   ha  ) ( resid  53  and name   hd1 ) 0.0 0.0 5.0
assign ( resid  53  and name   ha  ) ( resid  53  and name   hd2 ) 0.0 0.0 5.0
assign ( resid  53  and name   ha  ) ( resid  53  and name   hg# ) 0.0 0.0 3.5
assign ( resid  53  and name   hn  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  50  and name   ha  ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  50  and name   hn  ) 0.0 0.0 6.0
assign ( resid  53  and name   hn  ) ( resid  50  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  51  and name   ha  ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  52  and name   hn  ) 0.0 0.0 3.5
assign ( resid  53  and name   hn  ) ( resid  52  and name   hb2 ) 0.0 0.0 3.5
assign ( resid  53  and name   hn  ) ( resid  52  and name   hg1 ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  52  and name   hg2 ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  53  and name   hd1 ) 0.0 0.0 5.0
assign ( resid  53  and name   hn  ) ( resid  54  and name   ha  ) 0.0 0.0 6.0
assign ( resid  53  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 6.0
assign ( resid  53  and name  hb1  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  53  and name  hb1  ) ( resid  53  and name   hg# ) 0.0 0.0 2.8
assign ( resid  53  and name  hb2  ) ( resid  53  and name   hg# ) 0.0 0.0 3.5
assign ( resid  53  and name  hd1  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  53  and name  hd1  ) ( resid  53  and name   hb2 ) 0.0 0.0 3.5
assign ( resid  53  and name  hd1  ) ( resid  53  and name   hg# ) 0.0 0.0 5.0
assign ( resid  53  and name  hd2  ) ( resid  48  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  53  and name  hd2  ) ( resid  53  and name   hb1 ) 0.0 0.0 3.5
assign ( resid  53  and name  hd2  ) ( resid  53  and name   hd1 ) 0.0 0.0 2.8
assign ( resid  53  and name  hd2  ) ( resid  53  and name   hg# ) 0.0 0.0 5.0
assign ( resid  54  and name   ha  ) ( resid  54  and name   hb1 ) 0.0 0.0 3.5
```

```
assign ( resid  54  and name   ha  ) ( resid  54  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  54  and name   ha  ) ( resid  55  and name   hd1  ) 0.0 0.0 3.5
assign ( resid  54  and name   ha  ) ( resid  55  and name   hd2  ) 0.0 0.0 3.5
assign ( resid  54  and name   ha  ) ( resid  55  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  50  and name    ha  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  51  and name    ha  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  52  and name    hn  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  52  and name   hb2  ) 0.0 0.0 6.0
assign ( resid  54  and name   hn  ) ( resid  53  and name    ha  ) 0.0 0.0 3.5
assign ( resid  54  and name   hn  ) ( resid  53  and name   hd2  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  54  and name    ha  ) 0.0 0.0 3.5
assign ( resid  54  and name   hn  ) ( resid  54  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  55  and name    ha  ) 0.0 0.0 6.0
assign ( resid  54  and name   hn  ) ( resid  55  and name   hd1  ) 0.0 0.0 6.0
assign ( resid  54  and name   hn  ) ( resid  55  and name   hg#  ) 0.0 0.0 6.0
assign ( resid  54  and name   hn  ) ( resid  56  and name    hn  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  57  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  54  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  54  and name  hb1  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  54  and name  hb1  ) ( resid  57  and name  hd2#  ) 0.0 0.0 6.0
assign ( resid  54  and name  hb2  ) ( resid  54  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  54  and name  hb2  ) ( resid  57  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  54  and name  hb2  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  54  and name  hb2  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  55  and name   ha  ) ( resid  55  and name   hd1  ) 0.0 0.0 5.0
assign ( resid  55  and name   ha  ) ( resid  55  and name   hd2  ) 0.0 0.0 5.0
assign ( resid  55  and name  hd1  ) ( resid  55  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  55  and name  hd1  ) ( resid  55  and name   hg2  ) 0.0 0.0 3.5
assign ( resid  55  and name  hd2  ) ( resid  55  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  55  and name  hd2  ) ( resid  55  and name   hd1  ) 0.0 0.0 2.8
assign ( resid  56  and name   ha  ) ( resid  56  and name    hg  ) 0.0 0.0 3.5
assign ( resid  56  and name   ha  ) ( resid  59  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  56  and name   ha  ) ( resid  59  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  56  and name   hg  ) ( resid  56  and name  hd2#  ) 0.0 0.0 2.8
assign ( resid  56  and name   hn  ) ( resid  54  and name    hn  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  54  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  55  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  55  and name   hd1  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  56  and name    ha  ) 0.0 0.0 3.5
assign ( resid  56  and name   hn  ) ( resid  56  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  56  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  57  and name    hn  ) 0.0 0.0 2.8
assign ( resid  56  and name   hn  ) ( resid  57  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  56  and name   hn  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 6.0
assign ( resid  56  and name   hn  ) ( resid  59  and name   hg1  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  85  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  56  and name   hn  ) ( resid  86  and name    ha  ) 0.0 0.0 6.0
assign ( resid  56  and name  hb#  ) ( resid  56  and name    hg  ) 0.0 0.0 2.8
assign ( resid  56  and name  hb#  ) ( resid  56  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  56  and name  hb#  ) ( resid  56  and name  hd2#  ) 0.0 0.0 2.8
assign ( resid  57  and name   ha  ) ( resid  57  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  57  and name   ha  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  57  and name   ha  ) ( resid  57  and name  hd2#  ) 0.0 0.0 3.5
assign ( resid  57  and name   hn  ) ( resid  54  and name    ha  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  54  and name    hn  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  54  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  54  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  57  and name   hn  ) ( resid  55  and name    ha  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  55  and name   hd1  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  56  and name    ha  ) 0.0 0.0 3.5
assign ( resid  57  and name   hn  ) ( resid  56  and name    hn  ) 0.0 0.0 2.8
```

```
assign ( resid  57  and name   hn  ) ( resid  56  and name   hd1# ) 0.0 0.0 6.0
assign ( resid  57  and name   hn  ) ( resid  56  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  57  and name    ha  ) 0.0 0.0 3.5
assign ( resid  57  and name   hn  ) ( resid  57  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  57  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  57  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  58  and name   hn  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  59  and name   hn  ) 0.0 0.0 6.0
assign ( resid  57  and name   hn  ) ( resid  85  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  57  and name   hn  ) ( resid  85  and name   hb2  ) 0.0 0.0 3.5
assign ( resid  57  and name   hb1 ) ( resid  57  and name   hd1# ) 0.0 0.0 3.5
assign ( resid  57  and name   hb1 ) ( resid  57  and name   hd2# ) 0.0 0.0 3.5
assign ( resid  57  and name   hd2# ) ( resid  57  and name   hd1# ) 0.0 0.0 2.8
assign ( resid  58  and name   ha  ) ( resid  58  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  57  and name    ha  ) 0.0 0.0 3.5
assign ( resid  58  and name   hn  ) ( resid  57  and name   hn  ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  57  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  57  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  57  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  58  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  59  and name   hn  ) 0.0 0.0 3.5
assign ( resid  58  and name   hn  ) ( resid  59  and name   hb2  ) 0.0 0.0 6.0
assign ( resid  58  and name   hn  ) ( resid  85  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  58  and name   hn  ) ( resid  86  and name    ha  ) 0.0 0.0 5.0
assign ( resid  59  and name   ha  ) ( resid  59  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  57  and name   hn  ) 0.0 0.0 6.0
assign ( resid  59  and name   hn  ) ( resid  57  and name   hd1# ) 0.0 0.0 6.0
assign ( resid  59  and name   hn  ) ( resid  57  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  58  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  59  and name    ha  ) 0.0 0.0 2.8
assign ( resid  59  and name   hn  ) ( resid  59  and name   hb2  ) 0.0 0.0 2.8
assign ( resid  59  and name   hn  ) ( resid  59  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  61  and name   hn  ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  80  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  80  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  59  and name   hn  ) ( resid  86  and name    ha  ) 0.0 0.0 3.5
assign ( resid  59  and name   hn  ) ( resid  86  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  59  and name   hd1 ) ( resid  56  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  59  and name   hd1 ) ( resid  56  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  59  and name   hd1 ) ( resid  59  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  59  and name   hd1 ) ( resid  59  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  59  and name   hd1 ) ( resid  59  and name   hg1  ) 0.0 0.0 3.5
assign ( resid  59  and name   hd1 ) ( resid  59  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  59  and name   hd2 ) ( resid  56  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  59  and name   hd2 ) ( resid  56  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  59  and name   hd2 ) ( resid  59  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  59  and name   hd2 ) ( resid  59  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  59  and name   hd2 ) ( resid  59  and name   hg2  ) 0.0 0.0 5.0
assign ( resid  60  and name   hn  ) ( resid  59  and name    ha  ) 0.0 0.0 3.5
assign ( resid  60  and name   hn  ) ( resid  59  and name   hn  ) 0.0 0.0 5.0
assign ( resid  60  and name   hn  ) ( resid  59  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  60  and name   hn  ) ( resid  59  and name   hd#  ) 0.0 0.0 5.0
assign ( resid  60  and name   hn  ) ( resid  60  and name   ha1  ) 0.0 0.0 3.5
assign ( resid  60  and name   hn  ) ( resid  60  and name   ha2  ) 0.0 0.0 3.5
assign ( resid  60  and name   hn  ) ( resid  61  and name   hn  ) 0.0 0.0 3.5
assign ( resid  60  and name   hn  ) ( resid  87  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  60  and name   ha1 ) ( resid  87  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  60  and name   ha2 ) ( resid  60  and name   ha1  ) 0.0 0.0 2.8
assign ( resid  60  and name   ha2 ) ( resid  87  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  61  and name   hn  ) ( resid  59  and name   hd#  ) 0.0 0.0 6.0
assign ( resid  61  and name   hn  ) ( resid  60  and name   hn  ) 0.0 0.0 5.0
assign ( resid  61  and name   hn  ) ( resid  60  and name   ha1  ) 0.0 0.0 5.0
assign ( resid  61  and name   hn  ) ( resid  60  and name   ha2  ) 0.0 0.0 5.0
```

```
assign ( resid  61  and name   hn  ) ( resid  61  and name   ha1  ) 0.0 0.0 3.5
assign ( resid  61  and name   hn  ) ( resid  61  and name   ha2  ) 0.0 0.0 5.0
assign ( resid  61  and name   hn  ) ( resid  62  and name    hn  ) 0.0 0.0 5.0
assign ( resid  61  and name   hn  ) ( resid  87  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  61  and name   hn  ) ( resid  87  and name   hb2  ) 0.0 0.0 5.0
assign ( resid  61  and name  ha2  ) ( resid  61  and name   ha1  ) 0.0 0.0 2.8
assign ( resid  61  and name  ha2  ) ( resid  62  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  62  and name   ha  ) ( resid  62  and name    hb  ) 0.0 0.0 3.5
assign ( resid  62  and name   hb  ) ( resid  62  and name  hg1#  ) 0.0 0.0 2.8
assign ( resid  62  and name   hb  ) ( resid  62  and name  hg2#  ) 0.0 0.0 2.8
assign ( resid  62  and name   hn  ) ( resid  58  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  61  and name    hn  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  61  and name   ha1  ) 0.0 0.0 2.8
assign ( resid  62  and name   hn  ) ( resid  61  and name   ha2  ) 0.0 0.0 3.5
assign ( resid  62  and name   hn  ) ( resid  62  and name    ha  ) 0.0 0.0 3.5
assign ( resid  62  and name   hn  ) ( resid  62  and name    hb  ) 0.0 0.0 3.5
assign ( resid  62  and name   hn  ) ( resid  62  and name  hg1#  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  63  and name    ha  ) 0.0 0.0 6.0
assign ( resid  62  and name   hn  ) ( resid  63  and name    hn  ) 0.0 0.0 6.0
assign ( resid  62  and name   hn  ) ( resid  64  and name   he1  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  72  and name    ha  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  72  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  73  and name    ha  ) 0.0 0.0 5.0
assign ( resid  62  and name   hn  ) ( resid  73  and name    hn  ) 0.0 0.0 3.5
assign ( resid  62  and name   hn  ) ( resid  74  and name    ha  ) 0.0 0.0 5.0
assign ( resid  62  and name  hg2#  ) ( resid  62  and name  hg1#  ) 0.0 0.0 2.8
assign ( resid  63  and name   ha  ) ( resid  63  and name   hb1  ) 0.0 0.0 3.5
assign ( resid  63  and name   hn  ) ( resid  61  and name   ha1  ) 0.0 0.0 6.0
assign ( resid  63  and name   hn  ) ( resid  62  and name    hb  ) 0.0 0.0 5.0
assign ( resid  63  and name   hn  ) ( resid  62  and name    hn  ) 0.0 0.0 5.0
assign ( resid  63  and name   hn  ) ( resid  62  and name  hg1#  ) 0.0 0.0 3.5
assign ( resid  63  and name   hn  ) ( resid  62  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  63  and name   hn  ) ( resid  63  and name    ha  ) 0.0 0.0 5.0
assign ( resid  63  and name   hn  ) ( resid  63  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  63  and name   hn  ) ( resid  64  and name    hn  ) 0.0 0.0 5.0
assign ( resid  63  and name   hn  ) ( resid  65  and name    hn  ) 0.0 0.0 6.0
assign ( resid  63  and name   hn  ) ( resid  72  and name    ha  ) 0.0 0.0 6.0
assign ( resid  63  and name   hn  ) ( resid  73  and name    hn  ) 0.0 0.0 6.0
assign ( resid  63  and name  hb2  ) ( resid  63  and name   hb1  ) 0.0 0.0 2.8
assign ( resid  64  and name   ha  ) ( resid  62  and name  hg1#  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  58  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  62  and name    ha  ) 0.0 0.0 6.0
assign ( resid  64  and name   hn  ) ( resid  62  and name  hg2#  ) 0.0 0.0 6.0
assign ( resid  64  and name   hn  ) ( resid  63  and name    ha  ) 0.0 0.0 2.8
assign ( resid  64  and name   hn  ) ( resid  63  and name    hn  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  63  and name   hb1  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  64  and name    ha  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  64  and name   he1  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  70  and name   hb2  ) 0.0 0.0 6.0
assign ( resid  64  and name   hn  ) ( resid  71  and name    ha  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  71  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  71  and name   hg#  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  72  and name    ha  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  73  and name    hn  ) 0.0 0.0 5.0
assign ( resid  64  and name   hn  ) ( resid  73  and name   hb1  ) 0.0 0.0 6.0
assign ( resid  64  and name  hb1  ) ( resid  71  and name   hg#  ) 0.0 0.0 5.0
assign ( resid  64  and name  he1  ) ( resid  62  and name    hb  ) 0.0 0.0 5.0
assign ( resid  64  and name  he1  ) ( resid  62  and name  hg1#  ) 0.0 0.0 3.5
assign ( resid  64  and name  he1  ) ( resid  62  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  64  and name  he1  ) ( resid  64  and name    ha  ) 0.0 0.0 6.0
assign ( resid  64  and name  he1  ) ( resid  64  and name   hb#  ) 0.0 0.0 5.0
assign ( resid  64  and name  he1  ) ( resid  64  and name   hb2  ) 0.0 0.0 5.0
```

```
assign ( resid  64  and name  he1  ) ( resid  71  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  64  and name  he1  ) ( resid  71  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  64  and name  he1  ) ( resid  73  and name  ha  ) 0.0 0.0 5.0
assign ( resid  64  and name  he1  ) ( resid  73  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  65  and name  ha  ) ( resid  50  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  65  and name  ha  ) ( resid  50  and name  hg1  ) 0.0 0.0 6.0
assign ( resid  65  and name  ha  ) ( resid  65  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  65  and name  ha  ) ( resid  66  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  65  and name  hn  ) ( resid  62  and name  hg1#  ) 0.0 0.0 6.0
assign ( resid  65  and name  hn  ) ( resid  63  and name  ha  ) 0.0 0.0 6.0
assign ( resid  65  and name  hn  ) ( resid  64  and name  hn  ) 0.0 0.0 3.5
assign ( resid  65  and name  hn  ) ( resid  65  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  65  and name  hn  ) ( resid  65  and name  hd21  ) 0.0 0.0 5.0
assign ( resid  65  and name  hn  ) ( resid  65  and name  hd22  ) 0.0 0.0 6.0
assign ( resid  65  and name  hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 6.0
assign ( resid  65  and name  hn  ) ( resid  71  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  65  and name  hb2  ) ( resid  65  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  65  and name  hd21  ) ( resid  49  and name  hn  ) 0.0 0.0 6.0
assign ( resid  65  and name  hd21  ) ( resid  49  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd21  ) ( resid  50  and name  hn  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd21  ) ( resid  50  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd21  ) ( resid  65  and name  ha  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd21  ) ( resid  65  and name  hn  ) 0.0 0.0 6.0
assign ( resid  65  and name  hd21  ) ( resid  65  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd21  ) ( resid  65  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd22  ) ( resid  49  and name  hn  ) 0.0 0.0 6.0
assign ( resid  65  and name  hd22  ) ( resid  49  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd22  ) ( resid  50  and name  hn  ) 0.0 0.0 6.0
assign ( resid  65  and name  hd22  ) ( resid  65  and name  ha  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd22  ) ( resid  65  and name  hn  ) 0.0 0.0 6.0
assign ( resid  65  and name  hd22  ) ( resid  65  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  65  and name  hd22  ) ( resid  65  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  65  and name  nd21  ) ( resid  65  and name  hd21  ) 0.0 0.0 3.5
assign ( resid  65  and name  nd21  ) ( resid  65  and name  hd22  ) 0.0 0.0 2.8
assign ( resid  65  and name  nd22  ) ( resid  65  and name  hd21  ) 0.0 0.0 3.5
assign ( resid  65  and name  nd22  ) ( resid  65  and name  hd22  ) 0.0 0.0 3.5
assign ( resid  66  and name  ha  ) ( resid  66  and name  hb  ) 0.0 0.0 3.5
assign ( resid  66  and name  ha  ) ( resid  66  and name  hg2#  ) 0.0 0.0 3.5
assign ( resid  66  and name  ha  ) ( resid  67  and name  ha  ) 0.0 0.0 5.0
assign ( resid  66  and name  ha  ) ( resid  67  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  66  and name  hb  ) ( resid  66  and name  hg2#  ) 0.0 0.0 2.8
assign ( resid  66  and name  hb  ) ( resid  71  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  66  and name  hb  ) ( resid  71  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  65  and name  ha  ) 0.0 0.0 2.8
assign ( resid  66  and name  hn  ) ( resid  65  and name  hn  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  65  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  66  and name  ha  ) 0.0 0.0 3.5
assign ( resid  66  and name  hn  ) ( resid  66  and name  hb  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 3.5
assign ( resid  66  and name  hn  ) ( resid  67  and name  hn  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  69  and name  hn  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  70  and name  ha  ) 0.0 0.0 2.8
assign ( resid  66  and name  hn  ) ( resid  70  and name  hn  ) 0.0 0.0 5.0
assign ( resid  66  and name  hn  ) ( resid  70  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  66  and name  hn  ) ( resid  71  and name  hn  ) 0.0 0.0 3.5
assign ( resid  66  and name  hn  ) ( resid  71  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  66  and name  hn  ) ( resid  71  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  67  and name  ha  ) ( resid  47  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  67  and name  ha  ) ( resid  66  and name  hg2#  ) 0.0 0.0 6.0
assign ( resid  67  and name  ha  ) ( resid  67  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  67  and name  hn  ) ( resid  66  and name  ha  ) 0.0 0.0 2.8
assign ( resid  67  and name  hn  ) ( resid  66  and name  hb  ) 0.0 0.0 6.0
assign ( resid  67  and name  hn  ) ( resid  66  and name  hn  ) 0.0 0.0 6.0
```

```
assign ( resid  67  and name  hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 6.0
assign ( resid  67  and name  hn  ) ( resid  67  and name  ha  ) 0.0 0.0 5.0
assign ( resid  67  and name  hn  ) ( resid  67  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  67  and name  hn  ) ( resid  67  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  67  and name  hn  ) ( resid  68  and name  hn  ) 0.0 0.0 6.0
assign ( resid  68  and name  hn  ) ( resid  35  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  68  and name  hn  ) ( resid  47  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  68  and name  hn  ) ( resid  67  and name  ha  ) 0.0 0.0 3.5
assign ( resid  68  and name  hn  ) ( resid  67  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  68  and name  hn  ) ( resid  67  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  68  and name  hn  ) ( resid  67  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  68  and name  hn  ) ( resid  68  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  68  and name  hn  ) ( resid  68  and name  ha2  ) 0.0 0.0 5.0
assign ( resid  68  and name  hn  ) ( resid  69  and name  hn  ) 0.0 0.0 5.0
assign ( resid  68  and name  ha2  ) ( resid  68  and name  ha1  ) 0.0 0.0 2.8
assign ( resid  69  and name  ha  ) ( resid  69  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  69  and name  ha  ) ( resid  69  and name  hb2  ) 0.0 0.0 2.8
assign ( resid  69  and name  hn  ) ( resid  35  and name  he#  ) 0.0 0.0 5.0
assign ( resid  69  and name  hn  ) ( resid  47  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  69  and name  hn  ) ( resid  50  and name  hn  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  65  and name  ha  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  66  and name  hn  ) 0.0 0.0 5.0
assign ( resid  69  and name  hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  67  and name  hn  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  67  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  67  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  67  and name  hg#  ) 0.0 0.0 6.0
assign ( resid  69  and name  hn  ) ( resid  68  and name  hn  ) 0.0 0.0 3.5
assign ( resid  69  and name  hn  ) ( resid  68  and name  ha1  ) 0.0 0.0 3.5
assign ( resid  69  and name  hn  ) ( resid  70  and name  ha  ) 0.0 0.0 5.0
assign ( resid  69  and name  hb2  ) ( resid  69  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  70  and name  ha  ) ( resid  65  and name  ha  ) 0.0 0.0 2.8
assign ( resid  70  and name  ha  ) ( resid  66  and name  ha  ) 0.0 0.0 5.0
assign ( resid  70  and name  ha  ) ( resid  66  and name  hg2#  ) 0.0 0.0 2.8
assign ( resid  70  and name  ha  ) ( resid  70  and name  hb2  ) 0.0 0.0 2.8
assign ( resid  70  and name  ha  ) ( resid  71  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  70  and name  ha  ) ( resid  71  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  50  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  70  and name  hn  ) ( resid  50  and name  hg2  ) 0.0 0.0 6.0
assign ( resid  70  and name  hn  ) ( resid  65  and name  ha  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  66  and name  hn  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  69  and name  ha  ) 0.0 0.0 3.5
assign ( resid  70  and name  hn  ) ( resid  69  and name  hn  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  69  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  70  and name  hn  ) ( resid  70  and name  ha  ) 0.0 0.0 3.5
assign ( resid  70  and name  hn  ) ( resid  70  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  70  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  71  and name  hn  ) 0.0 0.0 5.0
assign ( resid  70  and name  hn  ) ( resid  71  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  70  and name  hn  ) ( resid  71  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  70  and name  hb2  ) ( resid  70  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  70  and name  hd#  ) ( resid  50  and name  ha  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  50  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  50  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  70  and name  hd#  ) ( resid  50  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  50  and name  hg2  ) 0.0 0.0 3.5
assign ( resid  70  and name  hd#  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  65  and name  ha  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  69  and name  ha  ) 0.0 0.0 6.0
assign ( resid  70  and name  hd#  ) ( resid  70  and name  ha  ) 0.0 0.0 5.0
assign ( resid  70  and name  hd#  ) ( resid  70  and name  hb1  ) 0.0 0.0 2.8
```

```
assign ( resid  70  and name  hd#  ) ( resid  70  and name  hb2  ) 0.0 0.0 2.8
assign ( resid  70  and name  he#  ) ( resid  50  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  50  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  50  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  50  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  51  and name  ha  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  57  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  57  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  70  and name  he#  ) ( resid  57  and name  hd2#  ) 0.0 0.0 3.5
assign ( resid  70  and name  he#  ) ( resid  70  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  70  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  70  and name  he#  ) ( resid  85  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  71  and name  hn  ) ( resid  50  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  71  and name  hn  ) ( resid  62  and name  hg1#  ) 0.0 0.0 6.0
assign ( resid  71  and name  hn  ) ( resid  63  and name  ha  ) 0.0 0.0 5.0
assign ( resid  71  and name  hn  ) ( resid  65  and name  ha  ) 0.0 0.0 3.5
assign ( resid  71  and name  hn  ) ( resid  66  and name  hn  ) 0.0 0.0 3.5
assign ( resid  71  and name  hn  ) ( resid  66  and name  hg2#  ) 0.0 0.0 3.5
assign ( resid  71  and name  hn  ) ( resid  70  and name  ha  ) 0.0 0.0 2.8
assign ( resid  71  and name  hn  ) ( resid  70  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  71  and name  hn  ) ( resid  70  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  71  and name  hn  ) ( resid  71  and name  ha  ) 0.0 0.0 3.5
assign ( resid  71  and name  hn  ) ( resid  71  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  71  and name  hn  ) ( resid  71  and name  hd#  ) 0.0 0.0 3.5
assign ( resid  71  and name  hn  ) ( resid  71  and name  hg#  ) 0.0 0.0 5.0
assign ( resid  71  and name  hn  ) ( resid  72  and name  ha  ) 0.0 0.0 5.0
assign ( resid  71  and name  hb#  ) ( resid  66  and name  hg2#  ) 0.0 0.0 3.5
assign ( resid  71  and name  hd1  ) ( resid  66  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  71  and name  hd2  ) ( resid  66  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  71  and name  hd2  ) ( resid  71  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  71  and name  hd2  ) ( resid  71  and name  hg#  ) 0.0 0.0 3.5
assign ( resid  71  and name  hg#  ) ( resid  66  and name  hg2#  ) 0.0 0.0 3.5
assign ( resid  72  and name  ha  ) ( resid  63  and name  ha  ) 0.0 0.0 3.5
assign ( resid  72  and name  ha  ) ( resid  72  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  72  and name  ha  ) ( resid  72  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  72  and name  ha  ) ( resid  72  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  72  and name  hn  ) ( resid  63  and name  ha  ) 0.0 0.0 6.0
assign ( resid  72  and name  hn  ) ( resid  71  and name  ha  ) 0.0 0.0 2.8
assign ( resid  72  and name  hn  ) ( resid  71  and name  hn  ) 0.0 0.0 5.0
assign ( resid  72  and name  hn  ) ( resid  71  and name  hd#  ) 0.0 0.0 5.0
assign ( resid  72  and name  hn  ) ( resid  72  and name  ha  ) 0.0 0.0 3.5
assign ( resid  72  and name  hn  ) ( resid  72  and name  hb#  ) 0.0 0.0 3.5
assign ( resid  72  and name  hn  ) ( resid  73  and name  hn  ) 0.0 0.0 5.0
assign ( resid  72  and name  hn  ) ( resid  80  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  72  and name  hn  ) ( resid  80  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  72  and name  hb1  ) ( resid  80  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  72  and name  hb1  ) ( resid  80  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  72  and name  hb2  ) ( resid  80  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  72  and name  hb2  ) ( resid  80  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  73  and name  ha  ) ( resid  62  and name  hb  ) 0.0 0.0 5.0
assign ( resid  73  and name  ha  ) ( resid  73  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  73  and name  ha  ) ( resid  73  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  73  and name  ha  ) ( resid  73  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  73  and name  ha  ) ( resid  73  and name  hg2  ) 0.0 0.0 3.5
assign ( resid  73  and name  hn  ) ( resid  61  and name  ha1  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  62  and name  ha  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  62  and name  hb  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  62  and name  hn  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  62  and name  hg1#  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  62  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  63  and name  ha  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  64  and name  hn  ) 0.0 0.0 5.0
assign ( resid  73  and name  hn  ) ( resid  64  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  73  and name  hn  ) ( resid  64  and name  he1  ) 0.0 0.0 5.0
```

```
assign ( resid  73  and name   hn  ) ( resid  72  and name   ha  ) 0.0 0.0 2.8
assign ( resid  73  and name   hn  ) ( resid  72  and name   hn  ) 0.0 0.0 5.0
assign ( resid  73  and name   hn  ) ( resid  72  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  73  and name   hn  ) ( resid  73  and name   ha  ) 0.0 0.0 3.5
assign ( resid  73  and name   hn  ) ( resid  73  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  73  and name   hn  ) ( resid  73  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  73  and name  hb2  ) ( resid  62  and name   hb  ) 0.0 0.0 3.5
assign ( resid  73  and name  hb2  ) ( resid  62  and name  hg1#  ) 0.0 0.0 5.0
assign ( resid  73  and name  hb2  ) ( resid  62  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  74  and name   ha  ) ( resid  61  and name  ha1  ) 0.0 0.0 2.8
assign ( resid  74  and name   ha  ) ( resid  61  and name  ha2  ) 0.0 0.0 5.0
assign ( resid  74  and name   ha  ) ( resid  74  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  74  and name   ha  ) ( resid  74  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  74  and name   ha  ) ( resid  75  and name   ha  ) 0.0 0.0 5.0
assign ( resid  74  and name   ha  ) ( resid  78  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  74  and name   hn  ) ( resid  61  and name  ha1  ) 0.0 0.0 6.0
assign ( resid  74  and name   hn  ) ( resid  72  and name   ha  ) 0.0 0.0 6.0
assign ( resid  74  and name   hn  ) ( resid  73  and name   ha  ) 0.0 0.0 2.8
assign ( resid  74  and name   hn  ) ( resid  73  and name   hn  ) 0.0 0.0 5.0
assign ( resid  74  and name   hn  ) ( resid  73  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  74  and name   hn  ) ( resid  73  and name  hg2  ) 0.0 0.0 6.0
assign ( resid  74  and name   hn  ) ( resid  74  and name   ha  ) 0.0 0.0 5.0
assign ( resid  74  and name   hn  ) ( resid  74  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  74  and name   hn  ) ( resid  74  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  74  and name   hn  ) ( resid  80  and name   hg  ) 0.0 0.0 6.0
assign ( resid  74  and name  hb2  ) ( resid  74  and name  hb1  ) 0.0 0.0 2.8
assign ( resid  75  and name   ha  ) ( resid  73  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  75  and name   ha  ) ( resid  75  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  75  and name   ha  ) ( resid  75  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  75  and name   ha  ) ( resid  75  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  75  and name   ha  ) ( resid  76  and name  hg1  ) 0.0 0.0 6.0
assign ( resid  75  and name  hb2  ) ( resid  75  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  75  and name  hd2  ) ( resid  75  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  75  and name  hg2  ) ( resid  75  and name  hg1  ) 0.0 0.0 2.8
assign ( resid  76  and name   ha  ) ( resid  76  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  76  and name   ha  ) ( resid  76  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  76  and name   ha  ) ( resid  76  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  76  and name   ha  ) ( resid  76  and name  hg1  ) 0.0 0.0 3.5
assign ( resid  76  and name   ha  ) ( resid  76  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  76  and name  hb2  ) ( resid  76  and name  hg2  ) 0.0 0.0 3.5
assign ( resid  76  and name  hd2  ) ( resid  76  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  77  and name  ha2  ) ( resid  77  and name  ha1  ) 0.0 0.0 2.8
assign ( resid  78  and name   ha  ) ( resid  74  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  78  and name   ha  ) ( resid  78  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  78  and name  hb1  ) ( resid  74  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  78  and name  hb2  ) ( resid  74  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  78  and name  hne2  ) ( resid  75  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  75  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  75  and name  hd1  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  75  and name  hg1  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  75  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  77  and name  ha2  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  78  and name   ha  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  78  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  78  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  87  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  78  and name  hne2  ) ( resid  88  and name   ha  ) 0.0 0.0 5.0
assign ( resid  79  and name   ha  ) ( resid  79  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  79  and name   ha  ) ( resid  79  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  79  and name   ha  ) ( resid  79  and name  hg1  ) 0.0 0.0 3.5
assign ( resid  79  and name   ha  ) ( resid  79  and name  hg2  ) 0.0 0.0 5.0
assign ( resid  79  and name   hn  ) ( resid  74  and name   hn  ) 0.0 0.0 6.0
assign ( resid  79  and name   hn  ) ( resid  78  and name   ha  ) 0.0 0.0 3.5
```

```
assign ( resid  79  and name   hn  ) ( resid  78  and name   hn  ) 0.0 0.0 5.0
assign ( resid  79  and name   hn  ) ( resid  78  and name   hb# ) 0.0 0.0 3.5
assign ( resid  79  and name   hn  ) ( resid  78  and name   he1 ) 0.0 0.0 5.0
assign ( resid  79  and name   hn  ) ( resid  79  and name   ha  ) 0.0 0.0 3.5
assign ( resid  79  and name   hn  ) ( resid  79  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  79  and name   hn  ) ( resid  79  and name   hb2 ) 0.0 0.0 3.5
assign ( resid  79  and name   hn  ) ( resid  79  and name   hg# ) 0.0 0.0 5.0
assign ( resid  79  and name   hn  ) ( resid  80  and name   hn  ) 0.0 0.0 5.0
assign ( resid  79  and name   hn  ) ( resid  80  and name   hb2 ) 0.0 0.0 6.0
assign ( resid  79  and name   hn  ) ( resid  87  and name   hb1 ) 0.0 0.0 6.0
assign ( resid  79  and name   hn  ) ( resid  88  and name   hn  ) 0.0 0.0 2.8
assign ( resid  79  and name   hn  ) ( resid  88  and name   hg11 ) 0.0 0.0 6.0
assign ( resid  79  and name   hn  ) ( resid  88  and name   hg12 ) 0.0 0.0 6.0
assign ( resid  79  and name   hn  ) ( resid  88  and name   hg2# ) 0.0 0.0 5.0
assign ( resid  79  and name   hg2 ) ( resid  79  and name   hg1 ) 0.0 0.0 2.8
assign ( resid  79  and name   he21 ) ( resid  79  and name   hb2 ) 0.0 0.0 6.0
assign ( resid  79  and name   he21 ) ( resid  79  and name   hg# ) 0.0 0.0 5.0
assign ( resid  79  and name   he21 ) ( resid  88  and name   hg12 ) 0.0 0.0 6.0
assign ( resid  79  and name   he22 ) ( resid  79  and name   hb2 ) 0.0 0.0 6.0
assign ( resid  79  and name   he22 ) ( resid  79  and name   hg# ) 0.0 0.0 5.0
assign ( resid  79  and name   he22 ) ( resid  88  and name   hg12 ) 0.0 0.0 6.0
assign ( resid  79  and name   he22 ) ( resid  88  and name   hg2# ) 0.0 0.0 6.0
assign ( resid  79  and name   ne21 ) ( resid  79  and name   he21 ) 0.0 0.0 2.8
assign ( resid  79  and name   ne21 ) ( resid  79  and name   he22 ) 0.0 0.0 2.8
assign ( resid  79  and name   ne22 ) ( resid  79  and name   he21 ) 0.0 0.0 2.8
assign ( resid  79  and name   ne22 ) ( resid  79  and name   he22 ) 0.0 0.0 2.8
assign ( resid  80  and name   ha  ) ( resid  80  and name   hg  ) 0.0 0.0 5.0
assign ( resid  80  and name   ha  ) ( resid  80  and name   hd1# ) 0.0 0.0 3.5
assign ( resid  80  and name   ha  ) ( resid  88  and name   hg11 ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  74  and name   hb2 ) 0.0 0.0 6.0
assign ( resid  80  and name   hn  ) ( resid  79  and name   ha  ) 0.0 0.0 3.5
assign ( resid  80  and name   hn  ) ( resid  79  and name   hn  ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  79  and name   ha  ) 0.0 0.0 3.5
assign ( resid  80  and name   hn  ) ( resid  79  and name   hn  ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  79  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  79  and name   hb2 ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  79  and name   hg# ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  80  and name   ha  ) 0.0 0.0 3.5
assign ( resid  80  and name   hn  ) ( resid  80  and name   hg  ) 0.0 0.0 5.0
assign ( resid  80  and name   hn  ) ( resid  80  and name   hb1 ) 0.0 0.0 3.5
assign ( resid  80  and name   hn  ) ( resid  80  and name   hb2 ) 0.0 0.0 2.8
assign ( resid  80  and name   hn  ) ( resid  82  and name   hd1 ) 0.0 0.0 6.0
assign ( resid  80  and name   hn  ) ( resid  88  and name   hg11 ) 0.0 0.0 6.0
assign ( resid  81  and name   ha  ) ( resid  81  and name   hb2 ) 0.0 0.0 3.5
assign ( resid  81  and name   ha  ) ( resid  88  and name   hg11 ) 0.0 0.0 5.0
assign ( resid  81  and name   hn  ) ( resid  80  and name   ha  ) 0.0 0.0 3.5
assign ( resid  81  and name   hn  ) ( resid  80  and name   hg  ) 0.0 0.0 6.0
assign ( resid  81  and name   hn  ) ( resid  80  and name   hn  ) 0.0 0.0 5.0
assign ( resid  81  and name   hn  ) ( resid  80  and name   hb2 ) 0.0 0.0 5.0
assign ( resid  81  and name   hn  ) ( resid  81  and name   hb2 ) 0.0 0.0 5.0
assign ( resid  81  and name   hn  ) ( resid  86  and name   hn  ) 0.0 0.0 5.0
assign ( resid  81  and name   hn  ) ( resid  86  and name   hb# ) 0.0 0.0 6.0
assign ( resid  81  and name   hn  ) ( resid  88  and name   hb  ) 0.0 0.0 6.0
assign ( resid  81  and name   hb1 ) ( resid  84  and name   hd1# ) 0.0 0.0 3.5
assign ( resid  81  and name   hb1 ) ( resid  84  and name   hd11 ) 0.0 0.0 5.0
assign ( resid  81  and name   hb1 ) ( resid  84  and name   hg2# ) 0.0 0.0 5.0
assign ( resid  81  and name   hb1 ) ( resid  86  and name   hb# ) 0.0 0.0 5.0
assign ( resid  81  and name   hb1 ) ( resid  88  and name   hg11 ) 0.0 0.0 5.0
assign ( resid  81  and name   hb2 ) ( resid  84  and name   hg2# ) 0.0 0.0 5.0
assign ( resid  81  and name   hb2 ) ( resid  86  and name   hb# ) 0.0 0.0 5.0
assign ( resid  82  and name   ha  ) ( resid  80  and name   hg  ) 0.0 0.0 5.0
assign ( resid  82  and name   ha  ) ( resid  80  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  82  and name   ha  ) ( resid  82  and name   hd1 ) 0.0 0.0 5.0
assign ( resid  82  and name   ha  ) ( resid  82  and name   hd2 ) 0.0 0.0 5.0
assign ( resid  82  and name   ha  ) ( resid  82  and name   hg# ) 0.0 0.0 3.5
```

```
assign ( resid  82  and name   ha  ) ( resid  84  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  82  and name  hd1  ) ( resid  80  and name   hg  ) 0.0 0.0 6.0
assign ( resid  82  and name  hd1  ) ( resid  84  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  82  and name  hd2  ) ( resid  82  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  82  and name  hd2  ) ( resid  84  and name  hd1#  ) 0.0 0.0 2.8
assign ( resid  82  and name  hd2  ) ( resid  84  and name  hg11  ) 0.0 0.0 5.0
assign ( resid  83  and name  hd21  ) ( resid  70  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  83  and name  hd21  ) ( resid  72  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  83  and name  hd21  ) ( resid  84  and name  hd1#  ) 0.0 0.0 6.0
assign ( resid  83  and name  hd22  ) ( resid  70  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  83  and name  hd22  ) ( resid  72  and name  hb2  ) 0.0 0.0 6.0
assign ( resid  83  and name  hd22  ) ( resid  83  and name  hn  ) 0.0 0.0 6.0
assign ( resid  83  and name  hd22  ) ( resid  84  and name  hd1#  ) 0.0 0.0 6.0
assign ( resid  83  and name  nd21  ) ( resid  83  and name  hd21  ) 0.0 0.0 2.8
assign ( resid  83  and name  nd21  ) ( resid  83  and name  hd22  ) 0.0 0.0 2.8
assign ( resid  83  and name  nd22  ) ( resid  83  and name  hd21  ) 0.0 0.0 2.8
assign ( resid  83  and name  nd22  ) ( resid  83  and name  hd22  ) 0.0 0.0 2.8
assign ( resid  84  and name   ha  ) ( resid  84  and name  hb  ) 0.0 0.0 3.5
assign ( resid  84  and name   ha  ) ( resid  84  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  84  and name   ha  ) ( resid  84  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  84  and name   ha  ) ( resid  86  and name  hb#  ) 0.0 0.0 6.0
assign ( resid  84  and name   hb  ) ( resid  84  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  84  and name   hn  ) ( resid  72  and name  hb#  ) 0.0 0.0 5.0
assign ( resid  84  and name   hn  ) ( resid  83  and name  ha  ) 0.0 0.0 5.0
assign ( resid  84  and name   hn  ) ( resid  84  and name  ha  ) 0.0 0.0 3.5
assign ( resid  84  and name   hn  ) ( resid  84  and name  hb  ) 0.0 0.0 5.0
assign ( resid  84  and name   hn  ) ( resid  84  and name  hd1#  ) 0.0 0.0 3.5
assign ( resid  84  and name   hn  ) ( resid  84  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  84  and name   hn  ) ( resid  85  and name  hn  ) 0.0 0.0 5.0
assign ( resid  85  and name   ha  ) ( resid  85  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  85  and name   ha  ) ( resid  85  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  85  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 6.0
assign ( resid  85  and name   hn  ) ( resid  84  and name  ha  ) 0.0 0.0 3.5
assign ( resid  85  and name   hn  ) ( resid  84  and name  hn  ) 0.0 0.0 5.0
assign ( resid  85  and name   hn  ) ( resid  84  and name  hd1#  ) 0.0 0.0 6.0
assign ( resid  85  and name   hn  ) ( resid  84  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  85  and name   hn  ) ( resid  85  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  85  and name   hn  ) ( resid  84  and name  hg2#  ) 0.0 0.0 5.0
assign ( resid  85  and name   hn  ) ( resid  85  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  85  and name   hn  ) ( resid  86  and name  hn  ) 0.0 0.0 5.0
assign ( resid  85  and name  hb1  ) ( resid  57  and name  hd1#  ) 0.0 0.0 5.0
assign ( resid  85  and name  hb1  ) ( resid  57  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  86  and name   ha  ) ( resid  59  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  86  and name   ha  ) ( resid  86  and name  hb#  ) 0.0 0.0 2.8
assign ( resid  86  and name   hn  ) ( resid  57  and name  hd2#  ) 0.0 0.0 6.0
assign ( resid  86  and name   hn  ) ( resid  80  and name   hg  ) 0.0 0.0 5.0
assign ( resid  86  and name   hn  ) ( resid  81  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  86  and name   hn  ) ( resid  84  and name  hb  ) 0.0 0.0 5.0
assign ( resid  86  and name   hn  ) ( resid  84  and name  hg2#  ) 0.0 0.0 6.0
assign ( resid  86  and name   hn  ) ( resid  85  and name  hn  ) 0.0 0.0 5.0
assign ( resid  86  and name   hn  ) ( resid  85  and name  hb1  ) 0.0 0.0 6.0
assign ( resid  86  and name   hn  ) ( resid  86  and name  ha  ) 0.0 0.0 3.5
assign ( resid  87  and name   ha  ) ( resid  80  and name  ha  ) 0.0 0.0 2.8
assign ( resid  87  and name   ha  ) ( resid  80  and name  hg  ) 0.0 0.0 5.0
assign ( resid  87  and name   ha  ) ( resid  80  and name  hb1  ) 0.0 0.0 5.0
assign ( resid  87  and name   ha  ) ( resid  80  and name  hd2#  ) 0.0 0.0 5.0
assign ( resid  87  and name   ha  ) ( resid  87  and name  hb1  ) 0.0 0.0 3.5
assign ( resid  87  and name   ha  ) ( resid  87  and name  hb2  ) 0.0 0.0 3.5
assign ( resid  87  and name   hn  ) ( resid  58  and name  ha  ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  59  and name  hn  ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  59  and name  hb2  ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  59  and name  hd#  ) 0.0 0.0 6.0
assign ( resid  87  and name   hn  ) ( resid  60  and name  ha1  ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  60  and name  ha2  ) 0.0 0.0 6.0
assign ( resid  87  and name   hn  ) ( resid  61  and name  hn  ) 0.0 0.0 5.0
```

```
assign ( resid  87  and name   hn  ) ( resid  80  and name   hb2 ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  86  and name   hn  ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  86  and name   hb# ) 0.0 0.0 5.0
assign ( resid  87  and name   hn  ) ( resid  87  and name   hb1 ) 0.0 0.0 3.5
assign ( resid  87  and name   hn  ) ( resid  87  and name   hb2 ) 0.0 0.0 3.5
assign ( resid  87  and name   hb2 ) ( resid  80  and name   hd2# ) 0.0 0.0 5.0
assign ( resid  88  and name   ha  ) ( resid  86  and name   hb# ) 0.0 0.0 6.0
assign ( resid  88  and name   ha  ) ( resid  88  and name   hg11 ) 0.0 0.0 3.5
assign ( resid  88  and name   ha  ) ( resid  88  and name   hg11 ) 0.0 0.0 6.0
assign ( resid  88  and name   ha  ) ( resid  88  and name   hg12 ) 0.0 0.0 5.0
assign ( resid  88  and name   hb  ) ( resid  88  and name   hg12 ) 0.0 0.0 3.5
assign ( resid  88  and name   hn  ) ( resid  78  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  79  and name   ha  ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  79  and name   hn  ) 0.0 0.0 2.8
assign ( resid  88  and name   hn  ) ( resid  79  and name   hn  ) 0.0 0.0 2.8
assign ( resid  88  and name   hn  ) ( resid  80  and name   ha  ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  87  and name   ha  ) 0.0 0.0 3.5
assign ( resid  88  and name   hn  ) ( resid  87  and name   hb1 ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  87  and name   hb2 ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  88  and name   ha  ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  88  and name   hb  ) 0.0 0.0 6.0
assign ( resid  88  and name   hn  ) ( resid  88  and name   hd1# ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  88  and name   hg11 ) 0.0 0.0 5.0
assign ( resid  88  and name   hn  ) ( resid  88  and name   hg12 ) 0.0 0.0 5.0
assign ( resid  88  and name   hg11 ) ( resid  88  and name   hg2# ) 0.0 0.0 3.5
assign ( resid  88  and name   hg12 ) ( resid  88  and name   hg11 ) 0.0 0.0 2.8
```

(2) Ambiguous Interproton Distance Restraints

(a) Protons With Two Possible Connections

```
assign (resid   5  and name   hn  ) ((resid   5  and name   hb  )
                                    or (resid  10  and name   hb  )) 0.0 0.0 3.5
assign (resid   6  and name   hn  ) ((resid   5  and name   hb  )
                                    or (resid  10  and name   hb  )) 0.0 0.0 5.0
assign (resid   7  and name   hn  ) ((resid   5  and name   hb  )
                                    or (resid  10  and name   hb  )) 0.0 0.0 6.0
assign (resid   7  and name   hn  ) ((resid  10  and name   hn  )
                                    or (resid  27  and name   hz  )) 0.0 0.0 5.0
assign (resid   7  and name   hn  ) ((resid  27  and name   hb1 )
                                    or (resid   8  and name   hb# )) 0.0 0.0 5.0
assign (resid   8  and name   ha  ) ((resid   9  and name   hb1 )
                                    or (resid  14  and name   hb2 )) 0.0 0.0 5.0
assign (resid   8  and name   hn  ) ((resid   7  and name   hg2 )
                                    or (resid  43  and name   he# )) 0.0 0.0 6.0
assign (resid   8  and name   hn  ) ((resid   8  and name   hb# )
                                    or (resid  20  and name   hb2 )) 0.0 0.0 3.5
assign (resid   8  and name   hn  ) ((resid  10  and name   hn  )
                                    or (resid  27  and name   hz  )) 0.0 0.0 5.0
assign (resid   9  and name   hd# ) ((resid   9  and name   hb1 )
                                    or (resid  14  and name   hb2 )) 0.0 0.0 2.8
assign (resid  10  and name   ha  ) ((resid  10  and name   hd1# )
                                    or (resid   7  and name   hb2 )) 0.0 0.0 3.5
assign (resid  10  and name   hn  ) ((resid  10  and name   hn  )
                                    or (resid  11  and name   hn  )) 0.0 0.0 2.8
assign (resid  11  and name   ha  ) ((resid  12  and name   hb1 )
                                    or (resid  12  and name   hg# )) 0.0 0.0 5.0
assign (resid  11  and name   hn  ) ((resid  12  and name   ha  )
                                    or (resid  15  and name   ha  )) 0.0 0.0 5.0
assign (resid  11  and name   hn  ) ((resid  12  and name   hd2 )
                                    or (resid  11  and name   hb# )) 0.0 0.0 2.8
assign (resid  12  and name   ha  ) ((resid  12  and name   hb1 )
                                    or (resid  12  and name   hg# )) 0.0 0.0 2.8
```

```
assign (resid  12  and name  hd1  ) ((resid  12  and name  hb1  )
                                    or (resid  12  and name  hg#  )) 0.0 0.0 2.8
assign (resid  12  and name  hd2  ) ((resid  12  and name  hb1  )
                                    or (resid  12  and name  hg#  )) 0.0 0.0 2.8
assign (resid  13  and name  hn   ) ((resid  12  and name  hb1  )
                                    or (resid  12  and name  hg#  )) 0.0 0.0 3.5
assign (resid  13  and name  hn   ) ((resid  12  and name  hd2  )
                                    or (resid  11  and name  hb#  )) 0.0 0.0 3.5
assign (resid  14  and name  ha   ) ((resid  14  and name  hb2  )
                                    or (resid   9  and name  hb1  )) 0.0 0.0 2.8
assign (resid  14  and name  ha   ) ((resid  14  and name  hb1  )
                                    or (resid  14  and name  hg   )) 0.0 0.0 2.8
assign (resid  14  and name  hn   ) ((resid  12  and name  ha   )
                                    or (resid  15  and name  ha   )) 0.0 0.0 5.0
assign (resid  14  and name  hn   ) ((resid  12  and name  hb1  )
                                    or (resid  12  and name  hg#  )) 0.0 0.0 5.0
assign (resid  14  and name  hn   ) ((resid  12  and name  hd2  )
                                    or (resid  11  and name  hb#  )) 0.0 0.0 3.5
assign (resid  14  and name  hb2  ) ((resid  14  and name  hb1  )
                                    or (resid  14  and name  hg   )) 0.0 0.0 2.8
assign (resid  15  and name  hn   ) ((resid  14  and name  hb1  )
                                    or (resid  14  and name  hg   )) 0.0 0.0 3.5
assign (resid  15  and name  hn   ) ((resid  15  and name  hb1  )
                                    or (resid  13  and name  hb1  )) 0.0 0.0 3.5
assign (resid  16  and name  hn   ) ((resid  14  and name  hb1  )
                                    or (resid  14  and name  hg   )) 0.0 0.0 5.0
assign (resid  16  and name  hn   ) ((resid  15  and name  hb1  )
                                    or (resid  13  and name  hb1  )) 0.0 0.0 5.0
assign (resid  17  and name  hn   ) ((resid  16  and name  hn   )
                                    or (resid  18  and name  hn   )) 0.0 0.0 5.0
assign (resid  17  and name  hn   ) ((resid  17  and name  ha   )
                                    or (resid  16  and name  ha2  )) 0.0 0.0 3.5
assign (resid  18  and name  hn   ) ((resid  15  and name  hb1  )
                                    or (resid  29  and name  hb#  )) 0.0 0.0 5.0
assign (resid  18  and name  hn   ) ((resid  17  and name  ha   )
                                    or (resid  16  and name  ha2  )) 0.0 0.0 3.5
assign (resid  18  and name  hn   ) ((resid  19  and name  ha   )
                                    or (resid  45  and name  ha   )) 0.0 0.0 5.0
assign (resid  18  and name  hn   ) ((resid  19  and name  hg2  )
                                    or (resid  32  and name  hb1  )) 0.0 0.0 5.0
assign (resid  19  and name  hn   ) ((resid  19  and name  hb1  )
                                    or (resid  30  and name  hb2  )) 0.0 0.0 3.5
assign (resid  20  and name  hn   ) ((resid  20  and name  hb1  )
                                    or (resid  15  and name  hb2  )) 0.0 0.0 3.5
assign (resid  21  and name  hn   ) ((resid  20  and name  hb2  )
                                    or (resid  27  and name  hb1  )) 0.0 0.0 5.0
assign (resid  21  and name  hn   ) ((resid  21  and name  hb   )
                                    or (resid  28  and name  hb2  )) 0.0 0.0 3.5
assign (resid  22  and name  hn   ) ((resid  22  and name  hb2  )
                                    or (resid  27  and name  hb1  )) 0.0 0.0 3.5
assign (resid  25  and name  hn   ) ((resid  24  and name  ha   )
                                    or (resid  23  and name  hb   )) 0.0 0.0 3.5
assign (resid  25  and name  hn   ) ((resid  24  and name  hb1  )
                                    or (resid  24  and name  hg1  )) 0.0 0.0 5.0
assign (resid  26  and name  hn   ) ((resid  24  and name  ha   )
                                    or (resid  23  and name  hb   )) 0.0 0.0 5.0
assign (resid  26  and name  hn   ) ((resid  26  and name  ha   )
                                    or (resid   6  and name  ha   )) 0.0 0.0 3.5
assign (resid  26  and name  hn   ) ((resid  26  and name  hb1  )
                                    or (resid   7  and name  hg2  )) 0.0 0.0 3.5
assign (resid  27  and name  hz   ) ((resid  14  and name  hb1  )
                                    or (resid  14  and name  hg   )) 0.0 0.0 5.0
assign (resid  27  and name  hz   ) ((resid  43  and name  hd2  )
```

```
assign (resid  27  and name  hd#  or (resid   7  and name  hb1  )) 0.0 0.0 5.0
                                  ) ((resid  26  and name   ha  )
assign (resid  27  and name  hd#  or (resid  13  and name   ha  )) 0.0 0.0 5.0
                                  ) ((resid  26  and name  hb1  )
assign (resid  27  and name  hd#  or (resid   7  and name  hg2  )) 0.0 0.0 3.5
                                  ) ((resid  43  and name  hd2  )
assign (resid  27  and name  he#  or (resid   7  and name  hb1  )) 0.0 0.0 3.5
                                  ) ((resid  26  and name  hb1  )
assign (resid  27  and name  he#  or (resid   7  and name  hg2  )) 0.0 0.0 5.0
                                  ) ((resid  43  and name  hd2  )
assign (resid  27  and name  hn   or (resid   7  and name  hb1  )) 0.0 0.0 3.5
                                  ) ((resid   7  and name  hb1  )
assign (resid  27  and name  hn   or (resid  43  and name  hd#  )) 0.0 0.0 5.0
                                  ) ((resid  26  and name   ha  )
assign (resid  27  and name  hn   or (resid   6  and name   ha  )) 0.0 0.0 3.5
                                  ) ((resid  26  and name  hb1  )
assign (resid  27  and name  hn   or (resid   7  and name  hg2  )) 0.0 0.0 3.5
                                  ) ((resid  28  and name  hg#  )
assign (resid  28  and name  hn   or (resid   7  and name  hg1  )) 0.0 0.0 5.0
                                  ) ((resid  21  and name   hn  )
assign (resid  28  and name  hn   or (resid   7  and name   hn  )) 0.0 0.0 3.5
                                  ) ((resid  27  and name  hb1  )
assign (resid  28  and name  hn   or (resid  20  and name  hb2  )) 0.0 0.0 5.0
                                  ) ((resid  28  and name  hb1  )
assign (resid  28  and name  hn   or (resid  43  and name  hb#  )) 0.0 0.0 5.0
                                  ) ((resid  28  and name  hb2  )
assign (resid  29  and name  hn   or (resid  21  and name   hb  )) 0.0 0.0 3.5
                                  ) ((resid  30  and name  hg2  )
assign (resid  30  and name  hn   or (resid  43  and name  hd#  )) 0.0 0.0 5.0
                                  ) ((resid  30  and name  hb2  )
assign (resid  31  and name  hn   or (resid  19  and name  hb1  )) 0.0 0.0 5.0
                                  ) ((resid  30  and name   ha  )
assign (resid  31  and name  hn   or (resid  44  and name   ha  )) 0.0 0.0 2.8
                                  ) ((resid  30  and name  hb2  )
assign (resid  31  and name  hn   or (resid  19  and name  hb1  )) 0.0 0.0 5.0
                                  ) ((resid  31  and name  hb1  )
assign (resid  32  and name  hn   or (resid  35  and name  hb1  )) 0.0 0.0 5.0
                                  ) ((resid  32  and name  hb2  )
assign (resid  33  and name  hn   or (resid  20  and name  hb1  )) 0.0 0.0 3.5
                                  ) ((resid  32  and name   hn  )
assign (resid  33  and name  hn   or (resid  34  and name   hn  )) 0.0 0.0 3.5
                                  ) ((resid  33  and name  hg#  )
assign (resid  34  and name  hn   or (resid  32  and name  hb1  )) 0.0 0.0 3.5
                                  ) ((resid  33  and name   hn  )
assign (resid  34  and name  hn   or (resid  35  and name   hn  )) 0.0 0.0 3.5
                                  ) ((resid  33  and name  hg#  )
assign (resid  34  and name  hn   or (resid  32  and name  hb1  )) 0.0 0.0 3.5
                                  ) ((resid  35  and name  hb2  )
assign (resid  35  and name  hd#  or (resid  68  and name  ha1  )) 0.0 0.0 5.0
                                  ) ((resid  32  and name  hb2  )
assign (resid  35  and name  hd#  or (resid  17  and name  hd1  )) 0.0 0.0 3.5
                                  ) ((resid  34  and name  ha2  )
assign (resid  35  and name  hd#  or (resid  32  and name   ha  )) 0.0 0.0 5.0
                                  ) ((resid  35  and name  hb1  )
assign (resid  35  and name  hd#  or (resid  47  and name  hb2  )) 0.0 0.0 2.8
                                  ) ((resid  47  and name  hb1  )
assign (resid  35  and name  he#  or (resid  67  and name  hg#  )) 0.0 0.0 5.0
                                  ) ((resid  34  and name  ha2  )
assign (resid  35  and name  he#  or (resid  32  and name   ha  )) 0.0 0.0 6.0
                                  ) ((resid  35  and name  hb1  )
assign (resid  35  and name  hn   or (resid  47  and name  hb2  )) 0.0 0.0 5.0
                                  ) ((resid  17  and name  hd1  )
                                  or (resid  32  and name  hb2  )) 0.0 0.0 5.0
```

```
assign (resid  35  and name  hn  ) ((resid  33  and name  hg#  )
                                   or (resid  32  and name  hb1  )) 0.0 0.0 5.0
assign (resid  35  and name  hn  ) ((resid  35  and name  hb2  )
                                   or (resid  68  and name  ha1  )) 0.0 0.0 3.5
assign (resid  35  and name  hn  ) ((resid  47  and name  hn  )
                                   or (resid  31  and name  hn  )) 0.0 0.0 5.0
assign (resid  35  and name  hn  ) ((resid  48  and name  hb  )
                                   or (resid  48  and name  hg12 )) 0.0 0.0 5.0
assign (resid  36  and name  hn  ) ((resid  48  and name  hb  )
                                   or (resid  48  and name  hg12 )) 0.0 0.0 5.0
assign (resid  38  and name  hn  ) ((resid  38  and name  hn  )
                                   or (resid  44  and name  hn  )) 0.0 0.0 2.8
assign (resid  38  and name  hn  ) ((resid  38  and name  ha1  )
                                   or (resid  37  and name  hb#  )) 0.0 0.0 2.8
assign (resid  38  and name  hn  ) ((resid  46  and name  hb2  )
                                   or (resid  36  and name  hg2  )) 0.0 0.0 5.0
assign (resid  39  and name  hn  ) ((resid  41  and name  hn  )
                                   or (resid  42  and name  hn  )) 0.0 0.0 5.0
assign (resid  39  and name  he#  ) ((resid  39  and name  hz  )
                                   or (resid  39  and name  hd#  )) 0.0 0.0 2.8
assign (resid  41  and name  hn  ) ((resid  41  and name  ha  )
                                   or (resid  40  and name  ha  )) 0.0 0.0 3.5
assign (resid  41  and name  hn  ) ((resid  41  and name  hn  )
                                   or (resid  42  and name  hn  )) 0.0 0.0 3.5
assign (resid  42  and name  hn  ) ((resid  42  and name  hn  )
                                   or (resid  41  and name  hn  )) 0.0 0.0 2.8
assign (resid  42  and name  hn  ) ((resid  42  and name  hb#  )
                                   or (resid  41  and name  hb#  )) 0.0 0.0 2.8
assign (resid  42  and name  hn  ) ((resid  42  and name  hg1  )
                                   or (resid  41  and name  hg#  )) 0.0 0.0 3.5
assign (resid  44  and name  ha  ) ((resid  44  and name  hb1  )
                                   or (resid  31  and name  hb1  )) 0.0 0.0 2.8
assign (resid  44  and name  hn  ) ((resid  44  and name  hb2  )
                                   or (resid  15  and name  hb2  )) 0.0 0.0 5.0
assign (resid  44  and name  hd21  ) ((resid  44  and name  hn  )
                                   or (resid  38  and name  hn  )) 0.0 0.0 5.0
assign (resid  44  and name  hd21  ) ((resid  44  and name  hb1  )
                                   or (resid  46  and name  hg#  )) 0.0 0.0 5.0
assign (resid  44  and name  hd22  ) ((resid  44  and name  hn  )
                                   or (resid  38  and name  hn  )) 0.0 0.0 5.0
assign (resid  44  and name  hd22  ) ((resid  44  and name  hb#  )
                                   or (resid  46  and name  hg#  )) 0.0 0.0 5.0
assign (resid  45  and name  hn  ) ((resid  44  and name  hb1  )
                                   or (resid  31  and name  hb1  )) 0.0 0.0 5.0
assign (resid  46  and name  hn  ) ((resid  46  and name  ha  )
                                   or (resid  37  and name  ha  )) 0.0 0.0 5.0
assign (resid  47  and name  hn  ) ((resid  47  and name  hb1  )
                                   or (resid  46  and name  hb#  )) 0.0 0.0 3.5
assign (resid  47  and name  hn  ) ((resid  47  and name  hb2  )
                                   or (resid  46  and name  hg#  )) 0.0 0.0 3.5
assign (resid  47  and name  hn  ) ((resid  48  and name  hb  )
                                   or (resid  48  and name  hg12 )) 0.0 0.0 6.0
assign (resid  48  and name  ha  ) ((resid  48  and name  hb  )
                                   or (resid  48  and name  hg12 )) 0.0 0.0 2.8
assign (resid  48  and name  hn  ) ((resid  48  and name  hb  )
                                   or (resid  48  and name  hg12 )) 0.0 0.0 3.5
assign (resid  49  and name  hn  ) ((resid  34  and name  ha1  )
                                   or (resid  50  and name  ha  )) 0.0 0.0 5.0
assign (resid  49  and name  hn  ) ((resid  48  and name  hb  )
                                   or (resid  48  and name  hg12 )) 0.0 0.0 5.0
assign (resid  50  and name  ha  ) ((resid  50  and name  hg1  )
                                   or (resid  52  and name  hb2  )) 0.0 0.0 3.5
assign (resid  50  and name  ha  ) ((resid  70  and name  hb1  )
```

```
                                                or (resid   53  and name   hd1  )) 0.0 0.0 5.0
assign (resid   50  and name    hn      ) ((resid   70  and name   hd#  )
                                                or (resid   70  and name   he#  )) 0.0 0.0 5.0
assign (resid   51  and name    ha      ) ((resid   57  and name   hb2  )
                                                or (resid   57  and name    hg  )) 0.0 0.0 5.0
assign (resid   51  and name    hn      ) ((resid   53  and name   hb2  )
                                                or (resid   52  and name   hb1  )) 0.0 0.0 5.0
assign (resid   51  and name    hn      ) ((resid   64  and name    ha  )
                                                or (resid   49  and name    ha  )) 0.0 0.0 3.5
assign (resid   51  and name    hn      ) ((resid   70  and name   hd#  )
                                                or (resid   70  and name   he#  )) 0.0 0.0 5.0
assign (resid   52  and name    ha      ) ((resid   52  and name   hb1  )
                                                or (resid   62  and name    hb  )) 0.0 0.0 2.8
assign (resid   52  and name    hn      ) ((resid   52  and name   hb1  )
                                                or (resid   62  and name    hb  )) 0.0 0.0 3.5
assign (resid   52  and name    hn      ) ((resid   52  and name   hb2  )
                                                or (resid   33  and name   hb#  )) 0.0 0.0 3.5
assign (resid   53  and name    hn      ) ((resid   53  and name    ha  )
                                                or (resid   52  and name    ha  )) 0.0 0.0 3.5
assign (resid   53  and name    hn      ) ((resid   53  and name   hd2  )
                                                or (resid   51  and name   hb#  )) 0.0 0.0 5.0
assign (resid   54  and name    ha      ) ((resid   55  and name   hb1  )
                                                or (resid   55  and name   hg1  )) 0.0 0.0 5.0
assign (resid   54  and name    ha      ) ((resid   57  and name   hb2  )
                                                or (resid   57  and name    hg  )) 0.0 0.0 5.0
assign (resid   54  and name    hb2     ) ((resid   57  and name   hb2  )
                                                or (resid   57  and name    hg  )) 0.0 0.0 3.5
assign (resid   54  and name    hn      ) ((resid   53  and name   hb1  )
                                                or (resid   53  and name   hb2  )) 0.0 0.0 3.5
assign (resid   54  and name    hn      ) ((resid   53  and name   hg#  )
                                                or (resid   56  and name    hg  )) 0.0 0.0 5.0
assign (resid   54  and name    hn      ) ((resid   54  and name   hb2  )
                                                or (resid   53  and name   hd1  )) 0.0 0.0 3.5
assign (resid   54  and name    hn      ) ((resid   55  and name   hd2  )
                                                or (resid   85  and name   hb1  )) 0.0 0.0 5.0
assign (resid   56  and name    hn      ) ((resid   53  and name   hd1  )
                                                or (resid   54  and name   hb2  )) 0.0 0.0 5.0
assign (resid   56  and name    hn      ) ((resid   55  and name    ha  )
                                                or (resid   61  and name   ha2  )) 0.0 0.0 3.5
assign (resid   56  and name    hn      ) ((resid   55  and name   hd2  )
                                                or (resid   85  and name   hb1  )) 0.0 0.0 3.5
assign (resid   56  and name    hn      ) ((resid   60  and name    hn  )
                                                or (resid   58  and name    hn  )) 0.0 0.0 6.0
assign (resid   57  and name    ha      ) ((resid   57  and name   hb2  )
                                                or (resid   57  and name    hg  )) 0.0 0.0 2.8
assign (resid   57  and name    hn      ) ((resid   70  and name   hd#  )
                                                or (resid   70  and name   he#  )) 0.0 0.0 6.0
assign (resid   58  and name    hn      ) ((resid   57  and name   hb2  )
                                                or (resid   57  and name    hg  )) 0.0 0.0 5.0
assign (resid   58  and name    hn      ) ((resid   59  and name   hb1  )
                                                or (resid   59  and name   hg#  )) 0.0 0.0 6.0
assign (resid   59  and name    hn      ) ((resid   58  and name    ha  )
                                                or (resid   60  and name   ha1  )) 0.0 0.0 3.5
assign (resid   59  and name    hn      ) ((resid   58  and name    hn  )
                                                or (resid   60  and name    hn  )) 0.0 0.0 2.8
assign (resid   59  and name    hn      ) ((resid   59  and name   hb1  )
                                                or (resid   59  and name   hg#  )) 0.0 0.0 2.8
assign (resid   59  and name    hn      ) ((resid   62  and name    ha  )
                                                or (resid   60  and name   ha2  )) 0.0 0.0 6.0
assign (resid   60  and name    hn      ) ((resid   59  and name   hb1  )
                                                or (resid   59  and name   hg#  )) 0.0 0.0 5.0
assign (resid   61  and name    hn      ) ((resid   59  and name   hb1  )
                                                or (resid   59  and name   hg#  )) 0.0 0.0 5.0
```

```
assign (resid  62  and name  hn  ) ((resid  73  and name  hb#  )
                                   or (resid  73  and name  hg1  )) 0.0 0.0 3.5
assign (resid  63  and name  ha  ) ((resid  62  and name  ha  )
                                   or (resid  60  and name  ha2  )) 0.0 0.0 5.0
assign (resid  63  and name  hn  ) ((resid  58  and name  hb#  )
                                   or (resid  72  and name  hb#  )) 0.0 0.0 5.0
assign (resid  63  and name  hn  ) ((resid  62  and name  ha  )
                                   or (resid  81  and name  hb2  )) 0.0 0.0 2.8
assign (resid  64  and name  hn  ) ((resid  64  and name  hn  )
                                   or (resid  71  and name  hn  )) 0.0 0.0 2.8
assign (resid  64  and name  hn  ) ((resid  64  and name  hb#  )
                                   or (resid  63  and name  hb2  )) 0.0 0.0 3.5
assign (resid  64  and name  hn  ) ((resid  70  and name  hd#  )
                                   or (resid  70  and name  he#  )) 0.0 0.0 5.0
assign (resid  65  and name  hn  ) ((resid  65  and name  ha  )
                                   or (resid  64  and name  ha  )) 0.0 0.0 2.8
assign (resid  65  and name  hn  ) ((resid  65  and name  hb2  )
                                   or (resid  64  and name  hb#  )) 0.0 0.0 3.5
assign (resid  66  and name  ha  ) ((resid  67  and name  hg#  )
                                   or (resid  49  and name  hb1  )) 0.0 0.0 5.0
assign (resid  66  and name  hn  ) ((resid  65  and name  hb2  )
                                   or (resid  70  and name  hb1  )) 0.0 0.0 3.5
assign (resid  66  and name  hn  ) ((resid  70  and name  hd#  )
                                   or (resid  70  and name  he#  )) 0.0 0.0 5.0
assign (resid  67  and name  hn  ) ((resid  67  and name  hb1  )
                                   or (resid  84  and name  hb  )) 0.0 0.0 5.0
assign (resid  69  and name  hn  ) ((resid  69  and name  hb1  )
                                   or (resid  34  and name  ha1  )) 0.0 0.0 3.5
assign (resid  69  and name  hn  ) ((resid  69  and name  hb2  )
                                   or (resid  67  and name  ha  )) 0.0 0.0 3.5
assign (resid  69  and name  hn  ) ((resid  70  and name  hn  )
                                   or (resid  47  and name  hn  )) 0.0 0.0 5.0
assign (resid  70  and name  ha  ) ((resid  70  and name  hb1  )
                                   or (resid  64  and name  hb#  )) 0.0 0.0 2.8
assign (resid  70  and name  he#  ) ((resid  57  and name  hb2  )
                                   or (resid  57  and name  hg  )) 0.0 0.0 5.0
assign (resid  70  and name  ha  ) ((resid  66  and name  hb  )
                                   or (resid  85  and name  ha  )) 0.0 0.0 5.0
assign (resid  70  and name  hn  ) ((resid  69  and name  hb2  )
                                   or (resid  67  and name  ha  )) 0.0 0.0 5.0
assign (resid  70  and name  hn  ) ((resid  70  and name  hd#  )
                                   or (resid  70  and name  he#  )) 0.0 0.0 3.5
assign (resid  70  and name  hn  ) ((resid  71  and name  hg#  )
                                   or (resid  50  and name  hb2  )) 0.0 0.0 5.0
assign (resid  71  and name  hd1  ) ((resid  71  and name  hg#  )
                                   or (resid  63  and name  hb2  )) 0.0 0.0 3.5
assign (resid  71  and name  hn  ) ((resid  65  and name  hb1  )
                                   or (resid  72  and name  hb#  )) 0.0 0.0 5.0
assign (resid  71  and name  hn  ) ((resid  70  and name  hn  )
                                   or (resid  72  and name  hn  )) 0.0 0.0 5.0
assign (resid  71  and name  hn  ) ((resid  70  and name  hd#  )
                                   or (resid  70  and name  he#  )) 0.0 0.0 3.5
assign (resid  71  and name  hn  ) ((resid  71  and name  hn  )
                                   or (resid  64  and name  hn  )) 0.0 0.0 2.8
assign (resid  72  and name  hn  ) ((resid  70  and name  hd#  )
                                   or (resid  70  and name  he#  )) 0.0 0.0 5.0
assign (resid  72  and name  hn  ) ((resid  71  and name  hb#  )
                                   or (resid  80  and name  hb2  )) 0.0 0.0 3.5
assign (resid  73  and name  hn  ) ((resid  71  and name  hg#  )
                                   or (resid  80  and name  hb1  )) 0.0 0.0 6.0
assign (resid  74  and name  hn  ) ((resid  73  and name  hb2  )
                                   or (resid  73  and name  hg1  )) 0.0 0.0 5.0
assign (resid  75  and name  ha  ) ((resid  75  and name  hd1  )
```

```
                                                or (resid   76  and name    hd1   )) 0.0 0.0 3.5
assign (resid   78  and name    hb2   ) ((resid   87  and name    hb1   )
                                                or (resid   74  and name    hb1   )) 0.0 0.0 3.5
assign (resid   80  and name    hn    ) ((resid   74  and name    hb1   )
                                                or (resid   87  and name    hb1   )) 0.0 0.0 5.0
assign (resid   80  and name    hn    ) ((resid   87  and name    ha    )
                                                or (resid   86  and name    ha    )) 0.0 0.0 5.0
assign (resid   81  and name    hn    ) ((resid   81  and name    hb1   )
                                                or (resid   82  and name    hd1   )) 0.0 0.0 5.0
assign (resid   82  and name    hd2   ) ((resid   82  and name    hb1   )
                                                or (resid   82  and name    hg#   )) 0.0 0.0 5.0
assign (resid   83  and name    hd21  ) ((resid   82  and name    hb#   )
                                                or (resid   82  and name    hg#   )) 0.0 0.0 5.0
assign (resid   83  and name    hd22  ) ((resid   82  and name    hb#   )
                                                or (resid   82  and name    hg#   )) 0.0 0.0 5.0
assign (resid   85  and name    hn    ) ((resid   84  and name    hb    )
                                                or (resid   82  and name    hb1   )) 0.0 0.0 5.0
assign (resid   85  and name    hn    ) ((resid   85  and name    ha    )
                                                or (resid   85  and name    hb2   )) 0.0 0.0 3.5
assign (resid   86  and name    ha    ) ((resid   82  and name    hg#   )
                                                or (resid   59  and name    hb2   )) 0.0 0.0 5.0
assign (resid   86  and name    hn    ) ((resid   84  and name    ha    )
                                                or (resid   80  and name    ha    )) 0.0 0.0 5.0
assign (resid   86  and name    hn    ) ((resid   85  and name    ha    )
                                                or (resid   85  and name    hb2   )) 0.0 0.0 2.8
assign (resid   87  and name    hn    ) ((resid   59  and name    hb1   )
                                                or (resid   59  and name    hg#   )) 0.0 0.0 5.0
assign (resid   87  and name    hn    ) ((resid   81  and name    hb1   )
                                                or (resid   61  and name    ha1   )) 0.0 0.0 5.0
```

(b) More Complex Ambiguous Interproton Distance Restraints

```
assign ((resid    4   and name     ha    ) or (resid   45  and name     ha   ))
       ((resid    4   and name     hb2   ) or (resid   13  and name     hb2  )) 0.0 0.0 2.8
assign ((resid    5   and name     ha    ) or (resid    3  and name     hb   ))
                                           (resid    5  and name     hb   ) 0.0 0.0 2.8
assign ((resid    5   and name     hb    ) or (resid   10  and name     hb   )
     or (resid   43   and name     hb#   ))  ((resid    5  and name    hd1#  )
     or (resid   10   and name    hg12   ) or (resid   43  and name     hd1  )) 0.0 0.0 3.5
assign ((resid    5   and name    hd1#   ) or (resid   10  and name    hg12  ))
       ((resid    5   and name    hg2#   ) or (resid   10  and name    hg11  )) 0.0 0.0 2.8
assign ((resid    5   and name    hg1#   ) or (resid   84  and name    hg11  )
                                           or (resid   84  and name    hg12  ))
       ((resid    5   and name    hd1#   ) or (resid   84  and name    hg2#  )) 0.0 0.0 2.8
assign  (resid    5   and name     hn    )  ((resid    5  and name    hg2#  )
                                           or (resid   10  and name    hg11  )) 0.0 0.0 3.5
assign  (resid    5   and name     hn    )  ((resid    3  and name     ha   )
     or (resid   24   and name     ha    ) or (resid   23  and name     hb   )) 0.0 0.0 6.0
assign ((resid    6   and name     ha    ) or (resid   49  and name     ha   ))
       ((resid    6   and name     hb1   ) or (resid   49  and name     hb1  )) 0.0 0.0 2.8
assign ((resid    6   and name     ha    ) or (resid   49  and name     ha   ))
       ((resid    6   and name     hb2   ) or (resid   49  and name     hb2  )) 0.0 0.0 2.8
assign ((resid    6   and name     hb2   ) or (resid   49  and name     hb2  ))
       ((resid    6   and name     hb1   ) or (resid   49  and name     hb1  )) 0.0 0.0 2.8
assign  (resid    7   and name     ha    )  ((resid   10  and name    hg11  )
                                           or (resid    5  and name    hg2#  )) 0.0 0.0 2.8
assign ((resid    7   and name     hb1   ) or (resid   43  and name     hd2  ))
                                           (resid   14  and name    hd2#  ) 0.0 0.0 3.5
assign ((resid    7   and name     hb2   ) or (resid   57  and name    hb12  )
                                           or (resid   57  and name      hg  ))
       ((resid    7   and name     hb1   ) or (resid   57  and name     hb1  )) 0.0 0.0 2.8
assign ((resid    7   and name     hg2   ) or (resid   46  and name     hg1  ))
       ((resid    7   and name     hg1   ) or (resid   46  and name     hb1  )) 0.0 0.0 2.8
```

```
assign ((resid   7   and name   hg2   ) or (resid   50   and name   hg2   ))
       ((resid   7   and name   hb1   ) or (resid   50   and name   hb1   )) 0.0 0.0 2.8
assign  (resid   7   and name    hn   )  ((resid    5   and name   hg2#   )
                                          or (resid   10   and name   hg11   )) 0.0 0.0 5.0
assign  (resid   7   and name    hn   )  ((resid    7   and name   hb2   )
                                          or (resid   10   and name   hd1#   )) 0.0 0.0 3.5
assign  (resid   7   and name    hn   )  ((resid   14   and name   hd1#   )
                                          or (resid   14   and name   hd2#   )) 0.0 0.0 6.0
assign ((resid   8   and name    ha   ) or (resid   66   and name    hb   ))
       ((resid   8   and name   hb2   ) or (resid   71   and name   hd2   )) 0.0 0.0 3.5
assign ((resid   8   and name   hb2   ) or (resid   71   and name   hd2   ))
       ((resid   8   and name   hb1   ) or (resid   20   and name   hb2   )
                                          or (resid   71   and name   hd1   )) 0.0 0.0 2.8
assign  (resid   8   and name    hn   )  ((resid    5   and name   hg2#   )
                                          or (resid   10   and name   hg11   )) 0.0 0.0 5.0
assign  (resid   8   and name    hn   )  ((resid   14   and name   hd1#   )
                                          or (resid   14   and name   hd2#   )) 0.0 0.0 5.0
assign ((resid   9   and name    ha   ) or (resid   43   and name    ha   )
                                          or (resid   16   and name   ha2   ))
       ((resid   9   and name   hb1   ) or (resid   14   and name   hb2   )) 0.0 0.0 3.5
assign ((resid   9   and name    ha   ) or (resid   43   and name    ha   )
     or (resid  16   and name   ha2   ))  (resid    9   and name   hg#   ) 0.0 0.0 2.8
assign ((resid   9   and name    ha   ) or (resid   43   and name    ha   )
     or (resid  16   and name   ha2   ))  (resid    9   and name   hd#   ) 0.0 0.0 3.5
assign ((resid  10   and name    hb   ) or (resid   80   and name   hb2   )
     or (resid  30   and name   hb1   ))  ((resid   10   and name   hg2#   )
     or (resid  80   and name   hd2#   ) or (resid   30   and name   hg1   )) 0.0 0.0 2.8
assign ((resid  10   and name    hb   ) or (resid   80   and name   hb2   ))
       ((resid  10   and name   hg11   ) or (resid   80   and name   hd1#   )) 0.0 0.0 2.8
assign ((resid  10   and name    hb   ) or (resid   71   and name   hb#   )
     or (resid  80   and name   hb2   ))  ((resid   10   and name   hd1#   )
     or (resid  71   and name   hg#   ) or (resid   80   and name    hg   )) 0.0 0.0 2.8
assign ((resid  10   and name   hd1#   ) or (resid   56   and name    hg   )
     or (resid  80   and name    hg   ))  ((resid   10   and name   hg11   )
     or (resid  56   and name   hd1#   ) or (resid   80   and name   hd1#   )) 0.0 0.0 2.8
assign ((resid  10   and name   hd1#   ) or (resid   80   and name    hg   ))
       ((resid  10   and name   hg2#   ) or (resid   80   and name   hd2#   )) 0.0 0.0 2.8
assign  (resid  10   and name    hn   )  ((resid   10   and name   hd1#   )
                                          or (resid    7   and name   hb2   )) 0.0 0.0 3.5
assign  (resid  10   and name    hn   )  ((resid   14   and name   hd1#   )
                                          or (resid   14   and name   hd2#   )) 0.0 0.0 5.0
assign ((resid  11   and name    ha   ) or (resid   37   and name    ha   ))
       ((resid  11   and name   hb#   ) or (resid   37   and name   hb1   )) 0.0 0.0 2.8
assign  (resid  11   and name    hn   )  ((resid   10   and name   hd1#   )
                                          or (resid    7   and name   hb2   )) 0.0 0.0 5.0
assign  (resid  11   and name    hn   )  ((resid   14   and name   hd1#   )
                                          or (resid   14   and name   hd2#   )) 0.0 0.0 5.0
assign ((resid  11   and name    hn   ) or (resid   84   and name    hn   ))
       ((resid  11   and name    hn   ) or (resid   84   and name    hn   )
                                          or (resid   10   and name    hn   )) 0.0 0.0 2.8
assign ((resid  12   and name    ha   ) or (resid   34   and name   ha2   ))
       ((resid  12   and name   hb2   ) or (resid   67   and name   hg#   )) 0.0 0.0 2.8
assign ((resid  13   and name    ha   ) or (resid   26   and name    ha   ))
       ((resid  13   and name   hb1   ) or (resid   26   and name   hb2   )) 0.0 0.0 2.8
assign ((resid  13   and name   hb2   ) or (resid   87   and name   hb2   ))
       ((resid  13   and name   hb1   ) or (resid   87   and name   hb1   )) 0.0 0.0 2.8
assign  (resid  13   and name    hn   )  ((resid   14   and name   hd1#   )
                                          or (resid   14   and name   hd2#   )) 0.0 0.0 5.0
assign  (resid  13   and name    hn   )  ((resid   14   and name   hb1   )
     or (resid  14   and name    hg   ) or (resid    5   and name   hg1#   )) 0.0 0.0 5.0
assign ((resid  14   and name   hb1   ) or (resid   14   and name    hg   ))
                                         (resid   14   and name   hd2#   ) 0.0 0.0 3.5
assign ((resid  14   and name   hb1   ) or (resid   14   and name    hg   ))
```

```
                                                  (resid  14  and name  hd1#  ) 0.0 0.0 3.5
assign  (resid  14  and name  hn  )  ((resid  14  and name  hd1#  )
                                       or (resid  14  and name  hd2#  )) 0.0 0.0 3.5
assign  (resid  14  and name  hn  )  ((resid  14  and name  hb1  )
     or (resid  14  and name  hg  ) or (resid   5  and name  hg1#  )) 0.0 0.0 3.5
assign ((resid  15  and name  hb2 ) or (resid  42  and name  hg2  )
     or (resid  72  and name  hb2 ))  ((resid  15  and name  hb1  )
     or (resid  42  and name  hg1 ) or (resid  72  and name  hb1  )) 0.0 0.0 2.8
assign  (resid  15  and name  hn  )  ((resid  14  and name  hd1#  )
                                       or (resid  14  and name  hd2#  )) 0.0 0.0 5.0
assign  (resid  16  and name  hn  )  ((resid  14  and name  hd1#  )
                                       or (resid  14  and name  hd2#  )) 0.0 0.0 5.0
assign ((resid  17  and name  ha  ) or (resid  25  and name  ha2  ))
       ((resid  17  and name  hb2 ) or (resid   3  and name  hg2#  )) 0.0 0.0 2.8
assign  (resid  17  and name  hn  )  ((resid  35  and name  hb1  )
     or (resid  15  and name  hb1 ) or (resid  13  and name  hb1  )) 0.0 0.0 5.0
assign ((resid  18  and name  ha2 ) or (resid  81  and name  hb2  )
     or (resid  82  and name  hd2 ))  ((resid  18  and name  ha1  )
     or (resid  81  and name  hb1 ) or (resid  82  and name  hd1  )) 0.0 0.0 2.8
assign ((resid  18  and name  ha1 ) or (resid  81  and name  hb1  ))
       ((resid  30  and name  hd# ) or (resid  88  and name  hg12 )) 0.0 0.0 5.0
assign ((resid  19  and name  ha  ) or (resid  55  and name  ha   ))
       ((resid  19  and name  hb1 ) or (resid  56  and name  hb#  )
     or (resid  55  and name  hg1 ) or (resid  55  and name  hb1  )
                                       or (resid  55  and name  hg2  )) 0.0 0.0 2.8
assign ((resid  19  and name  ha  ) or (resid  55  and name  ha   ))
       ((resid  19  and name  hg2 ) or (resid  55  and name  hb2  )
                                       or (resid  52  and name  hg2  )) 0.0 0.0 2.8
assign ((resid  19  and name  hg1 ) or (resid  36  and name  hg2  )
     or (resid  73  and name  hg2 ) or (resid  79  and name  hg1  )
     or (resid  24  and name  hb2 ))  ((resid  19  and name  hb1  )
     or (resid  36  and name  hb1 ) or (resid  73  and name  hb1  )
     or (resid  79  and name  hb1 ) or (resid  24  and name  hg1  )) 0.0 0.0 2.8
assign  (resid  20  and name  ha  )  ((resid  21  and name  hg1#  )
                                       or (resid  21  and name  hg2#  )) 0.0 0.0 5.0
assign ((resid  20  and name  hb2 ) or (resid   8  and name  hb1  ))
                                       (resid  20  and name  hb1  ) 0.0 0.0 2.8
assign  (resid  20  and name  hn  )  ((resid  21  and name  hg1#  )
                                       or (resid  21  and name  hg2#  )) 0.0 0.0 6.0
assign ((resid  21  and name  ha  ) or (resid  62  and name  ha   ))
       ((resid  21  and name  hg2# ) or (resid  21  and name  hg1#  )
                                       or (resid  62  and name  hg2#  )) 0.0 0.0 2.8
assign ((resid  21  and name  hb  ) or (resid  84  and name  hb   )
                                       or (resid  88  and name  hb   ))
       ((resid  21  and name  hg1# ) or (resid  21  and name  hg2#  )
     or (resid  84  and name  hd1# ) or (resid  88  and name  hg2#  )) 0.0 0.0 2.8
assign  (resid  21  and name  hn  )  ((resid  21  and name  hg1#  )
                                       or (resid  21  and name  hg2#  )) 0.0 0.0 3.5
assign ((resid  22  and name  ha  ) or (resid  31  and name  ha   ))
       ((resid  22  and name  hb1 ) or (resid  31  and name  hb1  )) 0.0 0.0 2.8
assign ((resid  22  and name  ha  ) or (resid  31  and name  ha   )
     or (resid  78  and name  ha  ))  ((resid  22  and name  hb2  )
     or (resid  31  and name  hb2 ) or (resid  78  and name  hb2  )) 0.0 0.0 2.8
assign ((resid  22  and name  hb2 ) or (resid  31  and name  hb2  ))
       ((resid  22  and name  hb1 ) or (resid  31  and name  hb1  )) 0.0 0.0 2.8
assign ((resid  22  and name  ha  ) or (resid  86  and name  ha   ))
       ((resid  27  and name  hb2 ) or (resid  70  and name  hb2  )) 0.0 0.0 5.0
assign  (resid  22  and name  hd21  )  ((resid  14  and name  hd1#  )
                                       or (resid  14  and name  hd2#  )) 0.0 0.0 6.0
assign  (resid  22  and name  hd21  )  ((resid  21  and name  hg1#  )
                                       or (resid  21  and name  hg2#  )) 0.0 0.0 6.0
assign  (resid  22  and name  hd22  )  ((resid  21  and name  hg1#  )
                                       or (resid  21  and name  hg2#  )) 0.0 0.0 6.0
```

```
assign (resid  22   and name   hn   ) ((resid  21   and name  hg1#  )
                                    or (resid  21   and name  hg2#  )) 0.0 0.0 3.5
assign ((resid  23   and name   ha   ) or (resid   6   and name    ha  ))
       ((resid  23   and name   hb   ) or (resid  24   and name    ha  )) 0.0 0.0 2.8
assign ((resid  23   and name   hb   ) or (resid  24   and name    ha  ))
                                       (resid   4   and name   hb1  ) 0.0 0.0 5.0
assign ((resid  23   and name   hb   ) or (resid  24   and name    ha  ))
                                       (resid   4   and name   hb2  ) 0.0 0.0 3.5
assign (resid  23   and name   hn   ) ((resid  23   and name    ha  )
    or (resid  22   and name   ha   ) or (resid  26   and name    ha  )) 0.0 0.0 2.8
assign (resid  23   and name   hn   ) ((resid  21   and name  hg1#  )
                                    or (resid  21   and name  hg2#  )) 0.0 0.0 5.0
assign ((resid  24   and name   ha   ) or (resid  23   and name    hb  )
    or (resid  68   and name   ha2  ) or (resid  88   and name    ha  ))
       ((resid  24   and name   hb2  ) or (resid  47   and name   hb1  )
                                    or (resid  79   and name   hg1  )) 0.0 0.0 2.8
assign ((resid  24   and name   ha   ) or (resid  23   and name    hb  ))
                                       (resid  24   and name   hg2  ) 0.0 0.0 3.5
assign ((resid  24   and name   hd1  ) or (resid  76   and name   hd1  ))
       ((resid  24   and name   hg2  ) or (resid  76   and name   hg2  )) 0.0 0.0 2.8
assign ((resid  24   and name   ha   ) or (resid  23   and name    hb  ))
                                       (resid  24   and name   hd2  ) 0.0 0.0 2.8
assign ((resid  24   and name   hd2  ) or (resid  55   and name   hd2  ))
       ((resid  24   and name   hb1  ) or (resid  55   and name   hg2  )) 0.0 0.0 2.8
assign ((resid  24   and name   hd1  ) or (resid  75   and name   hd1  )
    or (resid  76   and name   hd1  ) or (resid  16   and name   ha1  ))
       ((resid  24   and name   hb2  ) or (resid  79   and name   hg1  )
                                    or (resid  42   and name   hb#  )) 0.0 0.0 5.0
assign ((resid  24   and name   ha   ) or (resid  23   and name    hb  )
                                    or (resid  88   and name    ha  ))
       ((resid  24   and name   hd1  ) or (resid  76   and name   hd1  )) 0.0 0.0 3.5
assign ((resid  27   and name   ha   ) or (resid  35   and name    ha  )
    or (resid  63   and name   ha   )) ((resid  22   and name   hb1  )
    or (resid  31   and name   hb1  ) or (resid  72   and name   hb2  )) 0.0 0.0 5.0
assign ((resid  27   and name   ha   ) or (resid  63   and name    ha  ))
       ((resid  27   and name   hb1  ) or (resid  63   and name   hb2  )) 0.0 0.0 2.8
assign ((resid  27   and name   ha   ) or (resid  63   and name    ha  ))
       ((resid   7   and name   hb2  ) or (resid  71   and name   hg#  )) 0.0 0.0 5.0
assign (resid  27   and name   ha   ) ((resid  21   and name  hg1#  )
                                    or (resid  21   and name  hg2#  )) 0.0 0.0 5.0
assign ((resid  28   and name   ha   ) or (resid  80   and name    ha  ))
       ((resid  28   and name   hb1  ) or (resid  80   and name   hb2  )) 0.0 0.0 2.8
assign (resid  28   and name   hn   ) ((resid  28   and name   hg#  )
    or (resid  40   and name   hg2  ) or (resid   7   and name   hg1  )) 0.0 0.0 3.5
assign (resid  28   and name   hn   ) ((resid  14   and name  hd1#  )
                                    or (resid  14   and name  hd2#  )) 0.0 0.0 6.0
assign (resid  28   and name   hn   ) ((resid  21   and name  hg1#  )
                                    or (resid  21   and name  hg2#  )) 0.0 0.0 5.0
assign (resid  29   and name   ha   ) ((resid  21   and name  hg1#  )
                                    or (resid  21   and name  hg2#  )) 0.0 0.0 5.0
assign ((resid  29   and name   hb#  ) or (resid  58   and name   hb1  ))
       ((resid  30   and name   hb1  ) or (resid  71   and name   hb#  )) 0.0 0.0 5.0
assign (resid  29   and name   hn   ) ((resid  21   and name  hg1#  )
                                    or (resid  21   and name  hg2#  )) 0.0 0.0 5.0
assign ((resid  30   and name   ha   ) or (resid  71   and name    ha  ))
       ((resid  30   and name   hb1  ) or (resid  71   and name   hb#  )) 0.0 0.0 2.8
assign ((resid  30   and name   ha   ) or (resid  71   and name    ha  ))
       ((resid  30   and name   hd#  ) or (resid  71   and name   hg#  )) 0.0 0.0 2.8
assign ((resid  30   and name   hb1  ) or (resid  80   and name   hb2  ))
       ((resid  30   and name   hd#  ) or (resid  80   and name   hb1  )) 0.0 0.0 2.8
assign ((resid  30   and name   hb1  ) or (resid  43   and name   hb#  ))
       ((resid  30   and name   hg2  ) or (resid  43   and name   hd2  )) 0.0 0.0 2.8
assign ((resid  32   and name   ha   ) or (resid  34   and name   ha2  ))
```

```
                                                (resid  33  and name    ha  ) 0.0 0.0 5.0
assign (resid  32  and name   hn  ) ((resid  32  and name   hb1 )
    or (resid  19  and name   hg# ) or (resid  33  and name   hg# )) 0.0 0.0 3.5
assign (resid  32  and name   hn  ) ((resid  19  and name   hn  )
    or (resid  33  and name   hn  ) or (resid  35  and name   hn  )) 0.0 0.0 5.0
assign (resid  34  and name   hn  ) ((resid  48  and name   hb  )
                                     or (resid  48  and name   hg12 )) 0.0 0.0 5.0
assign (resid  34  and name   hn  ) ((resid  48  and name   hg11 )
    or (resid  48  and name   hg2# ) or (resid  17  and name   hb1 )) 0.0 0.0 5.0
assign ((resid  35  and name   ha  ) or (resid  27  and name   ha  ))
       ((resid  67  and name   hb1 ) or (resid  21  and name   hb  )) 0.0 0.0 6.0
assign (resid  35  and name   ha  ) ((resid  48  and name   hg11 )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign ((resid  35  and name   ha  ) or (resid  63  and name   ha  ))
       ((resid  48  and name   hd1# ) or (resid  62  and name   hg1# )) 0.0 0.0 5.0
assign (resid  35  and name   ha  ) ((resid  48  and name   hb  )
                                     or (resid  48  and name   hg12 )) 0.0 0.0 5.0
assign ((resid  35  and name   hd# ) or (resid  64  and name   he1 ))
       ((resid  67  and name   hb2 ) or (resid  73  and name   hb2 )) 0.0 0.0 2.8
assign (resid  35  and name   hn  ) ((resid  48  and name   hd1# )
                                     or (resid  17  and name   hb1 )) 0.0 0.0 5.0
assign (resid  35  and name   hn  ) ((resid  33  and name   hb# )
    or (resid  36  and name   hb2 ) or (resid  19  and name   hb2 )
                                     or (resid  59  and name   hb2 )) 0.0 0.0 5.0
assign (resid  35  and name   hn  ) ((resid  48  and name   hg11 )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign (resid  36  and name   ha  ) ((resid  48  and name   hg11 )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign (resid  36  and name   hg2 ) ((resid  48  and name   hg1# )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign (resid  36  and name   hg1 ) ((resid  48  and name   hg1# )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign (resid  36  and name   hb1 ) ((resid  48  and name   hg1# )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 3.5
assign ((resid  36  and name   hg2 ) or (resid  73  and name   hg2 )
    or (resid  79  and name   hg1 ) or (resid  47  and name   hb1 )
    or (resid  42  and name   hb# ) or (resid  24  and name   hb2 ))
      ((resid  36  and name   hb2 ) or (resid  73  and name   hg1 )
    or (resid  79  and name   hb2 ) or (resid  24  and name   hb1 )) 0.0 0.0 2.8
assign ((resid  36  and name   hg1 ) or (resid  73  and name   hb2 )
    or (resid  76  and name   hg2 ))   ((resid  36  and name   hb2 )
    or (resid  73  and name   hg1 ) or (resid  76  and name   hg1 )) 0.0 0.0 2.8
assign ((resid  36  and name   hg1 ) or (resid  73  and name   hb2 )
    or (resid  24  and name   hg2 ) or (resid  76  and name   hg2 ))
      ((resid  36  and name   hb1 ) or (resid  73  and name   hb1 )
    or (resid  24  and name   hg1 ) or (resid  76  and name   hb1 )) 0.0 0.0 2.8
assign (resid  36  and name   hn  ) ((resid  48  and name   hg11 )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign (resid  36  and name   hn  ) ((resid  35  and name   hb1 )
    or (resid  46  and name   hg# ) or (resid  47  and name   hb2 )) 0.0 0.0 3.5
assign (resid  36  and name   hn  ) ((resid  36  and name   hg# )
    or (resid  46  and name   hb# ) or (resid  47  and name   hb1 )) 0.0 0.0 3.5
assign ((resid  37  and name   ha  ) or (resid  31  and name   ha  ))
                                     (resid  37  and name   hb2 ) 0.0 0.0 2.8
assign (resid  37  and name   hn  ) ((resid  48  and name   hg11 )
                                     or (resid  48  and name   hg2# )) 0.0 0.0 5.0
assign ((resid  38  and name   ha2 ) or (resid  85  and name   hb2 ))
       ((resid  38  and name   ha1 ) or (resid  85  and name   hb1 )) 0.0 0.0 2.8
assign ((resid  39  and name   ha  ) or (resid  12  and name   ha  ))
       ((resid  43  and name   ha  ) or (resid   9  and name   ha  )) 0.0 0.0 5.0
assign ((resid  40  and name   ha  ) or (resid  59  and name   ha  ))
       ((resid  40  and name   hb# ) or (resid  59  and name   hb1 )) 0.0 0.0 2.8
assign ((resid  40  and name   ha  ) or (resid  59  and name   ha  ))
```

```
      ((resid  40   and name   hg2   ) or (resid  59   and name   hb2   )) 0.0 0.0 2.8
assign ((resid  41   and name   hg2   ) or (resid  42   and name   hg1   )
     or (resid  47   and name   hb2   ))  ((resid  41   and name   hb#   )
     or (resid  42   and name   hb#   ) or (resid  47   and name   hb1   )) 0.0 0.0 2.8
assign ((resid  42   and name    ha   ) or (resid  28   and name    ha   ))
      ((resid  42   and name   hg1   ) or (resid  41   and name   hg2   )) 0.0 0.0 5.0
assign ((resid  42   and name    ha   ) or (resid  28   and name    ha   ))
                                        (resid  43   and name   hd1   ) 0.0 0.0 5.0
assign ((resid  43   and name    ha   ) or (resid  57   and name    ha   ))
      ((resid  38   and name   ha1   ) or (resid  85   and name   hb1   )) 0.0 0.0 5.0
assign ((resid  44   and name    hn   ) or (resid   3   and name    hn   ))
      ((resid  44   and name    hn   ) or (resid  38   and name    hn   )
                                       or (resid   3   and name    hn   )) 0.0 0.0 2.8
assign ((resid  44   and name    hn   ) or (resid   3   and name    hn   ))
      ((resid  43   and name   hg#   ) or (resid   3   and name  hg2#   )) 0.0 0.0 5.0
assign  (resid  44   and name    hn   )  ((resid  42   and name    ha   )
     or (resid  15   and name    ha   ) or (resid  28   and name    ha   )
                                       or (resid  32   and name    ha   )) 0.0 0.0 5.0
assign ((resid  45   and name    ha   ) or (resid  30   and name    ha   ))
                                        (resid  45   and name   hb2   )  0.0 0.0 3.5
assign ((resid  45   and name    ha   ) or (resid  30   and name    ha   ))
                                        (resid  45   and name   hb1   )  0.0 0.0 2.8
assign  (resid  45   and name    hn   )  ((resid  45   and name   hb#   )
     or (resid  32   and name   hb2   ) or (resid  17   and name   hd1   )) 0.0 0.0 3.5
assign  (resid  45   and name    hn   )  ((resid  45   and name    ha   )
     or (resid  44   and name    ha   ) or (resid  30   and name    ha   )
                                       or (resid  15   and name    ha   )) 0.0 0.0 2.8
assign ((resid  46   and name    ha   ) or (resid  54   and name    ha   ))
      ((resid  18   and name   ha1   ) or (resid  53   and name    ha   )
                                       or (resid  59   and name    ha   )) 0.0 0.0 5.0
assign ((resid  47   and name    ha   ) or (resid  18   and name   ha1   ))
                                        (resid  35   and name   hb2   )  0.0 0.0 5.0
assign ((resid  47   and name    ha   ) or (resid  81   and name   hb1   ))
      ((resid  67   and name    ha   ) or (resid  69   and name   hb2   )
                                       or (resid  83   and name   hb#   )) 0.0 0.0 5.0
assign ((resid  47   and name   hb1   ) or (resid  73   and name   hg1   ))
      ((resid  33   and name   hb#   ) or (resid  73   and name   hg2   )
     or (resid  76   and name   hg1   ) or (resid  79   and name   hb2   )) 0.0 0.0 2.8
assign  (resid  48   and name    ha   )  ((resid  48   and name  hg11   )
                                       or (resid  48   and name  hg2#   )) 0.0 0.0 2.8
assign ((resid  48   and name    hb   ) or (resid  48   and name  hg12   ))
                                        (resid  48   and name  hd1#   )  0.0 0.0 2.8
assign ((resid  48   and name    hb   ) or (resid  48   and name  hg12   ))
      ((resid  48   and name  hg11   ) or (resid  48   and name  hg2#   )) 0.0 0.0 2.8
assign  (resid  48   and name    hn   )  ((resid  48   and name  hg11   )
                                       or (resid  48   and name  hg2#   )) 0.0 0.0 3.5
assign  (resid  49   and name    hn   )  ((resid  48   and name  hg11   )
                                       or (resid  48   and name  hg2#   )) 0.0 0.0 5.0
assign ((resid  50   and name   hg1   ) or (resid  36   and name   hb2   ))
                                        (resid  48   and name  hd1#   )  0.0 0.0 5.0
assign ((resid  50   and name   hg1   ) or (resid  59   and name   hb2   ))
      ((resid  50   and name   hb1   ) or (resid  86   and name   hb#   )) 0.0 0.0 2.8
assign  (resid  50   and name    hn   )  ((resid  48   and name  hg11   )
                                       or (resid  48   and name  hg2#   )) 0.0 0.0 5.0
assign  (resid  51   and name    hn   )  ((resid  50   and name   hb1   )
     or (resid  48   and name    hb   ) or (resid  48   and name  hg12   )) 0.0 0.0 5.0
assign ((resid  52   and name    ha   ) or (resid  82   and name   hd1   ))
      ((resid  49   and name   hb1   ) or (resid  82   and name   hb2   )) 0.0 0.0 5.0
assign ((resid  52   and name    ha   ) or (resid  82   and name    ha   ))
      ((resid  52   and name   hb2   ) or (resid  82   and name   hb1   )
                                       or (resid  82   and name   hg#   )) 0.0 0.0 2.8
assign ((resid  53   and name    ha   ) or (resid  59   and name    ha   ))
      ((resid  53   and name   hb1   ) or (resid  75   and name   hg1   )
```

```
                                                  or (resid   59  and name    hg1  )) 0.0 0.0 2.8
assign ((resid   53  and name     ha  ) or (resid   59  and name     ha  ))
       ((resid   53  and name    hb2  ) or (resid   59  and name    hg2  )) 0.0 0.0 2.8
assign ((resid   53  and name    hb2  ) or (resid   75  and name    hb1  )
     or (resid   71  and name    hb#  ) or (resid   80  and name    hb2  ))
       ((resid   53  and name    hb1  ) or (resid   75  and name    hg1  )) 0.0 0.0 2.8
assign ((resid   53  and name    hd1  ) or (resid   70  and name    hb1  ))
                                              (resid   53  and name    hb1  ) 0.0 0.0 2.8
assign ((resid   53  and name    hd2  ) or (resid   71  and name    hd1  ))
       ((resid   53  and name    hb2  ) or (resid   71  and name    hb#  )) 0.0 0.0 3.5
assign (resid   53  and name     hn   )  ((resid   53  and name    hg#  )
     or (resid   53  and name    hb1  ) or (resid   53  and name    hb2  )
                                                  or (resid   52  and name    hb1  )) 0.0 0.0 3.5
assign ((resid   54  and name    hb1  ) or (resid   73  and name    hg2  ))
       ((resid   57  and name    hb2  ) or (resid   80  and name     hg  )) 0.0 0.0 3.5
assign (resid   55  and name    hd1   )  ((resid   55  and name    hb1  )
     or (resid   55  and name    hg1  ) or (resid   56  and name    hb#  )) 0.0 0.0 2.8
assign (resid   55  and name    hb2   )  ((resid   55  and name    hb1  )
     or (resid   55  and name    hg1  ) or (resid   55  and name    hg2  )) 0.0 0.0 2.8
assign ((resid   55  and name    hd1  ) or (resid   24  and name    hd2  ))
       ((resid   55  and name    hb1  ) or (resid   55  and name    hg1  )
                                                  or (resid   24  and name    hg1  )) 0.0 0.0 2.8
assign ((resid   56  and name     ha  ) or (resid   88  and name     ha  ))
       ((resid   56  and name   hd2#  ) or (resid   88  and name   hg2#  )) 0.0 0.0 2.8
assign ((resid   56  and name     ha  ) or (resid   88  and name     ha  ))
       ((resid   56  and name   hd1#  ) or (resid   88  and name   hd1#  )
                                                  or (resid   80  and name   hd1#  )) 0.0 0.0 2.8
assign ((resid   56  and name     ha  ) or (resid   24  and name     ha  ))
       ((resid   56  and name    hb#  ) or (resid   24  and name    hg1  )) 0.0 0.0 2.8
assign ((resid   56  and name   hd2#  ) or (resid   88  and name   hg2#  ))
       ((resid   56  and name   hd1#  ) or (resid   88  and name   hd1#  )) 0.0 0.0 2.8
assign (resid   56  and name     hn   )  ((resid   56  and name    hb#  )
     or (resid   55  and name    hb1  ) or (resid   55  and name    hg1  )
                                                  or (resid   55  and name    hg2  )) 0.0 0.0 2.8
assign (resid   56  and name     hn   )  ((resid   56  and name     hg  )
     or (resid   57  and name    hb2  ) or (resid   57  and name     hg  )) 0.0 0.0 3.5
assign ((resid   57  and name    hb2  ) or (resid   57  and name     hg  ))
                                              (resid   57  and name   hd2#  ) 0.0 0.0 2.8
assign ((resid   57  and name    hb2  ) or (resid   57  and name     hg  ))
                                              (resid   57  and name   hd1#  ) 0.0 0.0 2.8
assign (resid   57  and name     hn   )  ((resid   57  and name    hb2  )
     or (resid   57  and name     hg  ) or (resid   56  and name     hg  )) 0.0 0.0 2.8
assign (resid   57  and name     hn   )  ((resid   56  and name    hb#  )
     or (resid   55  and name    hb1  ) or (resid   55  and name    hg1  )
                                                  or (resid   55  and name    hg2  )) 0.0 0.0 3.5
assign (resid   58  and name     hn   )  ((resid   58  and name     ha  )
     or (resid   55  and name    hd2  ) or (resid   85  and name    hb1  )) 0.0 0.0 3.5
assign (resid   59  and name     hn   )  ((resid   56  and name     hg  )
     or (resid   57  and name    hb2  ) or (resid   57  and name     hg  )) 0.0 0.0 5.0
assign ((resid   62  and name     ha  ) or (resid   60  and name    ha2  ))
                                              (resid   62  and name   hg1#  ) 0.0 0.0 2.8
assign (resid   62  and name     hn   )  ((resid   62  and name   hg2#  )
                                                  or (resid   80  and name   hd2#  )) 0.0 0.0 3.5
assign (resid   63  and name     hn   )  ((resid   63  and name    hb2  )
     or (resid   64  and name    hb#  ) or (resid   51  and name    hb#  )) 0.0 0.0 3.5
assign ((resid   64  and name     ha  ) or (resid   65  and name     ha  ))
       ((resid   64  and name    hb#  ) or (resid   65  and name    hb2  )) 0.0 0.0 2.8
assign ((resid   64  and name    hb2  ) or (resid   78  and name    hb2  ))
       ((resid   64  and name    hb1  ) or (resid   78  and name    hb1  )) 0.0 0.0 2.8
assign (resid   65  and name     hn   )  ((resid   66  and name     hn  )
                                                  or (resid   65  and name   hd21  )) 0.0 0.0 5.0
assign ((resid   67  and name     ha  ) or (resid   83  and name    hb#  )
     or (resid   75  and name    hd2  ) or (resid   76  and name    hd2  ))
```

```
      ((resid  67   and name   hg#   ) or (resid  76   and name   hb2   )) 0.0 0.0 2.8
assign ((resid  67   and name    ha   ) or (resid  83   and name   hb#   )
                                        or (resid  76   and name   hd2   ))
      ((resid  67   and name   hb2   ) or (resid  76   and name   hg2   )) 0.0 0.0 2.8
assign ((resid  67   and name   hb2   ) or (resid  24   and name   hg2   ))
      ((resid  67   and name   hb1   ) or (resid  24   and name   hb1   )) 0.0 0.0 2.8
assign ((resid  67   and name   hg#   ) or (resid  82   and name   hb2   ))
      ((resid  67   and name   hb1   ) or (resid  82   and name   hb1   )) 0.0 0.0 2.8
assign ((resid  69   and name   hb1   ) or (resid  34   and name   ha1   )
    or (resid  50   and name    ha   )) (resid  48   and name  hg11   ) 0.0 0.0 5.0
assign ((resid  69   and name   hb1   ) or (resid  34   and name   ha1   ))
                                        (resid  48   and name  hg2#   ) 0.0 0.0 5.0
assign (resid  69   and name    hn    ) ((resid  69   and name    ha   )
    or (resid  68   and name   ha2   ) or (resid  66   and name    hb   )) 0.0 0.0 3.5
assign (resid  72   and name    hn    ) ((resid  71   and name   hg#   )
                                        or (resid  88   and name  hg12   )) 0.0 0.0 5.0
assign (resid  73   and name    hn    ) ((resid  73   and name   hg1   )
    or (resid  73   and name   hb2   ) or (resid  63   and name   hb1   )) 0.0 0.0 3.5
assign ((resid  75   and name   hd2   ) or (resid  76   and name   hd2   ))
                                        (resid  76   and name   hg1   ) 0.0 0.0 3.5
assign ((resid  75   and name   hd1   ) or (resid  76   and name   hd1   ))
                                        (resid  76   and name   hg1   ) 0.0 0.0 3.5
assign ((resid  75   and name   hd1   ) or (resid  76   and name   hd1   ))
                                        (resid  75   and name   hg1   ) 0.0 0.0 5.0
assign ((resid  75   and name   hd2   ) or (resid  76   and name   hd2   ))
                                        (resid  75   and name   hb2   ) 0.0 0.0 5.0
assign ((resid  75   and name   hd2   ) or (resid  76   and name   hd2   ))
                                        (resid  73   and name   hg2   ) 0.0 0.0 5.0
assign ((resid  75   and name   hd2   ) or (resid  76   and name   hd2   ))
      ((resid  75   and name   hd1   ) or (resid  76   and name   hd1   )) 0.0 0.0 2.8
assign ((resid  75   and name   hb2   ) or (resid  19   and name   hg2   ))
      ((resid  75   and name   hb1   ) or (resid  30   and name   hb1   )) 0.0 0.0 2.8
assign ((resid  75   and name    ha   ) or (resid  83   and name    ha   ))
      ((resid  75   and name   hb1   ) or (resid  71   and name   hb#   )
                                        or (resid  80   and name   hb2   )) 0.0 0.0 3.5
assign ((resid  75   and name   hd1   ) or (resid  76   and name   hd1   ))
                                        (resid  75   and name   hb2   ) 0.0 0.0 5.0
assign ((resid  75   and name   hd1   ) or (resid  76   and name   hd1   ))
      ((resid  75   and name   hb1   ) or (resid  10   and name    hb   )
    or (resid  28   and name   hb1   ) or (resid  30   and name   hb1   )
    or (resid  80   and name   hb2   ) or (resid  71   and name   hb#   )) 0.0 0.0 5.0
assign ((resid  76   and name   hd1   ) or (resid  75   and name   hd1   )
    or (resid  16   and name   ha1   )) ((resid  76   and name   hb2   )
                                        or (resid  12   and name   hb2   )) 0.0 0.0 5.0
assign ((resid  76   and name   hd1   ) or (resid  75   and name   hd1   )
    or (resid  16   and name   ha1   ) or (resid  24   and name   hd1   ))
      ((resid  76   and name   hb1   ) or (resid  19   and name   hb1   )
                                        or (resid  24   and name   hg1   )) 0.0 0.0 3.5
assign (resid  80   and name    ha    ) ((resid  80   and name   hb1   )
                                        or (resid  88   and name  hg12   )) 0.0 0.0 3.5
assign ((resid  80   and name   hb1   ) or (resid  88   and name  hg12   )
    or (resid  30   and name   hd#   )) ((resid  80   and name  hd2#   )
    or (resid  88   and name  hg2#   ) or (resid  30   and name   hg1   )) 0.0 0.0 2.8
assign ((resid  80   and name   hb1   ) or (resid  88   and name  hg12   ))
      ((resid  80   and name  hd1#   ) or (resid  88   and name  hd1#   )) 0.0 0.0 3.5
assign ((resid  80   and name    ha   ) or (resid  28   and name    ha   ))
      ((resid  80   and name  hd2#   ) or (resid  88   and name  hg2#   )
    or (resid  21   and name  hg1#   ) or (resid  21   and name  hg2#   )) 0.0 0.0 3.5
assign (resid  80   and name    hn    ) ((resid  80   and name  hd1#   )
                                        or (resid  88   and name  hd1#   )) 0.0 0.0 5.0
assign (resid  80   and name    hn    ) ((resid  80   and name  hd2#   )
                                        or (resid  88   and name  hg2#   )) 0.0 0.0 5.0
assign ((resid  81   and name    ha   ) or (resid   6   and name    ha   )
```

```
     or (resid  23    and name     ha    ) or (resid  49    and name     ha    )
     or (resid  64    and name     ha    ) or (resid  65    and name     ha    ))
       ((resid  81    and name    hb1    ) or (resid  25    and name    ha1    )
     or (resid  52    and name     ha    ) or (resid  82    and name    hd1    )) 0.0 0.0 2.8
assign  (resid  81    and name     hn    )   ((resid  81    and name     ha    )
     or (resid  87    and name     ha    ) or (resid  86    and name     ha    )) 0.0 0.0 3.5
assign  (resid  81    and name     hn    )   ((resid  80    and name    hb1    )
                                             or (resid  88    and name   hg12    )) 0.0 0.0 5.0
assign  (resid  81    and name     hn    )   ((resid  80    and name   hd2#    )
                                             or (resid  88    and name   hg2#    )) 0.0 0.0 5.0
assign  (resid  81    and name     hn    )   ((resid  80    and name   hd1#    )
                                             or (resid  88    and name   hd1#    )) 0.0 0.0 3.5
assign  (resid  82    and name     ha    )   ((resid  82    and name    hb1    )
     or (resid  84    and name     hb    ) or (resid  67    and name    hb1    )) 0.0 0.0 2.8
assign ((resid  82    and name     ha    ) or (resid  76    and name     ha    ))
       ((resid  82    and name    hb2    ) or (resid  76    and name    hb2    )) 0.0 0.0 2.8
assign ((resid  83    and name     ha    ) or (resid  75    and name     ha    ))
       ((resid  69    and name     ha    ) or (resid  76    and name     ha    )) 0.0 0.0 5.0
assign ((resid  83    and name     ha    ) or (resid  75    and name     ha    ))
       ((resid  83    and name    hb#    ) or (resid  75    and name    hd2    )) 0.0 0.0 2.8
assign  (resid  84    and name     ha    )   ((resid  84    and name   hg11    )
                                             or (resid  84    and name   hg12    )) 0.0 0.0 3.5
assign  (resid  84    and name     hb    )   ((resid  84    and name   hg11    )
                                             or (resid  84    and name   hg12    )) 0.0 0.0 3.5
assign ((resid  84    and name    hg11    ) or (resid  84    and name   hg12    ))
                                                (resid  84    and name   hd1#    )  0.0 0.0 2.8
assign  (resid  84    and name     hn    )   ((resid  84    and name   hg11    )
                                             or (resid  84    and name   hg12    )) 0.0 0.0 3.5
assign ((resid  85    and name     ha    ) or (resid  82    and name    hd2    ))
       ((resid  56    and name    hd1#    ) or (resid  80    and name   hd1#    )
                                             or (resid  88    and name   hd1#    )) 0.0 0.0 3.5
assign  (resid  85    and name     hn    )   ((resid  84    and name   hg11    )
                                             or (resid  84    and name   hg12    )) 0.0 0.0 5.0
assign  (resid  86    and name     ha    )   ((resid  80    and name   hd2#    )
                                             or (resid  84    and name   hd1#    )) 0.0 0.0 5.0
assign  (resid  86    and name     hn    ).....((resid  87    and name     hn    )
     or (resid  81    and name     hn    ) or (resid  58    and name     hn    )) 0.0 0.0 3.5
assign  (resid  86    and name     hn    )   ((resid  80    and name   hd1#    )
                                             or (resid  80    and name   hd2#    )) 0.0 0.0 3.5
assign  (resid  86    and name     hn    )   ((resid  86    and name    hb#    )
     or (resid  84    and name    hg11    ) or (resid  84    and name   hg12    )) 0.0 0.0 3.5
assign  (resid  87    and name     ha    )   ((resid  88    and name   hd1#    )
                                             or (resid  80    and name   hd1#    )) 0.0 0.0 3.5
assign  (resid  87    and name     hn    )   ((resid  80    and name   hd2#    )
                                             or (resid  84    and name   hd1#    )) 0.0 0.0 5.0
assign  (resid  87    and name     hn    )   ((resid  87    and name     ha    )
     or (resid  86    and name     ha    ) or (resid  81    and name     ha    )) 0.0 0.0 2.8
assign  (resid  87    and name     hn    )   ((resid  80    and name   hd1#    )
                                             or (resid  88    and name   hd1#    )) 0.0 0.0 5.0
assign ((resid  88    and name     ha    ) or (resid  56    and name     ha    )
     or (resid  68    and name    ha2    ) or (resid  23    and name     hb    )
     or (resid  24    and name     ha    ))   ((resid  88    and name     hb    )
     or (resid  24    and name    hb1    ) or (resid  55    and name    hg2    )
                                             or (resid  67    and name    hb1    )) 0.0 0.0 2.8
assign  (resid  88    and name     hn    )   ((resid  88    and name   hg2#    )
     or (resid  80    and name    hd2#    ) or (resid  56    and name   hd2#    )) 0.0 0.0 5.0
```

(3) Phi (ϕ) Angle Restraints

J-coupling values are shown after each restraint.

set message=off echo=off end

```
restraints
  dihedral
    reset
! based on HMQCJ data (500 MHz); values for 500 MHz given
! alpha
assign (resid  7 and name  C  ) (resid   8 and name  N  )
       (resid  8 and name  CA ) (resid   8 and name  C  ) 1    -60.0 30.0 2 !4.4
assign (resid 49 and name  C  ) (resid  50 and name  N  )
       (resid 50 and name  CA ) (resid  50 and name  C  ) 1    -60.0 30.0 2 !2.3
assign (resid 56 and name  C  ) (resid  57 and name  N  )
       (resid 57 and name  CA ) (resid  57 and name  C  ) 1    -60.0 30.0 2 !3.9
assign (resid 58 and name  C  ) (resid  59 and name  N  )
       (resid 59 and name  CA ) (resid  59 and name  C  ) 1    -60.0 30.0 2 !2.1
assign (resid 64 and name  C  ) (resid  65 and name  N  )
       (resid 65 and name  CA ) (resid  65 and name  C  ) 1    -60.0 30.0 2 !4.6
assign (resid 79 and name  C  ) (resid  80 and name  N  )
       (resid 80 and name  CA ) (resid  80 and name  C  ) 1    -60.0 30.0 2 !4.6
! beta
assign (resid  4 and name  C  ) (resid   5 and name  N  )
       (resid  5 and name  CA ) (resid   5 and name  C  ) 1    -120.0 30.0 2 !8.5
assign (resid 12 and name  C  ) (resid  13 and name  N  )
       (resid 13 and name  CA ) (resid  13 and name  C  ) 1    -120.0 30.0 2 !8.5
assign (resid 18 and name  C  ) (resid  19 and name  N  )
       (resid 19 and name  CA ) (resid  19 and name  C  ) 1    -120.0 30.0 2 !9.0
assign (resid 20 and name  C  ) (resid  21 and name  N  )
       (resid 21 and name  CA ) (resid  21 and name  C  ) 1    -120.0 30.0 2 !9.4
assign (resid 22 and name  C  ) (resid  23 and name  N  )
       (resid 23 and name  CA ) (resid  23 and name  C  ) 1    -140.0 30.0 2 !8.7
assign (resid 28 and name  C  ) (resid  29 and name  N  )
       (resid 29 and name  CA ) (resid  29 and name  C  ) 1    -140.0 30.0 2 !8.9
assign (resid 29 and name  C  ) (resid  30 and name  N  )
       (resid 30 and name  CA ) (resid  30 and name  C  ) 1    -120.0 30.0 2 !9.0
assign (resid 34 and name  C  ) (resid  35 and name  N  )
       (resid 35 and name  CA ) (resid  35 and name  C  ) 1    -100.0 30.0 2 !9.5
assign (resid 35 and name  C  ) (resid  36 and name  N  )
       (resid 36 and name  CA ) (resid  36 and name  C  ) 1    -120.0 30.0 2 !8.1
assign (resid 45 and name  C  ) (resid  46 and name  N  )
       (resid 46 and name  CA ) (resid  46 and name  C  ) 1    -120.0 30.0 2 !9.2
assign (resid 48 and name  C  ) (resid  49 and name  N  )
       (resid 49 and name  CA ) (resid  49 and name  C  ) 1    -120.0 30.0 2 !8.6
assign (resid 53 and name  C  ) (resid  54 and name  N  )
       (resid 54 and name  CA ) (resid  54 and name  C  ) 1    -120.0 30.0 2 !9.3
assign (resid 55 and name  C  ) (resid  56 and name  N  )
       (resid 56 and name  CA ) (resid  56 and name  C  ) 1    -120.0 30.0 2 !8.7
assign (resid 61 and name  C  ) (resid  62 and name  N  )
       (resid 62 and name  CA ) (resid  62 and name  C  ) 1    -120.0 30.0 2 !9.9
assign (resid 63 and name  C  ) (resid  64 and name  N  )
       (resid 64 and name  CA ) (resid  64 and name  C  ) 1    -120.0 30.0 2 !9.3
assign (resid 70 and name  C  ) (resid  71 and name  N  )
       (resid 71 and name  CA ) (resid  71 and name  C  ) 1    -140.0 30.0 2 !8.2
assign (resid 71 and name  C  ) (resid  72 and name  N  )
       (resid 72 and name  CA ) (resid  72 and name  C  ) 1    -120.0 30.0 2 !8.4
assign (resid 72 and name  C  ) (resid  73 and name  N  )
       (resid 73 and name  CA ) (resid  73 and name  C  ) 1    -120.0 30.0 2 !9.1
assign (resid 78 and name  C  ) (resid  79 and name  N  )
       (resid 79 and name  CA ) (resid  79 and name  C  ) 1    -120.0 30.0 2 !8.2
assign (resid 87 and name  C  ) (resid  88 and name  N  )
       (resid 88 and name  CA ) (resid  88 and name  C  ) 1    -140.0 30.0 2 !9.4
  end
end
```

(4) Hydrogen Bonding Restraints

Several fake hydrogen bonds were also used, where NOE data supported their incorporation, during the middle of the structure calculation process to facilitate identification of incorrectly assigned NOEs.

```
!HBONDS from ²H₂O and H₂O data
! 19-30 and 30-19
assign (resid  30 and  name O  )(resid  19 and  name N  )  3.3 1.5 0.0
assign (resid  30 and  name O  )(resid  19 and  name HN )  2.3 0.5 0.0
assign (resid  19 and  name O  )(resid  30 and  name N  )  3.3 1.5 0.0
assign (resid  19 and  name O  )(resid  30 and  name HN )  2.3 0.5 0.0
! 23-26
assign (resid  26 and  name O  )(resid  23 and  name N  )  3.3 1.5 0.0
assign (resid  26 and  name O  )(resid  23 and  name HN )  2.3 0.5 0.0
! 28-21
assign (resid  21 and  name O  )(resid  28 and  name N  )  3.3 1.5 0.0
assign (resid  21 and  name O  )(resid  28 and  name HN )  2.3 0.5 0.0
! 36-46 and 46-36
assign (resid  46 and  name O  )(resid  36 and  name N  )  3.3 1.5 0.0
assign (resid  46 and  name O  )(resid  36 and  name HN )  2.3 0.5 0.0
assign (resid  36 and  name O  )(resid  46 and  name N  )  3.3 1.5 0.0
assign (resid  36 and  name O  )(resid  46 and  name HN )  2.3 0.5 0.0
! 62-73
assign (resid  73 and  name O  )(resid  62 and  name N  )  3.3 1.5 0.0
assign (resid  73 and  name O  )(resid  62 and  name HN )  2.3 0.5 0.0
assign (resid  62 and  name O  )(resid  73 and  name N  )  3.3 1.5 0.0
assign (resid  62 and  name O  )(resid  73 and  name HN )  2.3 0.5 0.0
! 64-71
assign (resid  71 and  name O  )(resid  64 and  name N  )  3.3 1.5 0.0
assign (resid  71 and  name O  )(resid  64 and  name HN )  2.3 0.5 0.0
assign (resid  64 and  name O  )(resid  71 and  name N  )  3.3 1.5 0.0
assign (resid  64 and  name O  )(resid  71 and  name HN )  2.3 0.5 0.0
! 66-69
assign (resid  69 and  name O  )(resid  66 and  name N  )  3.3 1.5 0.0
assign (resid  69 and  name O  )(resid  66 and  name HN )  2.3 0.5 0.0
! 88-79
assign (resid  79 and  name O  )(resid  88 and  name N  )  3.3 1.5 0.0
assign (resid  79 and  name O  )(resid  88 and  name HN )  2.3 0.5 0.0
!FAKE HBONDS based on NOEs
! 21-28
!assign (resid  28 and  name O  )(resid  21 and  name N  )  3.3 1.5 0.0
!assign (resid  28 and  name O  )(resid  21 and  name HN )  2.3 0.5 0.0
! 26-23
!assign (resid  23 and  name O  )(resid  26 and  name N  )  3.3 1.5 0.0
!assign (resid  23 and  name O  )(resid  26 and  name HN )  2.3 0.5 0.0
! 38-44 and 44-38
!assign (resid  44 and  name O  )(resid  38 and  name N  )  3.3 1.5 0.0
!assign (resid  44 and  name O  )(resid  38 and  name HN )  2.3 0.5 0.0
!assign (resid  38 and  name O  )(resid  44 and  name N  )  3.3 1.5 0.0
!assign (resid  38 and  name O  )(resid  44 and  name HN )  2.3 0.5 0.0
! 69-66
!assign (resid  66 and  name O  )(resid  69 and  name N  )  3.3 1.5 0.0
!assign (resid  66 and  name O  )(resid  69 and  name HN )  2.3 0.5 0.0
! 79-88
!assign (resid  88 and  name OT1 )(resid  79 and  name N  )  3.3 1.5 0.0
!assign (resid  88 and  name OT1 )(resid  79 and  name HN )  2.3 0.5 0.0
! 81-86 and 86-81
!assign (resid  86 and  name O  )(resid  81 and  name N  )  3.3 1.5 0.0
!assign (resid  86 and  name O  )(resid  81 and  name HN )  2.3 0.5 0.0
!assign (resid  81 and  name O  )(resid  86 and  name N  )  3.3 1.5 0.0
!assign (resid  81 and  name O  )(resid  86 and  name HN )  2.3 0.5 0.0
```

(5) Calcium Binding Restraints

```
assign (resid   4 and   name    OD2 ) (resid  101 and   name  CAL ) 2.4 0.2 0.2
assign (resid   6 and   name    OD2 ) (resid   8  and   name  HN  ) 0.0 0.0 2.3 !HBOND
assign (resid   6 and   name    OD2 ) (resid   8  and   name  N   ) 0.0 0.0 3.3 !HBOND
assign (resid   7 and   name    OE1 ) (resid  101 and   name  CAL ) 2.4 0.2 0.2
assign (resid   7 and   name    OE2 ) (resid  101 and   name  CAL ) 2.4 0.2 0.2
assign (resid  22 and   name    OD1 ) (resid  101 and   name  CAL ) 2.4 0.2 0.2
assign (resid   5 and   name     O  ) (resid  101 and   name  CAL ) 2.4 0.2 0.2
assign (resid  23 and   name     O  ) (resid  101 and   name  CAL ) 2.4 0.2 0.2
assign (resid  47 and   name    OD2 ) (resid  102 and   name  CAL ) 2.4 0.2 0.2
assign (resid  49 and   name    OD2 ) (resid  51  and   name  HN  ) 0.0 0.0 2.3 !HBOND
assign (resid  49 and   name    OD2 ) (resid  51  and   name  N   ) 0.0 0.0 3.3 !HBOND
assign (resid  50 and   name    OE1 ) (resid  102 and   name  CAL ) 2.4 0.2 0.2
assign (resid  50 and   name    OE2 ) (resid  102 and   name  CAL ) 2.4 0.2 0.2
assign (resid  65 and   name    OD1 ) (resid  102 and   name  CAL ) 2.4 0.2 0.2
assign (resid  48 and   name     O  ) (resid  102 and   name  CAL ) 2.4 0.2 0.2
assign (resid  66 and   name     O  ) (resid  102 and   name  CAL ) 2.4 0.2 0.2
```

Appendix V

^{1}H and ^{15}N Chemical Shift and J-Coupling Values for Calcium Saturated cbEGF12-13, pH 6.5 and 33°C

Stereospecific assignments made based on floating chirality are marked with an asterisk. It should be noted, however, that the floating chirality protocol is designed purely to ensure consistency of NMR data with stereoassignments and not to identify stereoassignments. The latter is not possible if a chiral centre has the IUPAC standard conformation in the resultant structures, as this could have resulted either because it was chosen based on favourable energetics or because the conformation was not swapped due to lack of any energetic preference.

Residue Name	Resdiue Number	Atom Identifier	Chemical Shift Position	J-coupling values (Hz)
Thr	1069	N	113.19	
Thr	1069	HN	8.12	
Thr	1069	H^{α}	4.25	
Thr	1069	H^{β}	4.13	
Thr	1069	$C^{\gamma 2}(H^{\gamma 2})_3$	1.15	
Asp	1070	N	122.83	6.2
Asp	1070	HN	8.27	
Asp	1070	H^{α}	4.64	
*Asp	1070	$H^{\beta 2/\beta 3}$	2.68/2.40	
Ile	1071	N	122.05	8.5
Ile	1071	HN	7.72	
Ile	1071	H^{α}	4.13	
Ile	1071	H^{β}	1.67	
Ile	1071	$H^{\gamma 12/\gamma 13}$	1.33	
Ile	1071	$C^{\delta 1}(H^{\delta 1})_3$	1.05	
Ile	1071	$C^{\gamma 2}(H^{\gamma 2})_3$	0.73	
Asp	1072	N	127.02	
Asp	1072	HN	8.58	
Asp	1072	H^{α}	4.81	
*Asp	1072	$H^{\beta 2/\beta 3}$	2.85/2.29	
Glu	1073	N	125.67	
Glu	1073	HN	9.66	
Glu	1073	H^{α}	3.50	
Glu	1073	$H^{\beta 2/\beta 3}$	1.29/1.47	
Glu	1073	$H^{\gamma 2/\gamma 3}$	2.04/2.49	
Cys	1074	N	112.32	4.4
Cys	1074	HN	7.63	
Cys	1074	H^{α}	4.17	
Cys	1074	$H^{\beta 2/\beta 3}$	3.04/3.12	
Arg	1075	H^{α}	4.03	
Arg	1075	$H^{\beta 2/\beta 3}$	1.60/1.81	
Arg	1075	$H^{\gamma 2}$	1.54	
Arg	1075	$H^{\gamma 3}$	1.54	
Arg	1075	$H^{\delta 2}$	3.15	
Arg	1075	$H^{\delta 3}$	3.15	
Ile	1076	N	117.39	6.3
Ile	1076	HN	7.24	
Ile	1076	H^{α}	3.77	
Ile	1076	H^{β}	1.70	

Ile	1076	$H^{\gamma 12/\gamma 13}$	0.72/1.08	
Ile	1076	$C^{\delta 1}(H^{\delta 1})_3$	1.46	
Ile	1076	$C^{\gamma 2}(H^{\gamma 2})_3$	0.82	
Ser	1077	N	113.41	
Ser	1077	HN	7.34	
Ser	1077	H^{α}	4.85	
Ser	1077	$H^{\beta 2/\beta 3}$	3.64	
Pro	1078	H^{α}	4.47	
Pro	1078	$H^{\beta 2/\beta 3}$	1.90/2.29	
Pro	1078	$H^{\gamma 2/\gamma 3}$	1.90	
Pro	1078	$H^{\delta 2/\delta 3}$	3.45/3.64	
Asp	1079	N	116.41	8.5
Asp	1079	HN	8.22	
Asp	1079	H^{α}	4.81	
*Asp	1079	$H^{\beta 2/\beta 3}$	2.69/2.58	
Leu	1080	N	121.17	
Leu	1080	HN	7.36	
Leu	1080	H^{α}	3.87	
Leu	1080	$H^{\beta 2/\beta 3}$	1.36/1.59	
Leu	1080	H^{γ}	1.36	
Leu	1080	$C^{\delta 1/\delta 2}(H^{\delta 1/\delta 2})_3$	0.34	
Cys	1081	N	111.60	6.8
Cys	1081	HN	8.54	
Cys	1081	H^{α}	4.47	
Cys	1081	$H^{\beta 2/\beta 3}$	2.57/2.75	
Gly	1082	N	109.58	
Gly	1082	HN	7.79	
Gly	1082	$H^{\alpha 2/\alpha 3}$	3.56/4.05	
Arg	1083	N	131.37	6.8
Arg	1083	HN	8.39	
Arg	1083	H^{α}	4.06	
Arg	1083	$H^{\beta 2/\beta 3}$	0.55/1.15	
Arg	1083	$H^{\gamma 2}$	0.95	
Arg	1083	$H^{\gamma 3}$	0.95	
Arg	1083	$H^{\delta 2/\delta 3}$	2.84/2.91	
Gly	1084	N	108.13	
Gly	1084	HN	7.85	
Gly	1084	$H^{\alpha 2/\alpha 3}$	3.89/4.18	
Gln	1085	N	118.38	9.0
Gln	1085	HN	8.28	
Gln	1085	H^{α}	4.52	
Gln	1085	$H^{\beta 2/\beta 3}$	1.85/2.00	
Gln	1085	$H^{\gamma 2/\gamma 3}$	2.27/2.33	
Gln	1085	$N^{\varepsilon 2}$	108.79	
Gln	1085	$H^{\varepsilon 21/\varepsilon 22}$	6.73/7.42	
Cys	1086	N	126.40	6.4
Cys	1086	HN	8.79	
Cys	1086	H^{α}	5.06	
Cys	1086	$H^{\beta 2/\beta 3}$	2.75/3.04	
Val	1087	N	131.99	9.4
Val	1087	HN	9.61	
Val	1087	H^{α}	4.20	
Val	1087	H^{β}	1.94	
Val	1087	$C^{\gamma 1/\gamma 2}(H^{\gamma 1/\gamma 2})_3$	0.84	
Asn	1088	N	126.29	5.2
Asn	1088	HN	9.02	
Asn	1088	H^{α}	4.88	
Asn	1088	$H^{\beta 2/\beta 3}$	2.66/3.02	

Asn	1088	$N^{\delta 2}$	109.19		
Asn	1088	$H^{\delta 21/\delta 22}$	6.76/7.57		
Thr	1089	N	118.12	8.7	
Thr	1089	HN	7.93		
Thr	1089	H^{α}	4.87		
Thr	1089	H^{β}	4.28		
Thr	1089	$C^{\gamma 2}(H^{\gamma 2})_3$	0.99		
Pro	1090	H^{α}	4.26		
Pro	1090	$H^{\beta 2/\beta 3}$	1.94/2.21		
Pro	1090	$H^{\gamma 2/\gamma 3}$	1.88/2.06		
Pro	1090	$H^{\delta 2/\delta 3}$	3.57/3.77		
Gly	1091	N	114.75		
Gly	1091	HN	9.03		
Gly	1091	$H^{\alpha 2/\alpha 3}$	3.91/4.05		
Asp	1092	N	120.86		
Asp	1092	HN	7.36		
Asp	1092	H^{α}	4.80		
Asp	1092	$H^{\beta 2/\beta 3}$	2.50/2.58		
Phe	1093	N	119.05	5.1	
Phe	1093	HN	8.64		
Phe	1093	H^{α}	5.11		
Phe	1093	$H^{\beta 2/\beta 3}$	3.04/3.39		
Phe	1093	$H^{\delta 1}$	6.88		
Phe	1093	$H^{\delta 2}$	6.88		
Phe	1093	$H^{\varepsilon 1}$	7.11		
Phe	1093	$H^{\varepsilon 2}$	7.11		
Phe	1093	H^{ζ}	7.20		
Glu	1094	N	118.17		
Glu	1094	HN	9.45		
Glu	1094	H^{α}	4.43		
*Glu	1094	$H^{\beta 2/\beta 3}$	1.94/1.71		
Glu	1094	$H^{\gamma 2/\gamma 3}$	2.03		
Cys	1095	N	117.08	8.9	
Cys	1095	HN	8.73		
Cys	1095	H^{α}	5.46		
Cys	1095	$H^{\beta 2/\beta 3}$	2.62		
Lys	1096	N	125.62	9.0	
Lys	1096	HN	9.19		
Lys	1096	H^{α}	4.55		
Lys	1096	$H^{\beta 2/\beta 3}$	1.69/1.84		
Lys	1096	$H^{\gamma 2/\gamma 3}$	0.83/1.27		
Lys	1096	$H^{\delta 2/\delta 3}$	1.42		
Cys	1097	N	125.16		
Cys	1097	HN	8.91		
Cys	1097	H^{α}	4.87		
Cys	1097	$H^{\beta 2/\beta 3}$	2.64/3.05		
Asp	1098	N	124.33	4.2	
Asp	1098	HN	8.70		
Asp	1098	H^{α}	4.46		
Asp	1098	$H^{\beta 2/\beta 3}$	2.34/2.78		
Glu	1099	N	119.31		
Glu	1099	HN	8.38		
Glu	1099	H^{α}	4.13		
Glu	1099	$H^{\beta 2}$	2.00		
Glu	1099	$H^{\beta 3}$	2.00		
Glu	1099	$H^{\gamma 2}$	2.35		
Glu	1099	$H^{\gamma 3}$	2.35		
Gly	1100	N	111.18		

Gly	1100	HN	8.72		
Gly	1100	$H^{\alpha 2/\alpha 3}$	3.31/4.47		
Tyr	1101	N	118.27	9.5	
Tyr	1101	HN	8.35		
Tyr	1101	H^{α}	5.10		
Tyr	1101	$H^{\beta 2/\beta 3}$	2.49/3.17		
Tyr	1101	$H^{\delta 1}$	6.72		
Tyr	1101	$H^{\delta 2}$	6.72		
Tyr	1101	$H^{\varepsilon 1}$	6.75		
Tyr	1101	$H^{\varepsilon 2}$	6.75		
Glu	1102	N	116.72	8.1	
Glu	1102	HN	9.42		
Glu	1102	H^{α}	4.77		
Glu	1102	$H^{\beta 2/\beta 3}$	1.89/1.98		
Glu	1102	$H^{\gamma 2/\gamma 3}$	2.12/2.16		
Ser	1103	N	119.51	7.6	
Ser	1103	HN	8.78		
Ser	1103	H^{α}	4.89		
Ser	1103	$H^{\beta 2/\beta 3}$	3.69/3.86		
Gly	1104	N	111.13		
Gly	1104	HN	8.14		
Gly	1104	$H^{\alpha 2/\alpha 3}$	3.81/4.07		
Phe	1105	N	123.34	5.7	
Phe	1105	HN	8.51		
Phe	1105	H^{α}	4.47		
Phe	1105	$H^{\beta 2}$	3.08		
Phe	1105	$H^{\beta 3}$	3.08		
Phe	1105	$H^{\delta 1}$	7.23		
Phe	1105	$H^{\delta 2}$	7.23		
Phe	1105	$H^{\varepsilon 1}$	7.33		
Phe	1105	$H^{\varepsilon 2}$	7.33		
Phe	1105	H^{ζ}	7.27		
Met	1106	N	122.00		
Met	1106	HN	8.60		
Met	1106	H^{α}	3.92		
Met	1106	$H^{\beta 2}$	1.86		
Met	1106	$H^{\beta 3}$	1.86		
Met	1106	$H^{\gamma 2/\gamma 3}$	1.91/2.03		
Met	1107	N	112.37	7.0	
Met	1107	HN	7.95		
Met	1107	H^{α}	4.00		
Met	1107	$H^{\beta 2}$	2.18		
Met	1107	$H^{\beta 3}$	2.18		
Met	1107	$H^{\gamma 2/\gamma 3}$	2.44/2.53		
Met	1108	N	119.31	7.0	
Met	1108	HN	7.97		
Met	1108	H^{α}	4.44		
Met	1108	$H^{\beta 2}$	2.21		
Met	1108	$H^{\beta 3}$	2.21		
Met	1108	$H^{\gamma 2/\gamma 3}$	2.53/2.65		
Lys	1109	N	119.77	5.8	
Lys	1109	HN	8.46		
Lys	1109	H^{α}	4.03		
Lys	1109	$H^{\beta 2}$	1.68		
Lys	1109	$H^{\beta 3}$	1.68		
Lys	1109	$H^{\gamma 2}$	1.15		
Lys	1109	$H^{\gamma 3}$	1.15		
Lys	1109	$H^{\delta 2/\delta 3}$	1.09/1.28		

Lys	1109	$H^{\varepsilon 2/\varepsilon 3}$	2.46/2.49	
Asn	1110	N	113.10	
Asn	1110	HN	8.11	
Asn	1110	H^{α}	4.64	
Asn	1110	$H^{\beta 2/\beta 3}$	2.64/2.74	
Asn	1110	$N^{\delta 2}$	110.04	
Asn	1110	$H^{\delta 21/\delta 22}$	6.65/7.46	
Cys	1111	N	117.75	6.4
Cys	1111	HN	7.98	
Cys	1111	H^{α}	4.52	
Cys	1111	$H^{\beta 2/\beta 3}$	2.71/2.82	
Met	1112	N	124.69	9.2
Met	1112	HN	9.20	
Met	1112	H^{α}	4.95	
Met	1112	$H^{\beta 2/\beta 3}$	2.07/2.14	
Met	1112	$H^{\gamma 2/\gamma 3}$	2.55/2.60	
Asp	1113	N	122.26	2.1
Asp	1113	HN	8.94	
Asp	1113	H^{α}	3.93	
Asp	1113	$H^{\beta 2/\beta 3}$	2.20/2.55	
Ile	1114	N	125.78	7.7
Ile	1114	HN	7.58	
Ile	1114	H^{α}	3.68	
Ile	1114	H^{β}	1.26	
Ile	1114	$H^{\gamma 12/\gamma 13}$	0.64/1.26	
Ile	1114	$C^{\delta 1}(H^{\delta 1})_3$	0.49	
Ile	1114	$C^{\gamma 2}(H^{\gamma 2})_3$	0.60	
Asp	1115	N	127.02	8.6
Asp	1115	HN	8.63	
Asp	1115	H^{α}	4.82	
*Asp	1115	$H^{\beta 2/\beta 3}$	2.84/2.27	
Glu	1116	N	126.19	2.3
Glu	1116	HN	9.60	
Glu	1116	H^{α}	3.35	
Glu	1116	$H^{\beta 2/\beta 3}$	1.23/1.40	
Glu	1116	$H^{\gamma 2/\gamma 3}$	1.99/2.49	
Cys	1117	N	112.84	5.6
Cys	1117	HN	7.33	
Cys	1117	H^{α}	4.12	
Cys	1117	$H^{\beta 2/\beta 3}$	3.03/3.07	
Gln	1118	N	117.60	6.6
Gln	1118	HN	7.18	
Gln	1118	H^{α}	3.90	
Gln	1118	$H^{\beta 2/\beta 3}$	1.78/2.00	
Gln	1118	$H^{\gamma 2/\gamma 3}$	2.16/2.32	
Gln	1118	$N^{\varepsilon 2}$	108.74	
Gln	1118	$H^{\varepsilon 21/\varepsilon 22}$	6.69/7.50	
Arg	1119	N	116.46	5.6
Arg	1119	HN	7.44	
Arg	1119	H^{α}	3.95	
Arg	1119	$H^{\beta 2/\beta 3}$	1.62/1.72	
Arg	1119	$H^{\gamma 2}$	1.51	
Arg	1119	$H^{\gamma 3}$	1.51	
Arg	1119	$H^{2/\delta 3}$	2.83/3.05	
Asp	1120	N	115.53	9.3
Asp	1120	HN	7.46	
Asp	1120	H^{α}	4.96	
Asp	1120	$H^{\beta 2/\beta 3}$	2.24/2.81	

Pro	1121	H^{α}	4.51	
Pro	1121	$H^{\beta 2/\beta 3}$	1.87/2.36	
Pro	1121	$H^{\gamma 2/\gamma 3}$	1.88/1.95	
Pro	1121	$H^{\delta 2/\delta 3}$	3.48/3.79	
Leu	1122	N	117.86	8.7
Leu	1122	HN	8.05	
Leu	1122	H^{α}	4.26	
Leu	1122	$H^{\beta 2}$	1.87	
Leu	1122	$H^{\beta 3}$	1.87	
Leu	1122	H^{γ}	1.50	
Leu	1122	$C^{\delta 1/\delta 2}(H^{\delta 1/\delta 2})_3$	0.74/0.86	
Leu	1123	N	121.17	3.9
Leu	1123	HN	7.20	
Leu	1123	H^{α}	3.99	
Leu	1123	$H^{\beta 2/\beta 3}$	1.30/1.47	
Leu	1123	H^{γ}	1.47	
Leu	1123	$C^{\delta 1/\delta 2}(H^{\delta 1/\delta 2})_3$	0.25/0.38	
Cys	1124	N	115.84	
Cys	1124	HN	8.83	
Cys	1124	H^{α}	3.74	
Cys	1124	$H^{\beta 2/\beta 3}$	2.65/2.86	
Arg	1125	N	121.17	2.1
Arg	1125	HN	7.74	
Arg	1125	H^{α}	3.98	
Arg	1125	$H^{\beta 2/\beta 3}$	1.79/1.98	
Arg	1125	$H^{\gamma 2/\gamma 3}$	1.68/1.75	
Arg	1125	$H^{2/\delta 3}$	3.18/3.21	
Gly	1126	N	114.18	
Gly	1126	HN	8.84	
Gly	1126	$H^{\alpha 2/\alpha 3}$	3.58/4.20	
Gly	1127	N	110.56	
Gly	1127	HN	8.39	
Gly	1127	$H^{\alpha 2/\alpha 3}$	3.93/4.52	
Val	1128	N	122.36	9.9
Val	1128	HN	8.60	
Val	1128	H^{α}	4.22	
Val	1128	H^{β}	1.77	
Val	1128	$C^{\gamma 1/\gamma 2}(H^{\gamma 1/\gamma 2})_3$	0.51/0.80	
Cys	1129	N	128.31	7.9
Cys	1129	HN	8.79	
Cys	1129	H^{α}	5.11	
Cys	1129	$H^{\beta 2/\beta 3}$	2.14/3.01	
His	1130	N	128.05	9.3
His	1130	HN	9.49	
His	1130	H^{α}	4.87	
His	1130	$H^{\beta 2/\beta 3}$	2.91/3.00	
His	1130	$H^{\varepsilon 1}$	6.77	
Asn	1131	N	125.41	4.6
Asn	1131	HN	9.36	
Asn	1131	H^{α}	4.83	
Asn	1131	$H^{\beta 2/\beta 3}$	2.61/2.96	
Asn	1131	$N^{\delta 2}$	121.01	
Asn	1131	$H^{\delta 21/\delta 22}$	7.88/8.07	
Thr	1132	N	117.50	
Thr	1132	HN	7.98	
Thr	1132	H^{α}	4.65	
Thr	1132	H^{β}	4.18	
Thr	1132	$C^{\gamma 2}(H^{\gamma 2})_3$	0.92	

Glu	1133	N	123.40	
Glu	1133	HN	8.68	
Glu	1133	H^{α}	3.82	
Glu	1133	$H^{\beta 2/\beta 3}$	1.94/2.09	
Glu	1133	$H^{\gamma 2}$	2.27	
Glu	1133	$H^{\gamma \beta}$	2.27	
Gly	1134	N	119.82	
Gly	1134	HN	9.11	
Gly	1134	$H^{\alpha 2/\alpha 3}$	3.08/4.26	
Ser	1135	N	112.16	5.4
Ser	1135	HN	6.55	
Ser	1135	H^{α}	4.35	
Ser	1135	$H^{\beta 2/\beta 3}$	3.32/3.82	
Tyr	1136	N	117.50	4.5
Tyr	1136	HN	8.96	
Tyr	1136	H^{α}	5.00	
*Tyr	1136	$H^{\beta 2/\beta 3}$	3.42/2.88	
Tyr	1136	$H^{\delta 1}$	6.68	
Tyr	1136	$H^{\delta 2}$	6.68	
Tyr	1136	$H^{\varepsilon 1}$	6.64	
Tyr	1136	$H^{\varepsilon 2}$	6.64	
Arg	1137	N	119.46	8.2
Arg	1137	HN	9.59	
Arg	1137	H^{α}	4.60	
Arg	1137	$H^{\beta 2}$	1.68	
Arg	1137	$H^{\beta 3}$	1.68	
Arg	1137	$H^{\gamma 2}$	1.44	
Arg	1137	$H^{\gamma \beta}$	1.44	
Arg	1137	$H^{\delta 2}$	3.00	
Arg	1137	$H^{\delta 3}$	3.00	
Cys	1138	N	118.79	8.4
Cys	1138	HN	8.98	
Cys	1138	H^{α}	5.72	
Cys	1138	$H^{\beta 2/\beta 3}$	2.60/2.66	
Glu	1139	N	124.53	9.1
Glu	1139	HN	9.25	
Glu	1139	H^{α}	4.65	
Glu	1139	$H^{\beta 2/\beta 3}$	1.90/2.12	
Glu	1139	$H^{\gamma 2/\gamma 3}$	2.04/2.24	
Cys	1140	N	125.10	
Cys	1140	HN	9.03	
Cys	1140	H^{α}	5.32	
Cys	1140	$H^{\beta 2/\beta 3}$	2.56/2.78	
Pro	1141	H^{α}	4.75	
Pro	1141	$H^{\beta 2/\beta 3}$	1.72/2.38	
Pro	1141	$H^{\gamma 2/\gamma 3}$	1.65/1.97	
Pro	1141	$H^{\delta 2/\delta 3}$	3.58/3.84	
Pro	1142	H^{α}	4.35	
Pro	1142	$H^{\beta 2/\beta 3}$	1.87/2.32	
Pro	1142	$H^{\gamma 2/\gamma 3}$	2.03/2.11	
Pro	1142	$H^{\delta 2/\delta 3}$	3.57/3.84	
Gly	1143	N	112.01	
Gly	1143	HN	8.83	
Gly	1143	$H^{\alpha 2/\alpha 3}$	3.62/4.12	
His	1144	N	119.51	
His	1144	HN	7.92	
His	1144	H^{α}	4.89	
His	1144	$H^{\beta 2/\beta 3}$	2.89/3.00	

His	1144	$H^{\varepsilon 1}$	6.82	
His	1144	$H^{\delta 2}$	7.21	
Gln	1145	N	118.69	8.2
Gln	1145	HN	8.80	
Gln	1145	H^{α}	4.63	
Gln	1145	$H^{\beta 2/\beta 3}$	1.90/2.00	
Gln	1145	$H^{\gamma 2/\gamma 3}$	2.21/2.28	
Gln	1145	$N^{\varepsilon 2}$	111.65	
Gln	1145	$H^{\varepsilon 21/\varepsilon 22}$	6.73/7.43	
Leu	1146	N	124.43	4.6
Leu	1146	HN	8.55	
Leu	1146	H^{α}	4.43	
Leu	1146	$H^{\beta 2/\beta 3}$	1.40/1.71	
Leu	1146	H^{γ}	1.50	
Leu	1146	$C^{\delta 1/\delta 2}(H^{\delta 1/\delta 2})_3$	0.75/0.83	
Ser	1147	N	121.84	6.1
Ser	1147	HN	8.82	
Ser	1147	H^{α}	4.85	
Ser	1147	$H^{\beta 2/\beta 3}$	3.94/4.18	
Pro	1148	H^{α}	4.33	
Pro	1148	$H^{\beta 2/\beta 3}$	1.96/2.30	
Pro	1148	$H^{\gamma 2/\gamma 3}$	1.99	
Pro	1148	$H^{\delta 2/\delta 3}$	3.90/4.18	
Asn	1149	N	108.96	
Asn	1149	HN	8.40	
Asn	1149	H^{α}	4.75	
Asn	1149	$H^{\beta 2/\beta 3}$	3.84	
Asn	1149	$N^{\delta 2}$	112.63	
Asn	1149	$H^{\delta 21/\delta 22}$	6.92/7.55	
Ile	1150	N	113.35	
Ile	1150	HN	7.36	
Ile	1150	H^{α}	4.40	
Ile	1150	H^{β}	1.95	
Ile	1150	$H^{\gamma 12/\gamma 13}$	1.29	
Ile	1150	$C^{\delta 1}(H^{\delta 1})_3$	0.81	
Ile	1150	$C^{\gamma 2}(H^{\gamma 2})_3$	1.05	
Ser	1151	N	116.36	
Ser	1151	HN	8.08	
Ser	1151	H^{α}	4.16	
Ser	1151	$H^{\beta 2/\beta 3}$	3.81/4.11	
Ala	1152	N	122.67	7.8
Ala	1152	HN	7.62	
Ala	1152	H^{α}	4.88	
Ala	1152	$C^{\beta}(H^{\beta})_3$	1.26	
Cys	1153	N	119.98	7.9
Cys	1153	HN	8.84	
Cys	1153	H^{α}	4.88	
Cys	1153	$H^{\beta 2/\beta 3}$	2.57/2.69	
Ile	1154	N	127.69	9.4
Ile	1154	HN	8.91	
Ile	1154	H^{α}	4.24	
Ile	1154	H^{β}	1.96	
Ile	1154	$H^{\gamma 12/\gamma 13}$	1.15/1.42	
Ile	1154	$C^{\delta 1}(H^{\delta 1})_3$	0.75	
Ile	1154	$C^{\gamma 2}(H^{\gamma 2})_3$	0.86	

Appendix VI

^{15}N-T_1, ^{15}N-T_2 and ^{15}N-Heteronuclear NOE Data for cbEGF12-13 at (a) 500MHz and (b) 750MHz, pH 6.5 and 35°C

(a) 500MHz Data

Residue	T_1 (ms)	σT_1 (ms)	T_2 (ms)	σT_2 (ms)	NOE	σNOE
D4	-	-	302.48	31.4741	-0.282	0.0754
I5	-	-	178.83	3.5817	0.415	0.0343
D6	470.37	40.2666	132.71	12.3996	0.63	0.0867
E7	460.41	18.8656	121.42	4.4964	0.667	0.0896
C8	419.99	11.4655	91.79	2.4183	0.69	0.0459
I10	480.08	4.3790	-	-	0.738	0.0227
D13	-	-	137.06	2.7615	0.593	0.0376
L14	-	-	-	-	0.596	0.0309
C15	478.24	12.5792	104.54	2.6336	0.621	0.0509
G16	-	-	120.18	3.2063	0.721	0.049
R17	453.93	23.6957	-	-	0.594	0.1013
G18	449.84	8.4990	-	-	0.688	0.0361
Q19	-	-	-	-	0.694	0.0414
C20	475.96	16.7641	73.14	2.4823	0.722	0.0706
V21	424.27	29.5206	57.94	4.6222	0.776	0.1495
N22	446.43	13.5523	105.80	3.0669	0.724	0.0546
T23	443.85	7.8802	133.16	2.4113	0.652	0.0391
G25	355.49	21.9892	109.75	7.5998	0.704	0.0759
D26	-	-	-	-	0.653	0.0336
F27	414.08	10.8021	135.83	3.4871	0.726	0.042
E28	458.30	19.1131	113.78	4.3756	0.684	0.0783
C29	482.39	11.6351	124.52	2.8994	0.661	0.0506
K30	468.82	14.2867	130.79	3.8316	0.704	0.0547
C31	-	-	127.55	9.8266	0.653	0.0724
D32	433.28	11.6391	128.44	3.3156	0.767	0.0454
E33	-	-	141.08	2.0303	0.676	0.0327
G34	417.19	12.0092	153.78	4.4929	0.711	0.0486
Y35	-	-	-	-	0.694	0.0414
E36	438.40	13.2617	134.32	4.0232	0.713	0.0545
S37	432.53	13.2826	126.37	3.7051	0.665	0.054
G38	-	-	160.10	3.5117	0.544	0.044
F39	-	-	124.63	4.2867	0.524	0.0508
M40	447.63	15.4285	128.90	3.0904	0.542	0.063
M41	-	-	62.38	3.8294	0.537	0.084
M42	-	-	-	-	0.538	0.0435
K43	496.52	19.2298	155.69	5.4781	0.61	0.063
C45	-	-	-	-	0.703	0.0455
M46	-	-	-	-	0.669	0.0553
D47	441.89	11.7161	132.21	3.5656	0.732	0.045
I48	417.19	13.4015	140.85	4.4832	0.705	0.0513

D49	417.19	15.8382	122.38	4.5233	0.752	0.0554
E50	437.45	11.0988	141.40	3.3591	0.694	0.0473
C51	429.55	10.1484	-	-	0.763	0.0386
Q52	480.08	4.3790	-	-	0.733	0.0197
R53	448.23	7.4337	-	-	0.696	0.034
D54	460.62	8.9111	150.88	2.9592	0.689	0.0364
L56	-	-	116.82	1.1737	0.731	0.0207
L57	494.32	8.5522	115.29	1.8608	0.72	0.0348
C58	445.24	21.0129	46.05	2.4195	0.627	0.089
R59	464.25	6.4659	131.32	1.7590	0.627	0.0282
G60	400.16	22.5781	103.41	6.0101	0.677	0.0677
G61	488.76	16.7219	103.82	3.6109	0.791	0.0614
V62	440.53	10.0914	126.50	2.8645	0.737	0.0431
C63	459.14	11.8052	132.35	3.2403	0.706	0.0484
H64	-	-	141.98	4.0319	0.694	0.049
N65	-	-	136.26	7.8907	0.707	0.0556
T66	-	-	-	-	0.683	0.0384
E67	448.43	44.6420	119.83	14.6182	0.85	0.0913
G68	416.67	14.5833	133.71	4.3800	0.715	0.0504
S69	-	-	-	-	0.73	0.0285
Y70	426.44	8.9107	133.28	2.9310	0.719	0.0333
R71	405.19	11.6565	130.70	3.5874	0.697	0.0514
C72	458.93	13.6899	88.21	2.5209	0.688	0.0546
E73	-	-	-	-	0.669	0.0553
C74	397.77	39.8722	138.29	16.4476	0.711	0.089
H78	579.71	14.1147	-	-	0.554	0.0718
Q79	-	-	-	-	0.559	0.0427
L80	483.56	9.1193	159.82	3.1418	0.542	0.035
S81	445.24	10.9029	99.02	2.3826	0.597	0.0489
N83	416.49	46.1421	-	-	-	-
S85	-	-	131.77	5.7993	0.582	0.051
A86	-	-	125.94	2.4110	0.588	0.0368
C87	480.31	14.7645	125.64	3.5835	0.655	0.0515
I88	485.20	14.5961	145.29	4.1794	0.543	0.0498

(b) 750MHz Data

Residue	T_1 (ms)	σT_1 (ms)	T_2 (ms)	σT_2 (ms)	NOE	σNOE
D4	901.71	21.1403	224.62	6.4076	-0.031	0.0231
I5	-	-	160.15	1.4107	0.545	0.0122
D6	747.38	22.9019	95.51	3.2567	0.76	0.0278
E7	630.91	18.3105	91.50	2.7293	0.818	0.0372
C8	646.83	9.2046	68.30	1.1056	0.77	0.0224
I10	634.92	4.0312	147.84	1.0054	0.845	0.0096
D13	-	-	125.11	1.0644	0.642	0.0127
L14	762.20	4.6475	113.68	0.6978	0.67	0.0122
C15	-	-	74.99	1.1189	0.742	0.0225
G16	-	-	93.66	1.3597	0.744	0.0207
R17	696.38	26.1869	35.03	1.8667	0.786	0.0549
G18	651.04	5.9340	126.07	1.2556	0.794	0.0154
Q19	698.81	5.8601	132.82	1.1467	0.847	0.0141

C20	654.02	13.2601	50.01	1.1628	0.826	0.0307
V21	648.09	42.0018	37.00	3.3409	0.801	0.0743
N22	651.47	9.3370	78.39	1.2414	0.802	0.0239
T23	625.00	5.4688	126.74	1.1084	0.764	0.0138
G25	643.50	21.1187	74.01	2.5418	0.745	0.0342
D26	-	-	101.99	0.7905	0.795	0.014
F27	-	-	129.74	1.5653	0.799	0.017
E28	664.89	15.9150	93.53	2.3268	0.778	0.0339
C29	689.18	7.5995	114.35	1.3599	0.761	0.0194
K30	710.23	11.6017	116.82	1.9652	0.813	0.0245
C31	686.34	20.2558	92.73	2.9064	0.802	0.0297
D32	615.38	7.1953	122.34	1.3919	0.812	0.0158
E33	674.76	4.0978	133.89	0.8067	0.787	0.0116
G34	621.12	8.1015	144.43	1.9190	0.784	0.0171
Y35	703.23	5.9345	132.82	1.1467	0.847	0.0141
E36	613.50	8.2803	126.02	1.7470	0.822	0.0189
S37	-	-	108.66	1.3932	0.805	0.0197
G38	672.04	7.6779	144.22	1.7263	0.68	0.0169
F39	649.35	8.4331	87.87	1.2584	0.668	0.0195
M40	-	-	119.10	2.5393	0.616	0.023
M41	730.46	24.5443	38.30	1.8978	0.614	0.0444
M42	-	-	89.53	1.0742	0.606	0.0173
K43	705.72	13.4470	93.21	1.8069	0.672	0.0262
C45	-	-	-	-	0.814	0.0139
M46	657.03	10.7922	109.52	1.8111	0.779	0.0234
D47	666.22	7.5455	122.43	1.3940	0.809	0.018
I48	611.25	8.2197	131.23	1.7911	0.819	0.0175
D49	677.97	9.6524	115.38	1.8504	0.816	0.0196
E50	621.50	8.1116	131.61	1.8015	0.82	0.0168
C51	669.79	6.2807	109.60	1.0691	0.822	0.0152
Q52	688.23	2.8420	114.30	0.4834	0.807	0.0072
R53	644.75	4.5727	144.11	1.0592	0.777	0.0127
D54	648.09	5.0402	140.13	1.0408	0.778	0.0129
L56	-	-	102.89	0.5187	0.782	0.0072
L57	717.88	4.6381	95.37	0.6731	0.783	0.0127
C58	630.52	27.0335	29.76	1.7097	0.762	0.0579
R59	691.09	4.2984	111.22	0.6804	0.718	0.0116
G60	688.23	17.5255	74.60	2.0201	0.771	0.0284
G61	697.35	19.4519	74.42	2.0936	0.831	0.0353
V62	632.11	5.5939	107.65	1.0199	0.772	0.0146
C63	-	-	123.09	1.4849	0.802	0.0175
H64	624.61	8.9732	118.11	1.7297	0.814	0.0176
N65	658.33	13.8687	115.35	2.5681	0.827	0.021
T66	-	-	-	-	0.814	0.0139
E67	683.53	22.8932	111.97	4.3755	0.794	0.028
G68	-	-	127.06	1.9536	0.804	0.0192
S69	-	-	135.96	1.1091	0.816	0.0106
Y70	-	-	124.47	1.1000	0.813	0.0129
R71	636.54	9.3191	122.10	1.7741	0.813	0.0164
C72	701.75	9.8492	58.41	0.8938	0.775	0.023
E73	657.03	10.7922	112.93	1.9130	0.779	0.0234
C74	739.64	26.8067	126.07	4.9430	0.754	0.0309
H78	-	-	-	-	0.452	0.048

Q79	764.53	10.5210	141.56	1.9038	0.631	0.0177
L80	-	-	138.10	0.9918	0.644	0.0124
S81	-	-	69.13	1.0656	0.71	0.0243
N83	674.31	55.4725	-	-	-	-
S85	647.25	11.7301	100.47	2.0189	0.695	0.0195
A86	-	-	102.32	0.8899	0.72	0.0133
C87	641.85	11.5351	103.96	1.8265	0.748	0.0273
I88	-	-	120.12	2.0778	0.701	0.0279