Re-engineering *Caenorhabditis elegans* SAS-6: from spirals to rings



Jodie Ford

Merton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Acknowledgements

I would like to thank my supervisor, Prof. John Vakonakis, his help and advice during my DPhil have been invaluable. Thank you to Dr. Phillip Stansfeld for his help with all things computers. Thanks to Merton college and the Medical Sciences Division for funding my DPhil, I am grateful for the opportunity. Thank you also to everyone in the lab, past and present, for all the troubleshooting. Special thanks to Jemma and Julia, my office craft buddies, and Beatrice and Natassa for their postdoc wisdom.

Thank you to Lucy and Lisa for the cycling-based distractions throughout the last 3 and a half years. I also wish to thank my Mom, Dad and Gran for their patience and support, and Meekho, Lanu, Baja, Rambo and Otie for always making me smile. Finally, I would like to thank my sister, Lucy, for being the best.

Abstract

Centrioles have conserved 9-fold symmetry. This symmetry is templated by a scaffold of SAS-6. SAS-6 oligomerises to form a stack of rings, each with 9-fold radial symmetry, that forms the cartwheel, which imparts symmetry to the organelle. However, crystallographic and electron microscopy studies suggest that *C. elegans* is unique in that it forms a spiral SAS-6 scaffold instead of a cartwheel. Thus, there are two competing architectures of SAS-6 scaffold. Inconclusive electron microscopy of *C. elegans* centrioles mean it is unclear which architecture is present *in vivo*. *C. elegans* is used as a model for centriole biology, it is therefore important to determine which SAS-6 oligomer structure is present at the centre of the centriole because a spiral scaffold would have implications for centriole duplication and elongation. To probe the nature of the *C. elegans* SAS-6 scaffold, I use a protein engineering approach to convert spiral-forming SAS-6 into a ring oligomer. Molecular dynamics simulations of SAS-6 are used to screen the architecture of the SAS-6 mutants. The conformation of the engineered SAS-6 oligomers *in vitro* is determined using atomic force microscopy. I show that there is flexibility in the architectures sampled by wild-type *C. elegans* SAS-6 and that two point mutations are sufficient to increase the number of rings formed by *C. elegans* SAS-6 by >20%.

Contents

List of Figures **xi**

List of Abbreviations **xv**

1 Introduction **1**
1.1 Centriole function and role in disease 1
1.2 Centriole architecture 3
1.3 Centriole duplication and assembly 7
1.4 Spindle assembly abnormal protein 6 (SAS-6) 14
1.5 Cartwheels and spirals 19
1.6 Implications of a spiral scaffold 23
1.7 Robustness of centriole symmetry 29
1.8 Aims and objectives 32

2 Methods **33**
2.1 Molecular dynamics simulations 33
2.2 Bioinformatics 34
2.3 Cloning 35
2.4 Protein expression and purification 40
2.5 Analytical Size Exclusion Chromatography 43
2.6 Circular Dichroism 43
2.7 Electron Microscopy 44
2.8 Atomic Force Microscopy (AFM) 45
2.9 Image analysis 45
2.10 Graphs and statistics 46

3 SAS-6 oligomer architecture **47**
3.1 Introduction 47
3.2 *C. elegans* SAS-6 spirals 48
3.3 Heterogeneity of SAS-6 oligomers 49
3.4 Molecular dynamics simulations of SAS-6 oligomers 54
3.5 Discussion 62

4 SAS-6 mutant design **63**
4.1 Introduction . . . 63
4.2 Sequence analysis . . . 64
4.3 SAS-6 mutations . . . 67
4.3.1 Negative design elements . . . 68
4.3.2 Positive design elements . . . 75
4.3.2.1 Salt-bridges . . . 75
4.3.2.2 Hydrophobic interactions . . . 76
4.3.2.3 Hydrogen bonds . . . 76
4.4 Discussion . . . 77

5 Molecular dynamics simulations of SAS-6 variants **79**
5.1 Introduction . . . 79
5.2 Control simulations . . . 80
5.3 CeSAS-6 mutant molecular dynamics simulations . . . 86
5.4 Discussion . . . 95

6 Atomic force microscopy of CeSAS-6 oligomers **97**
6.1 Introduction . . . 98
6.2 Electron microscopy of SAS-6 oligomers . . . 98
6.3 Atomic force microscopy optimisation of SAS-6 oligomers . . . 103
6.3.1 Optimisation of imaging conditions . . . 103
6.3.1.1 Control samples . . . 103
6.3.1.2 Construct optimisation . . . 105
6.3.2 Troubleshooting CeSAS-6 variant AFM . . . 108
6.3.2.1 Concentration optimisation . . . 121
6.3.3 Image analysis and metrics . . . 124
6.4 AFM of CeSAS-6 variants . . . 126
6.4.1 Ring diameter . . . 126
6.4.2 Backbone width . . . 130
6.4.3 Spoke separation . . . 132
6.4.4 Curvature . . . 135
6.4.5 Backbone periodicity . . . 137
6.5 *In silico* analysis of CeSAS-6APVG . . . 139
6.5.1 Molecular dynamics simulations of CeSAS-6APVG . . . 139
6.5.2 CeSAS-6APVG dihedral angle analysis . . . 141
6.6 Round II CeSAS-6APVG variants . . . 145
6.6.1 Molecular dynamics simulations of round II CeSAS-6 variants 145
6.6.2 AFM and troubleshooting of round II CeSAS-6APVG variants 149
6.7 Discussion . . . 152

7 Discussion **155**
7.1 Evolution of spirals . . . 155
7.2 Interconversion from spiral to ring . . . 158
7.3 Future work . . . 162
7.4 Conclusion . . . 162

Appendices

A Image analysis and modelling scripts **169**
A.1 Threaded models . . . 170
A.2 Image analysis scripts . . . 172

B Crystallisation attempts CeSAS-6APVG **173**
B.1 Introduction . . . 173
B.2 Methods . . . 173
B.3 Initial screening . . . 175
B.4 Crystallisation condition optimisation . . . 175

C Characterisation of RBM14-STIL interaction **177**
C.1 Introduction . . . 177
C.2 Methods . . . 181
C.2.1 Bioinformatic analysis . . . 181
C.2.2 Cloning . . . 181
C.2.3 Protein purification . . . 183
C.2.4 Analytical size exclusion chromatography . . . 184
C.2.5 NMR . . . 184
C.2.6 Fluorescence polarisation . . . 185
C.3 RNA binding protein 14 (RBM14) . . . 185
C.4 RBM14 and STIL interaction studies . . . 188
C.4.1 NMR . . . 188
C.4.2 Fluorescence polarisation . . . 191

References **193**

List of Figures

1.1 Overview of centriole architecture. 5
1.2 Comparison of centriole architecture across species. 6
1.3 Centriole assembly and duplication cycle. 9
1.4 The cell cycle and centrosome duplication cycle. 13
1.5 Structure of SAS-6. 16
1.6 Interactions at the SAS-6 hinge. 17
1.7 Comparison of spiral and cartwheel scaffolds. 20
1.8 Electron microscopy of centrioles. 22
1.9 XLF and XRCC4 have structurally similar N-terminal domains to SAS-6. 27

3.1 SAS-6 hinge. 49
3.2 SAS-6 oligomer symmetry. 51
3.3 SAS-6 oligomer models. 52
3.4 Variation in oligomer symmetry produced by flexibility in the C-C SAS-6 dimer interface. 56
3.5 Variation in oligomer pitch/turn produced by flexibility in the C-C SAS-6 dimer interface. 57
3.6 Variation in oligomer radius produced by flexibility in the C-C SAS-6 dimer interface. 58
3.7 Flexibility in α-3 position. 59
3.8 SAS-6 oligomer symmetry heterogeneity 61

4.1 SAS-6 covariance analysis. 66
4.2 Location of mutations on SAS-6 dimer. 67
4.3 Mutations introduced to CeSAS-6 (i) 69
4.4 Mutations introduced to CeSAS-6 (ii) 70
4.5 CeSAS-6 Mutants (i) 71
4.6 CeSAS-6 Mutants (ii) 72
4.7 CeSAS-6 Mutants (iii) 73
4.8 CeSAS-6 Mutants (iii) 74

5.1 Control molecular dynamics simulations 81

5.2 Molecular dynamics simulations of CeSAS-6 V167G C-C dimer. . . 83
5.3 Molecular dynamics simulation of LmSAS-6 with G142V mutation. 85
5.4 Stability of molecular dynamics simulations. 87
5.5 Symmetry sampled by CeSAS-6 variants (i). 88
5.6 Symmetry sampled by CeSAS-6 variants (ii). 89
5.7 Radii sampled by CeSAS-6 variants (i). 90
5.8 Radii sampled by CeSAS-6 variants (ii). 91
5.9 Pitch/turn sampled by CeSAS-6 variants (i). 92
5.10 Pitch/turn sampled by CeSAS-6 variants (ii). 93
5.11 Rotation of G142V. 96

6.1 SAS-6 constructs. 99
6.2 Electron microscopy of CeSAS-6 oligomers. 100
6.3 AFM height profiles of SAS-6 control samples. 104
6.4 CeSAS-6 protein purification. 106
6.5 CeSAS-6 construct optimisation for AFM. 107
6.6 CeSAS-6 spirals are fragile. 108
6.7 CeSAS-6 protein purification. 111
6.8 AFM height traces of unsuccessful CeSAS-6 variants (i). 112
6.9 AFM height traces of unsuccessful CeSAS-6 variants (ii). 113
6.10 AFM height traces of unsuccessful CeSAS-6 variants (iii). 114
6.11 Optimisation of ionic strength used for adsorption. 115
6.12 CD of CeSAS-6 variants. 116
6.13 Analytical size exclusion chromatography of SAS-6 mutants. 118
6.14 AFM concentration optimisation. 120
6.15 CeSAS-6APVG protein purification. 122
6.16 SAS-6 concentration and oligomer conformation. 123
6.17 SAS-6 architecture analysis metrics. 125
6.18 AFM height profiles of SAS-6. 127
6.19 AFM image analysis. 128
6.20 SAS-6 ring diameter. 129
6.21 SAS-6 backbone width. 131
6.22 SAS-6 oligomer spoke separation. 133
6.23 Curvature of SAS-6 oligomers. 136
6.24 Backbone periodicity of SAS-6 oligomers. 138
6.25 Molecular dynamics simulations of CeSAS-6APVG C-C dimer. . . . 140
6.26 Dihedral angles sampled by residues neighbouring the hinge. 142
6.27 CeSAS-6APVG models with forced ring-like dihedrals. 144
6.28 Symmetry sampled by round II CeSAS-6APVG variants. 146
6.29 Radii sampled by round II CeSAS-6APVG variants. 147

6.30 Pitch/turn sampled by round II CeSAS-6APVG variants. 148
6.31 AFM height traces of round II CeSAS-6APVG variants. 150
6.32 CD round II CeSAS-6APVG E54R H175E. 151

7.1 Phylogenetic tree of SAS-6 variants. 157
7.2 Mechanism of ring formation – end joining. 163
7.3 Mechanism of ring formation – loop out. 165
7.4 Mechanism of ring formation – unwinding of the spiral. 166

A.1 Example script used for comparative modelling. 170
A.2 Sequence alignments used to generate threaded models. 171
A.3 Example scripts used in image analysis. 172

B.1 Constructs for crystallisation trials. 174
B.2 Crystallisation trials of CeSAS-6APVG. 176

C.1 RBM14 and STIL protein domains. 179
C.2 RBM14 constructs and purification. 187
C.3 Analytical size exclusion chromatography of RBM14 and STIL. . . 188
C.4 NMR interaction studies of RBM14 and STIL. 190
C.5 Fluorescence polarisation studies of RBM14 and STIL. 192

List of Abbreviations

AFM . . . . . Atomic force microscopy

CD . . . . . Circular dichroism

Cep135 . . . . Centrosomal protein 135 kDa

CeSAS-6 . . . *C. elegans* SAS-6

CeSAS-6VG . *C. elegans* SAS-6 V167G

CeSAS-6APVG *C. elegans* SAS-6 A166P V167G

CPAP . . . . . centrosomal P4.1-associated protein

CrSAS-6 . . . *Chlamydomonas reinhardtii* SAS-6

DmSAS-6 . . . *Drosophila melanogaster* SAS-6

DrSAS-6 . . . *Danio rerio* SAS-6

EM . . . . . Electron microscopy

LmSAS-6 . . . *Leishmania major* SAS-6

PCM . . . . . Pericentriolar material

Plk4 . . . . . Polo-like kinase 4

RBM-14 . . . RNA binding protein 14

SAS-4 . . . . . Spindle assembly abnormal protein 4

SAS-5 . . . . . Spindle assembly abnormal protein 5

SAS-6 . . . . . Spindle assembly abnormal protein 6

SAS-7 . . . . . Spindle assembly abnormal protein 7

SPD-2 . . . . . spindle-defective protein 2

STIL . . . . . SCL-interrupting locus protein

SUMO . . . . small ubiquitin-related modifier

TBG-1 . . . . Tubulin gamma 1

XLF . . . . . XRCC4-like factor

XRCC4 . . . . x-ray repair cross-complimenting protein 4

ZYG-1 . . . . . zygote defective protein 1

Work in this thesis is included in the following publication:

Ford, J. E., Stansfeld, P. J. and Vakonakis, I. Coupling Form and Function: How the Oligomerisation Symmetry of the SAS-6 Protein Contributes to the Architecture of Centriole Organelles, *Symmetry-Basel*, 9.5, 2017.

"...an organising role to the hub-and-spokes placing the cartwheel before the horse"

— R. V. Dippell *1968* [1]

1 Introduction

Contents

1.1 Centriole function and role in disease 1
1.2 Centriole architecture 3
1.3 Centriole duplication and assembly 7
1.4 Spindle assembly abnormal protein 6 (SAS-6) 14
1.5 Cartwheels and spirals 19
1.6 Implications of a spiral scaffold 23
1.7 Robustness of centriole symmetry 29
1.8 Aims and objectives 32

1.1 Centriole function and role in disease

The centriole was first described by Theodor Boveri more than 140 years ago [2]. In that time many advances have been made in our understanding of its structure and function. However, its description as the '*terra incognita*' of the cell still holds somewhat [3]. The centriole is a conserved organelle in eukaryotes but is not found in higher plants or yeast [4]. Two centrioles, together with the pericentriolar material (PCM), an amorphous cloud of protein, comprise the centrosome. The centrosome functions as the main microtubule organising centre in the cell, directing formation of the mitotic spindle thereby ensuring correct chromosome segregation during cell division. Indeed, loss of centrosomes disrupts chromosome segregation

and leads to loss of cell viability [5, 6] in *Xenopus leavis* [7], *Caenorhabditis elegans* [8, 9] and *Drosophila melanogaster* [10]. While cell division is possible without centrosomes in vertebrates, it is slow [11] and leads to defects in cytokinesis and cell cycle progression [5, 6].

The centrosome also plays a role in many other cellular processes. For example, the centrosome contributes to cell polarity. In *C. elegans* embryos the centrosome contributed by the sperm cell triggers cell polarity and thus determines the anterior-posterior axis of the embryo [12, 13]. The centrosome further contributes to cell polarity through intracellular trafficking via its interaction with the Golgi apparatus [14, 15].

The semi-conservative nature of centriole duplication plays an important role in asymmetric cell division. Each centrosome is made up of an older centriole from the previous cell generation (the mother), and a daughter centriole. Centrioles are contributed by sperm cells and can persist through many cell generations; as such, they can act as a non-genetic carrier of information. The inheritance of faulty centrosomes can thus impact development [16]. Centriole duplication produces centrosomes that contain centrioles of different ages and the resulting centrosome asymmetry provides a marker for asymmetric cell division [17, 18]. Asymmetric cell division is important in brain development; in humans, impaired centriole assembly causes autosomal recessive primary microcephaly, a neurodevelopmental disorder characterised by reduced brain size [19]. Centrosomes also act as centres of cellular signalling. They aid in control of cell cycle progression with their maturation being linked to the action of several cell cycle components [20]. Additionally, in DNA damage recognition, the centrosome acts as a localisation point for proteins involved in DNA damage signalling pathways [21].

Disruption of centrosome structure and number can lead to disease [22]. Abnormal centrosome amplification, for example as a result of loss of p53, leads to multipolar spindle assembly causing chromosome instability and aneuploidy which is linked to tumorigenesis [23]. Often, tumours display centrosome abnormalities

though it unknown whether this is a byproduct of other cancerous phenotypes or a contributing factor [24].

In G0 phase mature mother centrioles can be converted to basal bodies that template cilia and flagella. Basal bodies lie just below the plasma membrane and nucleate the growth of axoneme microtubules. It is the distal appendages of mature centrioles which are responsible for the docking of the centriole on the plasma membrane [25]. In multi-ciliated cell types the strict control of replicating the centriole only once per cell cycle must be broken in order to produce many cenrtioles to template cillia. This is achieved via the use of deuterosomes. Deutersomes nucleate formation of many centrioles which are then released into the cytoplasm from where they are able to dock onto the plasma membrane [26, 27]. Defects in motile cilia and flagella lead to a variety of diseases, such as bronchitis and sinusitis as a result of dysfunction of mucus-clearing cilia, and male sterility as a result of sperm immotility [28, 29]. Impairment of the signalling function of primary cilia also gives rise to disease. For example, kidney disease can result from perturbation of the mechanosensation role of primary cilia in the kidney [30].

1.2 Centriole architecture

Centrioles are cylindrical organelles with canonical 9-fold radial symmetry (Fig. 1.1). The characteristic 9-fold symmetry of the centriole was first observed in the 1950s. Centriole size varies between species and cell type. In vertebrates, the centriole is approximately 250 nm in diameter and 500 nm in length (Fig. 1.2 B). Centrioles in *C. elegans* are smaller: 110-175 nm in diameter and 150-200 nm long [31–33] (Fig. 1.2 D). *Chlamydomonas reinhardtii* (green algae) centrioles are similar in width to vertebrate centrioles but slightly longer at 550 nm (Fig. 1.2 A). Further demonstrating the diversity that exists in centriole size, *D. melanogaster* somatic cell centrioles are very short at 100–150 nm in length, but approximately 250 nm in diameter, whereas centrioles in *D. melanogaster* sperm cells can reach up to 1 µm in length (Fig. 1.2 C).

The outer wall of the centriolar cylinder is formed by microtubule 'blades'. Each blade comprises a triplet of microtubules in vertebrates and *D. melanogaster* germline cells (or a doublet in *D. melanogaster* embryos) [34–36] and singlets in *C. elegans* (doublet microtubules are found in cilia of sensory neurons)[37, 38] (Fig. 1.2).

In most species the centriole is arranged around a cartwheel-shaped protein scaffold at its core. The 'cartwheel', so called because it comprises a central hub from which 9 spokes emanate radially, influences the symmetry of the organelle [39, 40]. A stack of rings, each with 9-fold symmetry, comprise the cartwheel (Fig. 1.1). In contrast to most species, *C. elegans* centrioles include a ring of electron density, termed the 'central tube', in place of the cartwheel stack [32]. This structure acts in an analogous manner to the cartwheel providing a scaffold around which the centriole forms. Immunolocalisaton electron microscopy and sub-diffraction fluorescence microscopy localised spindle assembly abnormal protein 6 (SAS-6) to the cartwheel at the centre of the centriole [41, 42]. SAS-6 is the focus of this thesis.

The centriole is polarised along its length with a proximal and distal end. The proximal end is the site of initiation of centriole assembly. Mature centrioles display appendages at the distal end which mediate membrane docking during ciliogenesis [43]. In vertebrates, the longitudinal axis can be divided into 3 sections: the distal end, where microtubule doublets are present, a central region, where microtubule triplets are located, and a proximal region, where the cartwheel is surrounded by microtubule triplets (Fig. 1.1) [31, 36, 44, 45]. This differs slightly in *C. elegans* where microtubule singlets and the central scaffold are maintained throughout the length of the centriole. This is also the case in *D. melanogaster* somatic cells.

In vertebrates, microtubule triplets are arranged such that the inner-most microtubule of each triplet (the A-tubule) is fully assembled and appended by 2 partially-assembled microtubules (the B- and C-tubules). The B-tubule attaches to the A-tubule and the C-tubule attaches to the B-tubule. Microtubule blades are linked via the A-C linker which joins the A-tubule to the C-tubule of the adjacent triplet (Fig. 1.1). The composition of this linker is unknown [46–49]. In *C. elegans* each microtubule is associated with an electron dense structure termed the

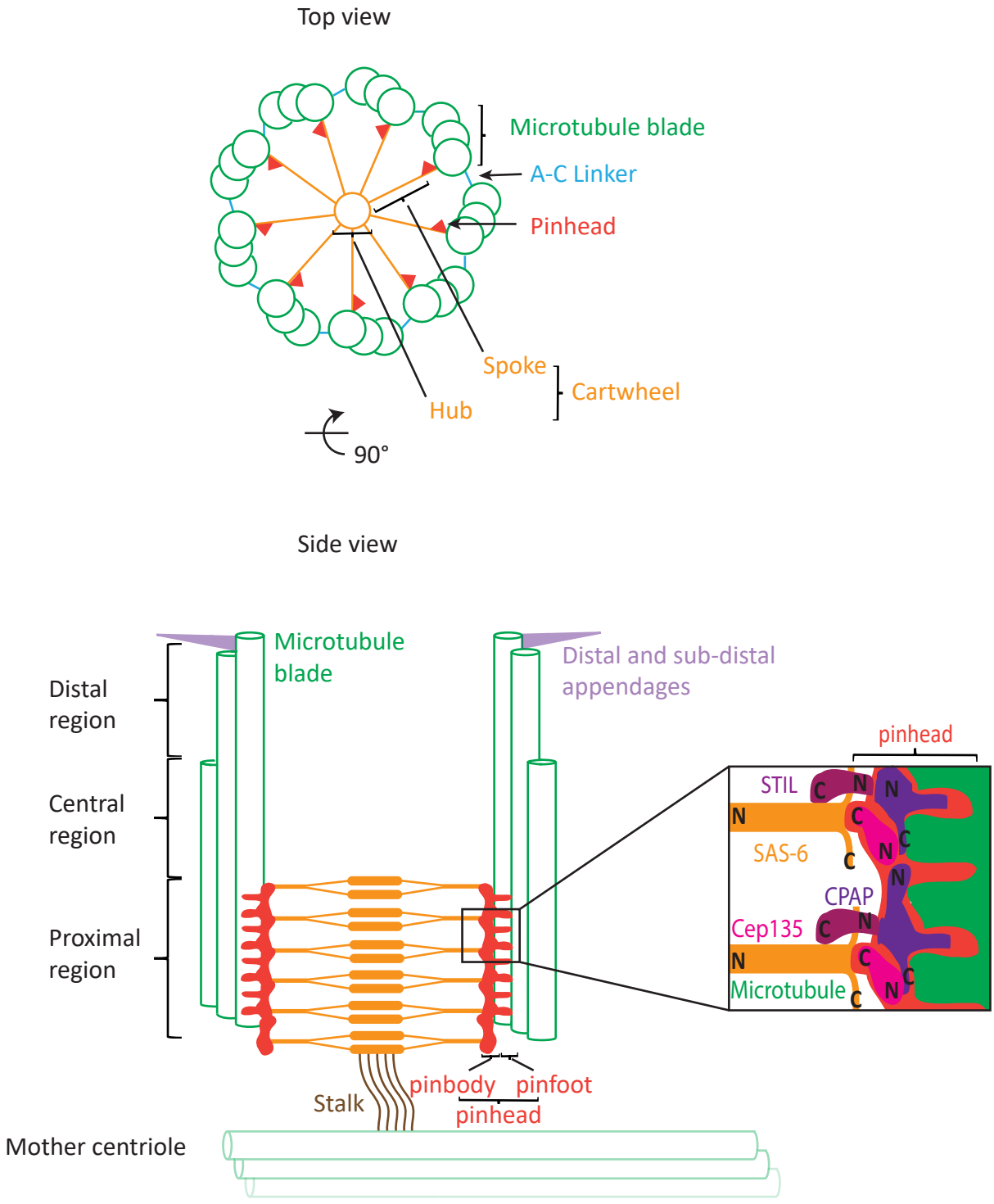


Figure 1.1: Overview of centriole architecture. Cartoon representation of a vertebrate centriole with a SAS-6 cartwheel scaffold (orange) interacting with the microtubule wall (green) via the pinhead (red). The A-C linker (blue) is shown. An enlarged view of the pinhead is shown illustrating the location of Cep135 (majenta), CPAP (dark purple) and STIL (light purple). The N and C termini of each pinhead protein is indicated. The three regions of the centriole are indicated in the lateral view, appendages (purple) and the stalk (brown), connecting mother and daughter centrioles are shown.

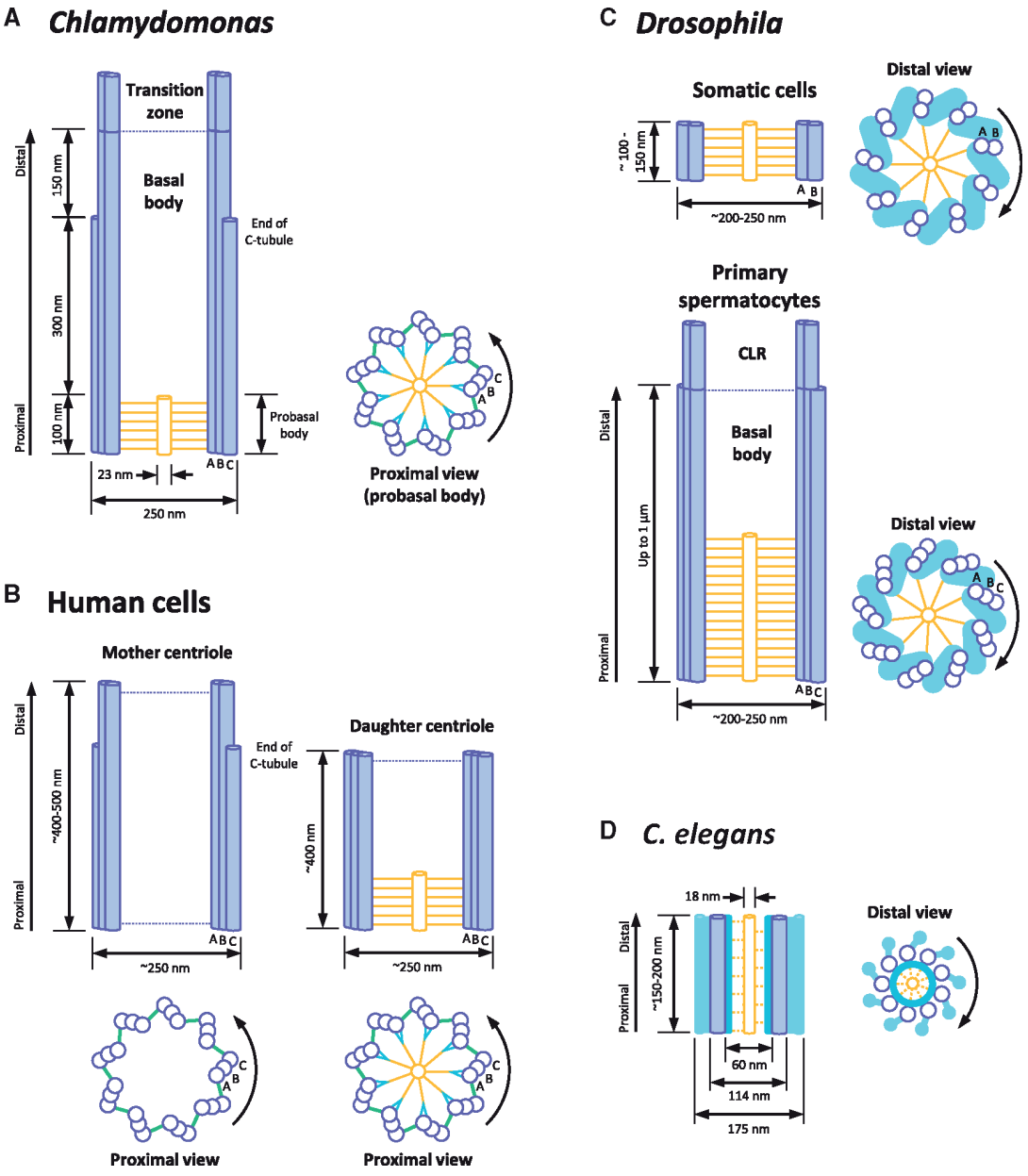


Figure 1.2: Comparison of centriole architecture across species. Schematic representation of centrioles from *C. reinhardtii* (A), *H. sapiens* (B), *D. melanogaster* (C) and *C. elegans* (D). The SAS-6 scaffold (orange) and mirotubule blades (dark blue) are shown. The central tube (blue) and paddlewheels (cyan) found in *C. elegans* centrioles are shown. Figure reprinted from The Journal of Biochemistry, 164, A. Gupta and D. Kitagawa, *Ultrastructural diversity between centrioles of eukaryotes*, 2018, with permission from Oxford University Press.

'paddlewheel' (Fig. 1.2 D). This was initially thought to be an equivalent structure to the appendages found in vertebrate centrioles [50], but has also been described as equivalent to the A-C linker between microtubules. Although the A-C linker alone is not sufficient to determine the symmetry of the organelle, it does dictate the angle between microtubule triplets thereby helping to reinforce centriole symmetry [35, 51]. The paddlewheels in *C. elegans* centrioles are twisted clockwise and act to provide polarity to the organelle as the triplet microtubules do in vertebrates [50].

A structure known as the 'pinhead' lies at the wall of the centriole between the spokes of SAS-6 and the circumferentially arranged microtubules (Fig. 1.1). The pinhead adds directionality along the proximal-distal axis of the centriole. It can be divided into a hook known as the 'pinbody' and 2 linkers termed 'pinfeet' which tilt towards the distal end of the centriole [46].

1.3 Centriole duplication and assembly

Aberrant duplication of centrioles can lead to diseases such as microcephaly and cancer [52]. It is therefore important that centriole duplication is strictly controlled and happens only once per cell cycle [53–55]. If duplication proceeds correctly, 4 centrioles will be present in each cell prior to mitosis as both centrioles in the centrosome can act as a template for centriole duplication. The mature centriole, from which the new centriole is assembled, is termed the 'mother' centriole, while the newly formed centriole is termed the 'daughter' centriole. The daughter centriole is assembled perpendicular to the mother centriole. A procentriole nucleates on the mother in early G1/S phase, the microtubule wall then assembles and the daughter centriole undergoes elongation in S and G2 phase. The centriole then matures and the new daughter and its mother separate from the other pair of centrioles for mitosis. In telophase/G1 phase the centrioles are disengaged when the physical link between the mother and daughter is broken [56–58]. The new centriole is then ready for licencing to duplicate if required in the next cell cycle.

Functional and genetic studies of spindle assembly defective mutants in *C. elegans* identified a set of 5 conserved proteins essential for centriole formation:

spindle-defective protein 2 (SPD-2), zygote defective protein 1 (ZYG-1), spindle assembly defective protein 4 (SAS-4), spindle assembly defective protein 5 (SAS-5) and spindle assembly defective protein 6 (SAS-6). All of these are coiled-coil proteins except the kinase ZYG-1. Functional equivalents were later identified in most eukaryotes [58–60].

Transmission electron microscopy studies identified SAS-7 as the main component of paddlewheels, a peripheral structure in *C. elegans*. It is the earliest acting component in centriole duplication [50] and makes the mother centriole competent for duplication (Fig. 1.3 A). SAS-7 is required for SPD-2 recruitment to the centriole in *C. elegans* [61]. No clear homologue of SAS-7 has been identified in other species, however its temporal localisation is similar to that of Cep-152 [62] which works with Cep-192 (SPD-2 in *C. elegans*) in PLK4 (ZYG-1 in *C. elegans*) recruitment in mammals. Although no sequence homology is shared between SAS-7 and Cep-295, Cep295 also works to recruit Cep192 (SPD-2) to centrioles [63] leading to the suggestion that SAS-7 is a functional analogue of Cep295 in vertebrates [61].

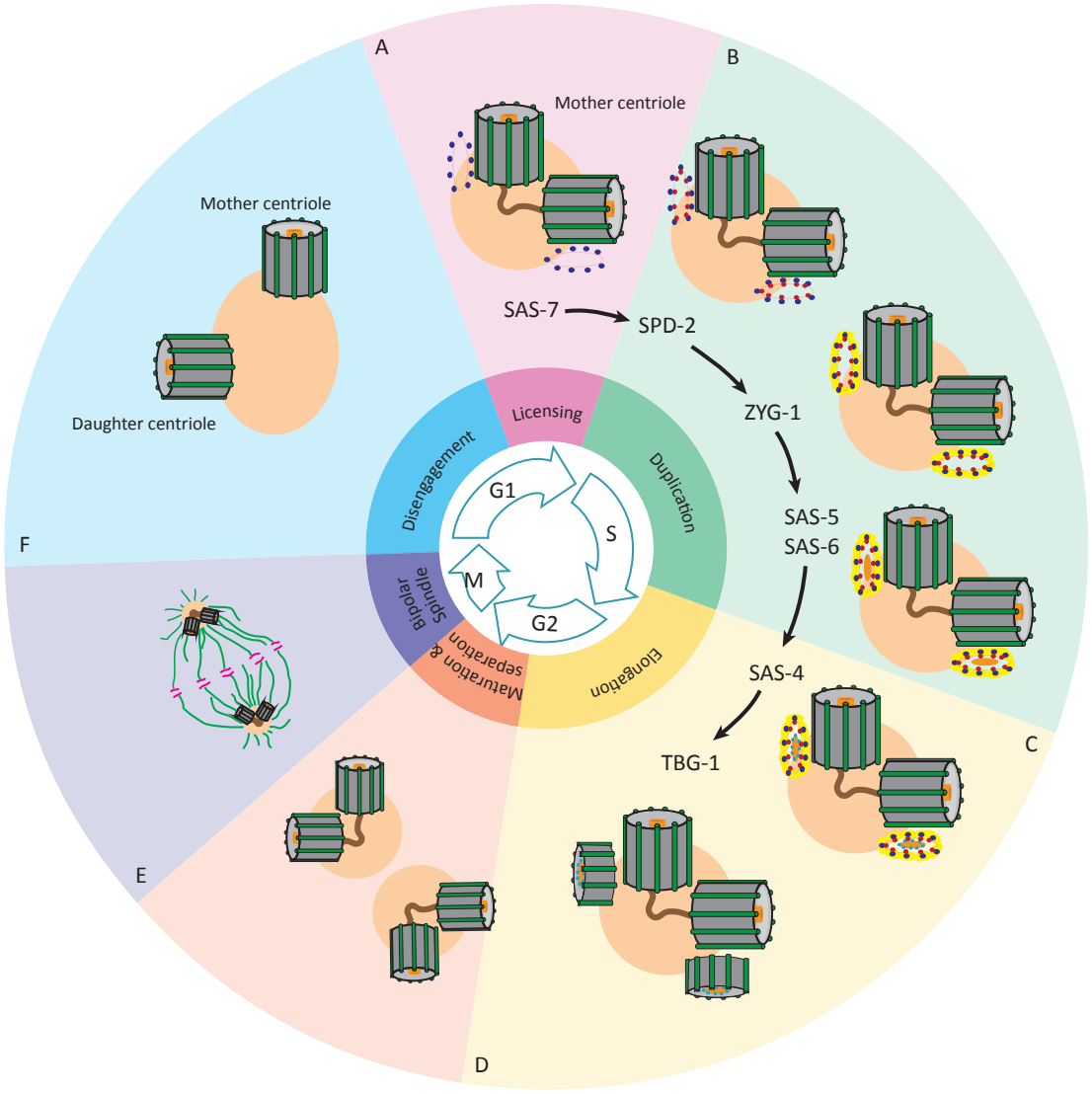


Figure 1.3: Centriole duplication cycle. Schematic representation of centriole duplication in *C. elegans*. The stages of centriole assembly as they occur at each phase of the cell cycle are indicated by coloured segments of the circle. A: Duplication is initiated in late G1/S phase following licensing by SAS-7 (blue). B: SPD-2 (red) is then recruited to the procentriole along with ZYG-1 (yellow). This is followed by SAS-5 and SAS-6 recruitment, forming the central tube (orange). C: SAS-4 (cyan) and TBG-1 (green) are then recruited. Procentriole elongation then occurs in late S/G2 phase. D: The two pairs of centrioles separate in G2/M phase, ready for spindle pole formation in mitosis (E). F: The stalk (brown) is broken down in late M/G2 phase, disengaging the mother and daughter centriole. The pericentriolar material is shown in light brown.

ZYG-1 is recruited by SPD-2 to the proximal end of the centriole where the procentriole forms (Fig. 1.3 B) [32, 35, 64, 65]. PLK4 forms a ring at the site of procentriole formation on the mother centriole. This ring organisation is achieved through the ability of PLK4 to self-assemble [66]. Autophosphorylation of an intrinsically disordered region in PLK4 regulates its ability to phase-separate. This condensation reduces the dissociation rate of PLK4 from the procentriole formation site, thereby causing a local increase in PLK4 concentration. This large PLK4 focus promotes trans-autophosphorylation of inactive PLK4, further increasing the levels of active PLK4 at a single locus on the mother centriole [67]. The resulting focus is the point at which SAS-5 (STIL in vertebrates) and SAS-6 are recruited to the mother centriole for procentriole formation. These proteins work together to initiate cartwheel formation [68]. ZYG-1 can also bind directly to SAS-6 [69, 70], and SAS-5 can homo-oligomerise to bridge several SAS-6 molecules together, binding to the coiled-coil region of SAS-6 (Fig. 1.5 A) [71]. It is only after recruitment of SAS-6 that ZYG-1 kinase activity is required for central tube or cartwheel formation. SAS-6/SAS-5 can only localise to the centriole when in complex together. SAS-5 freely exchanges with the cytoplasm but SAS-6 remains at the centre of the daughter centriole [72]. SAS-6/SAS-5 complex formation is an example of one of the stringent control mechanisms on centriole duplication [73]. SAS-6/SAS-5 complex formation is linked to the cell cycle by the action of Cdc6 which is recruited to centrosomes by cyclin A. Cdc6 binds SAS-6, inhibiting SAS-6/SAS-5 complex formation. In G1/S phase ZYG-1 (PLK4) phosphorylates Cdc6 releasing it from SAS-6, allowing SAS-5 to bind and centriole duplication to take place [74]. Once SAS-6/SAS-5 have been recruited scaffold assembly can take place. SAS-6 assembles to form the central tube in *C. elegans*. In other organisms SAS-6 rings stack at the core of the centriole forming the cartwheel [75]. RNAi studies in *C. elegans* embryos determined SAS-6 is required for centriole duplication [76, 77] demonstrating the importance of this scaffold.

In G2/S phase SAS-4 (CPAP in vertebrates) is recruited to the outer wall of the daughter centriole and triggers TBG-1 (γ-tubulin in vertebrates) recruitment

for microtubule assembly (Fig. 1.3 C) [78]. This stabilises the central tube of SAS-6/SAS-5 by creating a network where the C-terminus of SAS-4 binds to SAS-5 and the N-terminus of SAS-4 binds to microtubules [9, 79, 80]. In addition to CPAP, Cep135 is recruited to the pinhead at this point in vertebrate centriole assembly, however no functional orthologue has been identified in *C. elegans* [81]. A schematic view of the pinhead is shown in Fig. 1.1. As the structure and exact location of the pinhead proteins is unknown they are represented in cartoon form, with the N- and C-termini of each protein indicated. The pinhead proteins interact to form a network where the C-terminus of Cep-135 interacts with the C-terminus of SAS-6 and the N-terminus of Cep-135 interacts with the C-terminus of CPAP [81]. CPAP links the pinhead to the microtuble blades via its C-terminus [82]. The C-terminus of CPAP also interacts with the N-terminus of STIL and the C-terminus of STIL interacts with the N-terminus of α-3 in SAS-6 [83]. The interactions between STIL and CPAP are further discussed in Appendix C.

Centrosome maturation begins with accumulation of pericentriolar material components, an amorphous cloud at the proximal end of the centriole that includes proteins required for microtubule nucleation and anchoring [84]. During prophase spindle fibres form and the two mother centrioles move away from each other (Fig. 1.3 D). In metaphase the bipolar spindle forms ready for chromosome segregation (Fig. 1.3 E). In anaphase in *H. sapiens* PLK4 levels reach their maximum, and activate ubiquitination complexes. STIL and SAS-6 dissociate from the centriole and are degraded by proteolysis [70, 85], resulting in complete loss of the cartwheel by the end of mitosis [86]. The degradation of SAS-6 restricts centriole reduplication. In *C. elegans*, *C. reinhardtii* and *D. melanogaster* the SAS-6 scaffold is maintained in mature centrioles [76, 77].

In telophase centriole disengagement occurs where the physical link (the stalk) between the two centrioles is lost through the action of a number of factors including separase (Fig. 1.3 F) [87]. This licences the centriole for duplication in the next cell cycle [88]. Further maturation steps then occur where distal and sub-distal appendages form on the daughter centriole [89].

The components and order of recruitment required for centriole duplication described here may differ slightly between organisms, but assembly of the SAS-6 scaffold is always a critical early step in centriole assembly [90].

The cell cycle and centrosome duplication cycle must be closely linked to ensure the correct number of centrosomes are produces, and at the correct time, in the cell cycle. One way the centrosome cycle is linked to the cell cycle is via the use of the same regulatory components. During late G1/S phase CDK2, in complex with Cyclin E and Cyclin A, is recruited to the centrosome and working with PLK4 acts to initiate procentriole formation [91] (Fig. 1.4). This step links the centrosome cycle to the cell cycle by localising regulatory components of each cycle to the same place. This is the second mechanism by which the cell cycle and centrosome cycle are linked. There are 5 Cyclins and CDKs that localise to the centrosome: CDK1, CDK2, Cyclin A, Cyclin B and Cyclin E [92]. Cyclin E recruitment to the centrosome at G1-S phase is required for DNA replication while CDK2 recruitment to the centriole is required to initiate the centrosome cycle [93]. Another cycle regulator that is involved in the centrosome cycle is PLK1. PLK1 is involved in centrosome maturation during G2 phase and centrosome separation along with NIMA-regulated kinase 2 (Nek2), Nek6, Nek7 and Nek9, cell cycle regulated kinases [94]. Aurora A, a cell cycle kinase, is also involved in centrosome maturation in G2 phase [95]. It is regulated by Nucleophosmin/B23 [91]. A further example of the close link between the cell cycle and centrosome cycle is the action of centrosome-bound CDK1 in G2 phase – with PLK1 it activates Nek9 which is involved in centrosome separation, and it initiates nuclear envelope breakdown and chromosome condensation marking entry into mitosis. This localisation of cell cycle factors by the centrosome has been proposed to increase the rate at which cells enter mitosis by increasing the local concentration of Cyclin B1-CDK1 and PLK1 [96].

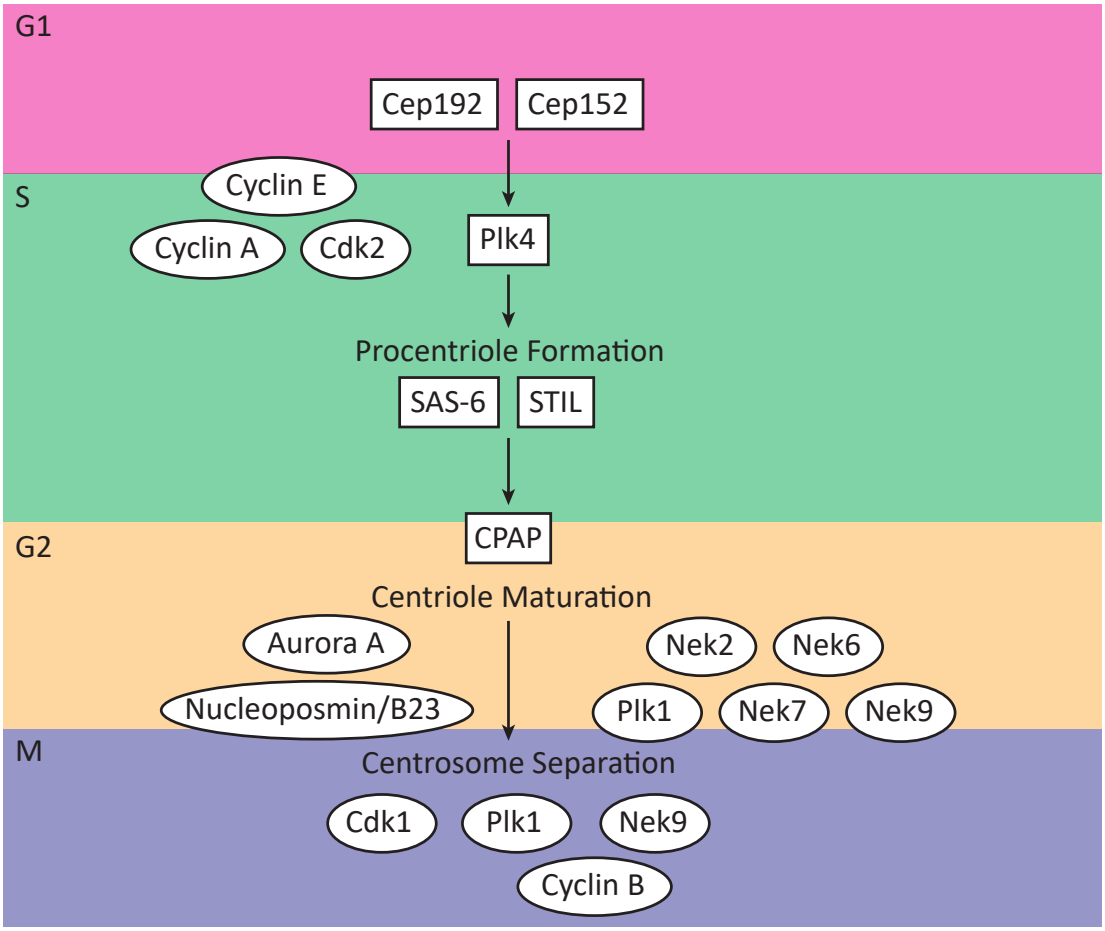


Figure 1.4: The cell cycle and centrosome duplication cycle. Schematic of the important proteins involved in the centrosome and cell cycles in mammalian cells. Protein components of the centriole are shown in boxes, cell cycle components that also play a role in the centrosome cycle are shown in elliptical boxes. Each phase of the cell cycle is indicated by the coloured background.

1.4 Spindle assembly abnormal protein 6 (SAS-6)

Crystallographic structures of SAS-6 revealed its structure is highly conserved across species, although it displays large sequence divergence. For example, *C. elegans* SAS-6 shares only 22% identity with *H. sapiens* SAS-6. SAS-6 comprises a globular N-terminal domain followed by a long coiled-coil domain and a disordered region at the C-terminus (Fig. 1.5 A). The Present in SAS-6 (PISA) domain at the N-terminus is the most conserved region of SAS-6 [97]. Mutations in this region in *H. sapiens* SAS-6 impair centriole formation and lead to autosomal recessive primary microcephaly [19]. SAS-6 dimerises strongly via the coiled-coil domain and forms oligomers via a second dimerisation event at the N-terminal domain (Fig. 1.5 B). The interaction between the N-terminus of one monomer with the N-terminus of the other monomer in the dimer (N-N interaction) forms the cartwheel hub. The long coiled-coil forms the spoke, projecting 40-45 nm from the hub. The N-N interaction is strengthened by a hydrophobic residue inserted into a pocket on the other monomer. In *C. elegans* this residue is I154 (Fig. 1.6 B), in *Leishmania major*, *C. reinhardtii* and *Danio rerio* it is a phenylalanine (Fig. 1.6 D). In *C. elegans* and *H. sapiens* an I154E mutation (or equivalent) abolishes centriole formation [98] and prevents formation of flagella in *C. reinhardtii* [41]. The N-N interface is weak with a dissociation constant of $\sim$110 μM in *C. elegans*, 50 μM in *D. rerio* and *C. reinhardtii* and 90 μM in *D. melanogaster* [40, 41]. In *C. elegans* there is a long loop between α2 and β5 which modulates N-N dimerisation and may be involved in the regulation of SAS-6 oligomerisation [99]. The disordered C-terminus of *C. reinhardtii* SAS-6 negatively regulates ring stacking *in vitro*, resulting in the stacking of rings where spokes from consecutive rings are out of register in procentriole formation. It is thought that microtubule blades then fix the cartwheel in place causing spoke registration [100].

Many crystallographic structures of SAS-6 variants are available of both of the coiled-coil and N-N dimerisation interfaces. SAS-6 crystallographic structures

of both interfaces are available for *C. elegans*, *D. rerio*, *C. reinhardtii* and *L. major*. Table 1.1 summaries the available crystallographic structures, detailing the species of origin, PDB code, construct, resolution, and the number of copies of SAS-6 in the asymmetric unit.

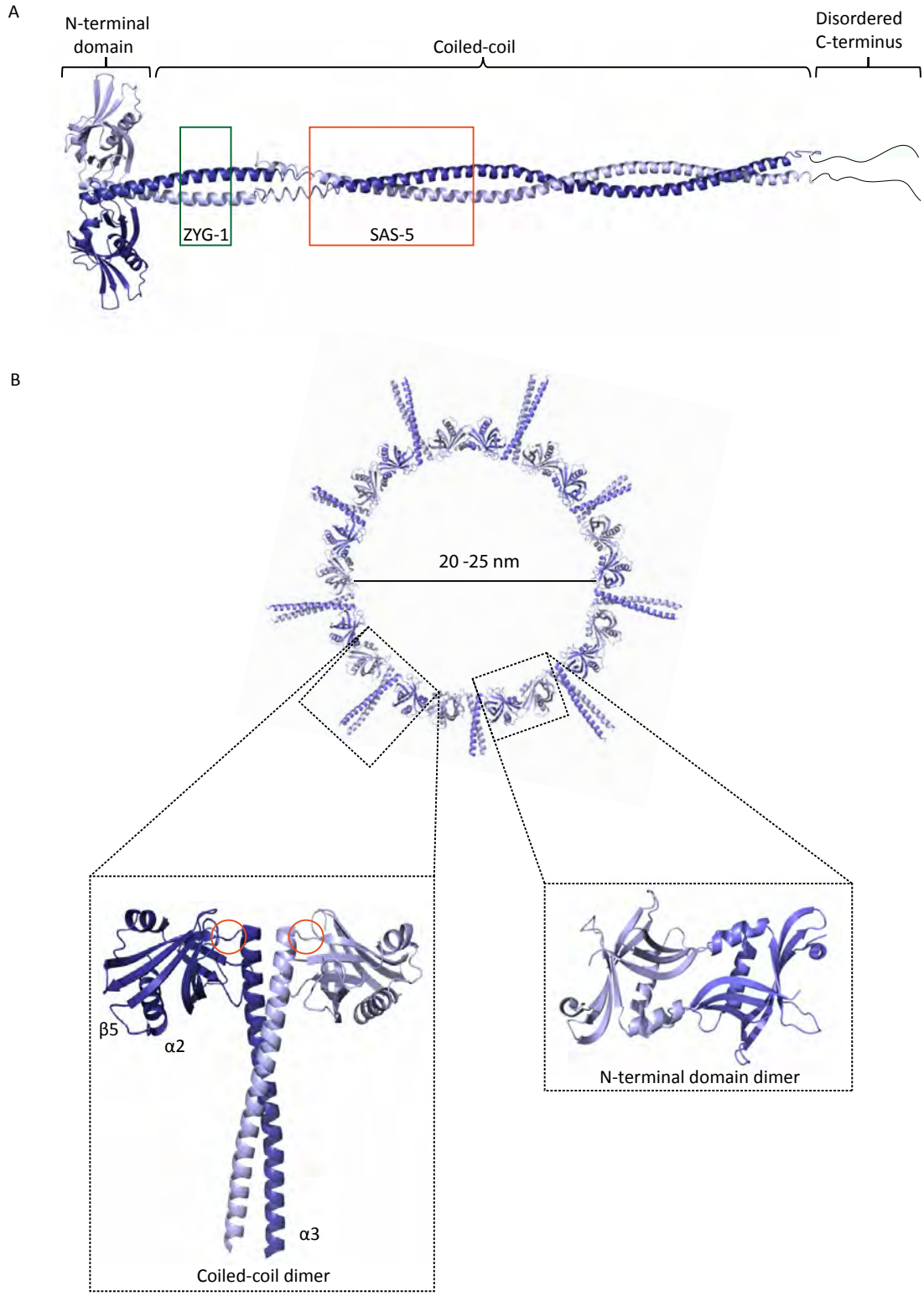


Figure 1.5: Structure of SAS-6. A: Full length SAS-6 dimer built from the crystallographic structures of the C-C dimer and the coiled coil (PDB codes: 4GFC and 4GKW). The N-terminal domain, coiled-coil and disordered C-terminus are indicated. The binding sites of SAS-5 and ZYG-1 are also shown. B: Shows the oligomerisation of SAS-6 coiled-coil and N-terminal domain dimerisation interfaces (*L. major* SAS-6 crystallographic structures used, PDB codes: 4CKP, 4CKN and 4CKM) to form a ring. The ring diameter is shown. The hinge is indicated by red circles on the coiled-coil dimer.

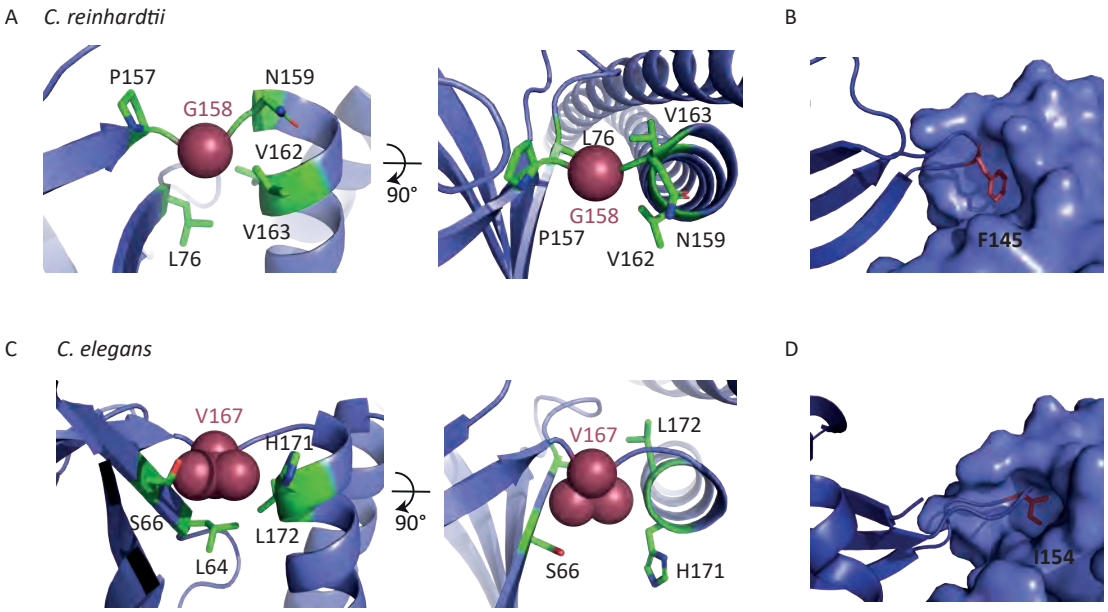


Figure 1.6: Interactions at the SAS-6 hinge. Enlarged view of the hinge region showing the residues surrounding the hinge of a ring-forming SAS-6 variant and a spiral forming SAS-6 variant. A: Residues (green) surrounding the *C. reinhardtii* SAS-6 hinge residue (G158) (red spheres) are shown. Residues N156, P157, V163, and L76 are shown in green. B: Enlarged view of the *C. reinhardtii* N-N SAS-6 dimer. F145 (red) that strengthens the N-N interaction is shown packing into hydrophobic pocket of the neighbouring N-terminal domain (surface). C: Residues (green) surrounding the *C. elegans* SAS-6 hinge residue (V167) (red spheres) are shown. Residues H171, L172, S66 and L64 are shown in green. D: Enlarged view of the *C. elegans* N-N SAS-6 dimer. I545 (red) that strengthens the N-N interaction is shown packing into hydrophobic pocket of the neighbouring N-terminal domain (surface). PDB codes: 4GFA, 3PYI, 3Q0X and 3Q0Y.

Table 1.1: Published SAS-6 crystallographic structures.

Species	**Interface**	**PDB code**	**Construct**	**Resolution (Å)**	**Copies**
C. elegans	C-C dimer	4GFC	1-215 ΔLoop 103-130 I154G	2.85	2
		4GFA	1-215 ΔLoop 103-130 ΔLoop 151-156	3.55	4
		4GKW	C-C fragment 256-342	3.3	2
	N-N dimer	4G79	1-142 ΔLoop 103-130	1.8	2
		3PYI	1-170 S123E	2.1	2
		4GEX	1-170 S123E I154W	2.8	4
		4GEU	1-170 S123E I154W	2.65	2
C. reinhardtii	C-C dimer	3Q0X	1-228 F154E	3.02	2
	N-N dimer	3Q0Y	1-161	2.1	6
		4TO7	1-161 Q93E F145W K146R (NN24)	1.25	1
		4TQ7	1-161 Q93E (NN10)	2.64	2
		4TPZ	1-161 F145W (NN2)	1.8	4
		4TTZ	1-161 G94E F145W Q147K (NN26)	2.5	4
		4TTY	1-161 G94D F145W Q147R (NN25)	1.9	2
		4TTX	1-161 G94C K146C (NN19)	2.5	6
		4TTW	1-161 K105C F145W (NN18)	1.2	2
		4U2J	1-161 Q93E F145W (NN27)	2.2	4
		4U2I	1-161 Q93E G94D K146R Q147R (NN23)	2.28	4
L. major	C-C dimer	4CKN	97-320 F257E	2.9	4
		4CKP	97-226	3.45	6
	N-N dimer	4CKM	97-320	2.15	2
D. rerio	C-C dimer	2Y3W	1-182 S102N	1.98	2
	N-N dimer	2Y3V	1-156	1.92	4
D. melanogaster	C-C dimer	5AL7	2-219 F143D	2.92	2

1.5 Cartwheels and spirals

In the majority of species studied, SAS-6 oligomers have been observed to form a ring with 9-fold symmetry. The symmetry of the SAS-6 scaffold imparts symmetry to the entire organelle. However, SAS-6 oligomers can sample symmetries that deviate from the canonical 9-fold. There is flexibility in both the N-N and coiled-coil (C-C) interfaces of SAS-6 which translates to variability in oligomer symmetry. This demonstrates that while the symmetry of the SAS-6 oligomer does influence the symmetry of the entire centriole, the connection to organelle symmetry is more complex.

The SAS-6 ring hub is 20-25 nm in diameter with a 40-45 nm long spoke. A stack of rings comprise the cartwheel. Cryo-electron tomography of *Trychonympha sp.* centrioles revealed the cartwheel has a periodicity (distance between adjacent hubs) of 8.5 nm at the hub. Spokes from neighbouring rings merge at approximately 20 nm from the hub (Fig. 1.7 A). This provides structural tension that maintains hub separation and gives directionality to the cartwheel [46]. Cartwheel stacks have different lengths dependent on species and cell cycle. For example, in *Trychonympha sp.* centrioles the cartwheel stack is 3-5 μm in length [46, 101, 102] constituting approximately 80% of total centriole length. In *C. reinhardtii* the cartwheel is remodelled dependent on cell cycle stage; it increases from approximately 100 nm to 160 nm during mitosis [103].

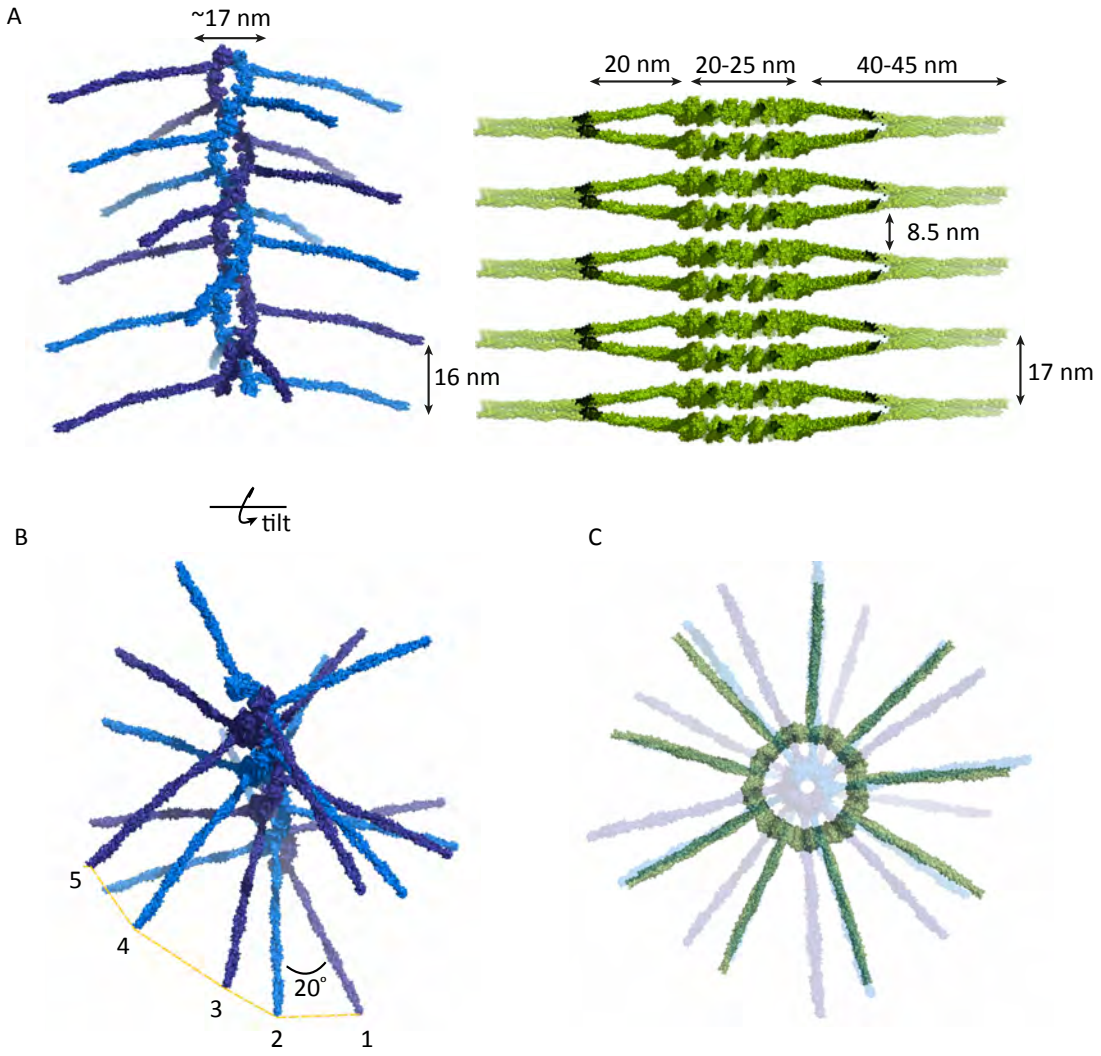


Figure 1.7: Comparison of spiral and cartwheel scaffolds A: Lateral view of two intertwined CeSAS-6 spirals (blue) and a cartwheel stack of SAS-6 rings (green). Distances are indicated by black arrows. B: Tilted view of the spiral in A, showing the screw axis of the spokes. C: Top view overlay of 2 intertwined spirals and a cartwheel stack (green). Figure adapted from [104], PDB codes: 4CKN, 4CKM, 4GFA, 3PYI and 4GKW.

C. elegans centrioles appear to be unique amongst the organisms studied by electron microscopy in that no such cartwheel has been observed at the centre of the centriole. Instead, a central tube is present which is approximately 60 nm in diameter and 110 nm long [32]. Building an oligomer from the crystallographic structures of *C. elegans* SAS-6 (CeSAS-6) N-N and C-C interfaces produces a spiral with approximate 4.5-fold symmetry and 33 nm pitch/turn (Fig. 1.7 A). Thus, two turns of the spiral would achieve 9-fold symmetry and four turns of the spiral corresponds well to the length of a *C. elegans* centriole. The secondary structure of CeSAS-6 is highly conserved between species, however there is a notable divergent feature in *C. elegans*; the angle of the coiled-coil relative to the N-terminal domain. This angle is determined by the steric effect of the residue that lies between the N-terminal domain and the coiled-coil (α3), termed the 'hinge' (Fig. 1.5 B). In nematodes the hinge residue is a valine (V167 in *C. elegans*, Fig. 1.6 C). In contrast, ring-forming SAS-6 variants have a glycine or alanine residue at the hinge (Fig. 1.6 A). The alternative position of α3 as a result of the hinge valine residue in *C. elegans* forces SAS-6 dimers out of the plane of the ring resulting in a spiral-shaped oligomer.

The electron lucent centre of fixed *C. elegans* centrioles observed by electron microscopy supports the presence of a spiral SAS-6 oligomer present *in vivo*, rather than a cartwheel (Fig. 1.8 A) [32]. The electron density of the spokes of a cartwheel stack is reinforced because spokes from all rings in the stack are in register (Fig. 1.8 B). In contrast, the architecture of the CeSAS-6 spiral is such that the spokes are splayed about a screw axis with a $\sim$20° offset [39], thus strong electron scattering from successive spokes does not accumulate explaining their absence in electron micrographs (Fig. 1.7 B and C).

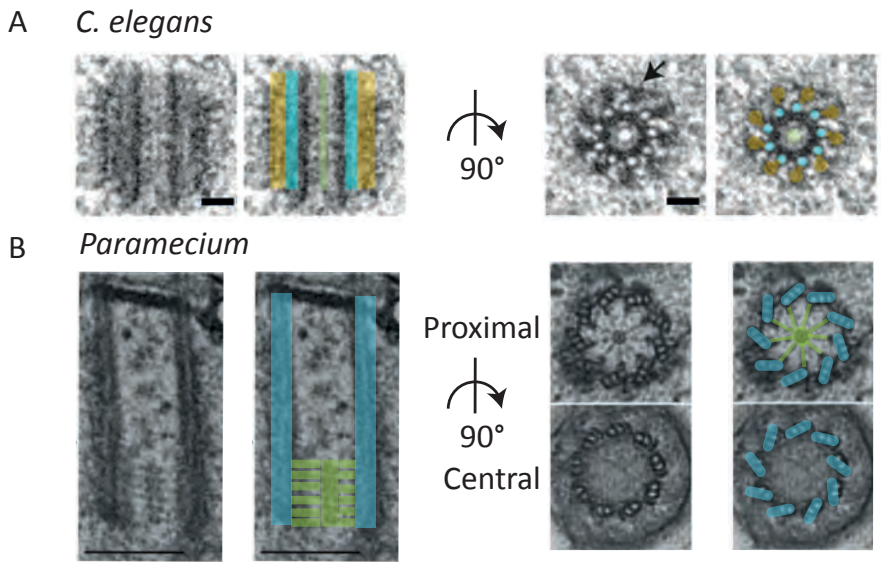


Figure 1.8: Electron microscopy of centrioles A: Electron micrographs of a *C. elegans* centriole, lateral and top views are shown. (Images adapted from [50]). B: Electron micrograph of a *Paramecium* centriole. (Images adapted from [52]). Microtubules (blue), SAS-6 scaffold (green) and paddlewheels (yellow) are highlighted.

Spirals have also been observed *in vitro* in preparations of purified CeSAS-6. Two CeSAS-6 spirals were observed to intertwine and come apart, achieving 9-fold symmetry with a single turn of the oligomer given that each spiral alone has 4.5-fold symmetry. The double spiral oligomer gives rise to spokes with an axial periodicity of 16 nm [39] (Fig. 1.7 A). This is comparable to *Trychonympha sp.* and *C. reinhardtii* centrioles, where the cartwheel stack has a periodicity of 17 nm at the end of the spokes distal to the hub (Fig. 1.7 A) [46]. Thus both spiral and ring scaffolds provide a similar axial periodicity. This periodicity matches well with that of two tubulin dimers (8 nm x2) suggesting that the spacing of spokes is important for the interaction with the microtubule wall, via the pinhead.

A smaller 'inner tube' at the centre of the centriole has been observed by electron microscopy in *C. elegans* embryos (Fig. 1.8 A) [32, 50]. This inner tube is 18 nm in diameter, similar to the diameter of the cartwheel hub present in other species (20-25 nm). The dimensions of the *C. elegans* SAS-6 spiral are incompatible with this value as the outer diameter of the C. elegans SAS-6 spiral oligomer is ∼17 nm [39, 50], suggesting this structure is responsible for the electron density of the inner

tube. Thus lends credence to the idea that in *C. elegans* a spiral SAS-6 oligomer is present at the centre of the centriole *in vivo* rather than a cartwheel.

As well as sharing geometrically similar contact points with the microtubule wall, the cartwheel stack and spiral scaffold share other features. Oligomers built from SAS-6 dimers are inherently symmetric, given the 2-fold symmetry of the dimer, whether they assemble into a cartwheel or a spiral. However, directionality in the ultrastructure is required for proper assembly. In centrioles with a cartwheel scaffold this is achieved by the stacking of rings. Directionality is imparted by the merging of spokes [97]. Additionally, pinfeet attached at the distal end of the spokes, mediating the interaction with the microtubule wall, are tilted towards the proximal end of the centriole, conferring polarity [46, 51]. The tilt of the microtubule blades around the circumference of the centriole also confers polarity; being tilted in a clockwise manner when viewed from the distal end of the centriole. Directionality is achieved from a spiral scaffold in an alternative manner. Intrinsically, the *C. elegans* SAS-6 spiral does not have polarity; however, the spiral does possess handedness. *C. elegans* centrioles are coated by singlet microtubules, thus the ultrastructure appears achiral [32, 38], however, the components of the spiral, SAS-6 dimers, are chiral therefore the structure, at the microscopic level, is chiral. Paddlewheels decorating the outer surface of the microtubule wall are arranged in a clockwise manner away from the proximal end (Fig. 1.8 A) [50] providing a more obvious source of chirality to the ultrastructure. In isolation the spiral has handedness and chirality but is non-polar; however, in combination with the paddlewheels and microtubules the ultrstructure is chiral and polar as is the cartwheel stack.

1.6 Implications of a spiral scaffold

The output of either a cartwheel stack or a spiral is largely similar, in terms of spoke periodicity at the microtubule wall (Fig. 1.7 A). However, a spiral scaffold has implications for both centriole assembly and elongation as well as stability.

A spiral scaffold may be less stable than a cartwheel stack as each end is free to move independently. However, since the spokes in a spiral scaffold are splayed

it is likely they do not merge with neighbouring spokes, therefore, twice as many independent points of contact with the microtubule wall will be present. This could act to stabilise the scaffold. Since spokes in a spiral scaffold are not supported by interaction with a second coiled-coil - as they are in a cartwheel - they will be more flexible. Thin filaments are seen in electron micrographs of *Phlyctochytrium* centrioles 20 nm from the cartwheel hub, corresponding to the same radial location of spoke merger observed in tomograms of *Trychonympha sp.* centrioles [105]. Similar but more irregular filaments are seen in *C. reinhardtii* centrioles [106, 107]. These filaments have been observed to stabilise the cartwheel stack in the absence of microtubules. Perhaps a similar, as yet, unidentified component is present in *C. elegans* centrioles circumventing the possible loss of stability resulting from a spiral scaffold. The shorter length of the *C. elegans* centriole compared to those with cartwheel scaffolds may have alleviated evolutionary pressure to maintain a more stable scaffold. The intertwining of two spirals would stabilise the scaffold, allowing for the more efficient (in terms of protein usage) spiral scaffold to arise. Another factor impacting centriole stability is the presence of microtubule blades. It has been observed that microtubule triplets act to stabilise the centriole. In *H. sapiens* retinal cells, δ- and ε-tubulin null mutant centrioles did not persist through the cell cycle. These mutant centrioles display singlet microtubules, instead of the usual triplet microtubule blades, and were torn apart during cell division [108]. These mutant cells produced new centrioles *de novo* every cell cycle. It is not known whether the stability conferred by triplet microtubules is due to increased intrinsic stability of microtubule triplets, or due to recruitment of additional stabilising factors that recognise only microtubule triplets. Organisms, such as *C. elegans*, with centrioles that display singlet microtubules may employ additional, as yet unidentified, stabilisation mechanisms that protect against centriolar wear and tear during the cell cycle.

Although spirals are less stable than rings because the maximum number of possible interactions are not fulfilled [109], many other oligomers exist with helical

symmetry. The architecture of the SAS-6 spiral oligomer resembles that of x-ray repair cross-complimenting protein 4 (XRCC4) and XRCC4 like factor (XLF) oligomers which are involved in DNA repair and non-homologous end joining. These proteins have a structurally similar N-terminal domain to SAS-6 and form homodimers via a coiled-coil domain (Fig. 1.9 A, B and C). As in SAS-6, the N-N interface is mediated by a hydrophobic interaction. XLF and XRCC4 can form a heterotetramer and oligomerise via their N-terminal domains in a similar manner to SAS-6. Their oligomerisation gives rise to a left-handed spiral with a pitch of 70-90 nm and a radius of 6 nm (Fig. 1.9 E). The spiral has 12-fold symmetry. Six XLF/XRCC4 spirals come together to form a super helical bundle that encases DNA, acting as a scaffold to bridge the two ends of DNA to be joined. Four crystallographic structures of XLF and XRCC4 tetramers are available. The pitch/turn of the oligomers built using each of these interfaces is variable. In a manner mirrored by SAS-6, it is the angle of the coiled-coil relative to the N-terminal domain that varies between each of the crystallographic structures resulting in elongation of the spiral (Fig. 1.9 D) [110–115]. At the equivalent to the hinge position in XLF, there is an alanine residue, similar to ring-forming SAS-6 variants, and in XRCC4 there is a valine residue similar to *C. elegans* SAS-6. This suggests the steric effect of the hinge residue is responsible for the positioning of the coiled-coil relative to the N-terminal domain, as it is in SAS-6.

A spiral scaffold would have implications for centriole duplication. The architecture of the spiral would allow for easy extension to seed growth of the daughter centriole [116]. Daughter centrioles always form with close contact to the mother, positioned orthogonally to the mother in most cases. In *Labyrinthula sp.*, a marine protist, 2 centrioles form end to end with a shared cartwheel stack [117]. These centrioles have 9-fold symmetry and triplet microtubules which twist in opposite directions on each centriole. Such bicentrioles are also found in several other species. The assembly process of bicentriole is not well understood; it is not known whether the cartwheels assemble in a bipolar manner or if the polarity of the organelle is determined by the microtubules or pinheads surrounding a continuous stack of cartwheels. These centrioles orientate orthogonally after assembly is complete [117]. If the cartwheel stack is continuous, not bipolar, a spiral scaffold would make it easy to assemble a centriole in this manner. Whilst it is an attractive proposal that the spiral could seed growth of bicentrioles during duplication, no such intermediate has been observed in *C. elegans*.

The mechanisms for controlling organelle length in centrioles containing spiral scaffolds must differ from that of centrioles containing cartwheel scaffolds. It is important to regulate centriole length, as over-elongation leads to the production of excess microtubules resulting in inaccurate chromosome segregation [118]. The stack height of the cartwheel or length of the spiral oligomer influences centriole height. As discussed, the height of the cartwheel stack varies between species and centriole maturity. Mature centrioles in *C. reinhardtii* have only 2-4 layers of rings [119, 120] corresponding to a height of 100 nm whereas, in *Trychonympha sp.* the stack is 3-5 μm in length [46, 101, 102]. How the height of the stack is determined *in vivo* is unknown. A first potential mechanism for limiting centriole length involves capping proteins. It can be postulated that the alternative scaffold architectures will require different capping mechanisms that limit scaffold height. Complete assembly of a ring contributes a discrete unit of scaffold, whereas the length of a spiral scaffold can be more variable because spiral elongation can progress in increments of a single SAS-6 dimer at a time. Indeed, *in vitro* CeSAS-6 spirals reach

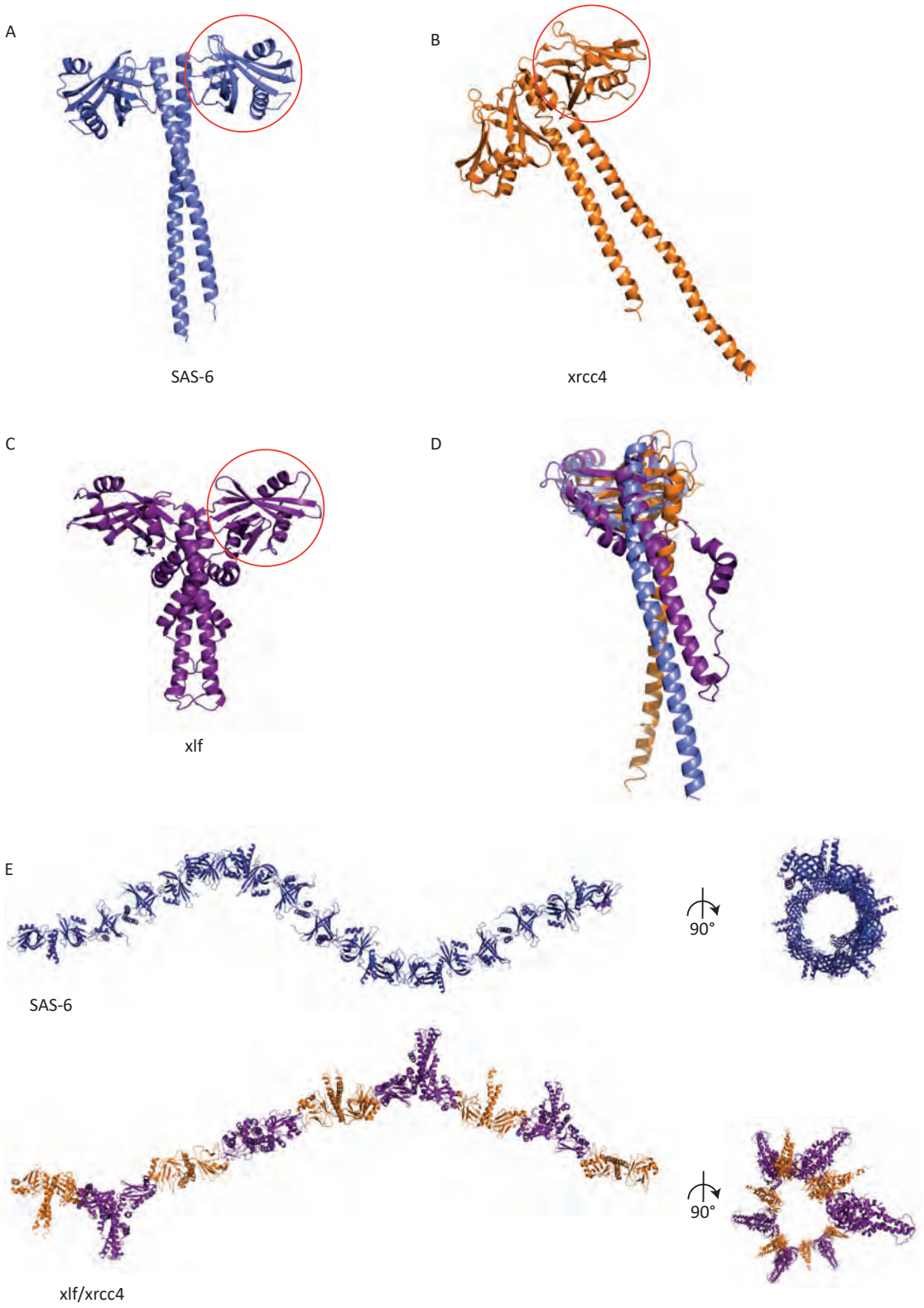


Figure 1.9: XLF and XRCC4 have structurally similar N-terminal domains to SAS-6. A: *C. reinhardtii* SAS-6 C-C dimer (PDB code: 4GFA), B: XRCC4 C-C dimer (PDB code: 1FU1), C: XLF C-C dimer (PDB code: 2R9A). An N-terminal domain is indicated by the red cicle. D: overlay of SAS-6 (blue), XRCC4 (orange) and XLF (purple) monomers (view rotated 90° to that shown in A). E: *C. elegans* SAS-6 oligomer (blue) (PDB code: 3PYI) and XLF/XRCC4 oligomer (purple/orange) (PDB code: 3RWR) shown laterally and in top view.

micrometres in length [39] suggesting a capping mechanism is present *in vivo*. A possible candidate for this capping role is SAS-4. Overexpression of SAS-4 (CPAP) leads to increased centriole length suggesting it acts as a molecular ruler. However, in contradiction to what would be expected of a capping protein of scaffold length, it acts as a scaffold for tubulin addition, causing an increase in centriole length by extending microtubules rather than SAS-6 scaffold length [9, 121]. In human cells the capping proteins CP110 and Cep97 act as negative regulators of centriole assembly, displaying a circular distribution at the distal end of the centriole lumen [33, 43, 121, 122]. The mechanism of centriole length regulation in *C. elegans* may be mediated through control SAS-6 scaffold length mediated by proteins with roles analogous to those of CP110, Cep97 or SAS-4.

Another mechanism by which centriole length may be controlled in *C. elegans* is via the action of ZYG-1. Through immunofluorescence microscopy in *D. melanogaster* embryos PLK4 (ZYG-1 in *C. elegans*) has been observed to act as a homeostatic clock to regulate centriole length. In *D. melanogaster*, centrioles do not undergo elongation during G2/M phase; instead, growth is halted in late S phase. As a result, *D. melanogaster* centrioles are short and the cartwheel stack covers the full length of the centriole. If the length of S phase is artificially extended, the growth rate of centrioles slows so that length is maintained. PLK4 regulates the rate of centriole growth by controlling the rate of SAS-6 recruitment, as well as determining the period of centriole growth [123]. Similar growth dynamics have been observed in *C. elegans* by monitoring the incorporation rate of GFP tagged SAS-6, where centriole growth rate is linear in S phase and then plateaus in mitosis [124]. This suggests that the dynamics of SAS-6 incorporation into centrioles are conserved in the early stages of centriole assembly. The action of ZYG-1 in *C. elegans* may therefore control spiral length through its control of SAS-6 recruitment to the centriole. In this manner a novel mechanism for central scaffold capping need not have arisen in *C. elegans* despite the alternative architecture of the scaffold.

It has also been suggested that the pool of available SAS-6 limits centriole length. Estimates of the number of SAS-6 molecules within the cell vary. Mass

spectrometry identified more than 500 copies of SAS-6 per U2OS cell [125]. However, immunofluorescence and semi-quantitative western blots on the same cell line predicted 185,000 copies in early S phase, which increased to 885,000 copies by late G2 [126]. Such large SAS-6 pools would negate the amount of protein as a controlling factor in centriole length. The most recent study using fluorescence microscopy and mass spectrometry predicted 276 SAS-6 molecules per human centriole, this would correspond to approximately 16 rings. This fits with the observation from electron micrographs that cartwheel stacks are 60-150 nm high, corresponding to 7–17 rings [127]. In organisms such as *H. sapiens* (where SAS-6 is degraded by proteolysis resulting in loss of the cartwheel from centrioles by the end of mitosis) there is a direct link between SAS-6 turnover and the centriole duplication cycle. However, in species such as *C. elegans* where the SAS-6 scaffold is maintained in mature centrioles [76, 77] any mechanism must distinguish between cytoplasmic SAS-6 and SAS-6 incorporated into the centriole. Thus scaffold length may be regulated by the action of capping proteins, control of the rate of SAS-6 recruitment to the centriole by ZYG-1, or by the pool of available SAS-6. At present it is not possible to determine which of these mechanisms or in what combination is responsible for scaffold length regulation in *C. elegans*.

1.7 Robustness of centriole symmetry

Flexibility exists in many large structures built by homo-oligomerisation. Two such examples are XLF/XRCC4 filaments and clathrin homo-oligomers. XLF/XRCC4 are involved in DNA repair and non-homologous end joining. XLF/XRCC4 form spiral-shaped hetero-oligomer filaments similar to SAS-6 spirals. These filaments display flexibility resulting in variation in spiral lumen size (Chapter 1.6) [111–115, 128]. Clathrin also displays flexibility in the angles adopted by the triskelion arms, assembling into vesicles with either 5- or 6-fold symmetry [129].

The variation in centriole size, as well as the number of microtubules within each blade, between species and cell type demonstrates flexibility in possible architecture built using the same components. As mentioned previously, *Trychonympha sp.*

centrioles display extremely long cartwheel stacks 3-5 μm in length, with a 400-600 nm cartwheel-less distal region. Unlike *H. sapiens* and *C. reinhardtii* centrioles that do not undergo large architectural changes throughout the cell cycle, *D. melanogaster* basal bodies in sperm cells display long centrioles up to 1 μm in length. These centrioles are decorated with triplet microtubules. However, in somatic cells the centriole is much shorter, 100-150 nm in length with a cartwheel stack along its full length [130, 131]. Another example of remodelling of centriole structure with developmental stage occurs in *Anisopteromalus calandrae* male wasp larvae. In the larval, stage centrioles do not have microtubules, instead they have 9 electron dense structures termed 'cogwheels' arranged around the circumference of the centriole. Microtubules appear at the nymph stage [132]. Despite their drastically disparate architectures, these centrioles display 9-fold symmetry [46, 101, 102, 130–132].

Protein engineering efforts on the *C. reinhardtii* SAS-6 N-N oligomerisation interface produced variants with symmetries ranging from 5- to 10-fold *in vitro*. However, when tested *in vivo* 9-fold symmetry was maintained in more than 70% of centrioles, demonstrating the robustness of centriole symmetry *in vivo*. One variant did cause centriolar symmetry variation *in vivo*, giving rise to 9- and 8-fold symmetric centrioles at a 2:1 ratio, despite producing 5-fold symmetric rings *in vitro* [133]. This disparity between oligomer symmetry *in vitro* and *in vivo* implies that centriole symmetry *in vivo* is determined by other factors in addition to SAS-6. Indeed, in *C. reinhardtii* Bld10 (Cep135 in vertebrates) null mutants, centrioles with variable symmetries are produced [107]. This suggests while SAS-6 plays a central role in centriole symmetry, Bld10 also contributes. Along with SAS-6 spoke length, Bld10 also influences centriole diameter, due to its location at the pinhead on the inner lumen of the microtubule wall. The diameter of the centriole directly impacts the number of microtubule blades that can be accommodated circumferentially. Truncating Bld12 (SAS-6) to generate cartwheel stacks with shorter spokes increases the abundance of centrioles displaying 8 microtubule blades [106]. Thus centriolar components that mediate centriole diameter influence centriolar symmetry.

9-fold symmetry is ubiquitous in centrioles but is not universally conserved. *Acerentomon microrhinus* centrioles in sperm cells display 14-fold symmetry with 14 microtubule doublets, though centrioles in somatic cells display 9-fold symmetry [134, 135]. The cartwheel hub in these centrioles is 57 nm in diameter but the overall centriole diameter is similar to that of *H. sapiens* and *C. reinhardtii* at 250 nm despite SAS-6 spoke length being maintained at approximately 50 nm. In some fungi centriole radial symmetry is lost altogether. For example, *Sciara centrioles* are 400 nm wide and 1 μm long, with 40-70 microtubule doublets or singlets [34, 136]. These rare exceptions demonstrate that centrioles that deviate from 9-fold symmetry exist and retain functionality however, 9-fold symmetry is overwhelmingly more common. Further understanding of the architecture of the scaffolds that underlie this symmetry would provide insight into the assembly and function of these organelles.

1.8 Aims and objectives

Although SAS-6 crystallographic structures and *in vitro* work provide strong evidence that a spiral scaffold exists in *C. elegans* centrioles, it has not been observed *in vivo*. Electron micrographs of *C. elegans* centrioles display an electron lucent core with no visible spokes, suggesting the presence of a spiral-shaped SAS-6 scaffold [32, 50]. This thesis aims to determine what is responsible for a SAS-6 spiral-to-ring oligomer conformational transition at the amino acid level. Protein engineering is used to probe the effects of amino acid variants on *C. elegans* SAS-6 architecture and the commutability of *C. elegans* SAS-6 in the adoption of a ring conformation.

Chapter 3 assesses the symmetry variability within *C. elegans*, *D. rerio*, *C. reinhardtii* and *Leishmania major* SAS-6 and uses molecular dynamics simulations to further establish the heterogeneity that exists in SAS-6 oligomers.

Chapter 4 discusses the identification of *C. elegans* SAS-6 variants designed to force adoption of a ring oligomer, through covariance analysis in combination with structural assessment of the α3-N-terminal domain interface.

Chapter 5 uses molecular dynamics simulations to determine the effect of the rationally designed SAS-6 variants on oligomer architecture.

In Chapter 6 atomic force microscopy is used to visualise SAS-6 oligomers *in vitro*. This chapter describes wild-type *C. elegans* SAS-6 rings and a variant successful in increasing the number of ring closure events in SAS-6 oligomers.

2 Methods

Contents

2.1 Molecular dynamics simulations 33
2.2 Bioinformatics 34
2.3 Cloning 35
2.4 Protein expression and purification 40
2.5 Analytical Size Exclusion Chromatography 43
2.6 Circular Dichroism 43
2.7 Electron Microscopy 44
2.8 Atomic Force Microscopy (AFM) 45
2.9 Image analysis 45
2.10 Graphs and statistics 46

All chemical reagents were purchased from Sigma unless otherwise stated.

2.1 Molecular dynamics simulations

The crystallographic structures of both the coiled-coil (C-C) dimer of: CeSAS-6 (PDB code: 4GFC/4GFA [39]), CrSAS-6 (PDB code: 3Q0X [40]), *L. major* SAS-6 (PBD code: 4CKN [41]), and *D. rerio* SAS-6 (PDB code: 2Y3W [41]) and the N-terminal dimerisation interface: CeSAS-6 (PDB code: 4G79 [39] and 3PYI [40]), CrSAS-6 (PDB code: 3Q0Y [40]), *L. major* SAS-6 (PBD code: 4CKM [137]) and *D. rerio* SAS-6 (PDB code: 2Y3V [41]) were used in the *in silico*

work presented (Table 1.1).

MODELLER 9.15 [138–141] was used to build missing loops of the crystallographic structures where required (loops were built onto: 4GFC, 4GFA, 2Y3W, 3PYI, and 4CKM). MODELLER 9.15 was also used to build homology models where the CeSAS-6 sequence was forced to adopt the structure of *L. major* SAS-6 and vice versa (Fig. A.1 A and A.2 A and B). Sequence alignments were manually edited and the resulting models manually checked for correct hinge residue positioning. For each variant, 15 models were built and the model with the lowest discrete optimised protein energy (DOPE) score used [142]. CeSAS-6 variants were built in PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). After adding hydrogen atoms to the structure, explicit water molecules were added in a box 1 nm away from the surface of the protein. NaCl was then added to neutralise charge in the system and give a final concentration of 150 mM. Energy minimisation was run for 50,000 steps to optimise the starting geometry. Molecular dynamics simulations were performed using GROMACS [143, 144] with an OPLS-AA/L [145] force field. All simulations were atomistic and run at 300K, 1 bar, for either 10 or 100 ns, in triplicate unless otherwise stated. Snapshots were taken every 2 fs. Snapshots from the molecular dynamics trajectory were extracted every 0.1-1 ns. Models of SAS-6 oligomers were then generated in PyMOL using a script written by Manuel Hilbert to calculate the symmetry, pitch/turn and radius of each oligomer.

To analyse dihedral angle changes in residues surrounding the hinge region, the dihedral angles of residues of interest were obtained for each snapshot of the molecular dynamics trajectory using GROMACS. To further explore the effect of hinge dihedral angles UCSF Chimera [146] was used to build models with altered psi and phi angles. This method was also used to build models that describe the distortion of protein oligomers as a result of AFM imaging.

2.2 Bioinformatics

Covariance analysis of the N-terminal domain α-3 interface was used to identify potential sites of mutation. Each SAS-6 variant (*C. elegans*, *D. rerio*, *L. major* and

C. reinhardtii) was used as a starting point. SAS-6 sequences were identified using protein BLAST (NCBI) and manually edited to confirm all sequences used were SAS-6 homologues. SAS-6 sequences were then aligned using Jackhmmer [147], manually edited and submitted to the consurf [148, 149] and EVcouplings [150] servers to identify pairs of evolutionarily conserved residues. Covariant analysis using *C. elegans* and *C. reinhardtii* as the query sequence did not identify covariant pairs at the N-terminal domain α-3 interface. However, covariant pairs were identified when *L. major* and *D. rerio* were used as the query sequence. 398 sequences were used in the *L. major* SAS-6 covariance analysis and 466 in the *D. rerio* SAS-6 analysis. A heat map of the EVcouplings data was generated using Origin (OriginLab, Northampton, MA). Interactions at the interface between the N-terminal domain and α-3 were also identified manually by visualisation in PyMOL and with PDBePISA [151]. SAS-6 variants were then designed based on this analysis.

To estimate the sequence divergence between SAS-6 protein homologues a phylogentic tree was built. Phyre2 [152] was used to predict the structures of SAS-6 homologues, which were then structurally aligned in PyMOL. Jalview [153] was used to visualise and manually edit the multiple sequence alignment. The alignment was then used to generate a phylogenetic tree in MEGA (7.0.14) [154]. Neighbour-Join and BioNJ algorithms were used to estimate the tree, which was then bootstrapped using a discrete gamma distribution (LG+G) to model the evolutionary rate.

For comparison of the N-terminal domain of SAS-6 and XLF/XRCC4, the crystallographic structure of XLF (PDB code: 2R9A [110]), XRCC4 (PDB code: 1FU1 [113]) and the XLF/XRCC4 interface (PDB code: 3RWR [to be published, Andres, N. S., Wyman, C., Modesti, M., and Junop, M. S.]) were used. Structural alignments were performed using PyMOL.

2.3 Cloning

CeSAS-6 1-215 (Uniprot ID O62479) carrying either I154W S123E, I154G Δloop 103-130, I154E Δloop, or Δloop 103-130 Δloop 151-156 mutations were obtained from Hilbert *et al.* 2013 [39] in pET15b vectors. CeSAS-6 S123E I154W 1-293 in a

pFloat2 vector was made by Kacper Rogala. A PSTCm1 vector encoding CrSAS-6 (Uniprot A9CQL4) (residues 1-159, carrying G94E, F145W and Q147K mutations) was gifted by the Steinmetz group (NN24 [133], PSI, Switzerland).

For electron microscopy experiments CeSAS-6 1-215 I154W S123E and CrSAS-6 1-159 G94E F145W Q147K were amplified by PCR using Phusion polymerase with the supplied buffer (New England Biolabs). Table 2.1 details the PCR protocol used for fragment amplification. Custom oligonucleotides were purchased from Sigma Aldrich. These fragments where then cloned into a pFloat2 vector [71]. The pFloat2 vector was amplified by PCR, and then digested by Dpn1 in CutSmart buffer (New England Biolabs) for 12 hours. The gene insert and vector PCR products were then purified following agarose gel purification using the zymoclean gel DNA recovery kit (Zymogen) according to the manufacturer's instructions. The amplified fragments of CeSAS-6 and CrSAS-6 were cloned into the pFloat2 vector with an N-terminal 6xHis-tag and a C-terminal human rhinovirus (HRV) 3C protease cleavage site, by Gibson assembly (New England BioLabs) using the Gibson Assembly Master Mix (New England Biolabs).

Table 2.1: PCR Protocol

Temperature	**Time**	**Cycles**	**Step**
98°C	30 seconds	1	initial denaturation
98°C	10 seconds	35	denaturation
50 - 72°C	15 seconds		primer annealing
72°C	20 seconds/kb		elongation
72°C	5 minutes	1	final elongation
4°C		1	storage

The assembly products were transformed by heat-shock into DH5-α competent *E. coli* cells (New England Biolabs). The transformed cells were plated onto Luria Bertani (LB) 1.5% agar plates with 50 μg/L kanamycin for selection of successful clones. Plates were incubated overnight at 37 °C. Single colonies were used to inoculate 10 mL of super optimal broth with catabolite repression (SOC) media and incubated at 37 °C, 180 rpm for 16 hours. Plasmid DNA from these cultures was extracted using the Qiagen Miniprep kit according to manufacturer's

instructions. The correct sequence of resulting plasmid DNA was confirmed by sequencing by Eurofins Genomics.

To generate constructs for AFM, a pFloat2 vector was modified to contain a C-terminal SUMO-tag. A SUMO-tag from a pre-existing vector was amplified by PCR using primers with overhangs for ligation into a pFloat2 vector by Gibson assembly (New England BioLabs). Mutations were introduced to the SUMO-tag by quickchange mutagenesis PCR. CeSAS-6 1-293 was amplified from plasmid DNA containing this gene made by Kacper Rogala. CeSAS-6 1-293 DNA fragments were cloned into the new vector with the C-terminal SUMO-tag using Gibson assembly (New England BioLabs) as described above. CeSAS-6 1-293 variant DNA was generated by quickchange PCR using CeSAS-6 1-293 C-SUMO-plasmid DNA as the template (Table 2.2). The resulting plasmid DNA was transformed in to DH5-α competent *E. coli* cells as described above. Plasmid DNA was extracted from overnight cultures using the Qiagen Miniprep kit. Correct sequences were then confirmed by sequencing by Eurofins Genomics. Round II CeSAS-6 1-293 variants were generated using the same primers as the first round of quickchange PCR, but instead using CeSAS-6APVG as template DNA. Transformation into DH5-α competent *E. coli* cells and confirmation of correct sequence was then carried out as described above. Table 2.3 details all SAS-6 constructs used in this thesis.

Table 2.2: Custom oligonucleotides used for cloning.

	Construct	Boundaries	Forward primer	Reverse primer
Vector amplification	pFloat2		TCTGTTCCAGGGGCCC	TTAGCAGCCGGATCTCTTA
Fragment amplification	CrSAS-6	1-159	TCTGTTCCAGGGGCCCATGCCGCTT CTTCTCGACG	TTAGCAGCCGGATCTC TTACACGT CTTCGCCTGGG
	SUMO-tag	1-285	GGCAGCGGCAGCAGCGATAGCGAAG TGAACCAGG	TTAGCAGCCGGATCTCTTAGATCTG TTCGCGATGCGC
	1-215	1-215	TCTGTTCCAGGGGCCCATGACTAGC AAAATTGCATTATTCGATC	TTAGCAGCCGGATCTCTTAAGACAGT TTTTCAACTTGTTTTTCCAAATC
	1-293	1-293	TCTGTTCCAGGGGCCCATGACTAGC AAAATTGC	GCTGCCGCTGCCCTGATCCACTTTT CCTTG
	1-190	1-190	TCTGTTCCAGGGGCCCATGACTAGC AAAATTGCATTATTCGATC	TTAGCAGCCGGATCTCTTAAAAAGT GGAGAGTTTCACTGC
Quick change mutagenesis	SUMO-tag E83K E84K	1-285	GCGAAATCTTTTTTAAAATTGAAGA AACCACCCCGCTGCG	CGCAGCGGGGTGGTTTCTTCAATTT TAAAAAAGATTTCGC
	V167G	1-293	CCATCAAGCTTCACGCAGTGCGCGG AGATCATTTAATTTC	GAAATTAAATGATCTCCGCGCACTG CGTGAAGCTTGATGG
	V167P	1-293	CCATCAAGCTTCACGCACCGCGCGG AGATCATTTAATTTCAC	GTGAAATTAAATGATCTCCGCGCGG TGCGTGAAGCTTGATGG
	V167G A166P	1-293	CTTTTCCATCAAGCTTCACCCGGTC CGCGGAGATC	GATCTCCGCGGACCGGGTGAAGCTTG ATGGAAAAG
	V167G E54R	1-293	GAAAAGGAGCTAAAATTCCGTATCA GCCGCTCCGATG	CATCGGAGGCGGCTGATACGGAATT TTACGTCCTTTTC
	V167G H175Y	1-293	CGCGGAGCTCATTTAATTTCATATC TTCTGAAAATTTCTTCCAGTGAAGC	GCTTGACTGGAACAAATTTTCAGAA GATATGAAATTAAATGATCTCCGCG
	V167G E62L	1-293	CCGCTCCGATGATTTTCTGTTTCTA TTCTCGGAAACG	CGTTTCCGAGAATAGAAACAGAAAA TCATCGGAGCGG
	V167G F61K	1-293	CGAAATCAGCCGCTCCGATGATAAA GAATTTCTATTCTCGGAAACG	CGTTTCCGAGAATAGAAATTCTTTA TCATCGGAGCGGCTGATTTCG
	V167G H175E	1-293	CGCGGAGATCATTTAATTTCAGAAC TTCTGAAAATTTCTTCCAGTC	GACTGGAACAAATTTTCAGAAGTTC TGAAATTAAATGATCTCCGCG
	V167G H175K L172K	1-293	GGCCGCGGAGATCATGAAATTTCAA AACTTCTGAAAATTTGTTCCAGTC	GACTGGAACAAATTTTCAGAAGTTT TGAAATTTCATGATCTCCGCGGCC

Quick change mutagenesis continued				
	V167G L64K	1-293	GCCGCTCCGATGATTTTGAATTTAA ATTCTCGGAAACGTTGAACAACG	CGTTGTTCAACGTTTCCGAGAATT TAAATTCAAAATCATCGGAGCGGC
	V167G I179K	1-293	CATTTAATTTCACATCTTCTGAAAA AATGTTCCAGTCAAGCAGTG	CACTGCTTGACTGGAACATTTTTTC AGAAGATGTGAAATTAAATG
	V167G I179E	1-293	CATTTAATTTCACATCTTCTGAAAG AATGTTCCAGTCAAGCAGTGAAAC	GTTTCACTGCTTGACTGGAACATTC TTTCAGAAGATGTGAAATTAAATG
	V167G S56L	1-293	GGAGCTAAAATTCGAAATCCTGCGC TCCGATGATTTTG	CAAAATCATCGGAGCGCAGGATTTC GAATTTTAGCTCC
	V167G S56E S58I	1-293	GCTAAAATTCGAAAATCGAACGCAT CGATGATTTTGAATTTCT	AGAAATTCAAAATCATCGATGCGTT CGATTTCGAATTTTAGC
	V167G S56R	1-293	GGAGCTAAAATTCGAAATCCGTCTC CGATGATTTTGAATTTCTCGG	CCGAGAAATTCAAAATCATCGGAGA CGGATTTCGAATTTTAGCTCC
	V167G S56E	1-293	GGAGCTAAAATTCGAAATCGAACGC TCCGATGATTTTG	CAAAATCATCGGAGCGTTCGATTTC GAATTTTAGCTCC
	sequencing primer		TTATTCGATCAAACACTCATCGCC	ATTATTGCTCAGCGGATGGC

Table 2.3: SAS-6 constructs

Species	Boundaries	Mutations	Tag
C. reinhardtii	1-266	Q39E F154W K146R	
C. elegans	1-215	S123E I154W	
	1-215	S123E I154E	
	1-215	I154E A166P V167G ΔLoop 103-130	
	1-215	I154G A166P V167G ΔLoop 103-130	
	1-215	A166P V167G ΔLoop 103-130 ΔLoop 151-156	
	1-190	I154E A166P V167G ΔLoop 103-130	
	1-190	I154G A166P V167G ΔLoop 103-130	
	1-190	A166P V167G ΔLoop 103-130 ΔLoop 151-156	
	1-293	S123E I154W	SUMO
	1-293	S123E I154W	SUMO E83K E84K
	1-293	S123E I154W V167G	SUMO
	1-293	S123E I154W V167P	SUMO
	1-293	S123E I154W V167G A166P	SUMO
	1-293	S123E I154W V167G E54R	SUMO
	1-293	S123E I154W V167G H175Y	SUMO
	1-293	S123E I154W V167G F61K	SUMO
	1-293	S123E I154W V167G H175E	SUMO
	1-293	S123E I154W V167G L172K H175K	SUMO
	1-293	S123E I154W V167G I179E	SUMO
	1-293	S123E I154W V167G S56E S58I I179K	SUMO
	1-293	S123E I154W V167G S56E	SUMO
	1-293	S123E I154W V167G L64K H175Y	SUMO
	1-293	S123E I154W V167G S56L F61L H175E	SUMO
	1-293	S123E I154W V167G S56R H175Y	SUMO
	1-293	S123E I154W V167G E54R H175Y	SUMO
	1-293	S123E I154W V167G A166P F61K	SUMO
	1-293	S123E I154W V167G A166P H175E	SUMO
	1-293	S123E I154W V167G A166P E54R H175E	SUMO
	1-293	S123E I154W V167G A166P H175K	SUMO

2.4 Protein expression and purification

Expression plasmids were transformed by heat-shock into Lemo21 (DE3), BL21 (DE3) (New England Biolabs), Rosetta (DE3) or Rosetta2 (DE3) (Thermo Fisher Scientific) *E. coli* cells for expression tests. *E. coli* transformation was carried out by incubating plasmid DNA with competent cells on ice for 30 minutes. Then cells are heat-shocked at 42 °C for 45 seconds before being incubated on ice for 2 minutes. The cells were then outgrown in LB media without antibiotic for 1 hour at 37 °C. Transformed cells then were plated onto 1.5% agar LB plates containing 50 μg/L kanamycin (with 34 μg/L chloroamphenicol for Lemo 21 (DE3) *E. coli*

cells) and incubated at 37 °C overnight. 100 mL LB medium containing appropriate antibiotics was inoculated with single colonies. The cultures were incubated at 37 °C, 180 rpm until an OD_{600} of 0.6 had been reached. A sample was taken prior to induction of protein expression by 0.25 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Protein expression was induced for either 6 hours at 37 °C or 18 hours at 18 °C. The cultures were then spun down for 10 minutes at 4000 rpm. A post-induction sample of the resulting supernatant and pellet was then taken. Expression was assessed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Protein expression was greatest in Lemo21 (DE3) *E. coli* cells which were used in larger scale cultures for protein expression.

For large scale protein expression a single colony of Lemo21 (DE3) *E. coli* was used to inoculate 100 mL of LB medium containing 50 μg/L kanamycin and 34 μg/L chloroamphenicol. This starter culture was then incubated at 37 °C, 180 rpm overnight. The starter culture was then used to inoculate flasks containing 1 L of LB medium. Protein expression was induced overnight for 18 hours at 18 °C with 0.25 mM (IPTG) once an OD_{600} of 0.6 had been reached. Cells were harvested by centrifugation at 4000 xg for 40 minutes at 4°C and the pellets resuspended in 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.5% v/v Triton X-100 buffer to a volume of $\sim$ 30 mL while keeping the cell suspension on ice. Lysis was carried out by sonication in a glass beaker on ice. Cells were subject to 2.5 W, 15 second pulses with a 30 second cooling period in between pulses, for a total of 6 minutes. The lysates were then clarified by centrifugation at 30,000 xg for 60 min at 4 °C.

Protein purification was carried out by immobilised metal affinity chromatography using a TALON column (GE Healthcare) at 4 °C on an ÄKTA fast protein liquid chromatography (FPLC) system. The TALON column was equilibrated with loading buffer (CeSAS-6 loading buffer: 20 mM Tris-HCl pH 7.5, 500 mM NaCl, CrSAS-6 loading buffer: 20 mM Tris-HCl, 10 mM imidazole) for a total of 10 column volumes. Following equilibration, cell lysate was loaded onto the column with a flow rate of 1.5 mL/minute. The column was washed to remove non-specific proteins with wash buffer (CeSAS-6 wash buffer: 20 mM Tris-HCl

pH 7.5, 500 mM NaCl, CrSAS-6 wash buffer: 20 mM Tris-HCl, 70 mM imidazole) for a total of 25 column volumes. Tagged protein was then eluted by a linear gradient of elution buffer (0-100% over 100 mL) containing imidazole (CeSAS-6 elution buffer: as wash buffer supplemented with 500 mM imidazole, CrSAS-6 elution buffer: as loading buffer supplemented with 200 mM imidazole). The eluted protein was assessed by SDS-PAGE. Fractions containing the protein of interest were pooled and samples were then dialysed into 20 mM Na_2HPO_4 pH7.2, 150 mM NaCl, 2 mM 1,4-Dithiothreitol (DTT) buffer at 4 °C overnight. Further purification was carried out by size exclusion chromatography using a Superdex 200 26/60 column (GE Healthcare Life Sciences) in 20 mM HEPES pH7.2, 150 mM NaCl, 2 mM DTT at 4 °C on an FPLC system. Fractions containing the protein of interest were identified by SDS-PAGE and the purity of the protein assessed. Fractions containing the protein were pooled and then concentrated by centrifugal ultrafiltration using a polyethersulphone (PES) membrane (Sartorius Vivaspin) at 4°C. The protein concentration determined by UV absorbance at 280 nm using an extinction coefficient estimated from the amino acid sequence and molecular weight. Samples were then aliquoted and flash frozen in liquid nitrogen and stored at -80 °C.

Samples for crystallisation trials were purified using the same protocol and buffers as described above with two additional steps. Following the affinity chromatography step, the 6xHis-tag was cleaved using HRV 3C protease during dialysis. Successful cleavage of the 6xHis tag was confirmed by SDS-PAGE. Prior to size exclusion chromatography, an additional anion exchange purification step was carried out (also at 4°C). A HiTrap Q column (GE Healthcare) was used. The protein sample was first dialysed into wash buffer (wash buffer: 20 mM Na_2HPO_4, pH8 and 20 mM NaCl). The sample was applied to a Q column equilibrated with wash buffer. The column was then washed with wash buffer for 10 column volumes. Protein was then eluted with a linear gradient of wash buffer supplemented with 1M NaCl (0-100% over 150 mL). Protein purity was assessed by SDS-PAGE. Fractions containing the protein of interest were pooled and dialysed into 20 mM Na_2HPO_4 pH7.2, 150 mM NaCl, 2 mM DTT buffer at 4 °C overnight. Large aggregates were removed

from the sample by centrifugation at 3000 xg for 30 minutes. Protein samples were then loaded onto a Superdex 75 26/60 column (GE Healthcare Life Sciences) using the protocol described above. Following size exclusion chromatography, protein purity was assessed by SDS-PAGE and the protein concentration determined by UV absorbance at 280 nm.

2.5 Analytical Size Exclusion Chromatography

A Superdex 200 increase 10/300 GL (GE Healthcare Life Sciences) column was used for analytical size exclusion chromatography experiments used to determine oligomerisation ability of CeSAS-6 variants. The same buffer was used as in size exclusion chromatography experiments (20 mM HEPES pH7.2, 150 mM NaCl, 2 mM DTT). Analytical size exclusion chromatography experiments were carried out at room temperature. Where the effect of freeze-thaw on samples was to be tested, the protein concentration was estimated by UV absorbance as descibed above, and the sample split in two, one frozen and defrosted on ice and the other applied directly to the column, as to ensure identical starting sample concentration. 250 μL of protein was loaded onto the column which had been equilibrated in 20 mM HEPES pH7.2, 150 mM NaCl, 2 mM DTT buffer.

2.6 Circular Dichroism

The secondary structure composition, correct folding and thermal stability of CeSAS-6 variants was determined by circular dichroism (CD). Chiral molecules, such as protein, can absorb circularly polarised light. The differential absorbance of left and right circularly polarised light is wavelength dependent. When measured over a range of wavelengths, this produces a CD spectrum. In the far UV region, the secondary structure composition of a protein can be inferred from characteristic features in the CD spectrum. α-helices are characterised by spectral minima at 208 and 220 nm, and a maxima at 192 nm. β-sheets are characterised by a spectral

minima at 217 nm and a maxima at 198 nm. Unfolded protein is characterised by a minima at $\sim$ 200 nm, and thus can be distinguished from folded protein.

The CD spectra of CeSAS-6 samples were measured on a Jasco J-815 spectropolarimeter. Samples were dialysed into 20 mM Na_2HPO_4, pH7.5, 100 mM NaF and diluted to 0.1 mg/mL in the same buffer prior to measurment. CD spectra were collected using a quartz cuvette with 1 mm path length. Spectra were recorded between 200 – 260 nm, at 18 °C. The buffer-only background has been subtracted from plotted data. The CD signal at 222 nm was measured between 10-80 °C, increasing the temperature at a rate of 1 °C/min to determine thermal stability of the sample. A second CD spectrum was measured following this heating process at 18 °C. Plotted data have been normalised for sample concentration and amino acid number.

2.7 Electron Microscopy

Protein samples were cross-linked using $BS(PEG)_5$, $BM(PEG)_3$ and DSS. For cross-linking with DSS the sample was dialysed into conjugation buffer (20 mM HEPES, pH 7.5, 150 mM NaCl). DSS was added at 3000 μM and 1200 μM and the sample incubated at 4°C for 2 hours. Quenching buffer (1 M Tris, pH 7.5, 150 mM NaCl) was added to stop the reaction, and then diluted to give a final concentration of 25 mM Tris. Cross-linking with $BM(PEG)_3$ was carried out at 1200 μM, 600 μM, 300 μM, 250 μM, 200 μM and 120 μM. Cross-linking with $BS(PEG)_5$ was carried out at 6000 μM, 3000 μM, 1500 μM, 1000 μM, 800 μM and 600 μM. Samples were first dialysed into conjugation buffer (20 mM Na_2HPO_4 pH7.2, 150 mM NaCl, 5 mM ethylenediaminetetraacetic acid (EDTA)) and incubated for 2 hours at 4 °C after addition of the cross-linker. To quench the reaction DTT was added to give a final concentration of 0.5 M. The samples were then dialysed into imaging buffer (20 mM HEPES, pH 7.5, 150 mM NaCl) to remove excess DTT. The success of the cross-linking reaction was assessed by SDS-PAGE.

Carbon coated grids were glow discharged using a Leica EM ACE200. Freshly purified (not freeze-thawed) protein samples were used for electron microscopy experiments. Protein samples were applied to the grids and incubated for 2

minutes. The grids were then blotted and stained with 2% w/v uranyl acetate for 30 seconds before being left to dry in the dark for 10 minutes. Spirals of CeSAS-6 assemblies were observed at 1.5 mg/mL protein concentration and cartwheels of CrSAS-6 were observed at 1 mg/mL (concentrations as used by Hilbert *et al.* 2013 [39]). Electron micrographs were taken in a Tecnai12 TEM at 120 kV on a Gatan OneView CMOS camera.

2.8 Atomic Force Microscopy (AFM)

Stocks of SAS-6 at 1.5 mg/mL were defrosted on ice and spun at 17,000 xg for 15 minutes. As quickly as possible the sample was diluted 1:100 using 20 mM Tris, pH7.2, 300 mM KCl buffer and 30 μL adsorbed onto freshly cleaved mica for 10 mins in 20 μL of adsorption buffer (20 mM Tris, pH7.2, 300 mM KCl) at room temperature. 10 mm mica discs were used for all AFM experiments. The sample was then gently washed with 30 μL imaging buffer (20 mM Tris, pH7.2, 150 mM KCl) 5 times prior to imaging.

Imaging conditions used were optimised based on those used by Hilbert *et al.* 2016 [133]. Images were collected at room temperature in PeakForce Tapping mode on a Nanoscope Multimode 8 (Bruker, Santa Barbara, California) with an MSNL-10 triangular cantilever D (Bruker), which has a spring constant of $\sim$0.01 Nm^{-1} and resonant frequency of $\sim$10 kHz. For high resolution images 512 pixels were scanned per line, 512 lines, at a line frequency of 0.94 Hz and a scan size of 1-2 μm. A vertical oscilation of 2 kHz and an amplitude of 25 nm was applied to the cantilever and an imaging force of 45-70 pN was used. Three biological replicates of CeSAS-6 were imaged by AFM. Two biological replicates of all CeSAS-6 mutants were imaged by AFM.

2.9 Image analysis

AFM height profiles were processed using Gwyddion [155]. Height profiles were levelled using mean plane subtraction and scars as a result of shifted scan lines were

corrected by subtraction of the mean height of each scan line. The colour range was then set manually for each image by defining the maximum and minimum height of the dynamic range. Images were then analysed using ImageJ [156]. The number of rings was determined manually by counting the number of completely closed rings out of all oligomers that were $\geqslant 4$ dimers in length. Curvature was measured by adapting an ImageJ plugin (Simple bissection tool by Olivier Burri at the BioImaging and Optics Platform) based on curvature measurements of butterfly claws [157]. Three points were picked manually on the inner arc of each oligomer, A, X and B (Figure 6.14 C). The curvature plugin was then used to calculate angle Y as shown in Figure 6.14 C. The length of arc AB was was determined by tracing the arc by hand. The ratio of arc length to angle Y was then calculated for each measurement. Backbone width, ring diameter and spoke separation were measured manually in ImageJ. Lines were traced across the backbone, bisecting a ring, or between spokes as shown in Figure 6.14 A. The intensity/height trace along each line was analysed using a python script for peak finding and full-width half maximum calculation in matplotlib (Fig. A.3 A and B) [158] to determine the distance between the two peaks in the intensity/height trace. Backbone intensity data were collected manually in ImageJ and the periodicity was then measured in MATLAB (MathWorks, Natick, Massachusetts) by FFT of the intensity data. Peaks in the FFT were identified using a peak finding script (as above) (Fig. A.3 B) to determine periodicity present in the intensity/height trace. A histogram of frequencies present in the FFT of backbone intensity/height traces was generated. To aid in visulaisation of the frequencies present, the smoothed probability function of the backbone periodicity histogram was calculated using MATLAB (MathWorks, Natick, Massachusetts).

2.10 Graphs and statistics

Graphs of molecular dynamics simulation data were plotted using matplotlib [158]. Statistical significance of curvature measurements were determined using a Wilcoxon rank sum test in R with a 95% confidence interval.

3

SAS-6 oligomer architecture

Contents

3.1 Introduction 47
3.2 *C. elegans* SAS-6 spirals 48
3.3 Heterogeneity of SAS-6 oligomers 49
3.4 Molecular dynamics simulations of SAS-6 oligomers . 54
3.5 Discussion 62

Contributions: Manual Hilbert wrote the PyMOL script used to build oligomers from snapshots of molecular dynamic trajectories.

3.1 Introduction

SAS-6 oligomers from *Leishmania major* (LmSAS-6), *Chlamydomonas reinhardtii* (CrSAS-6) and *Danio rerio* (DrSAS-6) have been observed to adopt a ring conformation. While variation in the symmetry sampled by these oligomers exists *in vitro*, the ring architecture is conserved. However, in *C. elegans* SAS-6 (CeSAS-6) oligomers adopt a spiral architecture. This has been observed in electron microscopy experiments and inferred from the crystallographic structure of this protein. In this chapter, the structural differences seen across SAS-6 variants are assessed with a focus on the range of symmetries sampled by the SAS-6 oligomers.

3.2 *C. elegans* SAS-6 spirals

C. elegans centrioles appear unique in that the SAS-6 scaffold adopts a spiral rather than a ring conformation. This alternative architecture can be inferred from crystallographic models of CeSAS-6. Although the CeSAS-6 sequence similarity to ring-forming SAS-6 variants is low, the tertiary structure is highly conserved. Structural analysis of CeSAS-6 shows that the only major difference is a rotation of α-helix 3 (α3) by 38° relative to the N-terminal domain (Fig. 3.1 B). The rotation of α3 means that addition of SAS-6 dimers to the growing oligomer no longer happens in a planar manner. This results in a spiral oligomer with a pitch of ∼33 nm. It is the steric effect of a bulky valine residue (V167) at the region linking α3 to the N-terminal domain, known as the 'hinge', that is responsible for this alternative orientation (Fig. 3.1 C). The hinge V167 residue packs agains L64, S66 and L172 (Fig. 3.1 D and E). A comparison of SAS-6 sequences across ring-forming variants shows that the hinge residue is small, typically an alanine or glycine (Fig. 3.1 A). The model CeSAS-6 spiral built from the crystallographic structures of the N-N and C-C dimers has an inferred symmetry of ∼5-fold. An alteration of the Ψ angle of V167 by 4° from 124° to 128° produces a model with 4.5-fold symmetry and 31 nm pitch/turn. Thus, two turns of the spiral give rise to 9-fold symmetry. Spirals of CeSAS-6 were also directly observed *in vitro* by negative stain electron microscopy; further, it was observed that 2 spirals could intertwine resulting in an oligomer with 9-fold radial symmetry per turn [39].

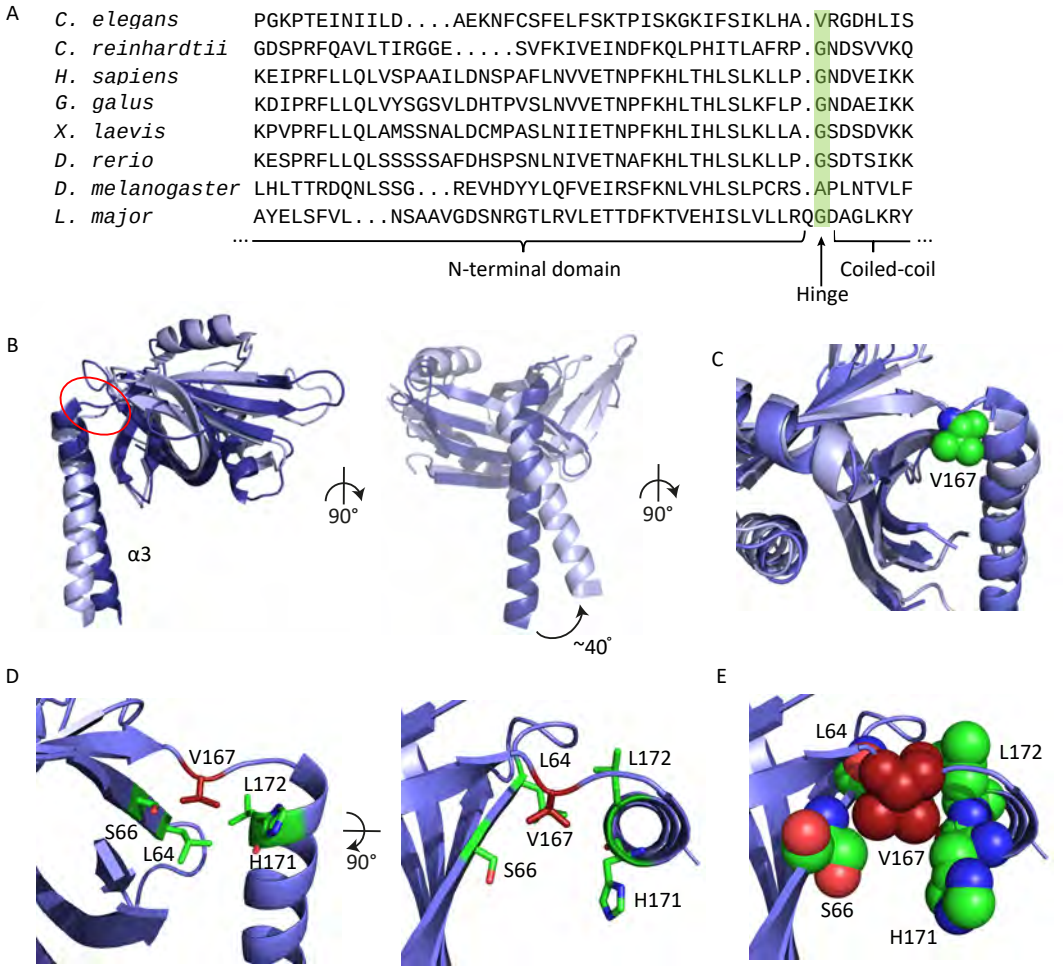


Figure 3.1: SAS-6 hinge. A: Multiple sequence alignment of SAS-6 variants with the hinge residue highlighted in green. B: Overlay of CeSAS-6 (grey, PDB code: 4GFA) and CrSAS-6 (blue, PDB code: 3Q0X) showing rotation of α3. The hinge is indicated by the red circle. C: Enlarged view of the hinge with the CeSAS-6 hinge V167 residue as green spheres. D: Residues L64, S66, H171 and L172 (green) pack against the hinge V167 residue (red). E: Side chains that interact with V167G (red) shown as spheres (green).

3.3 Heterogeneity of SAS-6 oligomers

By building *in silico* models of SAS-6 oligomers using available crystallographic structures the range of symmetries, radius and pitch/turn sampled by oligomers can be inferred. All combinations of SAS-6 N-N interface and C-C interface (C-C) dimer crystallographic structures were used. For example, there are 3 crystallographic structures of the CeSAS-6 C-C interface and 4 of the N-N interface, meaning 12

oligomer models can be generated (Fig. 3.2 A). These CeSAS-6 oligomer models produce right-handed spirals that display variable symmetry ranging from 4.7- to 5.6-fold. The pitch/turn of oligomers built from crystallographic structures ranged from 30 to 40.6 nm (Fig. 3.3 A and B) and the radius ranged from 3 to 3.4 nm. This suggests there is flexibility in the N-N and C-C interfaces.

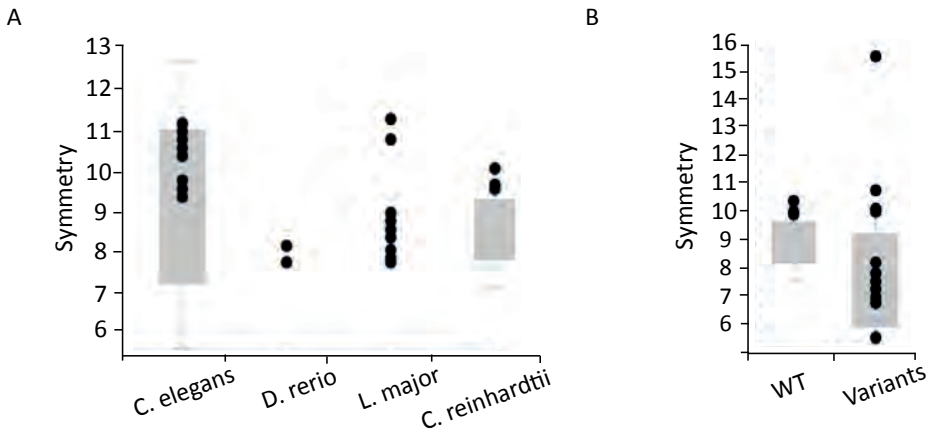
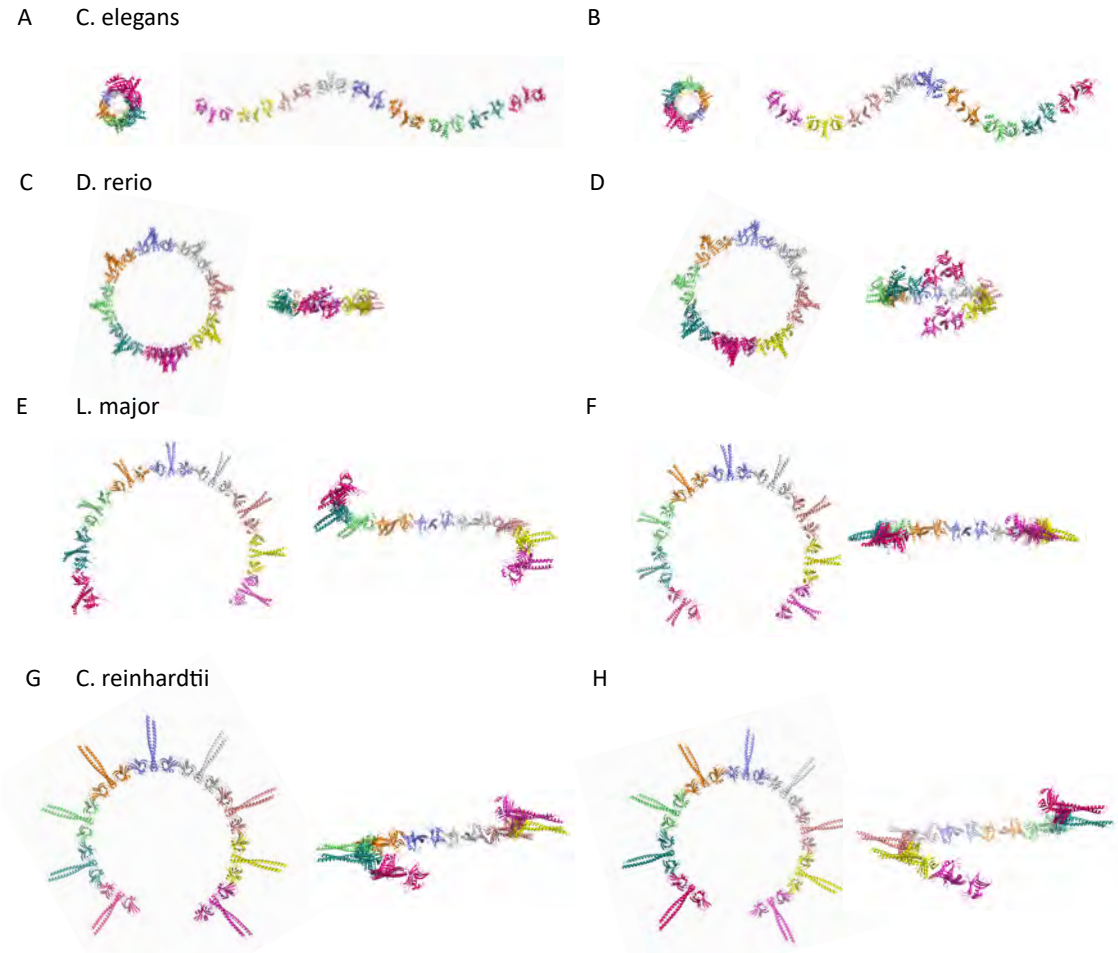


Figure 3.2: SAS-6 oligomer symmetry. A: Variation in symmetries inferred from C-C and N-N dimer crystallographic structures (black circles) or directly visualised by electron microscopy (grey boxes, standard deviation is indicated by bars) from *C. elegans*, *D. rerio*, *L. major* and *C. reinhardtii* SAS-6 variants. Values of CeSAS-6 oligomer symmetry presented in the figure have been doubled, representing two turns of the spiral or a single turn of two intertwined spirals, to allow easy comparison to the ring-forming SAS-6 variants. B: Symmetry variation in wild type CrSAS-6 and engineered variants inferred from crystallographic structures (black circles) or directly visualised by microscopy (grey boxes, standard deviation is indicated by bars). Figure adapted from [104].



	Direction	Symmetry	Pitch/turn (nm)	Radius (nm)
C. elegans (as A)	Right-hand	4.5	38.6	3.2
C. elegans (as B)	Right hand	4.7	30	3.3
D. rerio (as C)	Left-hand	8.2	0.8	10.3
D. rerio (as D)	Right-hand	7.8	8.1	9.7
L. major (as E)	Left-hand	11.3	27.9	12.9
L. major (as F)	Left-hand	10.8	2.6	13.1
C. reinhardtii (as G)	Left-hand	10.3	12.4	12.1
C. reinhardtii (as H)	Left-hand	9.8	14.7	11.5

Figure 3.3: SAS-6 oligomer models. A and B: CeSAS-6 right-handed spiral oligomers, C and D: DrSAS-6 spiral oligomers, E and F: LmSAS-6 spiral oligomers, G and H: CrSAS-6 spiral oligomers. Each oligomer is shown in a top view on the left and a lateral view (90° rotation of top view) on the right. The table summaries the helical direction, symmetry, pitch/turn and radius of each of the oligomers A-H.

Similar flexibility is observed in ring-forming SAS-6 variants. Of the 3 ring-forming SAS-6 variants with both N-N and C-C interface crystallographic structures available, all displayed heterogeneity in the architecture sampled by oligomers with at least one oligomer model from each species adopting a spiral architecture. The 3 models generated from CrSAS-6 crystallographic structures had radial symmetries ranging from 9.6- to 10.1-fold (Fig. 3.2 A). The oligomers generated from these interfaces were not planar; instead CrSAS-6 oligomers adopt a shallow left-handed spiral conformation (Fig. 3.3 G and H). The pitch/turn of these spirals ranged from 12.5 to 19.6 nm. The expected inner radius of a 9-fold symmetric ring is ∼11.5 nm. The radius of CrSAS-6 spiral oligomers ranged from 11.1 to 12.1 nm.

Similarly, LmSAS-6 oligomers also displayed heterogeneity. The inferred radial symmetries from LmSAS-6 dimer models ranged from 7.8- to 11.3-fold (Fig. 3.2 A, 3.3 E and F) and oligomers adopted a ring or a left-handed spiral architecture. The pitch/turn of these spirals ranged from 4.9 to 12.3 nm and the radius ranged from 9.3 to 13.1 nm.

The 2 oligomers built from DrSAS-6 crystallographic structures adopted either an 8-fold symmetric ring (almost planar, 0.8 nm pitch/turn) with a radius of 10.3 nm, or a right-handed spiral with shallow pitch/turn. This spiral has 7.8-fold symmetry, 8.1 nm pitch/turn and a 9.7 nm radius (Fig. 3.3 C and D).

The variation in symmetry observed in these models could be an artefact of crystallisation or real flexibility within the dimerisation interfaces of SAS-6. Indeed, greater variability in symmetry than that implied by the crystallographic structures was observed *in vitro*. In negative stain electron microscopy (EM) experiments CeSAS-6 oligomers displayed symmetries ranging from 7- to 11-fold [40]. Similar symmetry variation was seen in atomic force microscopy (AFM) and rotary metal shadowing experiments of CrSAS-6 oligomers, where symmetries ranging from 7.8- to 9.4-fold were observed [40, 133, 159] (Fig. 3.2 B). This level of variability suggests SAS-6 oligomer heterogeneity is real and not an artefact of crystallisation, and is representative of flexibility at the C-C and N-N interfaces.

Further variation in oligomer symmetry was introduced by rational engineering of the N-N interface (Fig. 3.2 B). These variants displayed symmetries ranging from 5- to 10-fold as determined by EM, AFM and crystallography [133].

3.4 Molecular dynamics simulations of SAS-6 oligomers

Further insight into symmetries sampled by SAS-6 oliogmers was obtained *in silico* by atomistic molecular dynamics simulations of models comprising either the N-N dimer or the C-C dimer of SAS-6. Molecular dynamics simulations of SAS-6 C-C dimers were run for 100 ns. At 1 ns intervals the C-C dimer structure was used to build an oligomer containing 9 dimers with a rigid N-N interface. The symmetry, pitch/turn and radius of the resulting oligomers are shown (Fig. 3.4, 3.5, 3.6). The symmetry of wild-type CeSAS-6 spirals fluctuates around 4.5-fold as expected for a single spiral (Fig. 3.4 A). The ring-forming SAS-6 variants; LmSAS-6, CrSAS-6 and DrSAS-6; showed greater variability in symmetries sampled by the oligomer. However, fluctuations centred around 9- to 10-fold symmetry (Fig. 3.4 B, C and D). As predicted by the spiral model built from the crystallographic structure of CeSAS-6, the pitch/turn of the spiral fluctuated about 30 nm ranging from 36 to 26 nm (Fig. 3.5 A). Ring-forming dimers simulated showed substantial fluctuations in pitch/turn. CrSAS-6 oligomer pitch/turn ranged from 20 to 65 nm, forming a left-handed spiral (Fig. 3.5 B). The pitch/turn of the LmSAS-6 oligomer ranged from 16 to 47 nm (Fig. 3.5 C). DrSAS-6 oligomer pitch/turn ranged from 0 to 72nm (Fig. 3.5 D). The radius of the spiral built from CeSAS-6 crystallographic models was 3.5 nm (Fig. 3.6 A). During the course of the 100 ns simulation the radius varied from 1.9 to 4.5 nm. The expected radius of 9-fold symmetric ring is approximately 11 nm. CrSAS-6 oligomer radius fluctuated around 20 nm ranging from 12 to 57 nm (Fig. 3.6 B). The radius of both LmSAS-6 and DrSAS-6 oligomers was closer to the expected value with reduced variability compared to CrSAS-6. LmSAS-6 radius centred around 10 nm ranging from 7.2 to 16.5 nm (Fig. 3.6 C). Similarly, DrSAS-6 oligomer radius ranged from 7 to 15 nm (Fig. 3.6 D). These simulations suggest

that SAS-6 oligomers display symmetry, pitch/turn and radius similar to canonical centriolar values but significant variation away from these values is observed.

The position of α-3 relative to the N-terminal domain varies slightly between crystallographic structures of a single SAS-6 variant. Comparison of the position of α-3 in the three availiable crystalloagraphic structures of CeSAS-6 shows its position varies by 9.2° (Fig. 3.7 A). A similar analysis of LmSAS-6 shows there is more variability in the position of α-3 in the 5 available C-C dimer crystallographic structures. In LmSAS-6 the position of α-3 varies by 12.5° (Fig. 3.7 B). The increased variation in symmetries sampled by SAS-6 C-C dimers seen in molecular dynamics simulations is reflected in an increased variation in the position of α-3 relative to the N-terminal domain. The range in symmetry sampled by each of the SAS-6 variant C-C dimers is summarised in table 3.1. Between the minimum and maximum symmetry sampled by CeSAS-6 during molecular dynamics simulations of the C-C dimer, the position of α-3 varies by 16.7° (Fig. 3.7 C). Similarly, the increased variability in symmetry sampled by LmSAS-6 during molecular dynamics simulations was also reflected by increased flexibility in the positions sampled by α-3. Between the minimum and maximum symmetries sampled by LmSAS-6 during molecular dynamics simulations, the position of α-3 varied by 27° (Fig. 3.7 D).

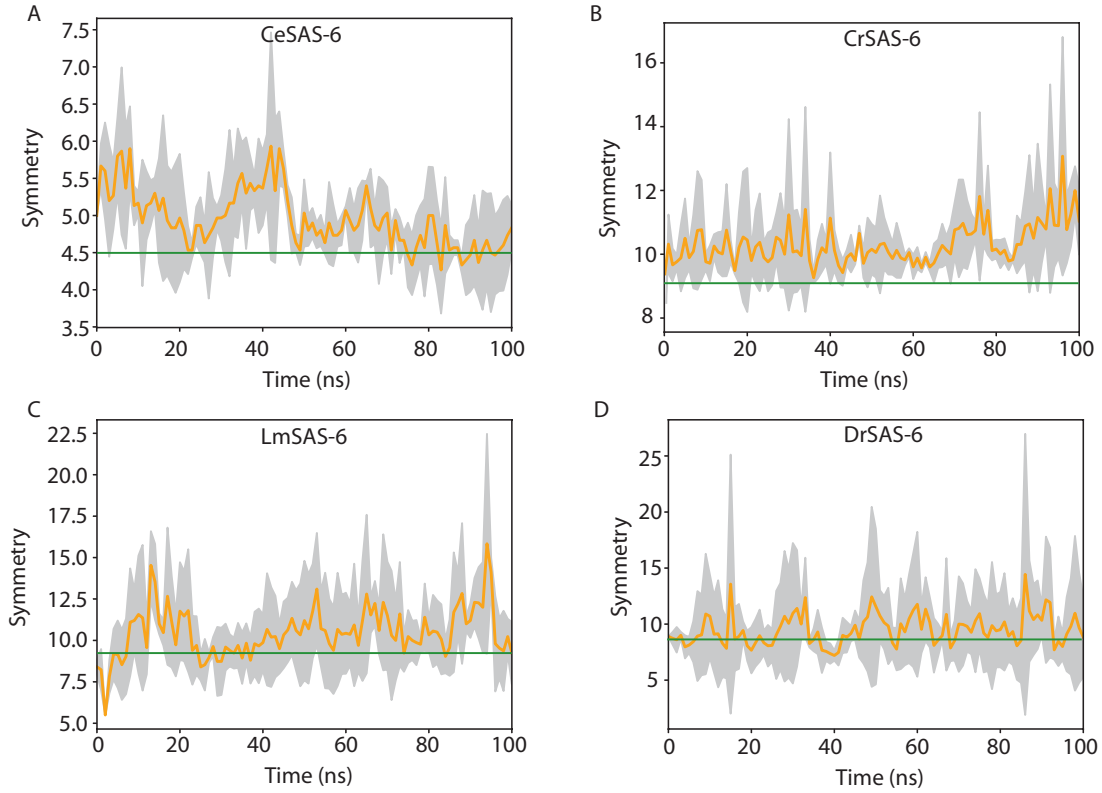


Figure 3.4: Variation in symmetry produced by flexibility in the C-C SAS-6 dimer interface. Symmetries sampled during 100 ns molecular dynamics simulations of A: CeSAS-6, B: CrSAS-6, C: LmSAS-6 and D: DrSAS-6. The mean (orange) and standard deviation (grey) of triplicate simulations are shown, green lines indicate canonical symmetry.

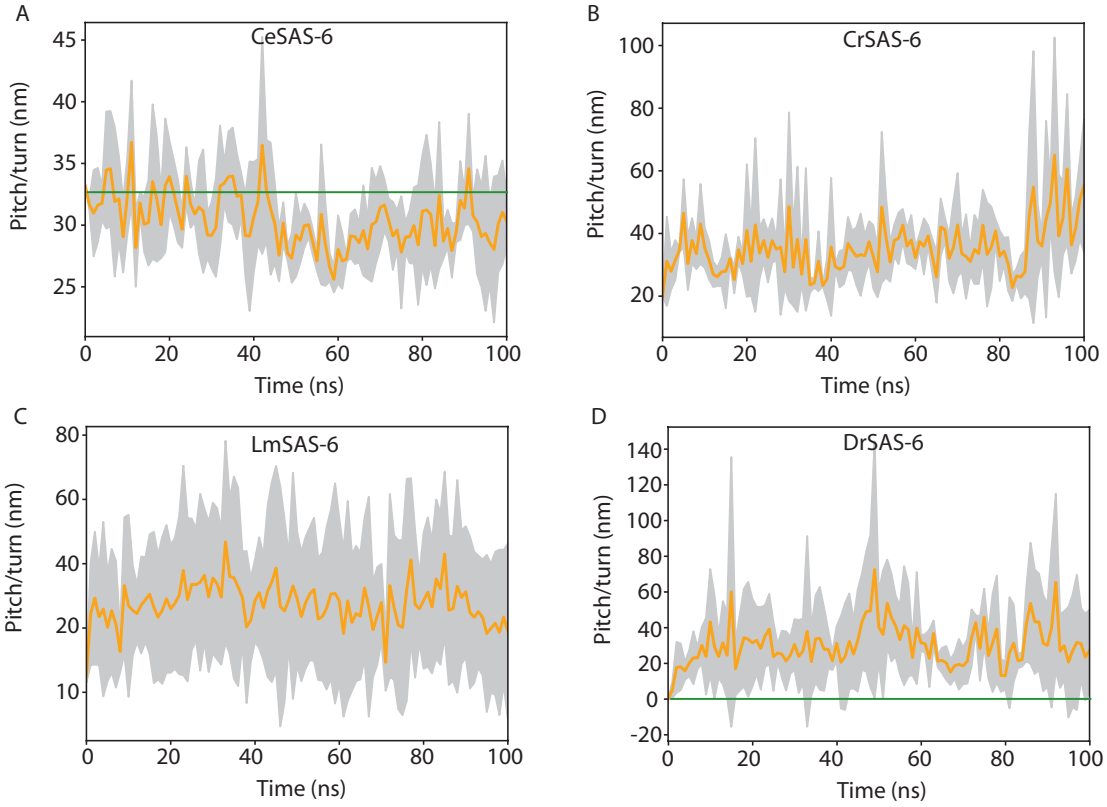


Figure 3.5: Variation in pitch/turn produced by flexibility in the C-C SAS-6 dimer interface. Pitch/turn sampled during 100 ns molecular dynamics simulations of A: CeSAS-6, B: CrSAS-6, C: LmSAS-6 and D: DrSAS-6. The mean (orange) and standard deviation (grey) of triplicate simulations are shown, green lines indicate canonical pitch/turn (0 nm for CrSAS-6 and LmSAS-6).

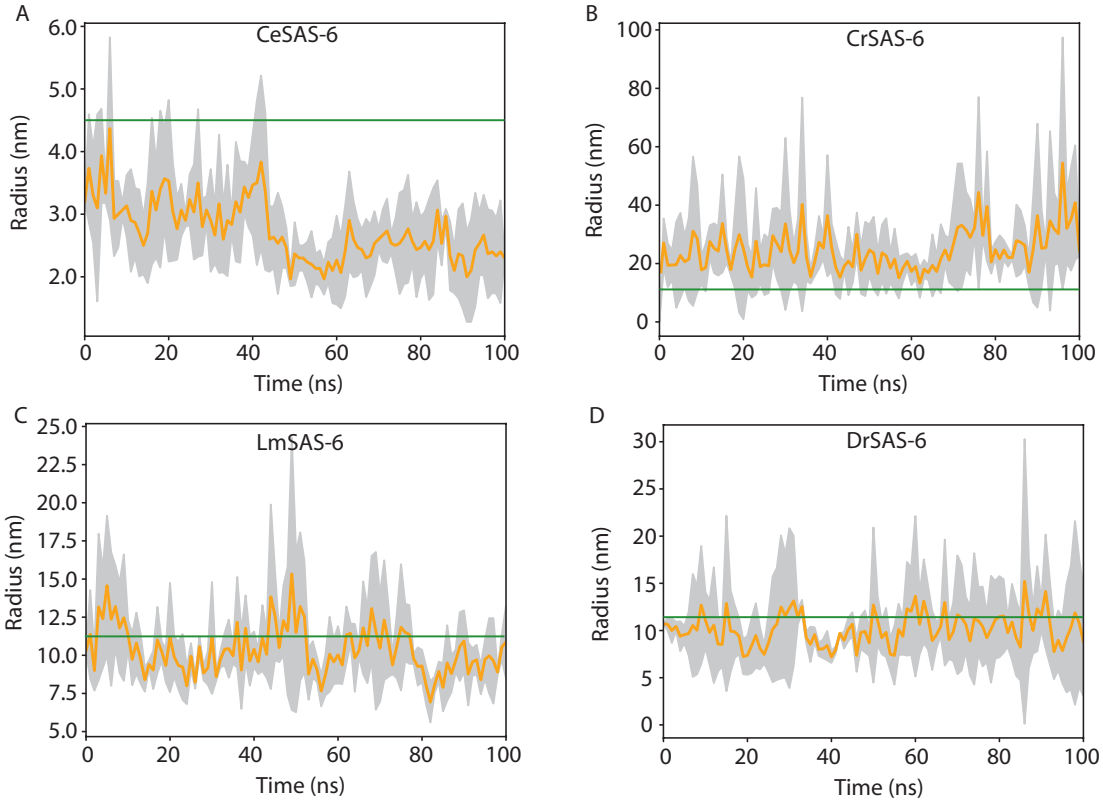


Figure 3.6: Variation in radius produced by flexibility in the C-C SAS-6 dimer interface. Radii sampled during 100 ns molecular dynamics simulations of A: CeSAS-6, B: CrSAS-6, C: LmSAS-6 and D: DrSAS-6. The mean (orange) and standard deviation (grey) of triplicate simulations are shown, green lines indicate canonical radius.

Table 3.1: Symmetry sampled by SAS-6

Species	**Crystallographic Structures**	**MD simulations**
C. elegans	4.5 - 5.4	4.3 - 5.9
C. reinhardtii	9.8 - 10.3	9.1 - 13.2
L. major	8.3 - 10.8	5.3 - 16.1
D. rerio	7.8 - 8.2	7.5 - 15

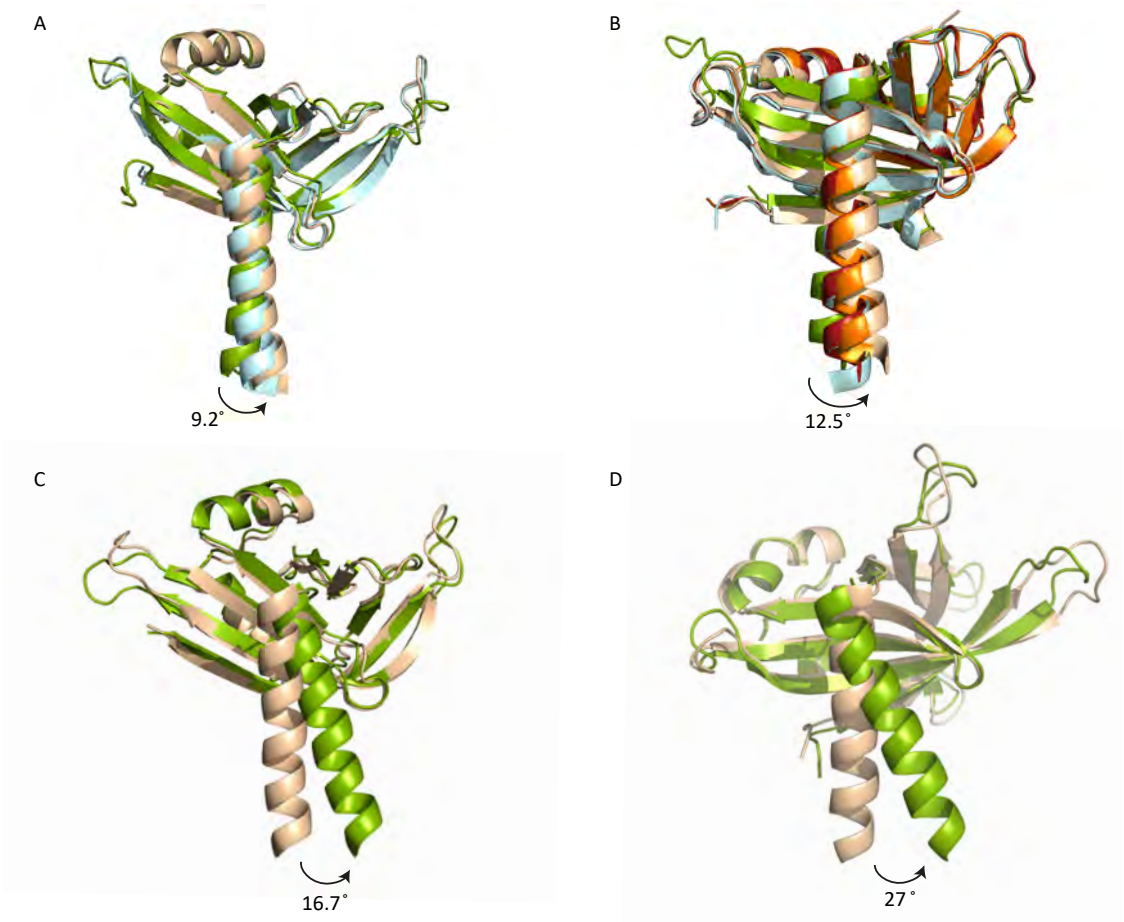


Figure 3.7: Flexibility in α-3 position. Structurally aligned SAS-6 monomers showing the range of movement of α-3 relative to the N-terminal domain. A: Crystallographic structures of CeSAS-6 (PDB code: 4GFC, chain A green; PDB code: 4GFA, chain A blue, chain C grey). B: Crystallographic structures of LmSAS-6 (PDB code: 4CKN, chain A green, chain C blue; PDB code: 4CKP, chain A grey, chain C orange, chain E red). C: Structural overlay of CeSAS-6 monomers taken from molecular dynamics simulations of the C-C interface when minimum symmetry (4.3-fold) was sampled (green) and when maximum symmetry (5.9-fold) was sampled (grey). D: Structural overlay of LmSAS-6 monomers taken from molecular dynamics simulations of the C-C interface when minimum symmetry (7.5-fold) was sampled (green) and when maximum symmetry (15-fold) was sampled (grey). Arrows indicate angle between α-3 at the extreme of its position.

The above simulations likely represent an underestimate of the flexibility within the oligomer, since only the C-C interface was varied. By combining independent molecular dynamics simulations of the N-N and C-C dimerisation interfaces the full range of oligomer variability can be sampled. Snap-shots of the trajectories of molecular dynamics simulations of the C-C and N-N interfaces were taken every nanosecond and oligomers built from each combination of interfaces, resulting in 10,000 distinct oligomer models. The mean symmetry sampled by CeSAS-6 models was 3.5±1.5-fold (Fig. 3.8 A). The symmetries sampled by oligomers resulting from this simulation deviated significantly compared to the mean value of oligomer symmetry inferred from CeSAS-6 crystallographic structures (p=<0.001, v=49496000) (p values were calculated using a Wilcoxon signed rank test). Significant numbers of models derived from all ring-forming SAS-6 variants sampled symmetries greater than 10-fold, with represented symmetries ranging from 4-fold to 21-fold. CrSAS-6 models had a mean symmetry of 6.7±2.5-fold (Fig. 3.8 B). Symmetry sampled by oligomers derived from the simulation were significantly different from the mean symmetry inferred from CrSAS-6 crystallographic structures (p=<0.001, v=19599000). LmSAS-6 models had a mean symmetry of 10±2.5-fold (Fig. 3.8 C). Again, symmetries sampled by LmSAS-6 oligomers derived from molecular dyanmics simulations deviated significantly from the mean symmetry inferred from LmSAS-6 crystallographic structures (p=<0.001, v=33038000). DrSAS-6 models had a mean symmetry of 7.5±1.5-fold (Fig. 3.8 D). Similarly, there was significant deviation in the symmetries sampled by DrSAS-6 oligomers resulting from molecular dynamics simulations when compared to the mean symmetry inferred from DrSAS-6 crystallographic structures (p=<0.001, v=5005800). This suggests that the variation seen in the crystallographic SAS-6 structures is an underestimate of the true flexibility within SAS-6 oligomers.

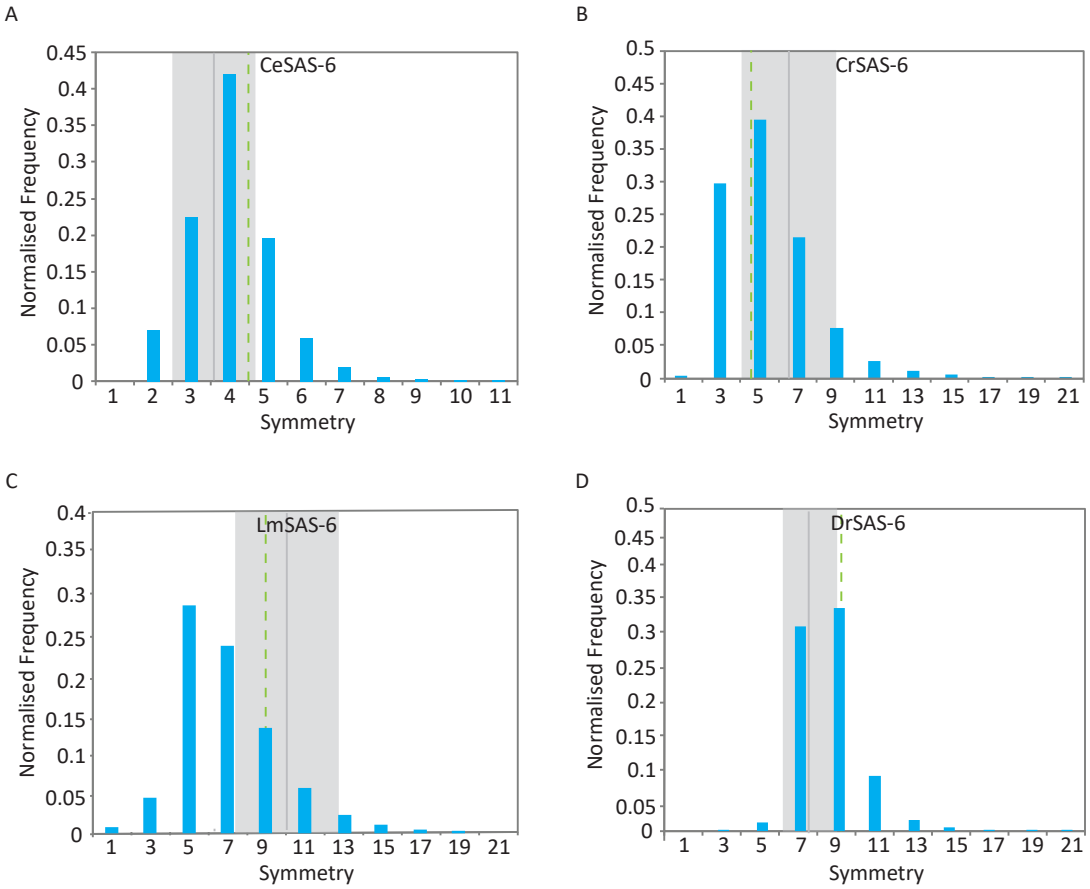


Figure 3.8: SAS-6 oligomer symmetry heterogeneity. Frequency of oligomer symmetry sampled over 100 ns molecular dynamics simulation when both N-N and C-C dimers are varied. Values derive from 10,000 models. A: CeSAS-6, B: CrSAS-6, C: LmSAS-6 and D: DrSAS-6. Mean and standard deviation are shown in grey, green lines indicate canonical symmetry.

3.5 Discussion

In this chapter the variation in symmetry represented by SAS-6 oligomers from *C. elegans*, *C. reinhardtii*, *L. major* and *D. rerio* was demonstrated through static modelling of the N-N and C-C interface from crystallographic structures and models generated through molecular dynamics simulations. There exists flexibility in both the N-N and C-C interfaces of SAS-6 and this flexibility is underestimated when oligomer symmetry is inferred from crystallographic structures alone. N-N and C-C interfaces of LmSAS-6, CrSAS-6 and DrSAS-6 variants are sufficiently flexible to sample both ring and shallow spiral conformations. However, the radius of these oligomers does not deviate greatly from the expected value for a 9-fold symmetric ring and the pitch/turn is generally shallower than that of CeSAS-6 spiral oligomers. Spirals built from LmSAS-6 and CrSAS-6 variants are left-handed. In contrast, DrSAS-6 spirals are right-handed, as are CeSAS-6 spirals. However, the high fold symmetry and large radius of this oligomer does not resemble a CeSAS-6 spiral. Thus, although heterogeneity exists in the architecture of *C. reinhardtii*, *L. major* and *D. rerio* SAS-6 oligomers, their architecture remains distinct from that of CeSAS-6.

Electron microscopy of *C. reinhardtii* centrioles suggest the shallow spiral oligomers resulting from the crystallographic models are artefacts, at least in this species [97]. *In vivo* additional factors must be involved in determining cartwheel symmetry, as demonstrated by the canonical 9-fold symmetry of centrioles in the presence of SAS-6 N-N interface variants engineered to sample alternative *in vitro* symmetries [133].

4

SAS-6 mutant design

Contents

4.1 Introduction 63
4.2 Sequence analysis 64
4.3 SAS-6 mutations 67
4.3.1 Negative design elements 68
4.3.2 Positive design elements 75
4.4 Discussion 77

4.1 Introduction

To re-engineer spiral-forming CeSAS-6 to adopt a ring conformation amino acid variants were designed. An analysis of the important interactions present at the interface between α3 and the N-terminal domain was carried out. Structural comparisons were made across CeSAS-6, CrSAS-6, LmSAS-6, DrSAS-6 and *Drosophila melanogaster* SAS-6 (DmSAS-6). An assessment of covariant residue pairs at the interface was also used to identify conserved interacting amino acids. In this manner, ring-specific and spiral-specific interactions were identified and could therefore be emulated, or removed, in CeSAS-6. The aim of the mutations was to move CeSAS-6 α3 to a position compatible with a ring-like oligomer conformation; a rotation of 30° relative to the N-terminal domain.

4.2 Sequence analysis

Although CeSAS-6 shares high structural similarity with other SAS-6 homologues, it shares only 22% sequence identity with *H. sapiens* SAS-6, for example. Similar sequence divergence is present between all SAS-6 homologues. EVcouplings [150] was used to identity covariant pairs within the SAS-6 family. First, a multiple sequence alignment was made. Then pairs of residues that are close to one another in space were identified. Since not all pairs that are close together will be covariant, a subgroup of residue pairs which change simultaneously are identified. Pairs of residues from this group are then scored. Scores are normalised for protein sequence length and the number of non-redundant sequences used in the alignment as this will influence the background of residue pairs which are close in space but either do not interact with one another or are not covariant. These evolutionary constraints are then mapped onto the 3D protein structure to generate a contact map. Only residue pairs above a given threshold are likely to be classified as covariant [160] as this is predictive of the most conserved residue-residue contacts.

A multiple sequence alignment was generated from SAS-6 homologues. For a reliable estimate of covariance the number of sequences used must ideally be 5x the number of residues in the target protein. 398 sequences were used in the LmSAS-6 alignment and 466 were used in the DrSAS-6 alignment. Different numbers of sequences were used in the alignments as only those which met threshold values of coverage and sequence identity were included. The number of sequences is lower than the ideal number for SAS-6 ($\sim$1000) and will reduce the reliability of the covariance analysis; however, it is above the minimum requirement of 0.3x the number of residues in the target sequence [160]. Larger numbers of sequences were not included by relaxing the threshold as this would increase the number of non-SAS-6 sequences used in the alignment, reducing reliability of the covariance analysis. At the interface between the N-terminal domain and α3, 2 covariant pairs were seen when DrSAS-6 was used as the target sequence and 1 pair when LmSAS-6 was used as the target sequence above a threshold score of 0.2. This threshold balances the false positives that would result from a threshold that is too low and covariant pairs

that would not be considered by a too stringent threshold. No covariant pairs were identified in this region when CeSAS-6 or CrSAS-6 sequences were used as the target sequence. This is likely a result of the non-ideal alignment of these targets meaning equivalent covariant pairs were not identified across all species, as would have been expected. While the pairs of residues identified to be covariant when LmSAS-6 and DrSAS-6 were used as the target sequence are not the same, they are structurally in similar locations, which provides some confidence in the covariance analysis.

In DrSAS-6 R47 and Y152 are covariant (Fig. 4.1 A and B). The pi stacking between these residues may influence α3 orientation making this pair a good candidate for rational engineering in CeSAS-6. L57 and G144 (hinge) in DrSAS-6 are also covariant (Fig. 4.1 A and C). The steric effect of the leucine and hydrogen bonding between its backbone amide and the carboxyl group on the backbone of G144 may also influence α3 orientation. In LmSAS-6, L186 and Q270 (hinge -1) are covariant (Fig. 4.1 D and E). In a manner similar to that of L57 and G144 in DrSAS-6, the steric effect of the leucine side chain contributes to positioning α3 in an orientation which may result in oligomers with a ring conformation.

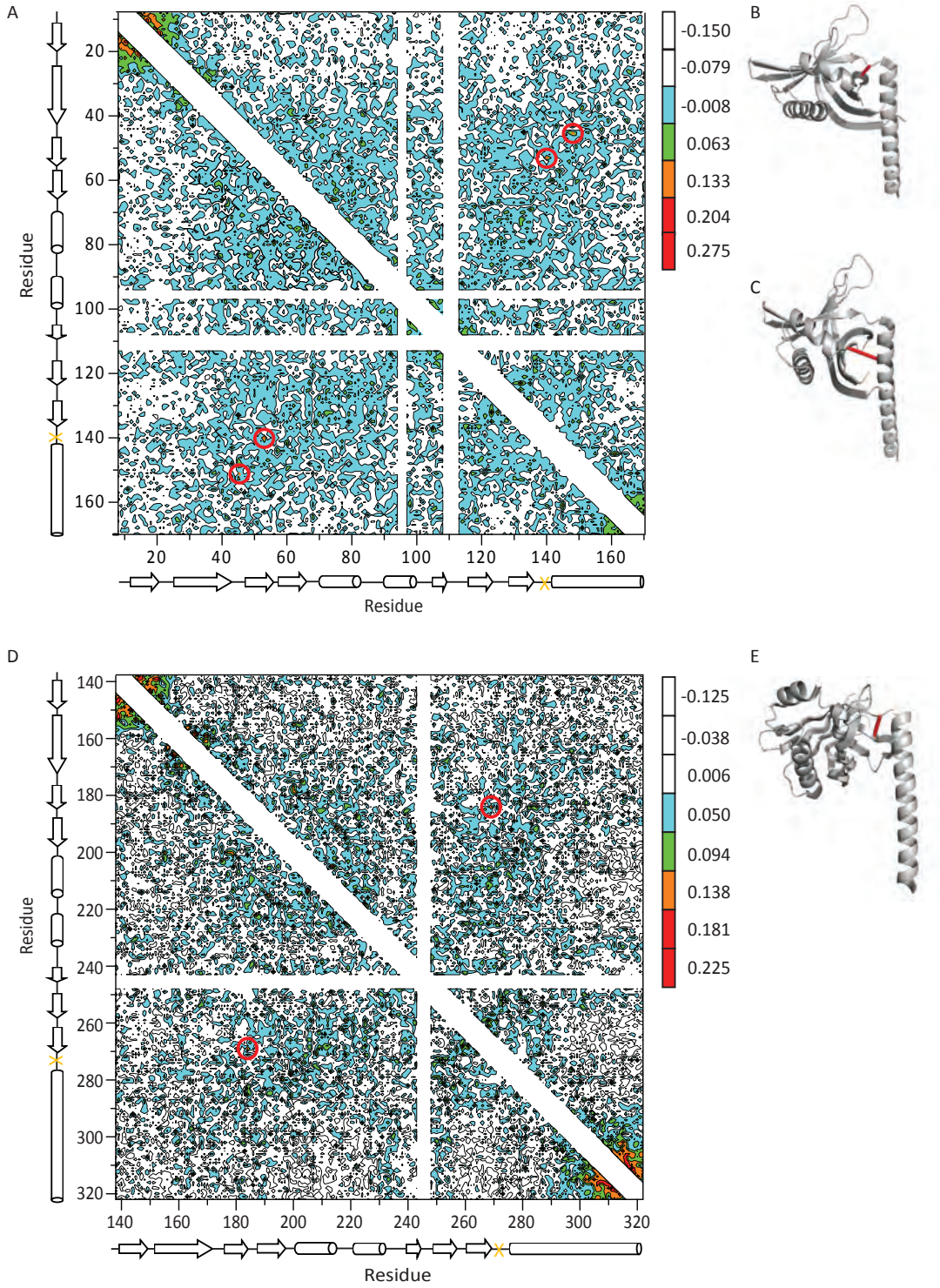


Figure 4.1: SAS-6 covariance analysis. A: Heatmap of covariant scores of each residue pair when DrSAS-6 was used as the starting sequence. Residues of interest highlighted by red circles. SAS-6 topology is shown along the axes with the hinge residue highlighted (yellow). The circles correspond to G144 and L57, and R47 and Y152. B: Covariant pair G144-L57 (red line) mapped onto the DrSAS-6 structure. C: Covariant pair R47-Y153 (red line) mapped onto the DrSAS-6 structure. D: Heatmap of covariant scores of each residue pair when LmSAS-6 was used as the starting sequence, L186-Q270 is highlighted by the red circles. E: Covariant pair L186-Q270 mapped onto the LmSAS-6 structure.

4.3 SAS-6 mutations

Structural analysis suggests that it is the steric effect of V167 at the hinge position in CeSAS-6 that is responsible for the alternative orientation of α3 leading to the formation of spiral oligomers. Thus, all CeSAS-6 variants were designed with a V167G mutation to remove the steric effect of this residue at the hinge and allow movement of α3. Mutations were introduced with the aim of bringing α3 back to the position seen in ring-forming SAS-6 variants. Therefore, mutations were introduced only at the interface between α3 and the N-terminal domain (Fig. 4.2). In addition to the covariance analysis discussed, mutations were also rationally engineered based on PDBePISA [151] analysis to identify interacting residues at the α3-N-terminal domain interface, in addition to manual inspection to identify invariant features. The mutations introduced can be split into 2 groups: those that aim to force rotation of α3 to adopt an orientation compatible with a ring-like oligomer conformations (positive design elements), and those that seek to 'undo' interactions seen in CeSAS-6 but not in ring-forming SAS-6 variants (negative design elements).

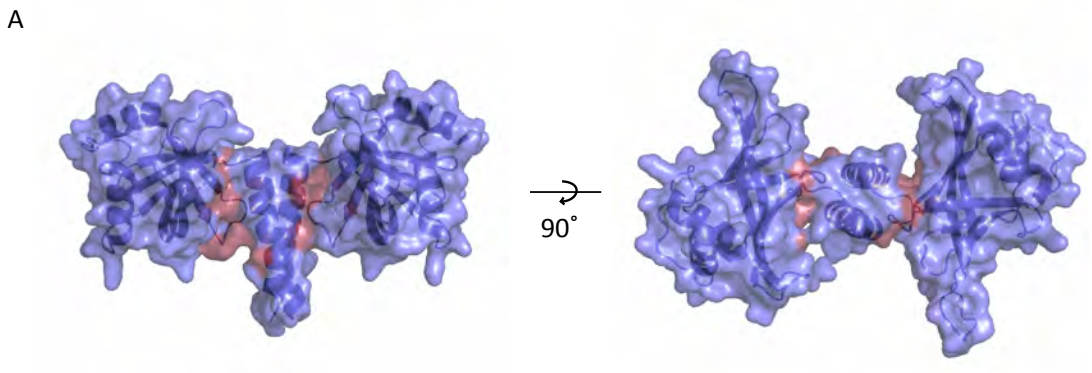


Figure 4.2: Location of mutations on SAS-6 dimer. CeSAS-6 C-C dimer is shown in blue with residues where mutations were introduced highlighted in red.

4.3.1 Negative design elements

In CeSAS-6, H175 on the coiled-coil interacts via a hydrogen bond with E54 in the N-terminal domain. Since an interaction at this location was not observed in ring-forming SAS-6 variants, a H175Y mutation was introduced to disrupt the pre-existing interaction (Fig. 4.4 G, 4.6 G).

Residues equivalent to L64 and V167 in CeSAS-6 were identified as covariant in ring-forming SAS-6. In CeSAS-6 these residues form a hydrophobic interaction. An L64K mutation was introduced to disrupt this interaction in combination with H175Y to disrupt the packing between L64 and H175 and mimic the conserved tyrosine at this position in ring-forming SAS-6 variants (Fig. 4.3 H, 4.6 H).

In CeSAS-6 F61 packs against C180. This interaction is not present in ring-forming variants, an F61K mutation was therefore introduced to disrupt this packing (Fig. 4.3 D, 4.5 D).

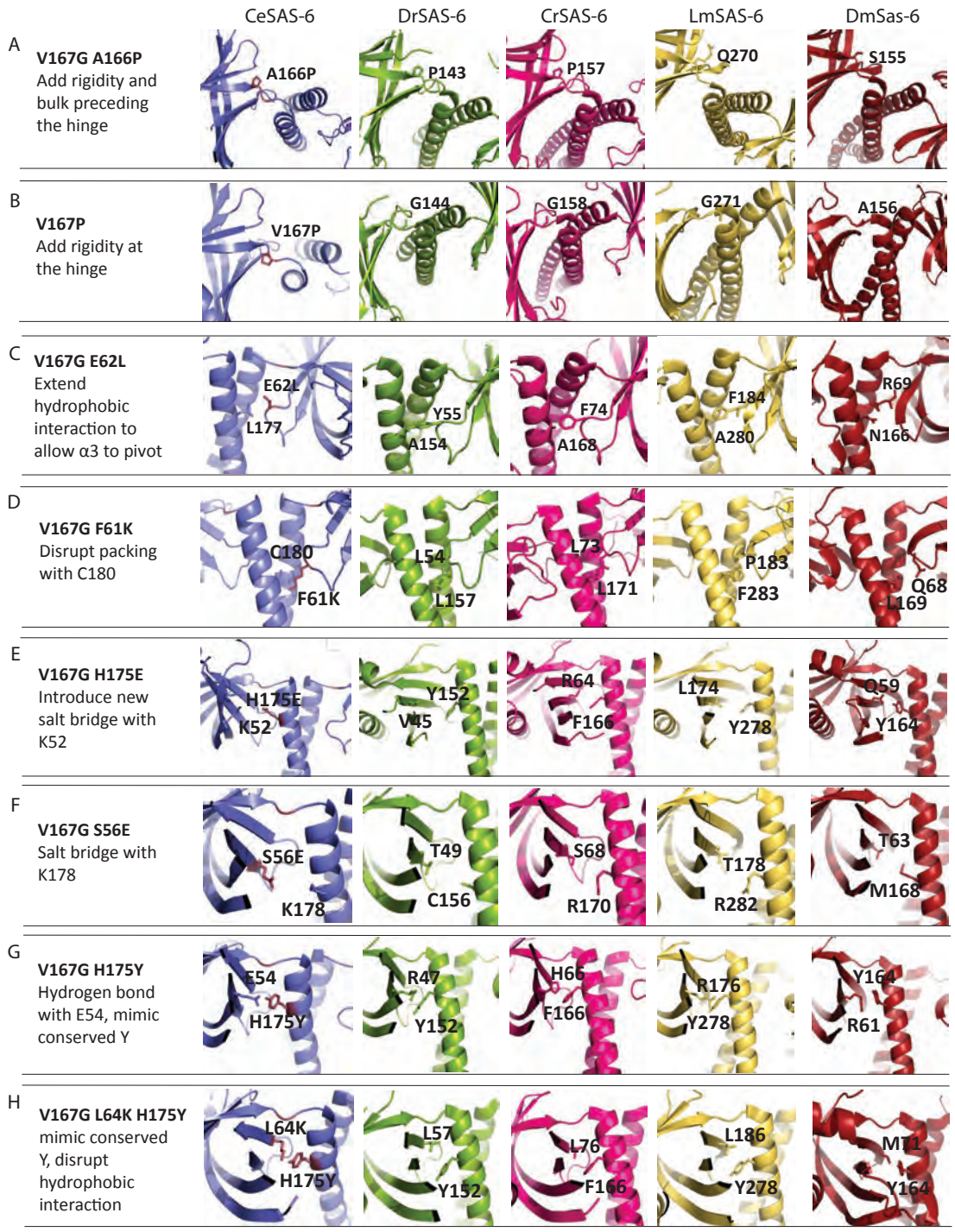


Figure 4.3: Mutations introduced to CeSAS-6 (i). Variants 1-8 are shown. C-C dimers of SAS-6 are shown. The reason for the mutation is noted in column 1, CeSAS-6 (blue) is shown, with mutated residues highlighted in red. The residues labeled indicate those in SAS-6 variants equivalent to those labelled in CeSAS-6, DrSAS-6 (green), CrSAS-6 (pink), LmSAS-6 (yellow) and DmSAS-6 (red).

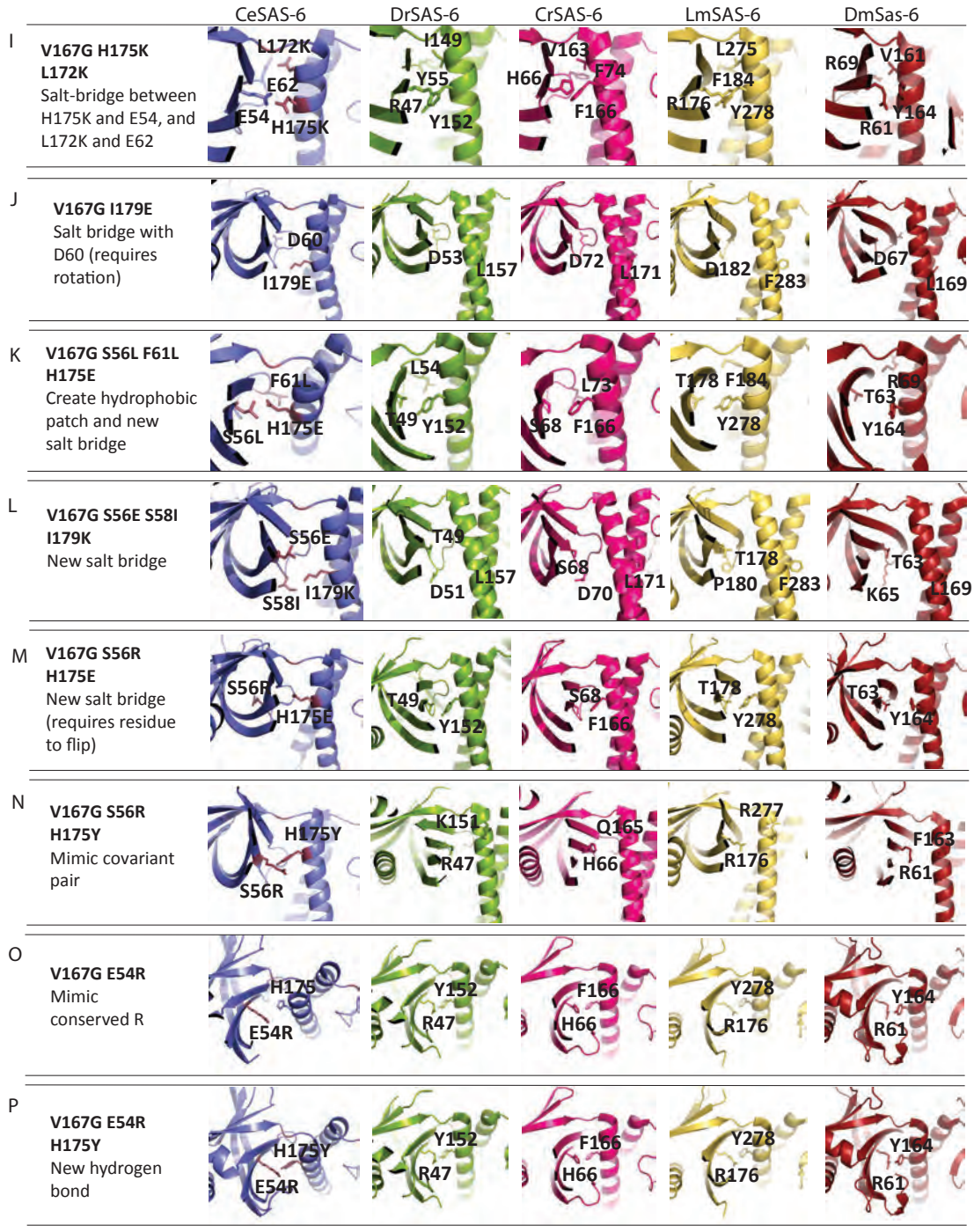


Figure 4.4: Mutations introduced to CeSAS-6 (ii). Variants 9-16 are shown. C-C dimers of SAS-6 are shown. The reason for the mutation is noted in column 1, CeSAS-6 (blue) is shown, with mutated residues highlighted in red. The residues labeled indicate those in SAS-6 variants equivalent to those labelled in CeSAS-6, DrSAS-6 (green), CrSAS-6 (pink), LmSAS-6 (yellow) and DmSAS-6 (red).

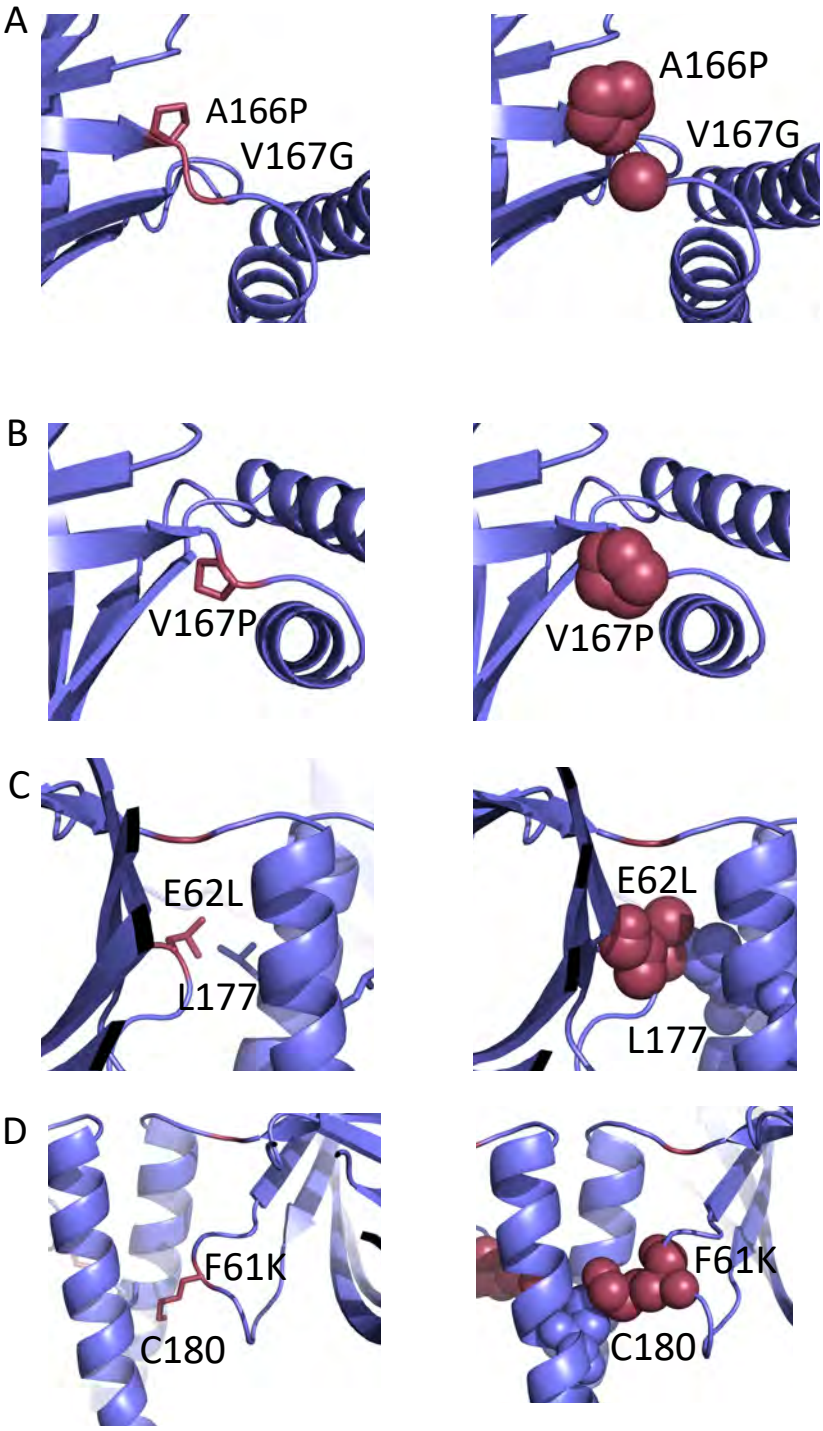


Figure 4.5: CeSAS-6 Mutants (i). The mutations (red) introduced into CeSAS-6 (blue) are shown. Interacting residues are shown in stick form (left column) and as spheres (right column). Labels indicate mutated residues and residues involved in the interaction. A: V167G A166P, B: V167P, C: V167G E62L, D: V167G F61K.

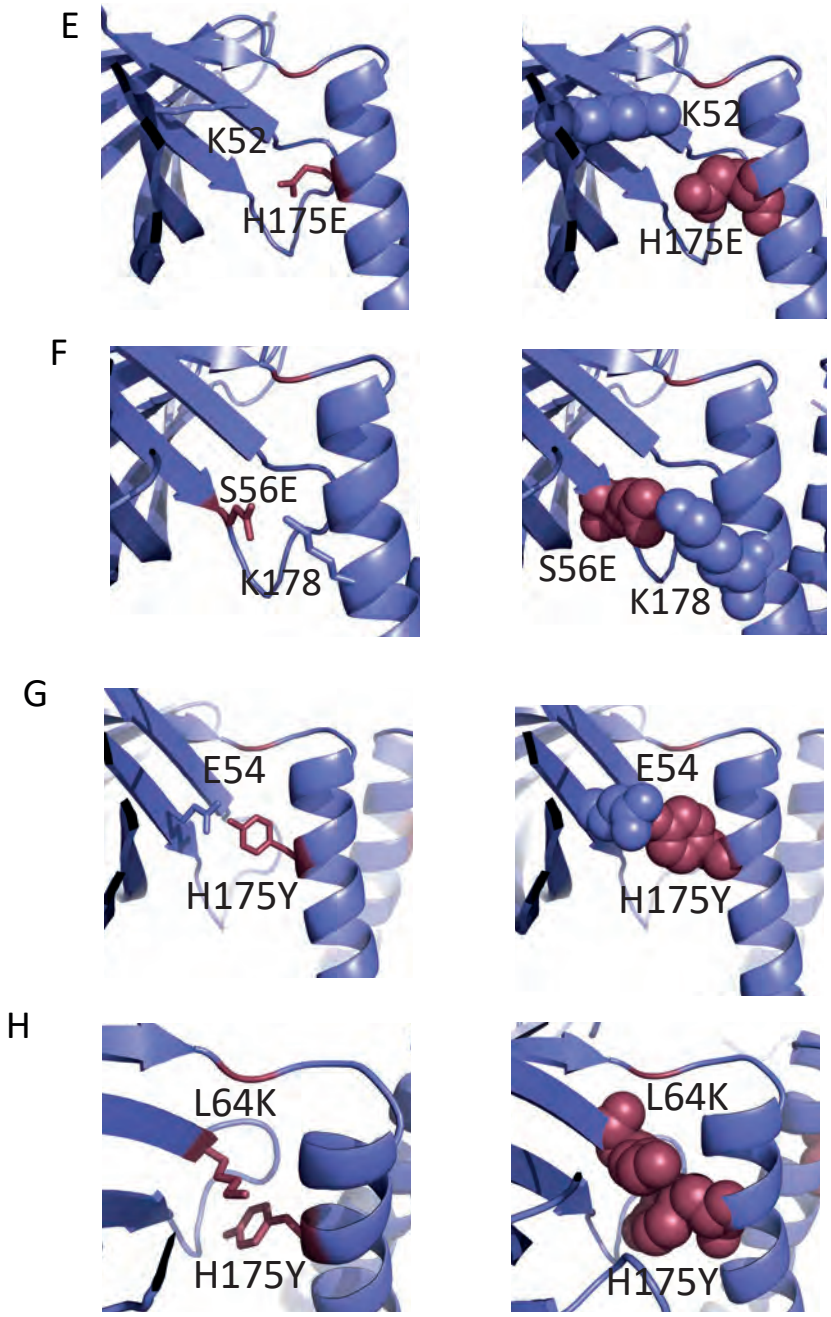


Figure 4.6: CeSAS-6 Mutants (ii). The mutations (red) introduced into CeSAS-6 (blue) are shown. Interacting residues are shown in stick form (left column) and as spheres (right column). Labels indicate mutated residues and residues involved in the interaction. E: V167G H175E, F: V167G S56E, G: V167G H175Y, H: V167G L64K H175Y.

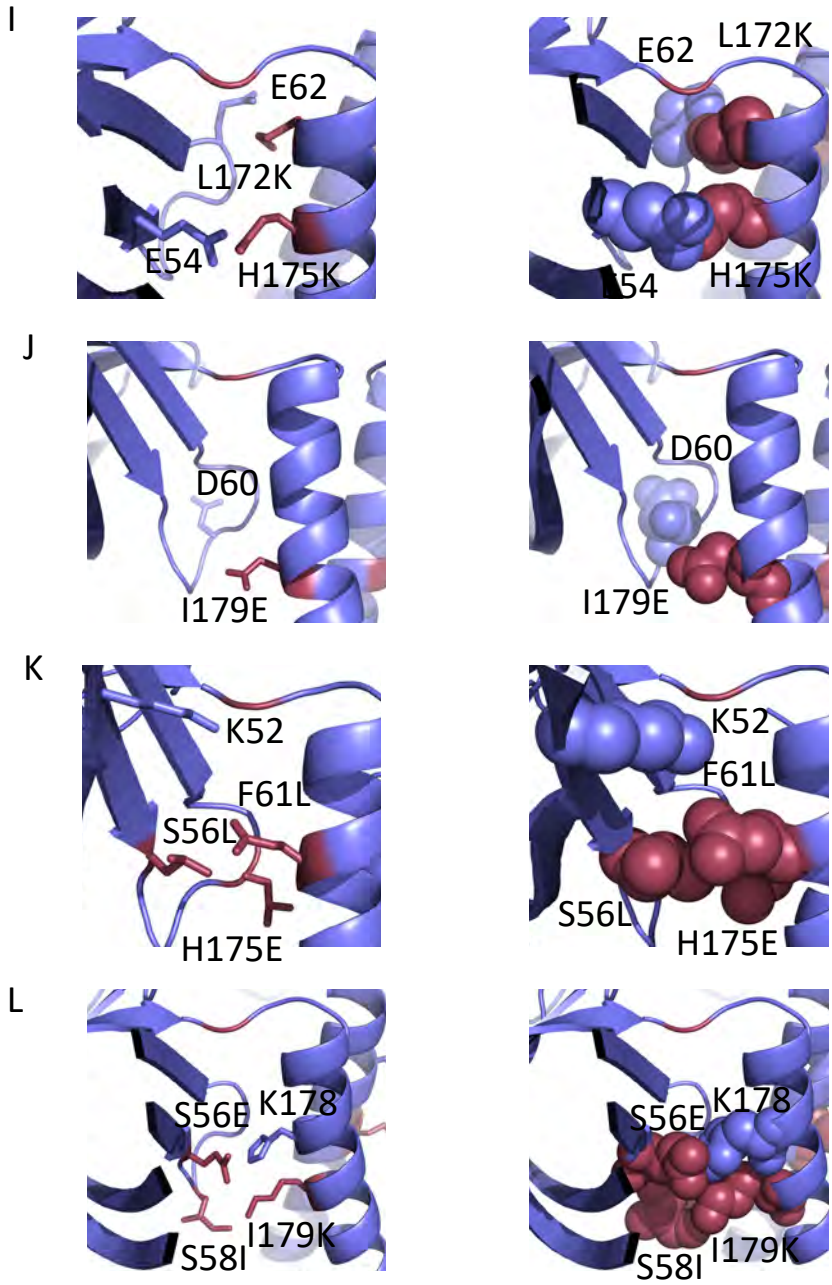


Figure 4.7: CeSAS-6 Mutants (iii). The mutations (red) introduced into CeSAS-6 (blue) are shown. Interacting residues are shown in stick form (left column) and as spheres (right column). Labels indicate mutated residues and residues involved in the interaction. I: V167G L172K H175K, J: V167G I179E, K: V167G S56L F61L H175E, L: V167G S56E S58I I179K.

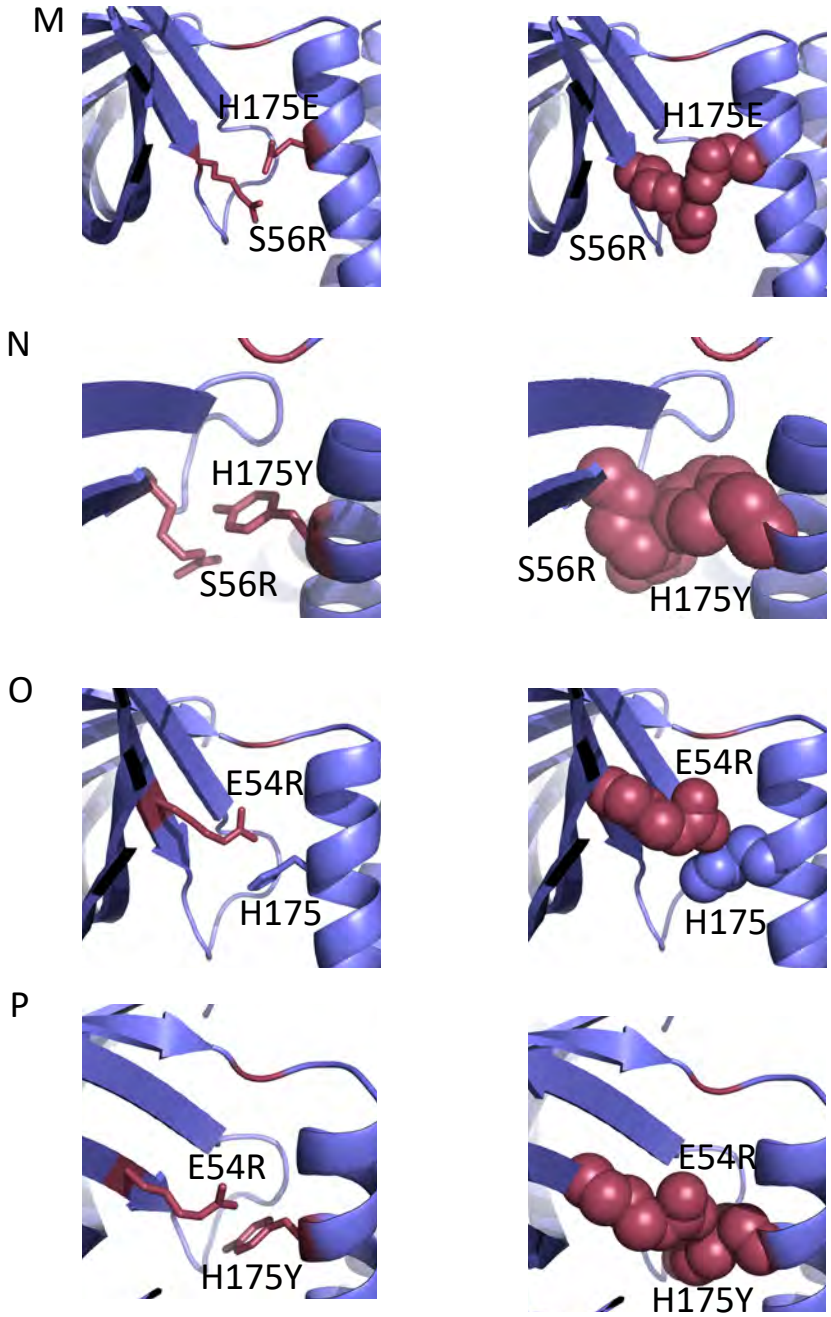


Figure 4.8: CeSAS-6 Mutants (iv). The mutations (red) introduced into CeSAS-6 (blue) are shown. Interacting residues are shown in stick form (left column) and as spheres (right column). Labels indicate mutated residues and residues involved in the interaction. M: V167G S56R H175E, N: V167G S56R H175Y, O: V167G E54R, P: V167G E54R H175Y.

4.3.2 Positive design elements

Residues equivalent to S56 and H175 in CeSAS-6 were identified as a covariant pair in ring-forming SAS-6 which interact via a hydrogen bond. This interaction was mimicked in CeSAS-6 by introducing S56R H175Y mutations (Fig. 4.4 N, 4.8 N). Pi stacking between these 2 residues would mimic that which occurs in ring-forming SAS-6 and promote orientation of α3 into a position compatible with a ring-like oligomer conformation.

Residues equivalent to L98 and A166 were also identified as covariant in ring-forming SAS-6. L98 is a conserved residue (L186 in LmSAS-6) and was therefore not altered. The hinge -1 position, is a conserved proline in ring-forming SAS-6 homologues. An A166P mutation was introduced to mimic this, adding rigidity prior to the hinge (Fig. 4.3 A, 4.5 A). Similarly, a V167P mutation was introduced to determine the effects of added rigidity at the hinge (Fig. 4.3 B, 4.5 B).

4.3.2.1 Salt-bridges

An S56R mutation was introduced in combination with a H175E mutation to introduce a new salt-bridge between the N-terminal domain and α3, which should rotate α3 in the desired direction (Fig. 4.4 M, 4.8 M). Similarly, a H174K mutation was introduced so a new salt-bridge with E54 could occur. In addition to the H175K mutation an L172K mutation was introduced to allow formation of a second salt-bridge with E62 (Fig. 4.4 I, 4.7 I). H175 was also mutated to a glutamic acid to allow the formation of a new salt-bridge with K52 (Fig. 4.3 E, 4.6 E). This interaction would pull α3 towards a position compatible with a ring-like oligomer conformation.

E54 in CeSAS-6 was mutated to an arginine to mimic the conserved arginine seen at this position in ring-forming SAS-6 variants (Fig. 4.4 O, 4.8 O). This mutation was made alone and in combination with H175Y (Fig. 4.4 P, 4.8 P) to emulate the conserved tyrosine residue seen at this position in ring-forming variants.

An S56E mutation was made to introduce a potential new salt-bridge with K178 (Fig. 4.3 F, 4.6 F). Large movement of α3 would be required to bring these two residues close enough to form a salt-bridge.

I179 was also mutated to a lysine in combination with an S56E mutation so a new salt-bridge could form between the 2 residues. This mutation was introduced in combination with S58I to break its existing interaction with H175 and allow α3 to move away from the position it adopts in CeSAS-6 (Fig. 4.4 L, 4.7 L).

4.3.2.2 Hydrophobic interactions

The H175E mutation (4.3.2.1) was also introduced in combination with S56L/F61L to introduce a hydrophobic interaction and disrupt the packing of F61 with C180 (Fig. 4.4 K, 4.7 K). Thus, a combinatorial approach of positive action; to promote rotation of α3 towards a position compatible with ring-like oligomer conformation; and removal of interactions only present in CeSAS-6 to allow such a movement was used.

In order to introduce a hydrophobic interaction with L172 and L176 on the coiled-coil an E62L mutation in the N-terminal domain was introduced (Fig. 4.3 C, 4.5 C). This should facilitate rotation of α3 by extending the hydrophobic patch with which L172 and L176 on α3 can interact.

4.3.2.3 Hydrogen bonds

An I179E mutation was introduced to force rotation of α3 by interaction with D60 creating a new hydrogen bond (Fig. 4.4 J, 4.7 J).

4.4 Discussion

The low sequence identity between SAS-6 homologues makes sequence alignment challenging. Therefore, a greater number of sequences could not be included in the covariance analysis. The alignment used had >75% coverage, which improves the reliability of the predicted evolutionary constraints [160]. Since 2 out of the 3 identified covariant interactions were already present in CeSAS-6, which perhaps demonstrates their structural importance but not their involvement in ring- or spiral-determining properties; this information was used to inform only 2 mutants and will therefore have minimal impact on the engineering process.

A range of interactions were engineered, some requiring significant rearrangements of SAS-6 secondary structure elements. For example, for S56R to form a salt-bridge with H175E, the S56R side chain would have to rotate $\sim 100°$. Since it is unlikely a single mutation will allow for spiral to ring conversion, several mutations that each subtly alter the position of $\alpha 3$ in combination is likely the most successful route for rational engineering of SAS-6. The majority of the engineered interactions are salt-bridges but hydrophobic interactions have also been introduced. For example, an S56L/F61L mutation was introduced to create a hydrophobic patch, in combination with H175E to introduce a new salt-bridge. Other mutations were made to mimic conserved residues found in ring-forming variants. For example, E54R mimics a conserved arginine at the equivalent position in ring-forming SAS-6 variants. The surrounding residues may not match those present in ring-forming variants but the geometry, of A166P for example, will be translated across.

5

Molecular dynamics simulations of SAS-6 variants

Contents

5.1 Introduction 79
5.2 Control simulations 80
5.3 CeSAS-6 mutant molecular dynamics simulations . . . 86
5.4 Discussion 95

5.1 Introduction

To screen CeSAS-6 mutants for their propensity to form spirals or rings prior to being tested *in vitro*, molecular dynamics simulations were used. To assess whether this method could capture a spiral to ring transition, control simulations were performed using models where a ring-forming sequence was threaded onto a spiral-forming structure or vice versa, and monitored for reversal back to the original conformation. The importance of the hinge residue in determining overall oligomer architecture was also assessed by molecular dynamics simulations by removing the steric effect at the hinge in spiral-forming SAS-6 or introducing steric hindrance at the hinge in ring-forming SAS-6.

5.2 Control simulations

To assess whether a spiral to ring transition or vice versa can be captured *in silico*; control simulations of homology models were performed. Here, the sequence of a ring-forming SAS-6 variant, in this case *L. major* was imposed onto the structure of spiral-forming CeSAS-6 (ring onto spiral) by homology modelling. The reverse was also modelled, where the sequence of spiral-forming CeSAS-6 was imposed onto the structure of ring-forming LmSAS-6 (spiral onto ring). Simulations of the resulting C-C dimer models were run for 100 ns and the symmetry, pitch/turn and radius of the resulting oligomers were sampled at 1 ns intervals, as described in Chapter 3.3.

In molecular dynamics simulations of the model where a spiral sequence was imposed onto a ring structure (spiral onto ring), the symmetry immediately reverted to spiral-like values (Fig. 5.1 B). It dropped from 8-fold to 5-fold within the first nanosecond and fluctuated about 4.5-fold symmetry for the remainder of the simulation (Fig. 5.1 D). The radius also dropped from 8.7 nm to 4 nm within the first nanosecond, with the oligomer sampling values close to those observed in the wild-type CeSAS-6 simulation (Fig. 3.6 A) for the remainder of the simulation (Fig. 5.1 F). The pitch/turn was lower than that of wild-type CeSAS-6 spirals but did not drop to 0 nm as would be expected for a ring (Fig. 3.5 A). The minimum pitch/turn sampled was 13 nm (Fig. 5.1 H). This demonstrates the inherent spiral-forming tendency of the CeSAS-6 sequence and that a transition from ring- to spiral-like geometry can be captured using this method.

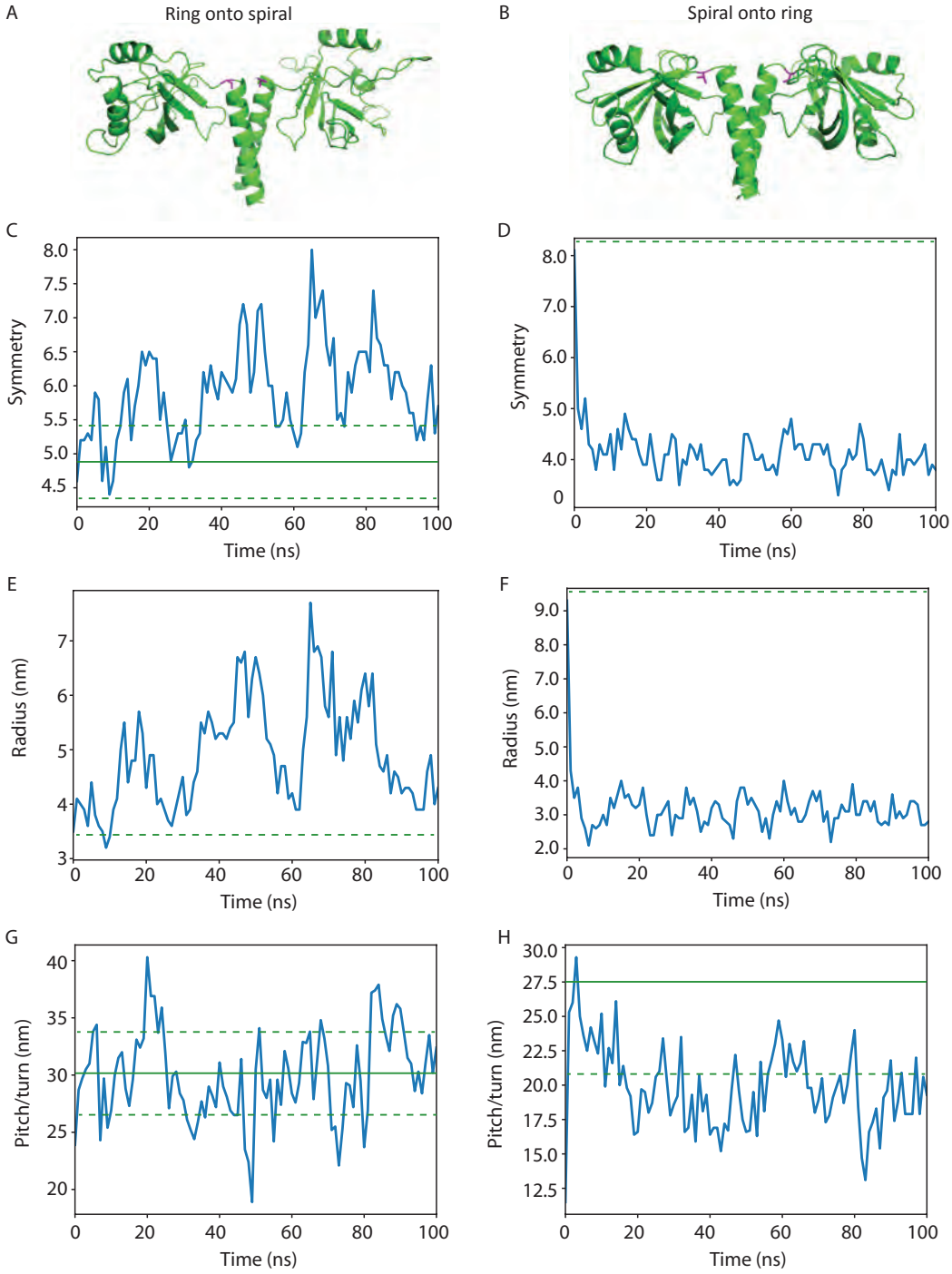


Figure 5.1: Control molecular dynamics simulations. A: LmSAS-6 sequence imposed onto CeSAS-6 structure C-C dimer by homology modelling (ring onto spiral). Hinge G142 residue is highlighted in pink. B: CeSAS-6 sequence imposed onto LmSAS-6 structure C-C dimer by homology modelling (spiral onto ring). Hinge V167 residue is highlighted in pink. The symmetry (D), radius (F), and pitch/turn (H) sampled by oligomers built from the ring onto spiral model are shown. The mean (green solid line) and standard deviation (dashed line) of wild-type CeSAS-6 is indicated in D, F and H. Similarly, symmetry (C), radius (E), and pitch/turn (G) are shown for the spiral onto ring model. The mean (green solid line) and standard deviation (dashed line) of wild-type LmSAS-6 is indicated in C, E and G. Simulations were run for 100 ns.

When the converse was simulated, a ring sequence imposed onto spiral structure (ring onto spiral), geometries sampled by the model tended towards ring-like characteristics (Fig. 5.1 A). Higher-order symmetries were briefly sampled, with a maximum symmetry of 7.9-fold and a mean symmetry of 5.9-fold across 100 ns (Fig. 5.1 C). This is a reduction compared to the 10-fold mean symmetry observed in the wild-type LmSAS-6 simulation (Fig. 3.4 C). The radius ranged from 3.2 to 7.7 nm; smaller than the 7.2 nm to 16.5 nm range sampled by wild-type LmSAS-6 (Fig. 5.1 E, 3.6 C). The pitch/turn was comparable to that seen in wild-type CeSAS-6 spirals with a mean of 29.9 nm. However, low pitch/turn was also briefly represented, with 8 nm being the minimum pitch sampled (Fig. 5.1 G). In contrast, the minimum pitch/turn sampled by wild-type CeSAS-6 was 26 nm (Fig. 3.5 A). The slower reversal towards geometry determined by the sequence rather than the artificially imposed structure indicates adoption of a spiral-forming structure by a ring-forming sequence is more favourable than the converse. This is likely due to the lack of a bulky residue at the hinge resulting in unfavourable geometry. The role of the hinge residue is further investigated below.

To assess the importance of the steric effect of the valine at the hinge position in CeSAS-6; the CeSAS-6 C-C dimer was simulated with a V167G mutation; thus, removing the steric effect at the hinge (Fig. 5.2 A). The symmetry, pitch/turn and radius of the oligomer models built at 1 ns simulation intervals were recorded. During a 100 ns simulation the symmetry fluctuated between 4.6- and 6.8-fold; although higher order symmetries were briefly sampled by one of the seeds, the mean symmetry fluctuates between values compatible with those of the wild-type (Fig. 5.2 B and 3.4 A). The pitch/turn fluctuated about 28 and 42 nm, similar to that of the wild-type CeSAS-6 spiral (Fig. 5.2 D and 3.5 A). Similarly, the radius of the spiral fluctuated around 2.6 to 5.2 nm, again compatible with that of the wild-type spiral (Fig. 5.2 C and 3.6 A). While the mean symmetry, radius and pitch/turn sampled were compatible with that of wild-type CeSAS-6, greater variability was observed. This indicates that in triplicate molecular dynamics simulations of 100 ns duration, removal of the steric effect at the hinge alone in spiral-forming SAS-6 is not sufficient

to promote geometries compatible with those of a ring-like oligomer. However, the increased variability in symmetry, pitch/turn and radius suggests a broader range of conformations is available to the dimer upon removal of the steric effect at the hinge.

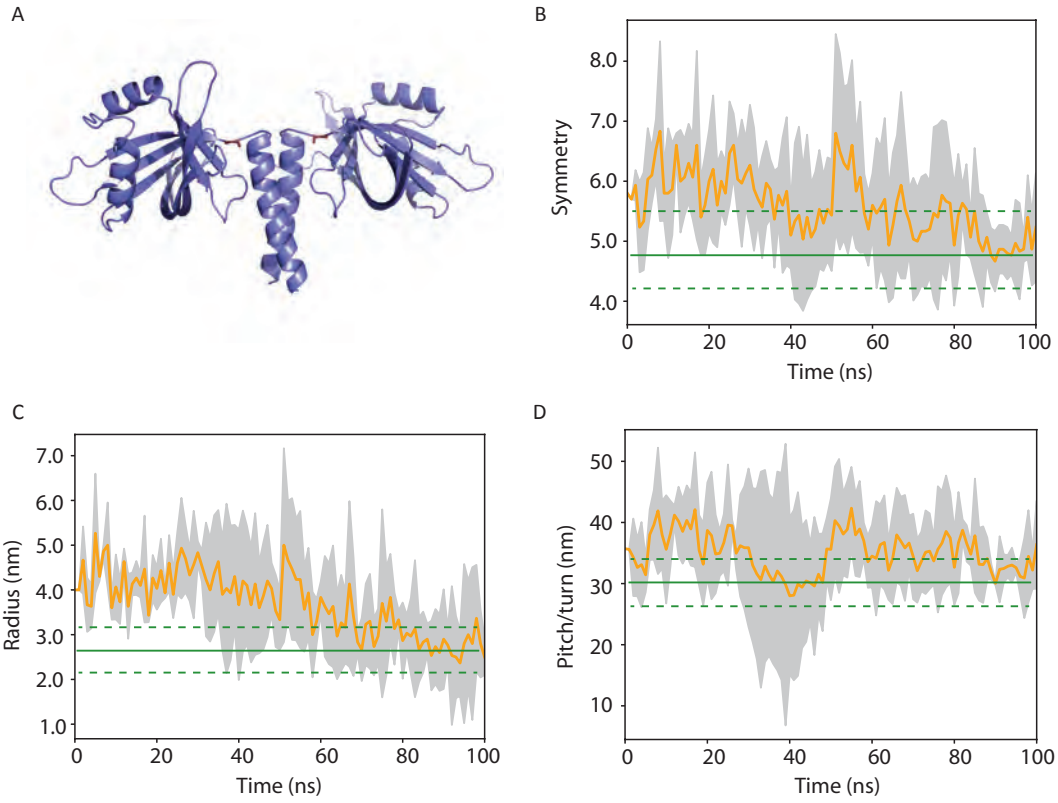


Figure 5.2: Molecular dynamics simulations of CeSAS-6 V167G C-C dimer. A: CeSAS-6 V167G (red) C-C dimer model. B: Symmetry, C: Radius and D: Pitch/turn sampled by oligomers built at 1 ns intervals. Mean (orange) and standard deviation derived from triplicate simulations (grey) is shown. The mean (solid green line) and standard deviation (dashed green line) is indicated for wild-type CeSAS-6 simulations.

To further assess the role of the hinge residue in oligomer geometry, the geometry of LmSAS-6 G142V (hinge) oligomers was determined by molecular dynamics simulations. In this model steric clash has been introduced at the hinge by the addition of a valine side chain (Fig. 5.3 A) but the remainder of the structure has not been modified. The symmetry of the oligomer built from this model at 0 ns has 8.2-fold symmetry. There is an overall downward trend in symmetries sampled, with the model stabilising after ∼60 ns and sampling symmetries fluctuating between 4.6- to 3.7-fold (Fig. 5.3 B). Similarly, the radius starts at 9.8 nm and drops to sample radii in the range 3.7 to 5.3 nm for the last 30 ns of the simulation (Fig. 5.3 C). The pitch/turn sampled by this model shows large fluctuations ranging from 0 to 20 nm (Fig. 5.3 D). The pitch/turn does not stabilise during the 100 ns simulation. The pitch/turn is generally greater than that sampled by wild-type LmSAS-6, but shallower than that sampled by wild-type CeSAS-6 (Fig. 3.5 A and C). Together, this suggests that while the introduction of a hinge valine residue into an otherwise ring-forming background is capable of causing a move towards more spiral-like geometries, it is not sufficient in isolation.

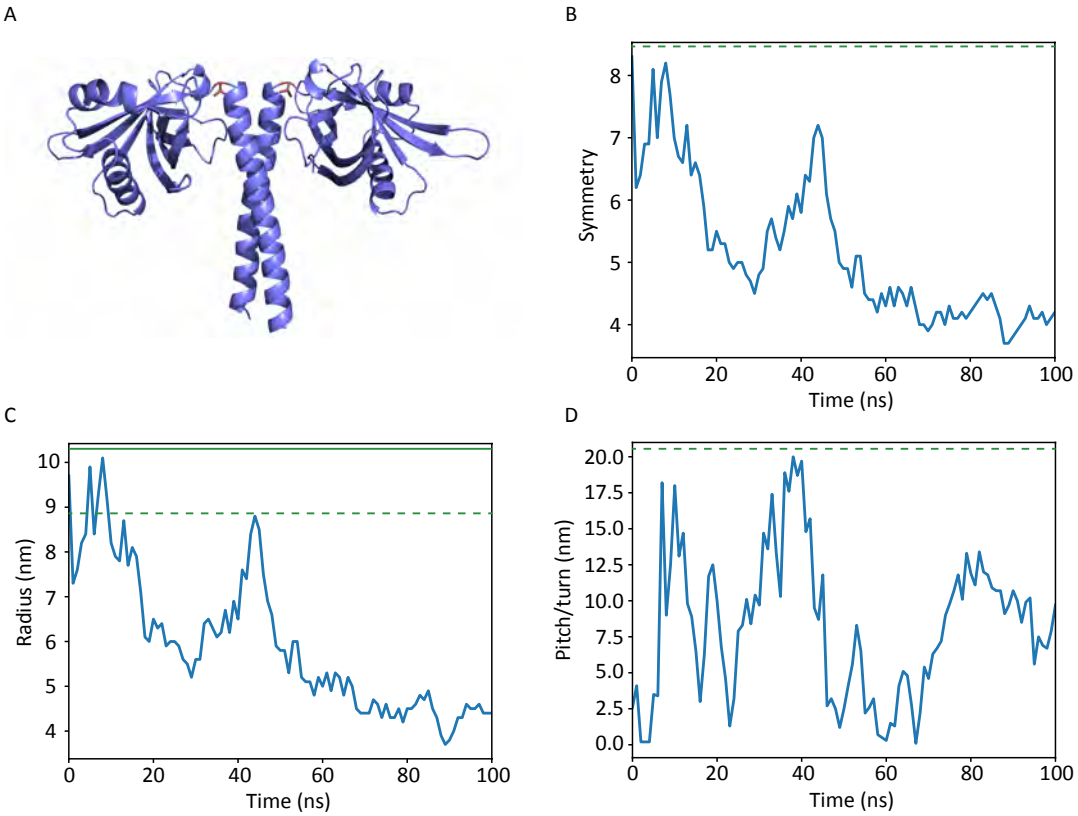


Figure 5.3: Molecular dynamics simulation of LmSAS-6 with G142V mutation. LmSAS-6 C-C dimer with hinge G142V mutation (red). B: Symmetry sampled by LmSAS-6 G142V model shown in A. Radius (C) and pitch/turn (D) of oligomers built from the model shown in panel A over a 100 ns simulation. Green lines indicate the mean (solid) and standard deviation (dashed) of wild-type LmSAS-6.

5.3 CeSAS-6 mutant molecular dynamics simulations

The 16 CeSAS-6 variants described in Chapter 4.3 were assessed for their ability to adopt a ring conformation *in silico* prior to being tested *in vitro*. The symmetry, pitch/turn and radius of the resulting oligomers were monitored over a 10 ns molecular dynamics simulation. Using molecular dynamics simulations of the CeSAS-6 C-C dimer as a model, the 3 seeds used in the simulation are convergent at 10 ns (Fig. 5.4 A) however, fluctuations in the backbone RMSD relative to the starting structure indicate 10 ns may not be sufficient to capture conformational changes as the system has not stabilised on this timescale (Fig. 5.4 B). CeSAS-6 mutant simulations were carried out at 10 ns duration. While changes in symmetry, radius, and pitch/turn were captured on this timescale for some mutatants it may be that these simulations were too short to capture such conformational changes. Therefore, extended simulations have been completed on some SAS-6 variants (see Chapter 6.5.1, 6.6.1). The majority of the CeSAS-6 mutants did not adopt more ring-like geometry as assessed using this method.

Five CeSAS-6 mutant variants showed elements of ring-forming characteristics. The most promising mutant was V167G S56L F61L H175E (Fig. 4.4 K) which displayed ring-like characteristics in symmetry, pitch/turn and radius. This mutant sampled higher order symmetries for approximately 3 ns of the simulation. The maximum symmetry sampled was 10-fold (Fig. 5.6 J). It also sampled much lower pitch/turn than the wild-type spiral. The minimum pitch/turn sampled was 9 nm (Fig. 5.10 J). The radius also increased compared to the wild-type spiral increasing from 4 nm to 12 nm (Fig. 5.8 J), similar to that of the expected 11.5 nm radius of a 9-fold symmetric ring.

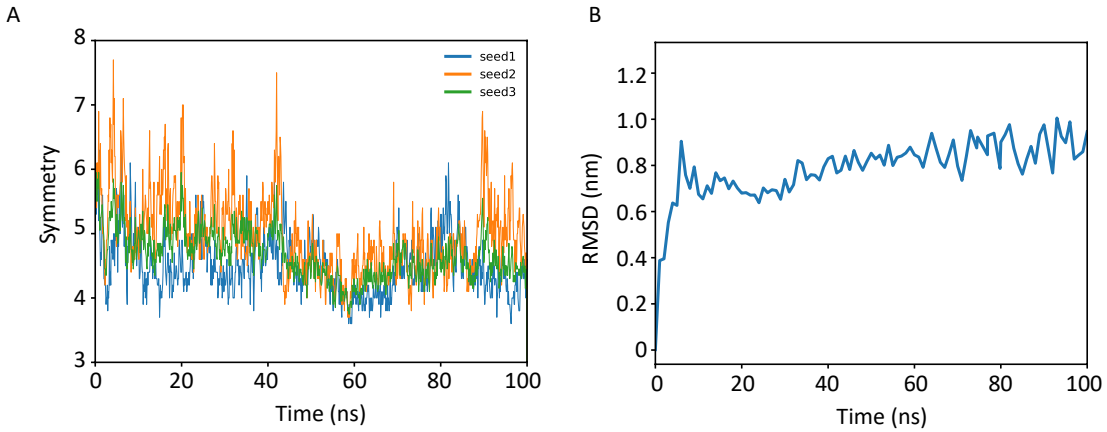


Figure 5.4: Stability of molecular dynamics simulations. Molecular dynamics simulations of CeSAS-6 C-C dimers run for 100 ns. A: Shows variation in symmetry derived from oligomers of CeSAS-6 dimers between 3 seeds over 100 ns. B: Root-mean-square deviation of the CeSAS-6 C-C dimer backbone at each time point compared to the starting structure over 100 ns.

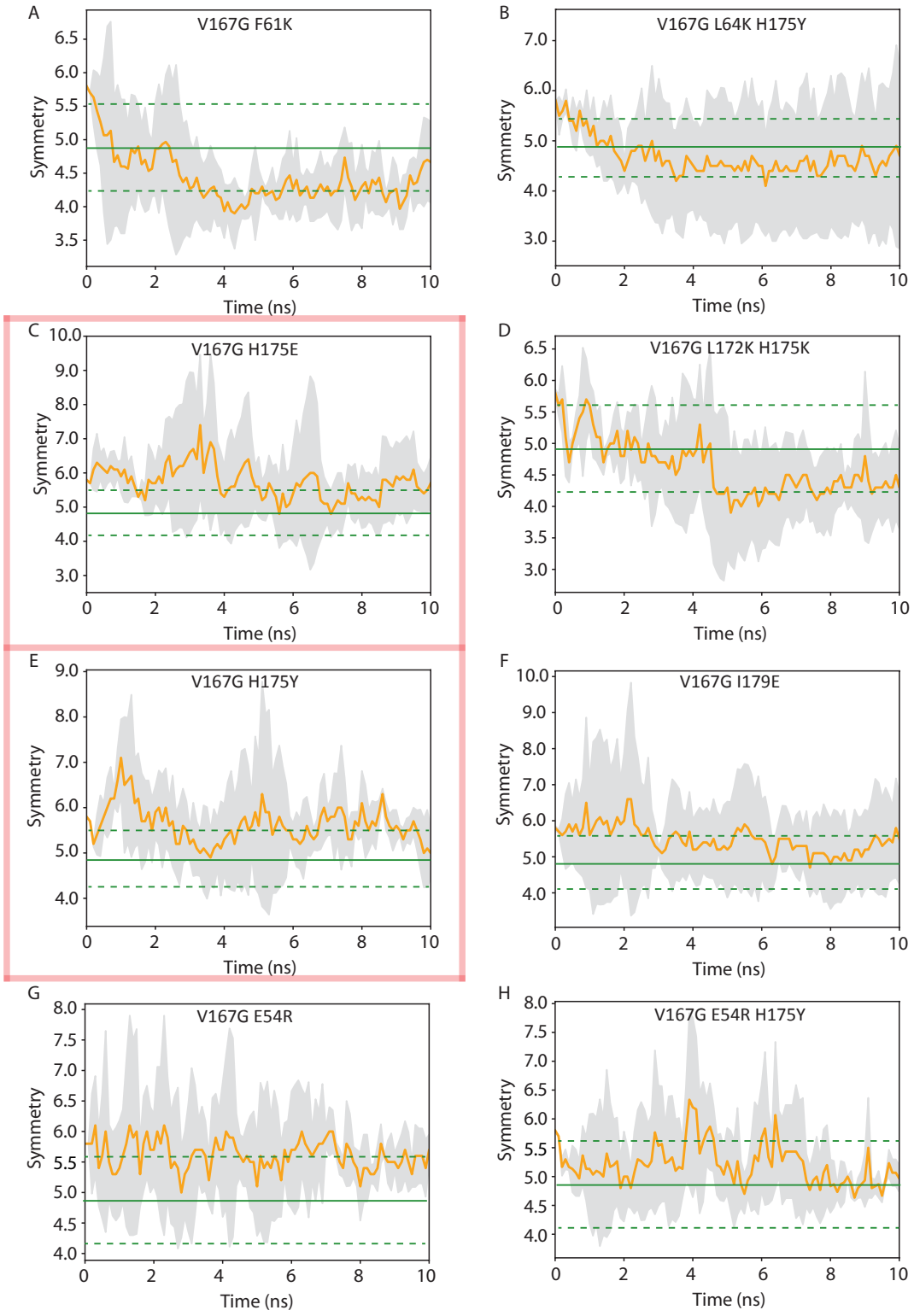


Figure 5.5: Symmetry sampled by CeSAS-6 variants (i). Symmetry variation in 10 ns molecular dynamics simulations of oligomers built from CeSAS-6 variant C-C dimer models. A: V167G F61K, B: V167G L64K H175Y, C: V167G H175E, D: V167G L172K H175K, E: V167G H175Y, F: V167G I179E, G: V167G E54R, and H: V167G E54R H175Y. The mean (orange) and standard deviation from triplicate simulations (grey) are shown. The mean (solid green line) and standard deviation (dashed green line) for wild-type CeSAS-6 are indicated. Pink boxes indicate more successful mutants.

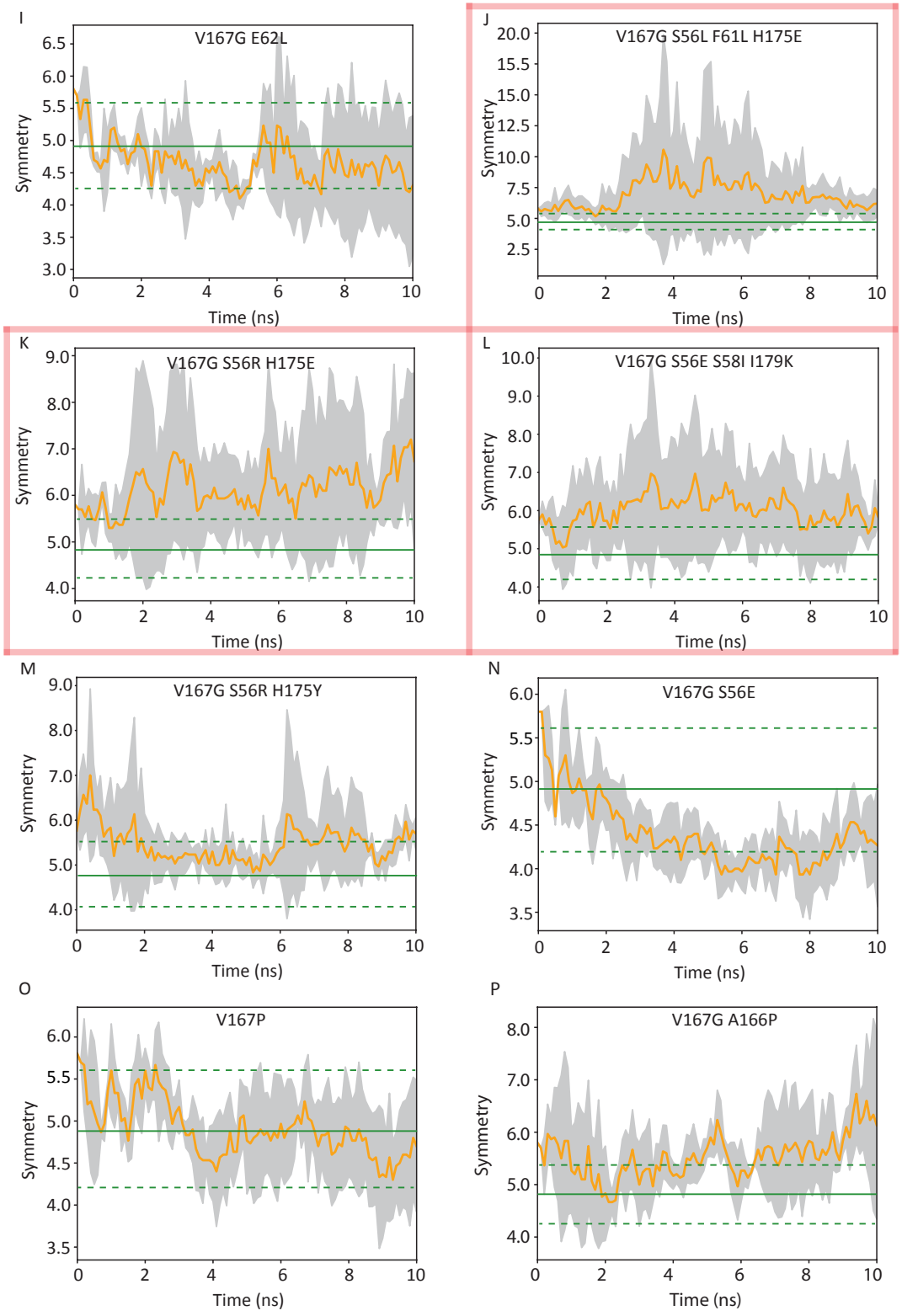


Figure 5.6: Symmetry sampled by CeSAS-6 variants (ii). Symmetry variation in 10 ns molecular dynamics simulations of oligomers built from CeSAS-6 variant C-C dimer models. I: V167G E62L, J: V167G S56L F61L H175E, K: V167G S56R H175E, L: V167G S56E S58I I179K, M: V167G S56R H175Y, N: V167G S56E, O: V167P, P: V167G A166P. The mean (orange) and standard deviation from triplicate simulations (grey) are shown. The mean (solid green line) and standard deviation (dashed green line) for wild-type CeSAS-6 are indicated. Pink boxes indicate more successful mutants.

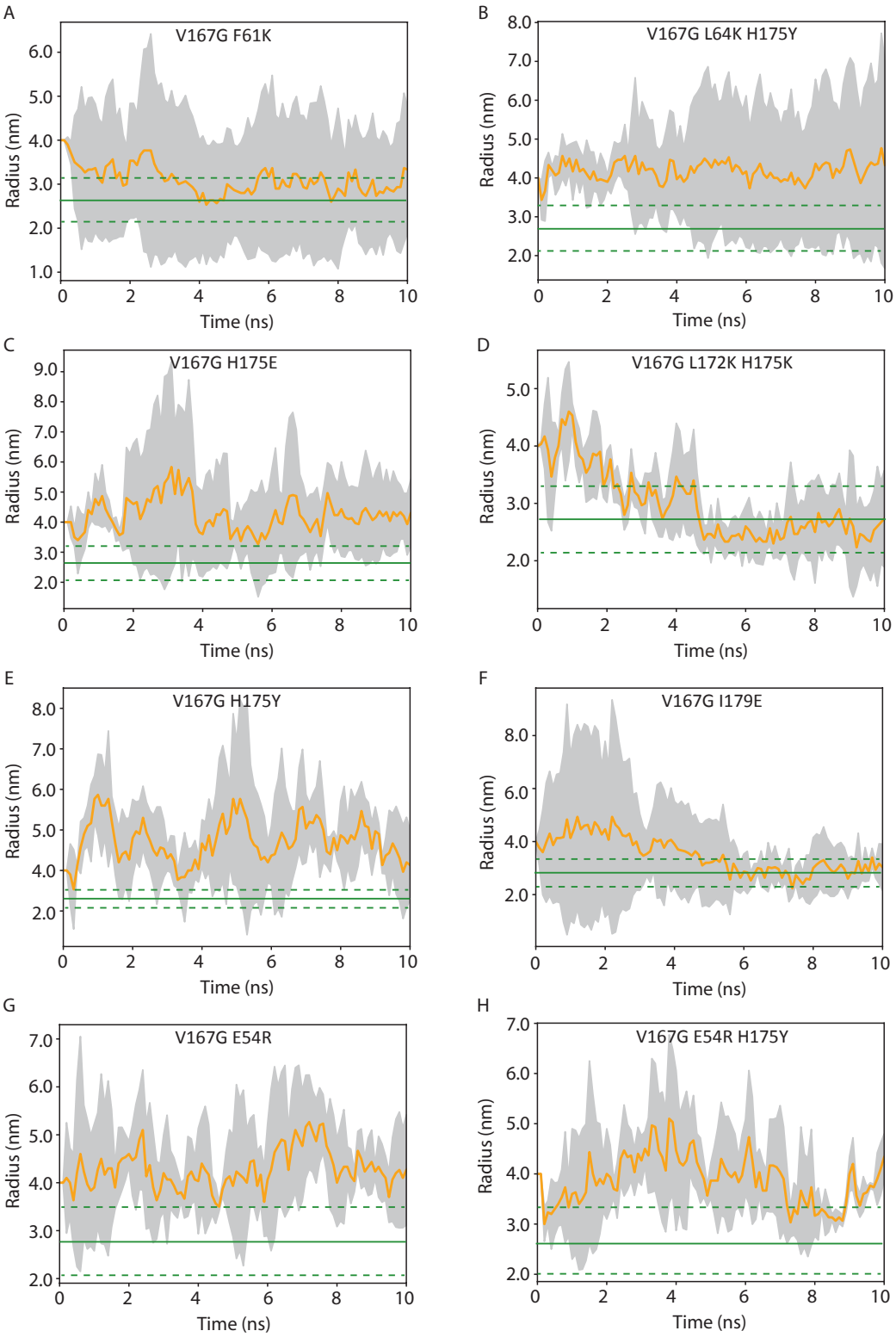


Figure 5.7: Radii sampled by CeSAS-6 variants (i). Radius of oligomers built from CeSAS-6 C-C dimer variants over a 10 ns molecular dynamics simulation. A: V167G F61K, B: V167G L64K, C: V167G H175E, D: V167G L172K H175K, E: V167G H175Y, F: V167G I179E, G: V167G E54R, and H: V167G E54R H175Y. The mean (orange) and standard deviation (grey) from triplicate simulations are shown. The mean (solid green line) and standard deviation (dashed green line) for wild-type CeSAS-6 are indicated.

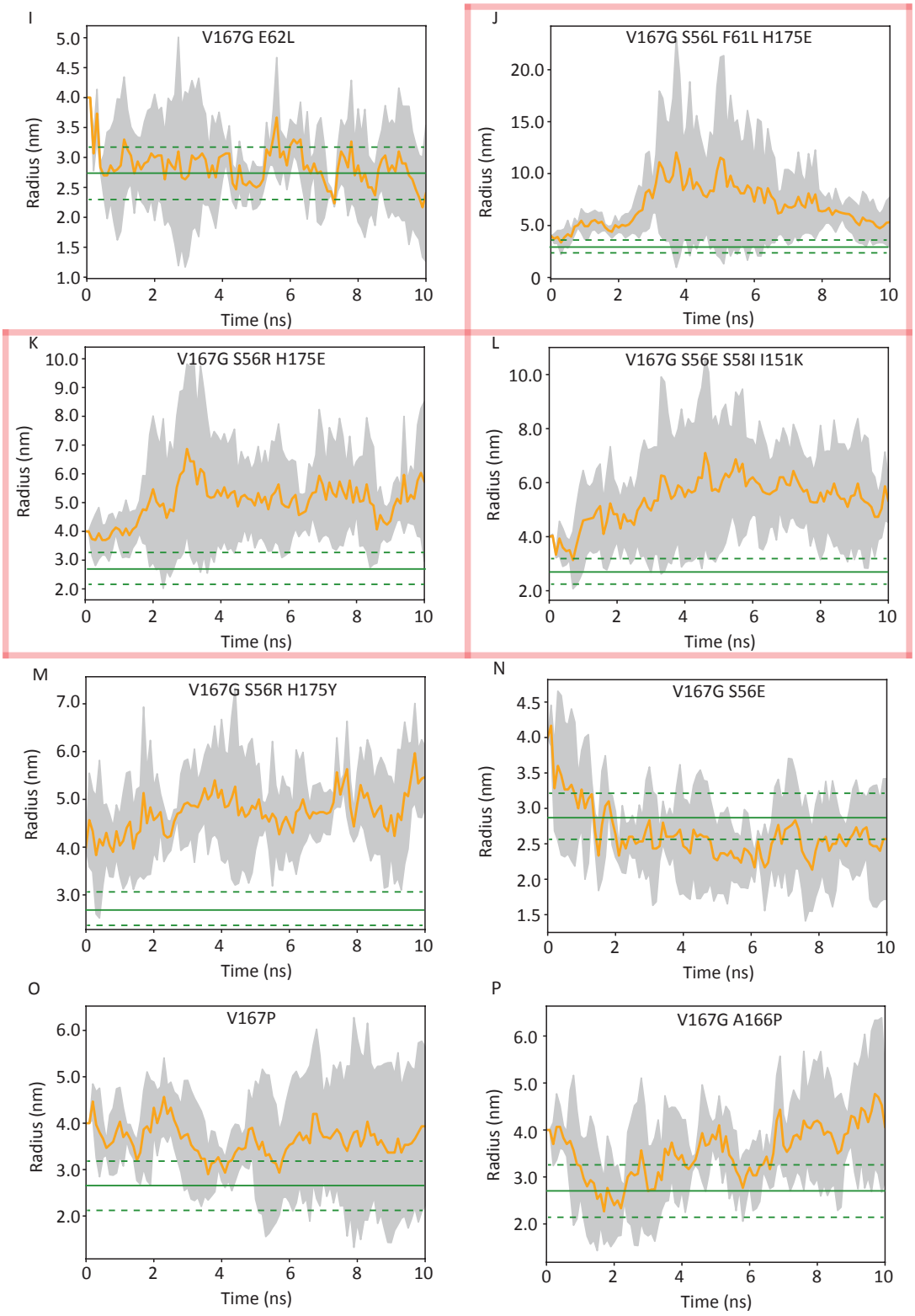


Figure 5.8: Radii sampled by CeSAS-6 variants (ii). Radius of oligomers built from CeSAS-6 C-C dimer variants over a 10 ns molecular dynamics simulation. I: V167G E62L, J: V167G S56L F61L H175E, K: V167G S56R H175E, L: V167G S56E S58I I179K, M: V167G S56R H175Y, N: V167G S56E, O: V167P, P: V167G A166P. The mean (orange) and standard deviation (grey) from triplicate simulations are shown. The mean (solid green line) and standard deviation (dashed green line) for wild-type CeSAS-6 are indicated. Pink boxes indicate more successful mutants.

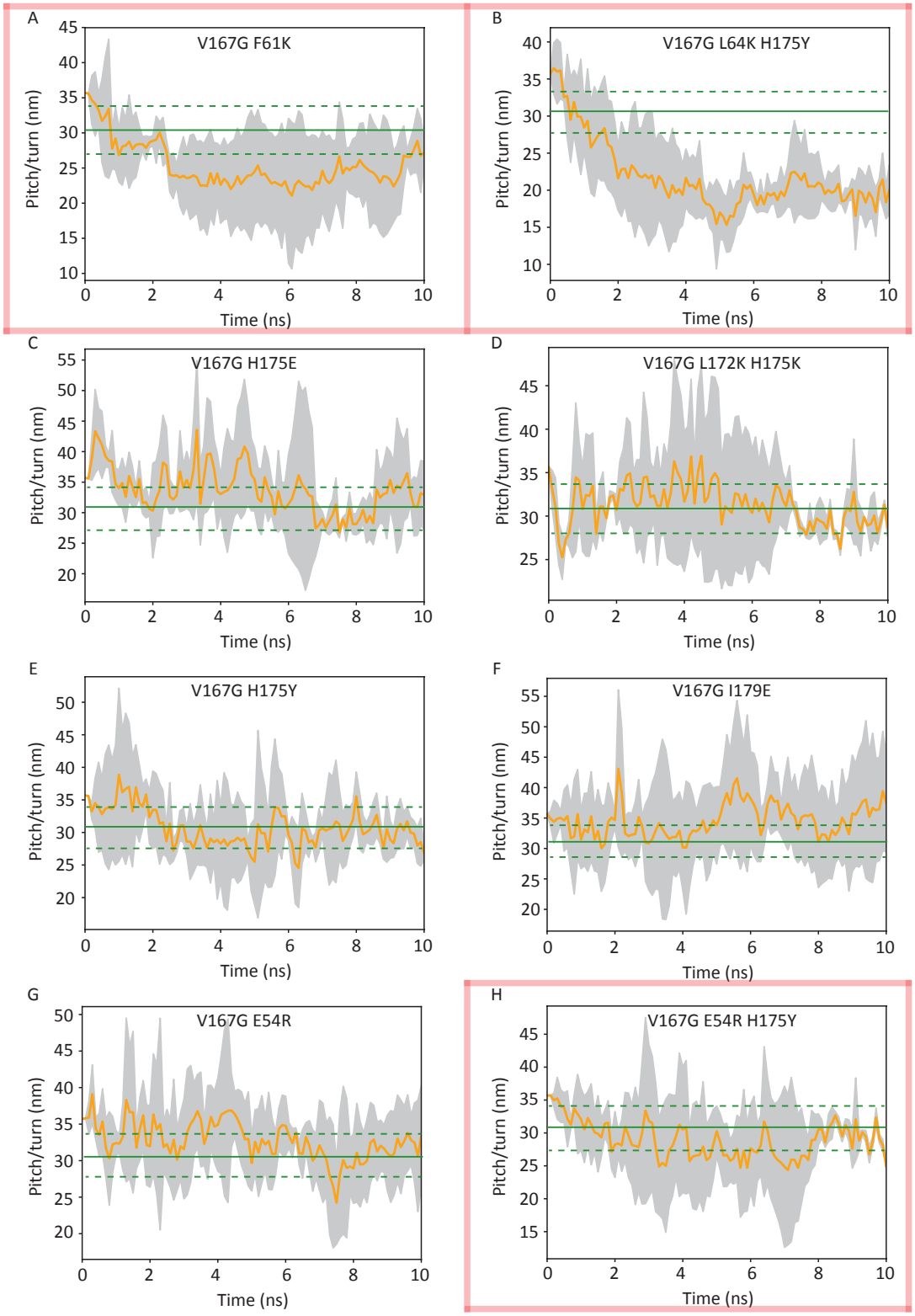


Figure 5.9: Pitch/turn sampled by CeSAS-6 variants (i). Pitch/turn of oligomers built from CeSAS-6 C-C dimer variants over a 10 ns molecular dynamics simulation. A: V167G F61K, B: V167G L64K, C: V167G H175E, D: V167G L172K H175K, E: V167G H175Y, F: V167G I179E, G: V167G E54R, and H: V167G E54R H175Y. The mean (orange) and standard deviation (grey) from triplicate simulations are shown. The mean (solid green line) and standard deviation (dashed green line) for wild-type CeSAS-6 are indicated. Pink boxes indicate more successful mutants.

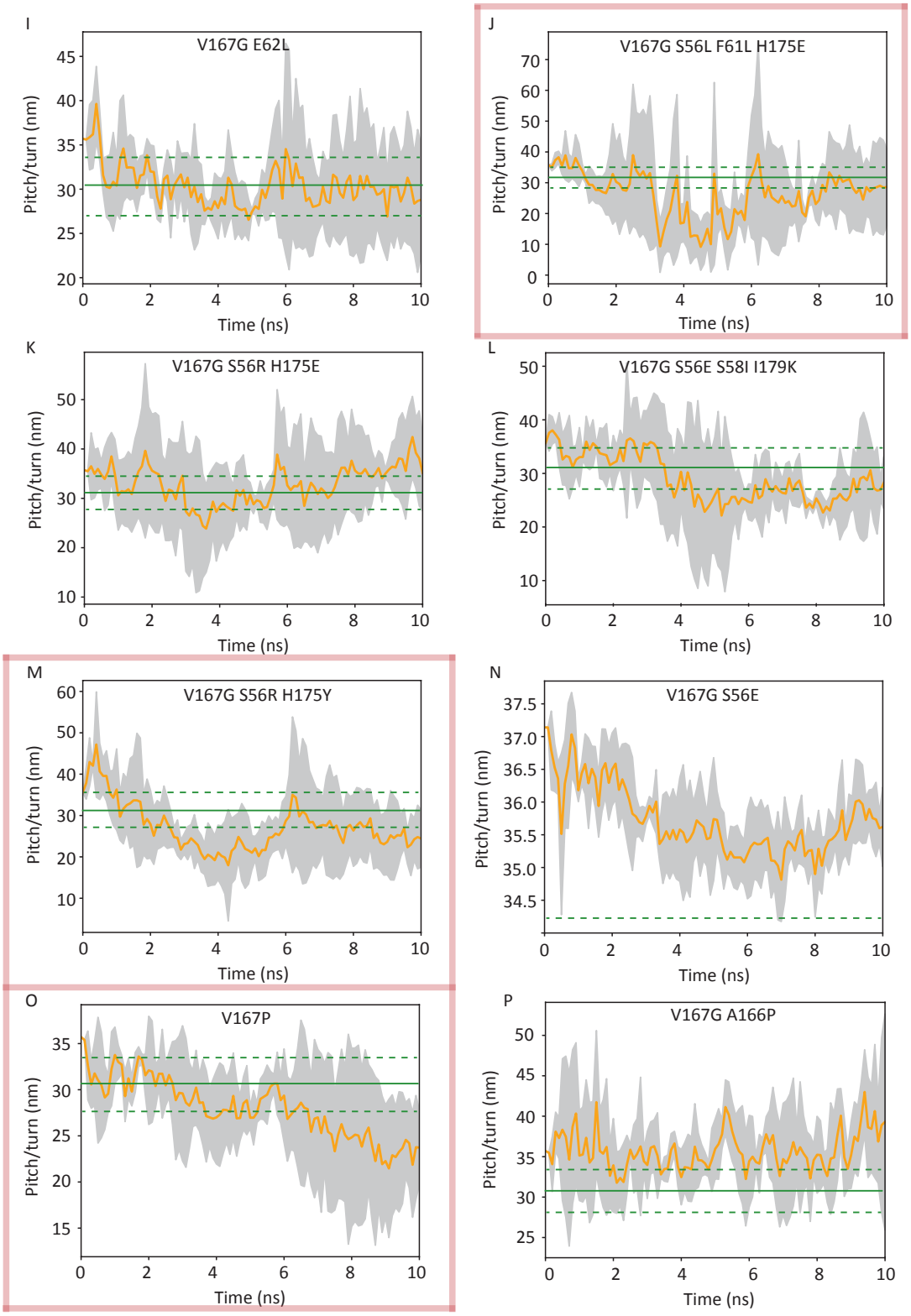


Figure 5.10: Pitch/turn sampled by CeSAS-6 variants (ii). Pitch/turn of oligomers built from CeSAS-6 C-C dimer variants over a 10 ns molecular dynamics simulation. I: V167G E62L, J: V167G S56L F61L H175E, K: V167G S56R H175E, L: V167G S56E S58I I179K, M: V167G S56R H175Y, N: V167G S56E, O: V167P, P: V167G A166P. The mean (orange) and standard deviation (grey) from triplicate simulations are shown. The mean (solid green line) and standard deviation (dashed green line) for wild-type CeSAS-6 are indicated. Pink boxes indicate more successful mutants.

Other SAS-6 mutant variants displayed only some ring-like elements; for example, 4 other mutants displayed an increase in symmetry over the 10 ns simulation. The V167G H175E mutant (Fig. 4.3 E) sampled a maximum symmetry of 7.4-fold (Fig. 5.5 C). However, pitch/turn and radius did not deviate from that of wild-type spirals (Fig. 5.7 C and 5.9 C). Similarly, the V167G H175Y mutant (Fig. 4.4 G) briefly sampled 7.1-fold symmetry but this mutation did not lead to changes in pitch/turn or radius (Fig. 5.5 E, 5.7 E and 5.9 E). The V167G S56R H175E variant (Fig. 4.4 M) displayed a slight upward trend in the symmetries it sampled with a maximum of 7-fold (Fig. 5.6 K). This variant did not cause a change in pitch/turn, but resulted in a slight increase in radius compared to wild-type spirals (Fig. 5.10 K and 5.8 K). The maximum radius sampled was 6.8 nm. The V167G S56E S58I I179K variant (Fig. 4.4 L) also sampled slightly increased symmetries compared to wild-type CeSAS-6 (Fig. 5.6 L). The maximum symmetry sampled was 6.9-fold. Additionally, the radius of the oligomer increased to 6.8 nm (Fig. 5.8 L).

A further 5 mutants produced oligomers which displayed very few ring-like elements; for example, decreased pitch/turn but spiral-like symmetry and radius. V167G F61K (Fig. 4.3 D, 5.9 A), V167G E54R H175Y (Fig. 4.4 P, 5.9 H) and V167P (Fig. 4.3 B, 5.10 O) had slightly reduced pitch/turn but did not sample pitch/turn shallower than 20 nm (minimum pitch/turn V167G F61K 21 nm, V167G E54R H147Y 24.5 nm, V167P 21.5 nm). V167G S56R H175Y (Fig. 4.4 N) produced oligomers with a minimum pitch/turn of 18 nm (Fig. 5.10 M) and V167G L64K H175Y (Fig. 4.3 H) produced oligomers with a minimum pitch/turn of 15 nm (Fig. 5.9 B). While the pitch/turn was decreased compared to the wild-type spiral, for none of the mutants tested was 0 nm sampled where a planar conformation could be adopted.

5.4 Discussion

The aim of the work presented in this chapter was to assess whether it is possible to capture *in silico* the effect of mutations introduced into CeSAS-6 on oligomer structure and pre-screen CeSAS-6 mutants prior to their testing *in vitro*. It is possible to detect a ring to spiral transition using this method as demonstrated by the control model where the CeSAS-6 sequence was imposed onto the LmSAS-6 structure by homology modelling. The rapid transition from ring- to spiral-like geometry was anticipated due to the unfavourable nature of the steric clash from the valine residue at the hinge. This indicated that conversion of oligomer geometry from spiral to ring would not be possible without the removal of the hinge valine. Therefore, all CeSAS-6 mutant variants were made upon a background of V167G. The reverse; a ring to spiral transition in a control model, was also captured but in a less striking manner. Increased symmetries and larger radii were sampled and pitch/turn was slightly reduced compared to that sampled by the wild-type spiral. This slow reversal back to the original ring conformation suggests that while the sequence drives adoption of ring geometry, a spiral conformation is not so unfavourable as to cause immediate reversal to a ring conformation. Thus, interconversion between the two conformations is achievable.

A single mutation in the context of an otherwise ring-like sequence is capable of causing spiral-like geometries to be adopted, as was shown in the LmSAS-6 G142V simulation. Although the steric effect introduced by the valine residue at the hinge causes the oligomer to adopt a more spiral-like geometry, the change happens over a longer period, taking ∼70 ns compared to the almost instantaneous reversal seen in the threaded models. A conformation intermediate between spiral- and ring-like geometry is accommodated by rotation of the hinge valine side chain by approximately 90° compared to its orientation in wild-type CeSAS-6 (Fig. 5.11). Thus, while the hinge residue is important, other components of the sequence influence oligomer conformation. Consistent with that assessment, removal of the steric effect of the hinge valine residue in CeSAS-6 by mutating it to a glycine did not cause ring-like geometries to be sampled. While removal of this bulky side

chain may have relaxed the constraint to maintain a spiral conformation, there was no additional pressure to adopt ring-like geometries.

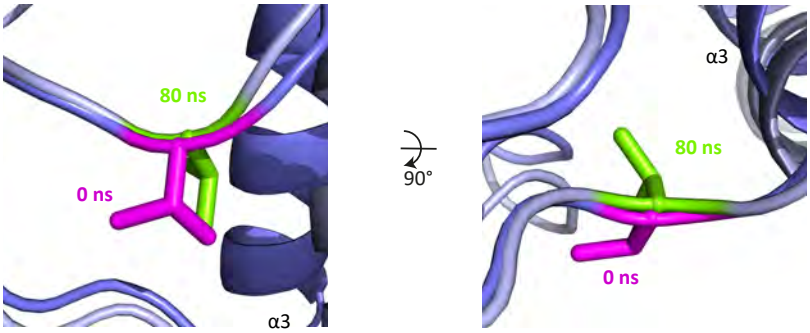


Figure 5.11: Rotation of G142V. Position of the G142V side chain in LmSAS-6 at 0 ns (pink) where the position is equivalent to that seen in CeSAS-6 and at 80 ns where the valine side chain has rotated by ∼90° (green).

Of the mutants simulated, none sampled strongly ring-like pitch/turn. This does not rule out these mutations as causing a spiral to ring transition. In the wild-type simulations (Chapter 3.3) ring-forming SAS-6 variant models did not consistently sample pitch/turn close to planar. This could be due to a limitation in the method used to model oligomers where C-C dimers are simulated out of the context of the ring, and no energy is associated with ring closure. Mutants can therefore not be ruled out if they do not sample planar pitch/turn. All mutants will therefore be screened *in vitro* with a preference for mutants with the most ring-like characteristics as identified by molecular dynamics simulations.

6

Atomic force microscopy of CeSAS-6 oligomers

Contents

6.1 Introduction 98
6.2 Electron microscopy of SAS-6 oligomers 98
6.3 Atomic force microscopy optimisation of SAS-6 oligomers 103
6.3.1 Optimisation of imaging conditions 103
6.3.2 Troubleshooting CeSAS-6 variant AFM 108
6.3.3 Image analysis and metrics 124
6.4 AFM of CeSAS-6 variants 126
6.4.1 Ring diameter 126
6.4.2 Backbone width 130
6.4.3 Spoke separation 132
6.4.4 Curvature 135
6.4.5 Backbone periodicity 137
6.5 *In silico* analysis of CeSAS-6APVG 139
6.5.1 Molecular dynamics simulations of CeSAS-6APVG . . . 139
6.5.2 CeSAS-6APVG dihedral angle analysis 141
6.6 Round II CeSAS-6APVG variants 145
6.6.1 Molecular dynamics simulations of round II CeSAS-6 variants 145
6.6.2 AFM and troubleshooting of round II CeSAS-6APVG variants 149
6.7 Discussion 152

Contributions: AFM data was collected with the help of Alexander Yon and

Phillip Haynes, Hoogenboom group, London Centre for Nanotechnology (LCN).

6.1 Introduction

To screen the CeSAS-6 variants for ring-forming ability, oligomers were imaged *in vitro*. Atomic force microscopy (AFM) was used to image the oligomers. Using tapping mode AFM, it is possible to generate a height trace of objects adsorbed onto a surface. Sub-nanometre resolution can be achieved. A high contrast image can be obtained and with less disturbance to the sample than in contact mode [161] which is important given the weak nature of the N-N oligomerisation interface of SAS-6, with a dissociation constant of ∼60 μM [40, 137]. The force experienced by samples in tapping-mode is 45 to 65 pN, making imaging of SAS-6 oligomers challenging. Previous studies have used AFM to visualise and monitor the assembly pathways of CrSAS-6 [133, 162]. In order to capture the assembly processes of SAS-6 oligomers a compromise between imaging speed and increased peak force was reached by using high-speed photo-thermal off-resonance (HS-PORT) AFM, where heat is used to deform the cantilever, allowing for increased frequency at which it can oscillate thereby increasing imaging speed. This earlier work demonstrated that there are multiple routes to ring formation, not just step-by-step addition of SAS-6 dimers. Here, the focus is on CeSAS-6 oligomer geometry, imaging speed is not a consideration, so tapping mode AFM was used instead allowing for reduced force experienced by the sample.

6.2 Electron microscopy of SAS-6 oligomers

Initial attempts to visualise CeSAS-6 oligomers were made using negative stain EM. A CrSAS-6 variant that forms strengthened rings (K_d ∼2.3 μM) with 5-fold symmetry was used as a positive control (plasmid gifted by the Steinmetz lab, PSI). This construct comprises residues 1-226 of *C. reinhardtii* SAS-6, encompassing the N-terminal domain and a coiled-coil with a length of approximately 9 nm

(Fig. 6.1 A). Rings formed by CrSAS-6 protein could be visualised by negative stain EM (Fig. 6.2 B).

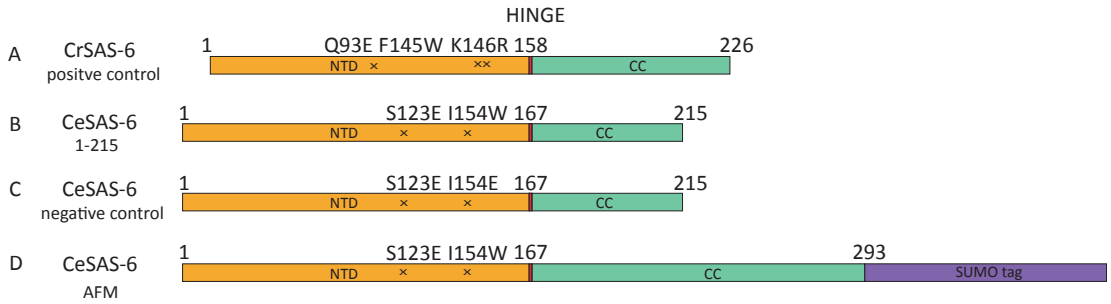


Figure 6.1: SAS-6 constructs. A: CrSAS-6 construct, positive control with Q93E F145W K146R mutations. B: 1-215 CeSAS-6 construct with S123E I154W mutations. C: 1-215 CeSAS-6 S123E I154W construct, negative control. D: 1-293 CeSAS-6 with S123E I154W mutations and a C-terminal SUMO tag. Mutations and hinge (red) residues are indicated. N-terminal domain (NTD), coiled-coil (CC) and SUMO-tag are shown.

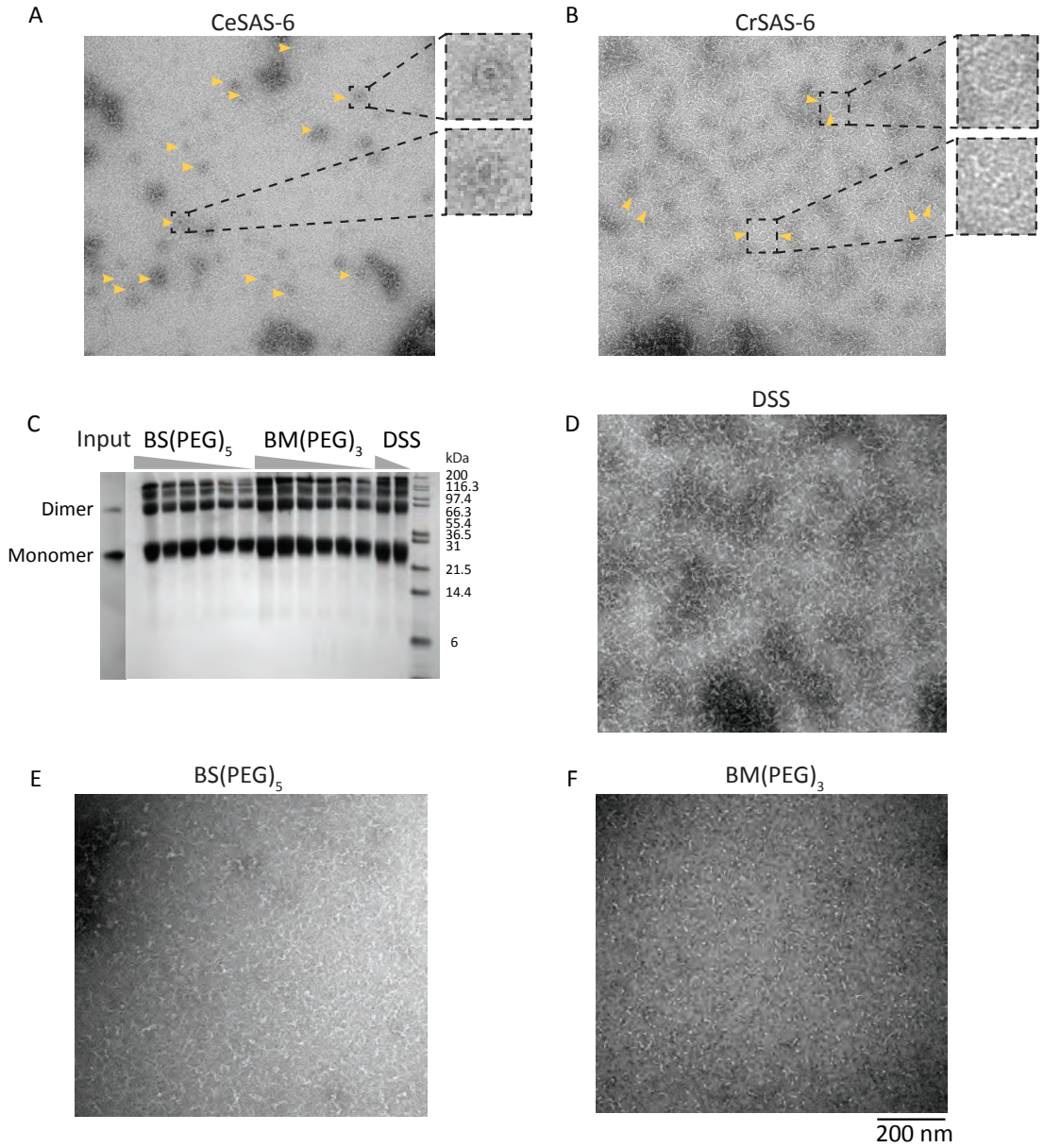


Figure 6.2: Electron microscopy of CeSAS-6 oligomers. Negative stain EM of A: CrSAS-6 and B: CeSAS-6. C: SDS-PAGE of cross-linked CeSAS-6. Lane 1 shows the input sample without cross-linking, lanes 2-7 shows cross-lining with $BS(PEG)_5$ at 6,000 μM, 3,000 μM, 1,500 μM, 1,000 μM, 800 μM and 600 μM. Lanes 8-13 show samples cross-linked with $BM(PEG)_3$ at 1,200 μM, 600 μM, 300 μM, 250 μM, 200 μM and 120 μM. Lanes 14 and 15 show samples cross-linked with DSS at 3,000 μM and 1,200 μM. EM images of CeSAS-6 cross-linked with D: 1,200 μM DSS, E: 1,000 μM $BS(PEG)_5$ and F: 200 μM $BM(PEG)_3$. Scale is indicated.

For visualisation of CeSAS-6 oligomers, a CeSAS-6 construct was used in similar negative stain experiments that comprised residues 1-215, producing a coiled-coil approximately 7 nm long (Fig. 6.1 B). This construct also contains an S123E mutation to improve solubility and an I154W mutation to strengthen N-N dimerisation (K_d $\sim$5 μM). Previous negative stain EM has been carried out on this construct by M. Hilbert *et al.* [39], noting that these mutations do not influence oligomer geometry and CeSAS-6 spirals could be visualised.

In our studies, given the weak dissociation constant of the N-N interface, in order to see oligomers a concentration of 1.5 mg/mL ($\sim$60 μM) was required. This resulted in a noisy background and low contrast images from over-absorbance of protein on the grid (Fig. 6.2 B). Methods where a reduced protein concentration was required were therefore assessed.

In order to reduce the concentration of protein required for oligomers to be visible, cross-linking of CeSAS-6 was carried out. Table 6.1 provides details on the cross-linkers used. A ladder of oligomers with increasing molecular weights was observed by SDS-PAGE, indicating successful cross-linking of CeSAS-6 by $BS(PEG)_5$, $BM(PEG)_3$ and DSS (Fig. 6.2 C). $BS(PEG)_5$ was used at 1,000 μM as no further increase in higher molecular weight oligomers was seen above this concentration. Similarly, $BM(PEG)_3$ was used at 200 μM and DSS at 1,200 μM. Cross-linking with $BS(PEG)_5$ or $BM(PEG)_3$ did not significantly improve background noise (Fig. 6.2 E and F). Improved contrast was achieved in negative stain EM images on DSS cross-linked samples; however, no long spirals were observed (Fig. 6.2 D). This meant it was not possible to determine spiral pitch/turn or alternative metrics for assessing oligomer architecture. An alternative imaging technique was therefore necessary.

Table 6.1: Cross-linkers used in negative stain electron microscopy experiments.

Cross-linker	**Reactive towards**	**Spacer arm length**
$BS(PEG)_5$	primary amines	21.7 Å
$BM(PEG)_3$	sulphydryl groups	17.8 Å
DSS	primary amines	11.4 Å

6.3 Atomic force microscopy optimisation of SAS-6 oligomers

6.3.1 Optimisation of imaging conditions

6.3.1.1 Control samples

AFM was used to image oligomers since adsorption of the protein onto a mica surface meant lower protein concentration was required and therefore improved signal to noise could be achieved. AFM of the same positive control used in EM (a ring-forming CrSAS-6 variant that has 6-fold symmetry) produced rings comprising 6 dimers as expected (Fig. 6.3 E and F). These images have clean background and high contrast. As a negative control a CeSAS-6 variant with an I154E mutation that disrupts N-N oligomerisation was used (Fig. 6.1 C). Here, the largest oligomer present was the C-C dimer (Fig. 6.3 A and B). This is shown in height traces, where a single object with a height of 4 nm can be seen when a trace across a monomer is drawn (Fig. 6.3 C), and a double peak can be seen when a line across a dimer is drawn (Fig. 6.3 D).

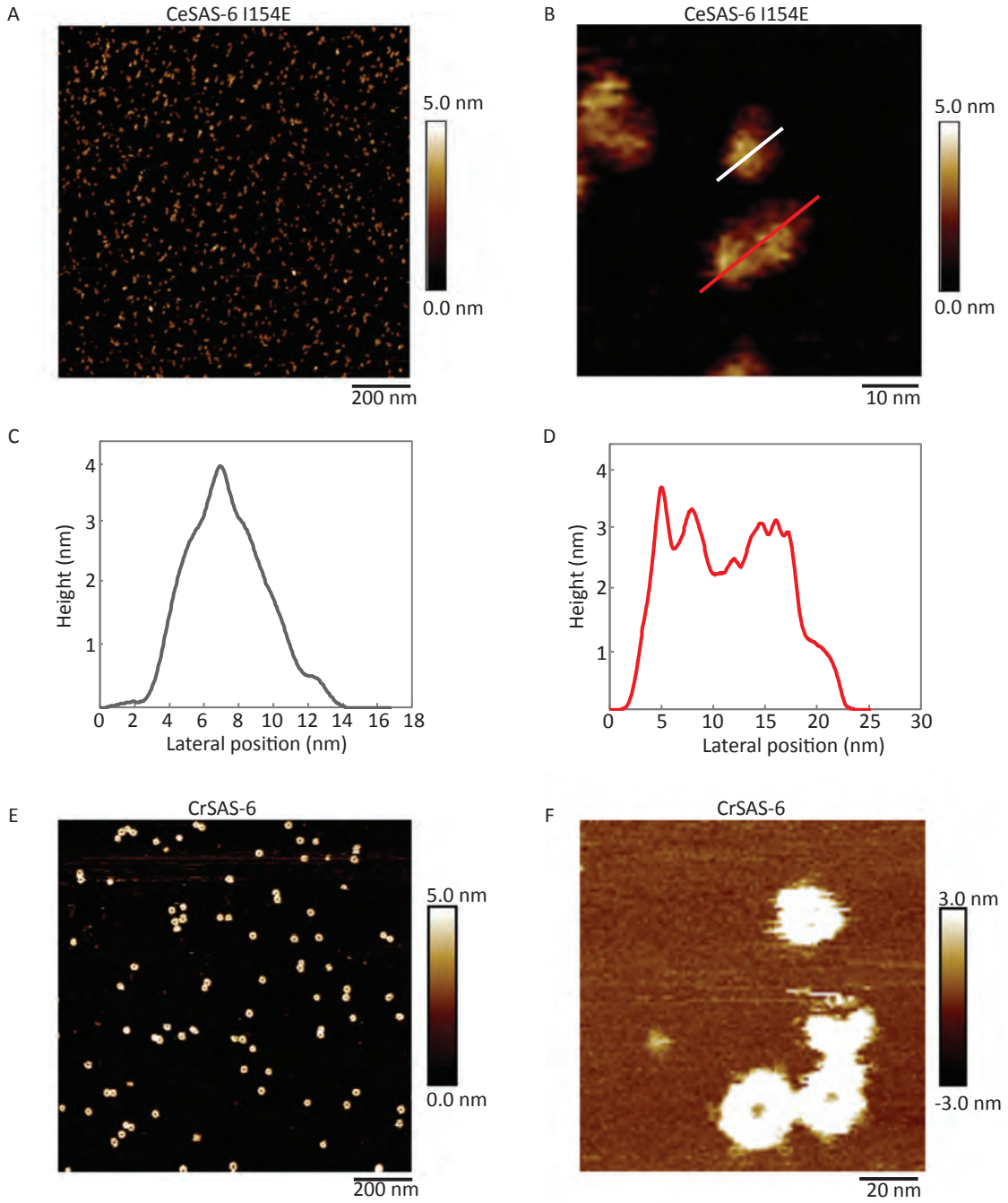


Figure 6.3: AFM height profiles of SAS-6 control samples. A and B: CeSAS-6 I154E negative control. Height trace of the monomer (white line) shown in C, and dimer (red line) shown in D. E and F: CrSAS-6 positive control. Scale and colour gradient are indicated for each image.

6.3.1.2 Construct optimisation

The CeSAS-6 construct used in EM experiments (residues 1-215 S123E I154W) was also used in AFM. Filamentous oligomers can been seen with good contrast (Fig. 6.5 A). CeSAS-6 rings were also observed but at low frequency (4.7% (9/192)). Given the short length of the coiled-coil in this construct, spokes are not visible. As this reduces the amount of information that can be obtained on the architecture of the oligomers, a construct was designed with an extended coiled-coil. This construct consists of CeSAS-6 residues 1-293 and possesses a SUMO tag at its C-terminus (Fig. 6.1 D). The SUMO tag adds density at the end of the coiled-coil and aids in mica adsorption (Fig. 6.5 C and D). Its size is such that it should not influence oligomer geometry (Fig. 6.5 B). An example of the purification steps for this construct is shown in Fig 6.4. In a further construct, the SUMO tag was engineered to contain 2 more positive charges to aid in adsorption to the negatively charged mica surface used for imaging, in order to stabilise the placement of spokes. To that end residues E83 and E84 located on a loop region of SUMO were mutated to lysine residues. Although in this sample aggregates were present, spokes could be resolved but they were not as well defined as when a wild-type SUMO-tag was used and no long oligomers were observed (Fig. 6.5 E and F). Thus, for further imaging we reverted to the 1-293 CeSAS-6 construct with wild-type SUMO tag at the C-terminus. CeSAS-6 oligomers are fragile, and despite the low imaging forces used oligomers were destroyed by the cantilever tip (Fig. 6.6). It was therefore not possible to image the same area multiple times.

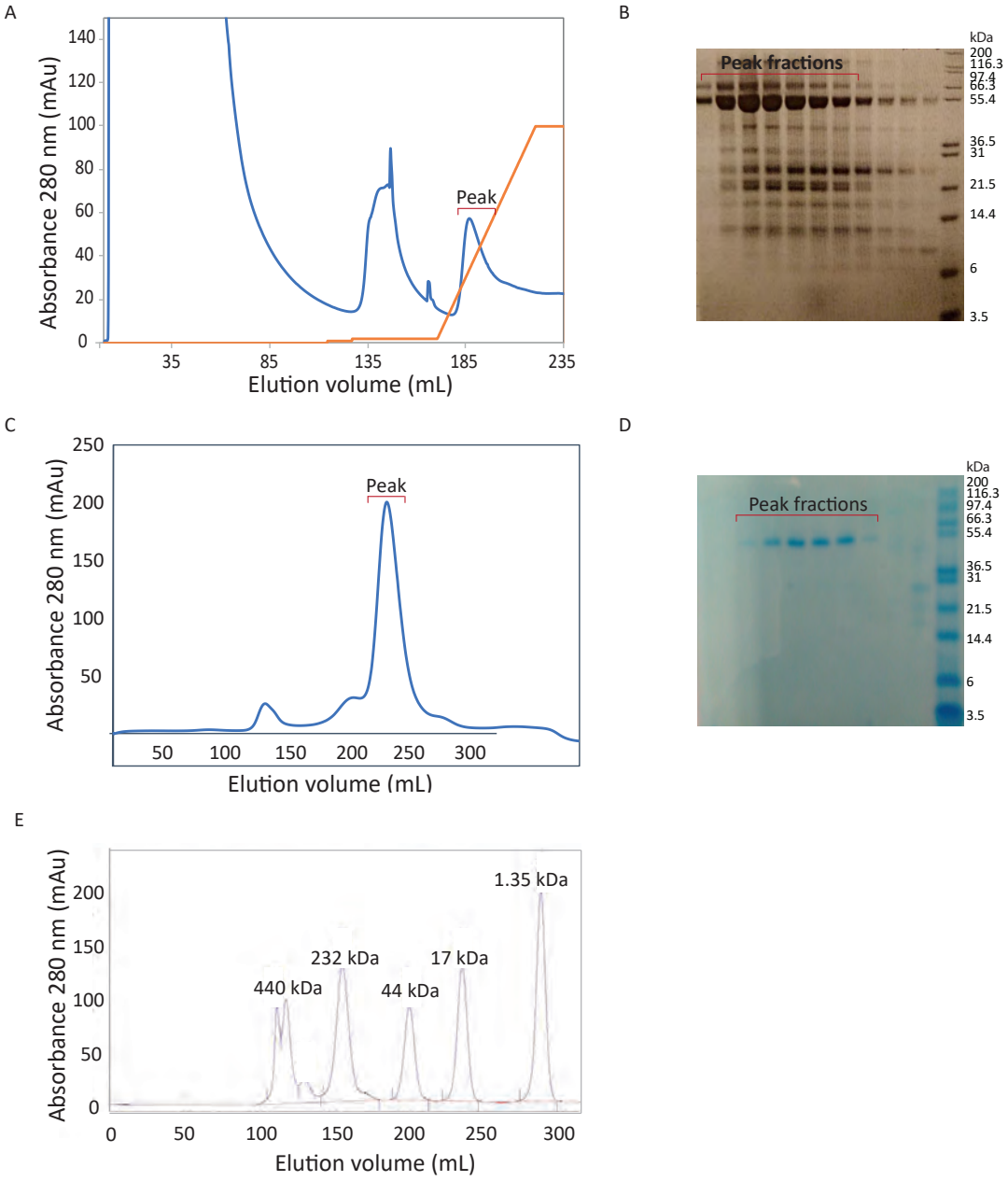


Figure 6.4: CeSAS-6 protein purification. A: Chromatogram of immobilised metal affinity purification step. Absorbance at 280 nm (blue) and imidazole gradient (orange) are shown. B: SDS-PAGE analysis of peak fractions resulting from metal affinity chromatography. C: Size exclusion chromatography (S200 26/60, 20 mM HEPES pH7.2, 150 mM NaCl, 2 mM DTT, 4 °C). D: SDS-PAGE analysis of peak fractions resulting from size exclusion chromatography shown in C. E: Calibration curve for S200 26/60 column, molecular weight corresponding to each peak is indicated.

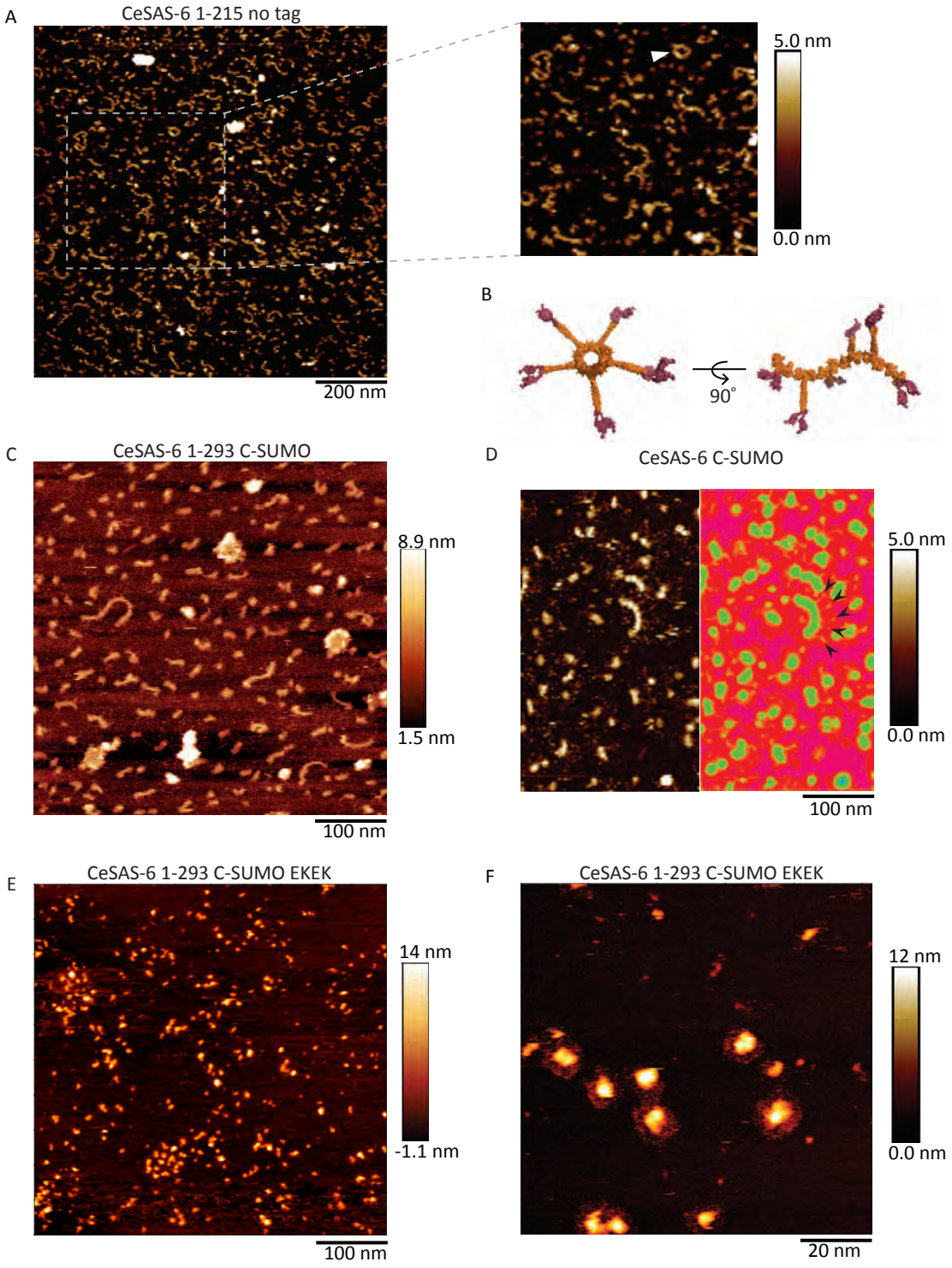


Figure 6.5: CeSAS-6 construct optimisation for AFM. Height profiles of A: CeSAS-6 1-215 (white arrow indicates a ring), B: model of CeSAS-6 1-293 (orange) with SUMO tag (pink). C: CeSAS-6 1-293 C-SUMO, D: second example of C with alternative colour gradient. Black arrows indicate SUMO tag. E: CeSAS-6 1-293 with engineered SUMO tag. F: enlarged image of disc E. Scale and colour gradient are indicated for each image.

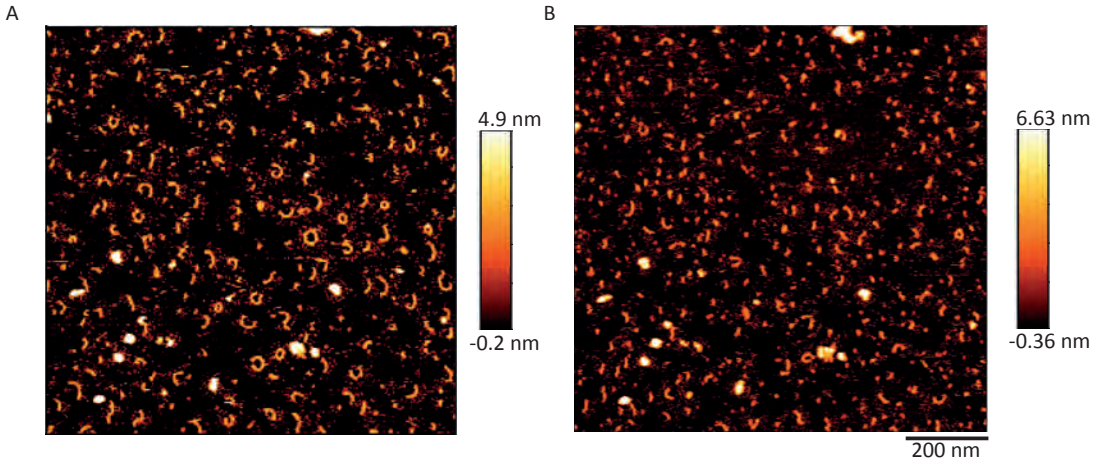


Figure 6.6: CeSAS-6 spirals are fragile. AFM height profiles of the same area of mica, the initial image is shown in A and the area imaged again in B. Scale bar and colour gradient are shown for each image.

6.3.2 Troubleshooting CeSAS-6 variant AFM

In AFM of all the mutants tested, except CeSAS-6VG and CeSAS-6APVG, no oligomerisation was observed. Rather, small aggregates, monomers and dimers can be seen (Fig. 6.8, 6.9, 6.10). SDS-PAGE analysis indicated protein samples were pure and of the expected molecular weight (Fig.6.7 B). Size exclusion chromatography did not give any indication that samples were aggregated, CeSAS-6, CeSAS-6APVG and an example mutant CeSAS-6 V167G H175E size exclusion chromatograms are shown in Fig. 6.7 A and C. For comparison, Table 6.2 summarises the elution volume from size exclusion chromatography of each mutant. Although the same concentration was used, differences in the amount of protein adsorbed onto the mica can be seen, for example CeSAS-6 V167P (Fig. 6.8 A) adsorbed more stongly than CeSAS-6 V167G S56R H175E (Fig. 6.10 O). The size of aggregates also varied between samples, CeSAS-6 V167G F61K displayed the largest aggregates (Fig. 6.8 D). To determine whether this was due to altered mobility on mica by the proteins, a range of KCl concentrations were tested (0 mM, 150 mM and 500 mM) to optimise adsorption of mutants where the overall charge of the protein had been changed. This was unsuccessful and oligomerisation, again, was not observed (Fig. 6.11 shows

a representative variant CeSAS-6 V167G E54R H175E). Therefore, ionic strength of the adsorption and imaging buffers used does not cause the lack of oligomerisation.

Table 6.2: CeSAS-6 variants used in AFM experiments

Mutation	**Molecular weight (kDa)**	**Elution volume (mL)**	**Yield (mg)**
S123E I154W	45.8	192	1.4
S123E I154W A166P V167G	45.8	189	0.6
S123E I154W V167P	45.8	196	0.8
S123E I154W V167G E62L	45.7	192	0.6
S123E I154W V167G F61K	45.7	189	1.6
S123E I154W V167G H175E	45.7	192	0.7
S123E I154W V167G S56E	45.7	193	0.5
S123E I154W V167G H175Y	45.8	192	0.5
S123E I154W V167G L64K H175Y	45.8	194	0.6
S123E I154W V167G L172K H175K	45.8	193	0.9
S123E I154W V167G I179E	45.8	189	0.5
S123E I154W V167G S56L F61L H175E	45.7	200	1.2
S123E I154W V167G S56E S58I I179K	45.8	196	0.8
S123E I154W V167G S56R H175E	45.8	191	1.0
S123E I154W V167G S56R H175Y	45.8	197	0.5
S123E I154W V167G E54R	45.8	192	0.8
S123E I154W V167G E54R H175Y	45.8	191	0.7
S123E I154W V167G E54R H175E	45.8	192	0.9
S123E I154W V167G S56E I179K	45.8	194	0.6
S123E I154W V167G L64K	45.7	199	0.6
S123E I154W A166P V167G F61K	45.8	195	0.6
S123E I154W A166P V167G H175E	45.8	194	0.6
S123E I154W A166P V167G E54R H175E	45.8	192	0.6
S123E I154W A166P V167G H175K	45.8	190	0.6

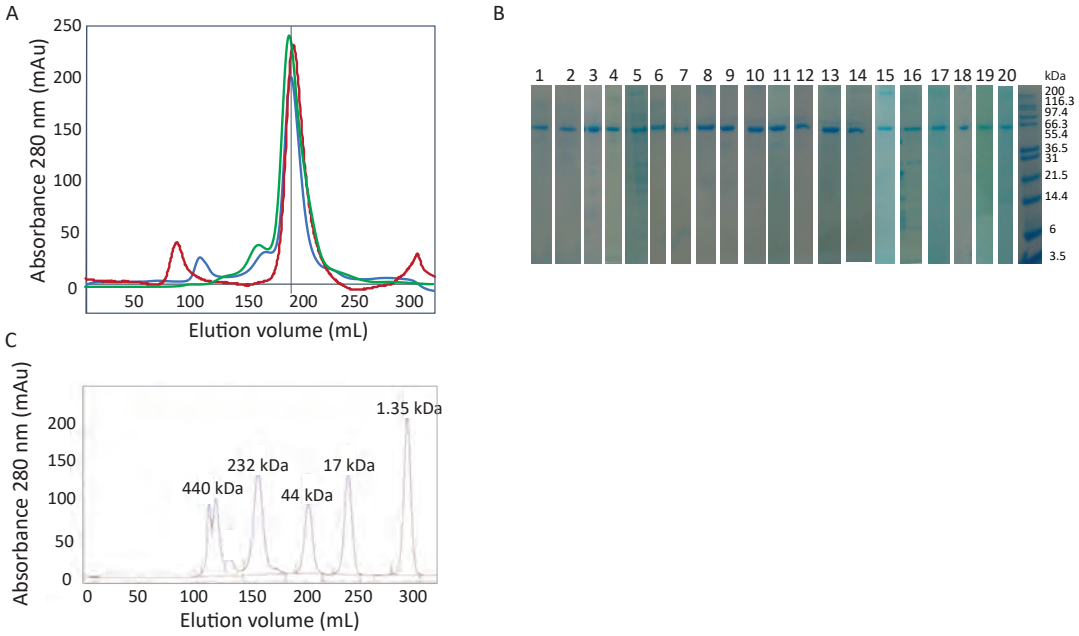


Figure 6.7: CeSAS-6 protein purification. A: Example size exclusion chromatograms of CeSAS-6 (blue), CeSAS-6APVG (green) and CeSAS-6 V167G H175E (red). B: SDS-PAGE analysis for each of the proteins used in AFM experiments following size exclusion chromatography. Lane 1:CeSAS-6, 2:CeSASAPVG, 3:V167P, 4:V167G E62L, 5:V167G F61K, 6:V167G H175Y, 7:V167G S56E, 8:V167G H175Y, 9:V167G L64K H175Y, 10: V167G L172K H175K, 11:V167G I179E, 12:V167G S56L F61L H175E, 13: V167G S56E S58I I179K, 14:V167G S56R H175E, 15:V167G S56R H175Y, 16: V167G E54R, 17: V167G E54R H175Y, 18: V167G E54R H175E, 19: V167G S56E I179K, 20:V167G L64K. C: Calibration curve for S200 26/60 column, molecular weight corresponding to each peak is indicated.

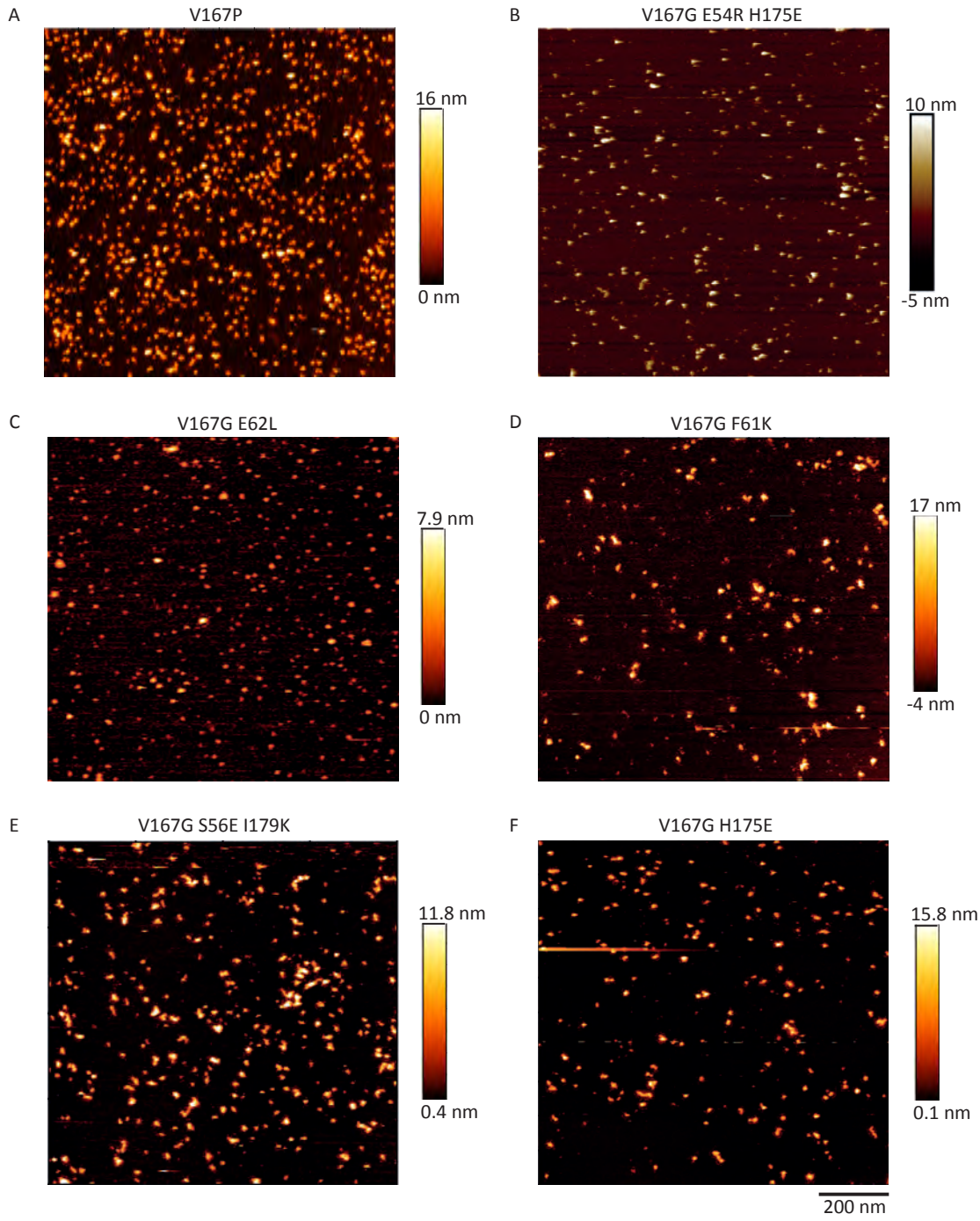


Figure 6.8: AFM height traces of unsuccessful CeSAS-6 variants (i). AFM height traces are shown for CeSAS-6 A: V167P, B: V167G E54R H175E, C: V167G E62L, D: V167G F61K, E: V167G S56E I179K and F: V167G H175E. Colour gradient and scale are indicated.

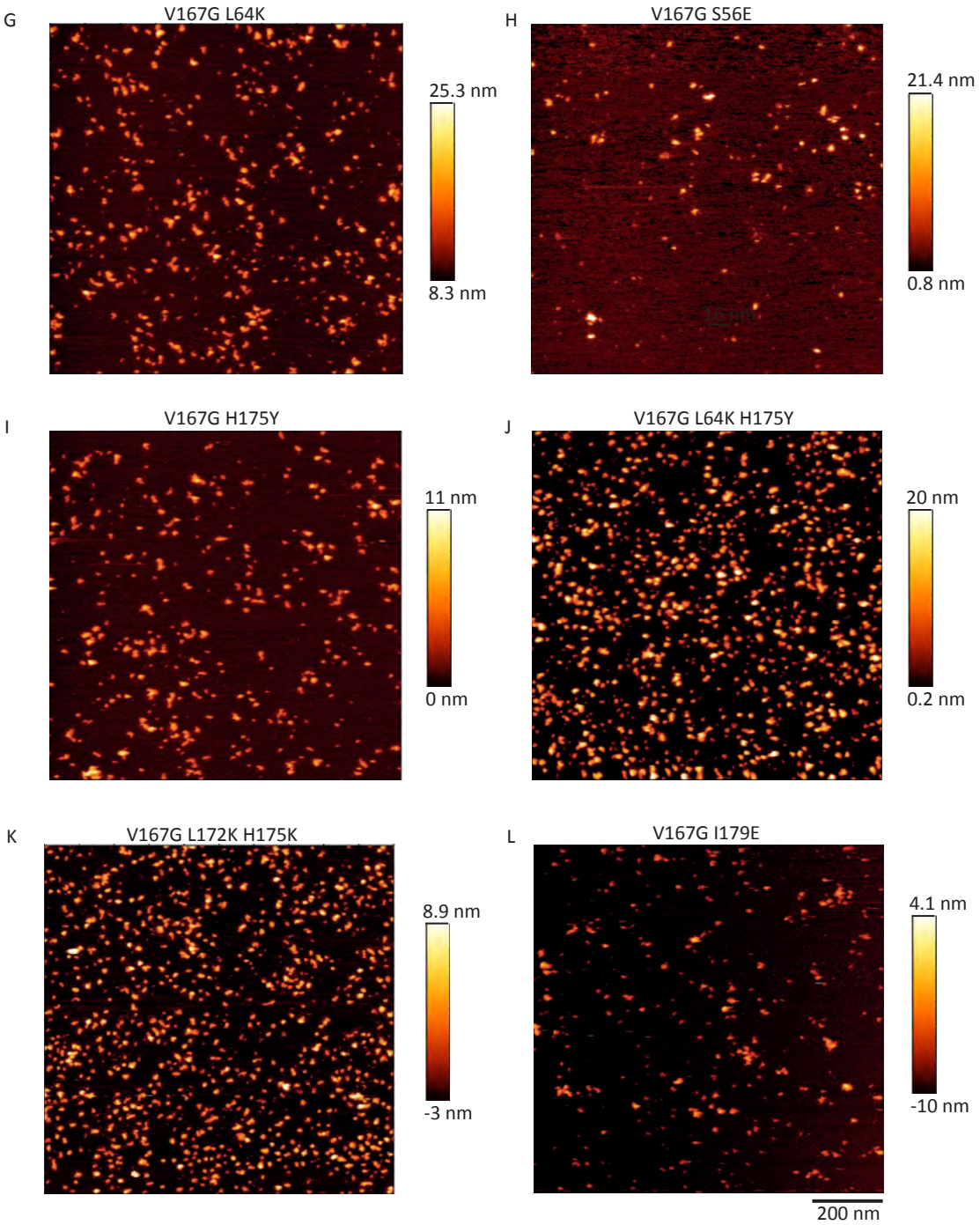


Figure 6.9: AFM height traces of unsuccessful CeSAS-6 variants (ii). AFM height traces are shown for CeSAS-6 G: V167G L64K, H: V167G S56E, I: V167G H175Y, J: V167G L64K H175Y, K: V167G L172K H175K and L: V167G I179E. Colour gradient and scale are indicated.

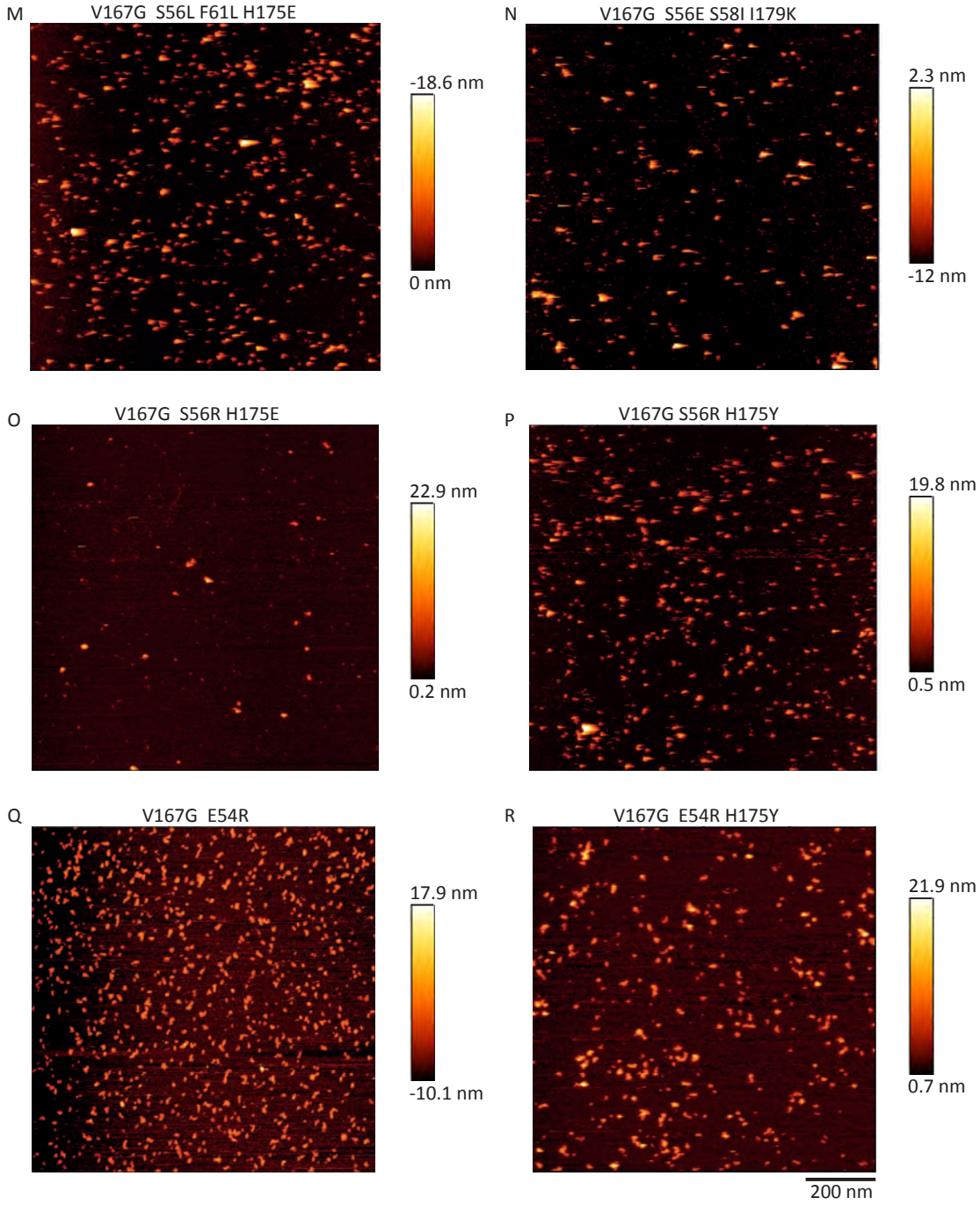


Figure 6.10: AFM height traces of unsuccessful CeSAS-6 variants (iii). AFM height traces are shown for CeSAS-6 M: V167G S56L F61L H175E, N: V167G S56E S58I I179K, O: V167G S56R H175E, P: V167G S56R H175Y, Q: V167G E54R, and R: V167G E54R H175Y. Colour gradient and scale are indicated.

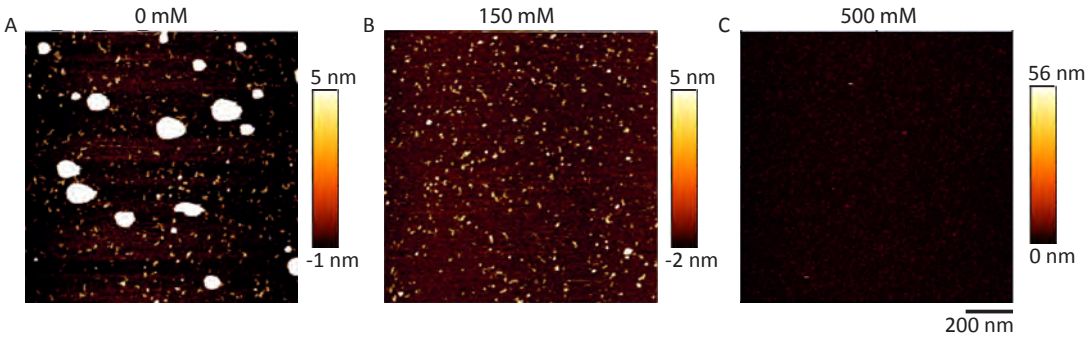


Figure 6.11: Optimisation of ionic strength used for adsorption. AFM height traces of CeSAS-6 V167G E54R H175E when 0 mM (A), 150 mM (B), or 500 mM KCl was used in adsorption buffer. Colour gradient and scale are indicated.

To determine if the lack of oligomerisation was a result of improper folding of CeSAS-6 variants, circular dichroism (CD) was carried out on 2 samples that give rise to oligomers: CeSAS-6 and CeSAS-6APVG, and a representative unsuccessful mutant: CeSAS-6 V167G H175E. The samples were correctly folded with the expected secondary structure composition prior to AFM experiments, as demonstrated by CD suggesting a mixture of α-helices and β-sheets. There is a minima at 207 nm for both CeSAS-6 and CeSAS-6APVG (Fig. 6.12 A and C) this shifted to 205 nm for CeSAS-6 V167G H175E (Fig. 6.12 E). CeSAS-6 V167G H175E also displays a less prominent minimum at 222 nm. This suggests a reduction in helical components but not complete unfolding. The melting temperature of these three constructs is comparable, at approximately 35 °C (Fig. 6.12 B, D and F). Therefore, it appears improper folding was not the cause of the lack of oligomerisation observed in AFM.

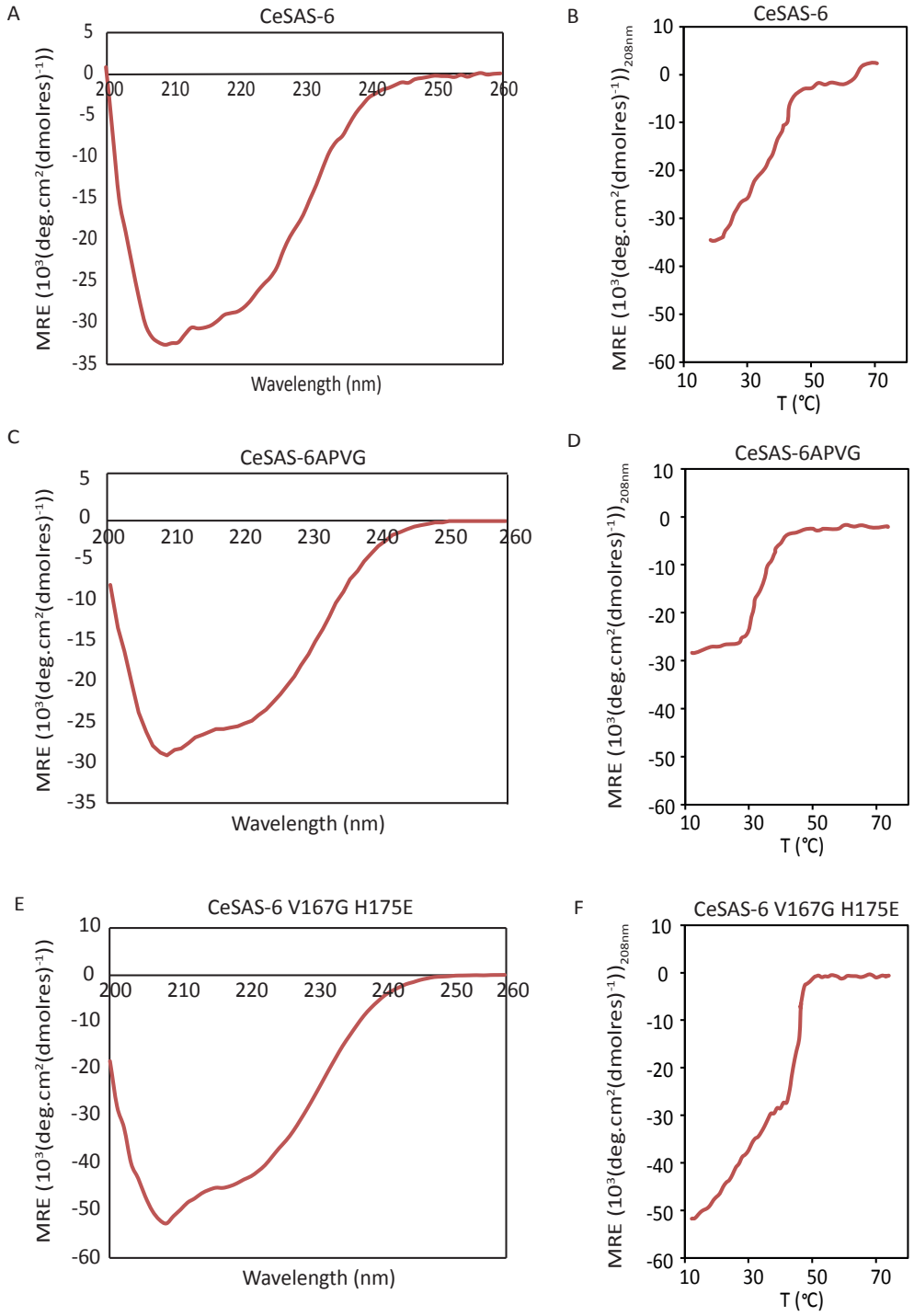


Figure 6.12: CD of CeSAS-6 variants. Circular dichroism spectra (mean residue ellipticity against wavelength) of CeSAS-6 (A), CeSAS-6APVG (C) and CeSAS-6 V167G H175E (E) are shown (red). The melting curves at 208 nm are shown for CeSAS-6 (B), CeSAS-6APVG (D) and CeSAS-6 V167G H175E (F).

To determine whether aggregation was the cause of the lack of oligomerisation observed in the unsuccessful CeSAS-6 variants, analytical size exclusion chromatography was performed. An S200 increase 10/300 column (GE healthcare) was used. Samples were run in 20 mM HEPES pH7.2, 150 mM NaCl, 2 mM DTT buffer at room temperature. CeSAS-6 and CeSAS-6APVG were used as control samples and CeSAS-6 H175E was used as a representative variant. Analytical size exclusion chromatography confirmed these samples were not substantially aggregated prior to AFM experiments (Fig. 6.13 A, B, C and D). A small aggregate peak at 8 mL is present in all 3 samples but the majority of the protein elutes after the void volume. The expected elution volume based on the curve of the molecular weight standard is approximately 13 mL for a ∼46 kDa protein (Fig. 6.13 F). The overlay of analytical size exclusion chromatograms of CeSAS-6, CeSAS-6APVG and CeSAS-6 H174E shows the elution volumes of the major peak is different for each variant (Fig. 6.13 E). CeSAS-6 eluted at the expected volume (Fig. 6.13 B). CeSAS-6APVG eluted earlier than expected, at approximately 11 mL (Fig. 6.13 C), possibly indicating the presence of higher order oligomers. CeSAS-6 H175E eluted later than expected at approximately 16 mL (Fig. 6.13 D). This may be attributable to the behaviour of coiled-coils in size exclusion chromatography, eluting at volumes away from the expected values for globular proteins; or due to degradation of the protein.

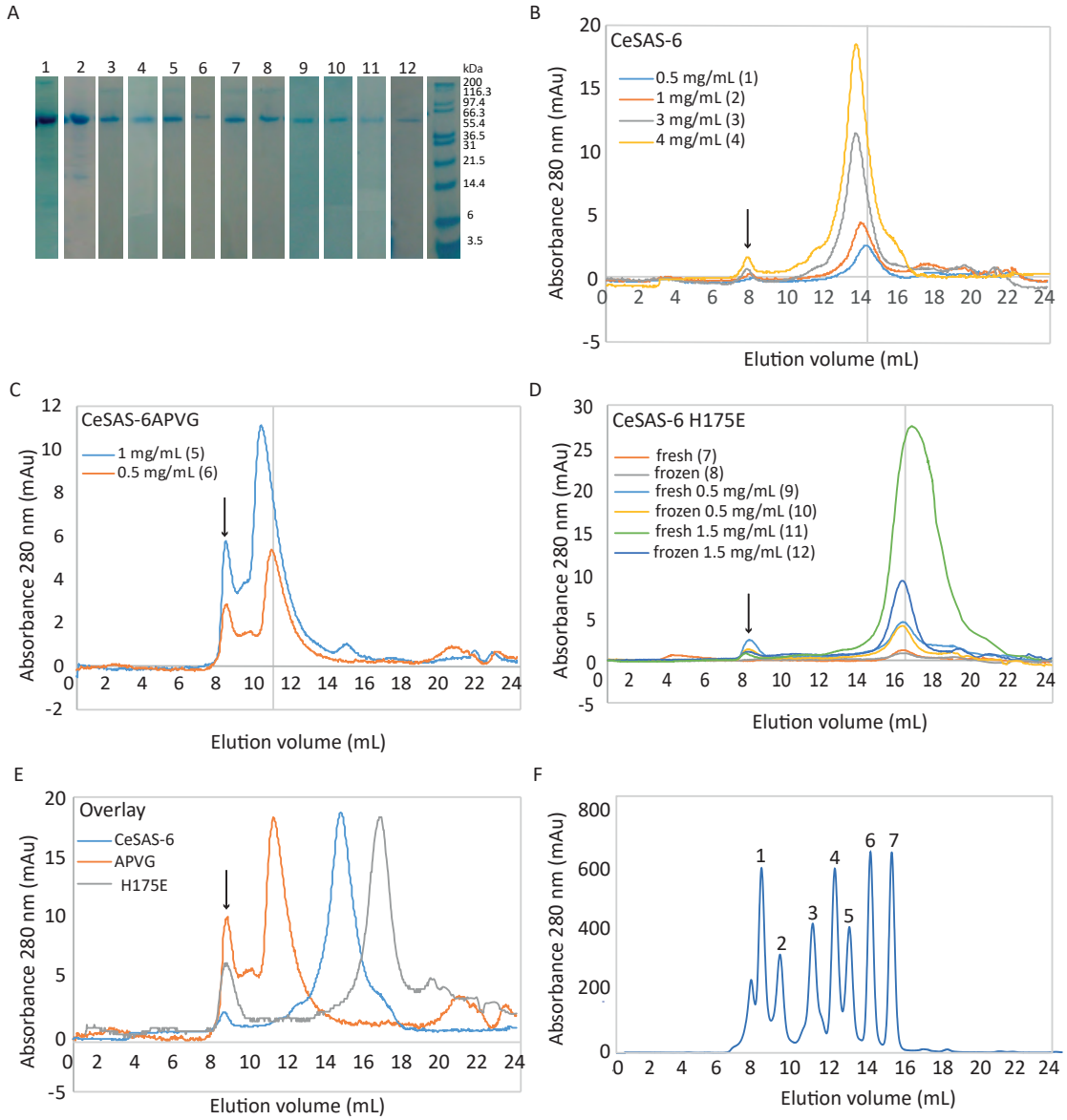


Figure 6.13: Analytical size exclusion chromatography of SAS-6 mutants. A: SDS-PAGE of SAS-6 samples used in analytical size exclusion chromatography experiments (ASEC). The corresponding lane (1-12) is indicated on each ASEC trace. B: CeSAS-6 after freeze-thaw at 0.5 mg/mL (blue), 1 mg/mL (red), 3 mg/mL (grey) and 4 mg/mL (yellow). C: CeSAS-6APVG after freeze-thaw at 0.5 mg/mL (red) and 1 mg/mL (blue). D: CeSAS-6 V167G H175E fresh sample without concentration (red), frozen sample without concentration (grey), fresh concentrated sample at 0.5 mg/mL (blue), frozen concentrated sample at 0.5 mg/mL (yellow), concentrated sample at 1.5 mg/mL (green) and frozen concentrated sample at 1.5 mg/mL (dark blue). E: overlay of CeSAS-6 (blue), CeSAS-6APVG (red) and CeSAS-6 V167G H175E (grey). Arrows indicate aggregate peaks. F: calibration curve for S200 increase 10/300 column. Peak number corresponds to a molecular weight standard: 1:670 kDa, 2:440 kDa, 3:158 kDa, 4:75 kDa, 5:44 kDa, 6:29 kDa, 7:13.7 kDa).

To determine whether the samples could oligomerise, analytical size exclusion chromatography was performed at a range of concentrations. With increasing concentration a peak shift to smaller elution volume could be seen in CeSAS-6 samples, suggesting the presence of higher molecular weight oligomers (Fig. 6.13 B). Similarly, concentration dependent peak shifts can be seen in CeSAS-6APVG size exclusion chromatography experiments (Fig. 6.13 C). Only slight peak shifts can be seen with increasing concentration of CeSAS-6 H175E (Fig. 6.13 D). This suggests the oligomerisation ability of this mutant may have been disrupted. The reduced helical component of this sample inferred from the CD spectra may account for the reduced dimerisation.

As freeze-thawed samples were used for AFM, analytical size exclusion chromatography was performed on both fresh and frozen CeSAS-6 V167G H175E to determine whether this was the cause of the lack of oligomerisation observed. The analytical size exclusion chromatography of CeSAS-6 and CeSAS-6APVG was performed with protein samples that had undergone freeze-thaw indicating this causes a small amount of aggregation but does not result in disruption of the protein secondary structure as to prevent dimerisation (Fig. 6.13 B and C). Similarly, freeze-thawed CeSAS-6 V167G H175E eluted at the expected volume and did not produce a greater proportion of eluent in the void volume compared to CeSAS-6 and CeSAS-6APVG (Fig. 6.13 D). At concentrations greater than 0.5 mg/mL, a slightly larger proportion of CeSAS-6 V167G H175E soluble protein is lost by freeze-thaw however, some soluble protein remains. This could result in less protein adsorbed onto the mica and may account for lack of oligomerisation observed. To test this, a range of CeSAS-6 V167G H175E concentrations were imaged by AFM however, no oligomerisation was observed (Fig. 6.14). This suggests that freeze-thaw of CeSAS-6 V167G H175E was not the cause of the lack of oligomerisation seen in AFM.

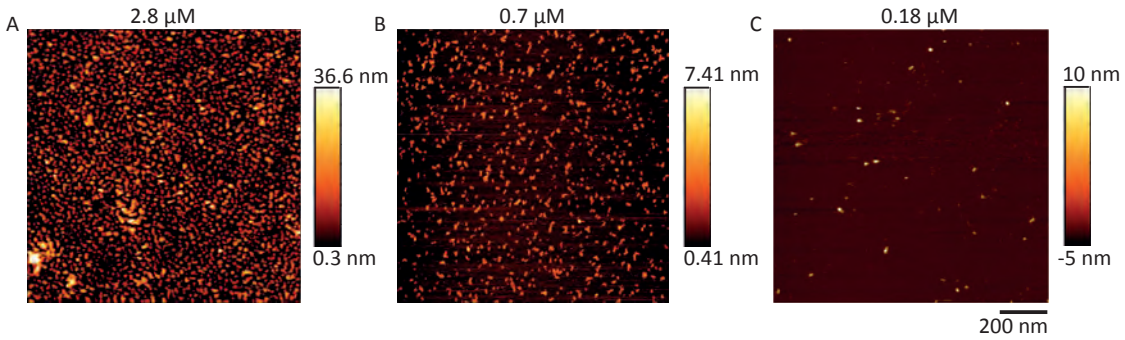


Figure 6.14: AFM concentration optimisation. AFM height traces of CeSAS-6 V167G H175E at 2.8 μM (4:1), 0.7 μM (1:1) and 0.18 μM (1:4). Scale and colour gradient are indicated.

Samples for AFM were concentrated by centrifugal ultrafiltration following size exclusion chromatography; thus, analytical size exclusion chromatography on both fresh and freeze-thawed, concentrated and not concentrated CeSAS-6 V167G H175E was carried out to determine if the concentration step caused aggregation (Fig. 6.13 D). Again, a small aggregation peak is present in the void volume, but it did not increase in size when the sample had been concentrated suggesting that this is not the reason for the lack of oligomerisation observed in AFM. There is error in the absorbance reading of the fresh 1.5 mg/mL CeSAS-6 V167G H175E sample resulting in a higher than expected absorbance reading at 280 nm, the amount of material lost by freeze-thaw at 1.5 mg/mL can therefore not be established.

6.3.2.1 Concentration optimisation

SAS-6 concentration was optimised for imaging so that the local concentration at the mica surface was sufficient for N-N oligomerisation to occur but not too high as to prevent individual oligomers from being resolved. Three concentrations were tested for each sample; CeSAS-6, and 2 CeSAS-6 variants: CeSAS-6 with V167G (CeSAS-6VG) and CeSAS-6 A166P V167G (CeSAS-6APVG). An example of CeSAS-6 APVG purification steps are shown in Fig. 6.15. For the 3 variants tested, at 1.5 µM oligomers could be seen but there was too much crowding to resolve individual oligomers (Fig. 6.16 A, D and G). At 0.8 μM oligomers could be seen with sufficient separation for image analysis (Fig. 6.16 B, E and H). 0.15 μM was too low for sufficient numbers of oligomerisation events to occur (Fig. 6.16 C, F and I). There were discrepancies in the amount of material adsorbed onto the mica. CeSAS-6VG adsorbed more strongly, therefore 0.15 μM was used in further experiments. Packing between oligomers at higher concentrations does not appear to influence oligomer architecture. Both rings and filamentous oligomers are observed at both 1.5 and 0.8 μM suggesting oligomer conformations present are not a result of interaction with other oligomers.

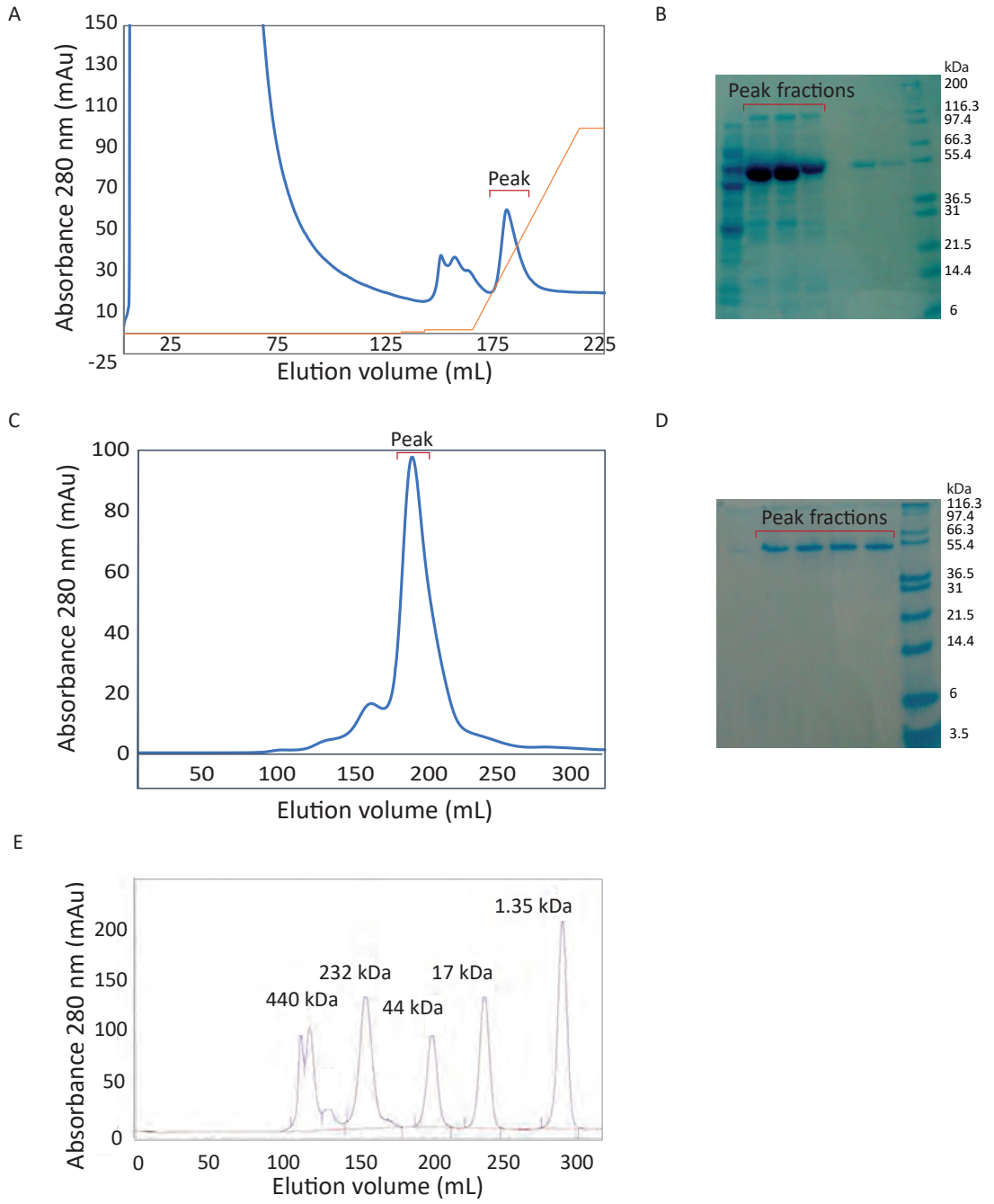


Figure 6.15: CeSAS-6APVG protein purification. A: Chromatogram of immobilised metal affinity purification step. Absorbance at 280 nm (blue) and imidazole gradient (orange) are shown. B: SDS-PAGE analysis of peak fractions resulting from metal affinity chromatography. C: Size exclusion chromatography (S200 26/60, 20 mM HEPES pH7.2, 150 mM NaCl, 2 mM DTT, 4 °C). D: SDS-PAGE analysis of peak fractions resulting from size exclusion chromatography shown in C. E: Calibration curve for S200 26/60 column, molecular weight corresponding to each peak is indicated.

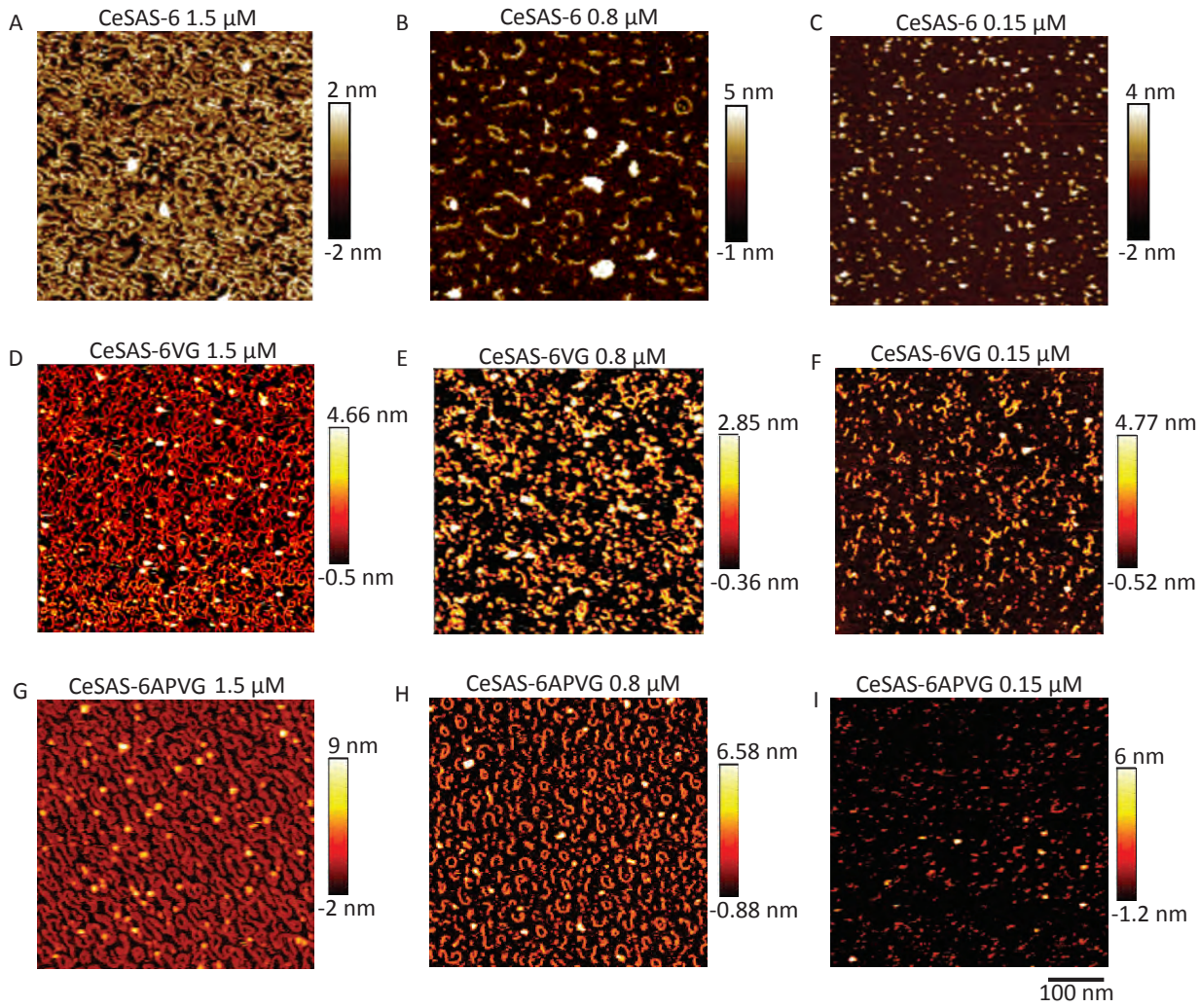


Figure 6.16: SAS-6 concentration and oligomer conformation. AFM height profiles of CeSAS-6 at 1.5 μM (A), 0.8 μM (B) and 0.15 μM (C), CeSAS-6VG at 1.5 μM (D), 0.8 μM (E) and 0.15 μM (F), and CeSAS-6APVG at 1.5 μM (G), 0.8 μM (H) and 0.15 μM (I). Scale bar and colour gradients are shown.

6.3.3 Image analysis and metrics

To analyse CeSAS-6 oligomer architecture from AFM images several metrics were used. Ring diameter was measured between the full width half-maxima of the height peaks at either side of the ring. The percentage of rings present in oligomers of at least 4 dimers was calculated manually. Spoke separation was used to assess the orientation of dimers as well as whether filaments existed as 2 intertwined spirals or spiral singlets (Fig. 6.17 A). This was also assessed by measuring backbone width (Fig. 6.17 A), equivalent to the width of the hub in ring-forming variants. The periodicity of the backbone was used to give an indication of the pitch of the oligomer present. To characterise how close oligomers were to undergoing ring closure, the curvature was measured. To do this 3 points on the inside of the arc of the filament being measured were picked (A, B and X); the centre of the circle the points lie on was then established by the intersect of the lines perpendicular to AX and BX [157]. The radius of the circle was established by drawing lines AC and BC. The angle Y (AC to BC) was then measured and curvature calculated as a ratio of the arc length between A and B, and angle Y (Fig. 6.17 B and C).

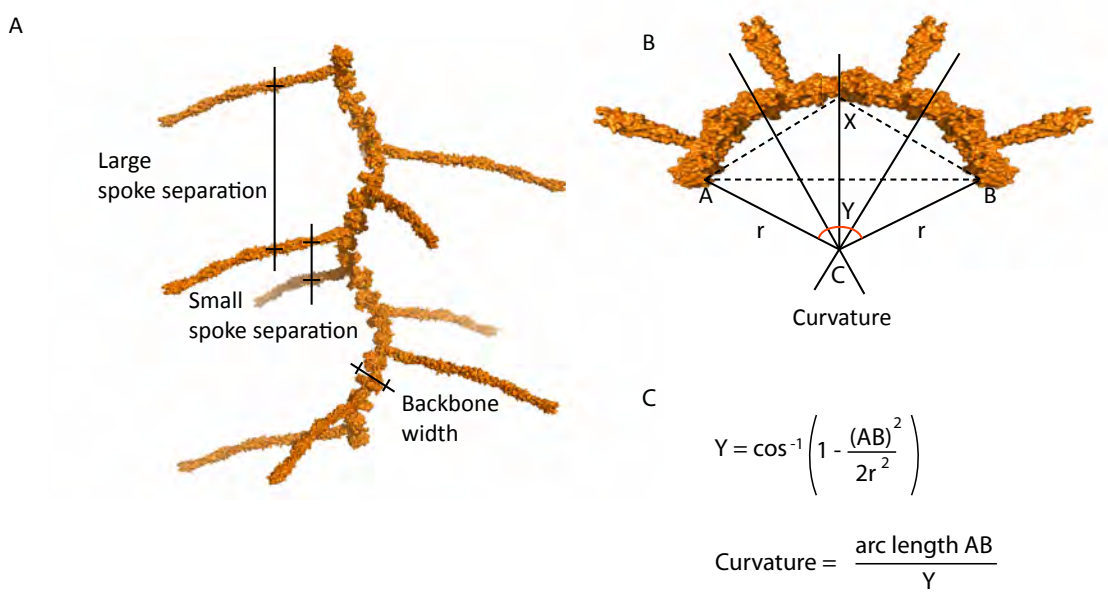


Figure 6.17: SAS-6 architecture analysis metrics. A: spoke separation and backbone width measurements and B: curvature measurement on SAS-6 filament (orange) in a 2D projection of a spiral. Where A, B and X are points on the arc of the filament to be measured, C is the centre of the circle A, B and X describe, and r is the radius of the circle.

6.4 AFM of CeSAS-6 variants

6.4.1 Ring diameter

Rings were observed in AFM images of CeSAS-6. Of oligomers comprising ≥4 dimers, 6.61% (25/378) were rings (Fig. 6.18 A and B). The expected diameter of a ring comprising 9 dimers is 23 nm. Ring diameter was calculated from AFM height traces (Fig. 6.19 C and D). The mean diameter of CeSAS-6 rings was 16±5 nm (n=25) (Fig. 6.20 A). This would correspond to a ring with approximate 6-fold symmetry. CeSAS-6VG also formed rings. 3.6% (10/278) of oligomers that were ≥4 dimers long were rings (Fig. 6.18 C and D). The mean diameter of these rings was 29±7 nm (n=10) corresponding to a ring with approximate 11-fold symmetry (Fig. 6.20 B). Ring closure events happened more frequently in CeSAS-6APVG (Fig. 6.18 E and F). 30% (41/140) of oligomers comprising ≥4 dimers were rings. The mean diameter to these rings was 20±5 nm (n=41) (Fig. 6.20 C) corresponding to a ring with approximate 8-fold symmetry.

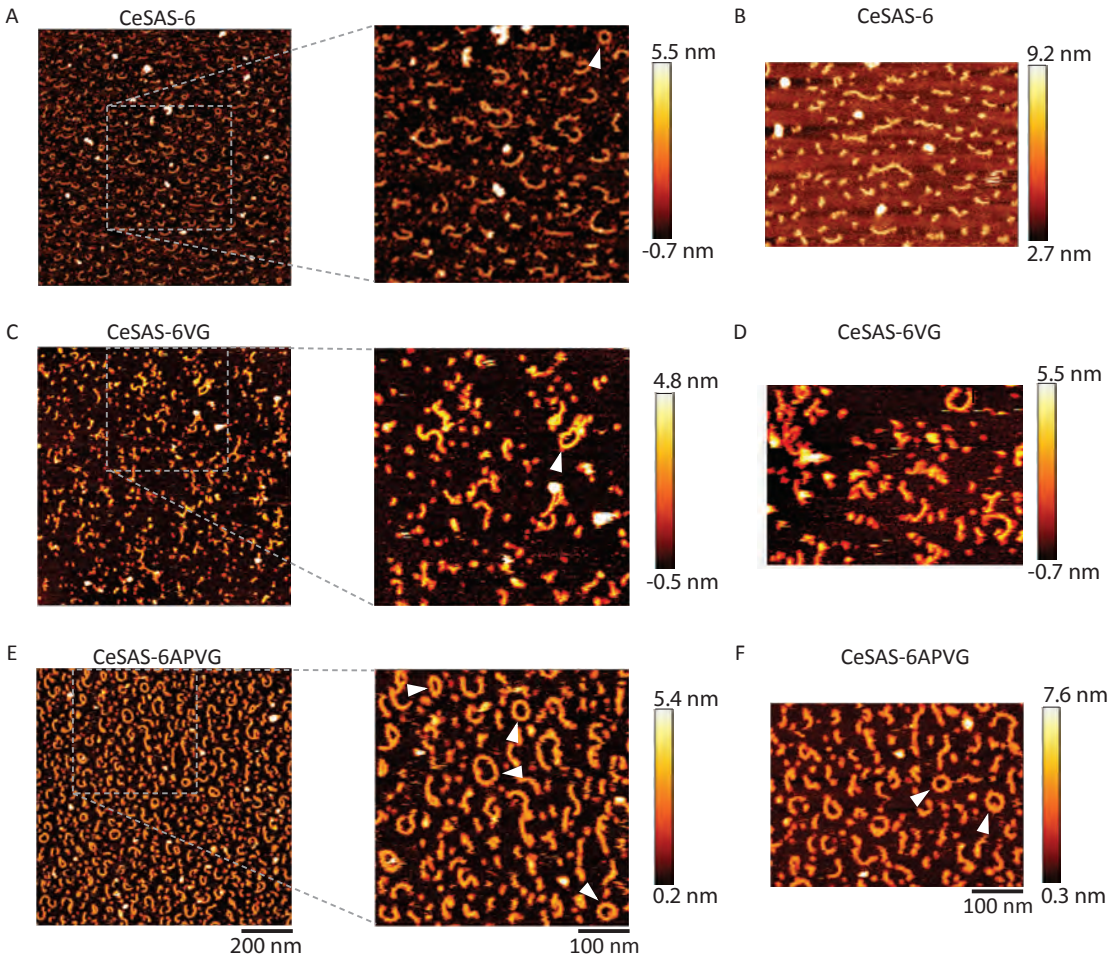


Figure 6.18: AFM height profiles of SAS-6. Representative images of CeSAS-6 (A and B), CeSAS-6VG (C and D) and CeSAS-6APVG (E and F). Scale and height colour gradients are shown, rings are indicated by white arrows.

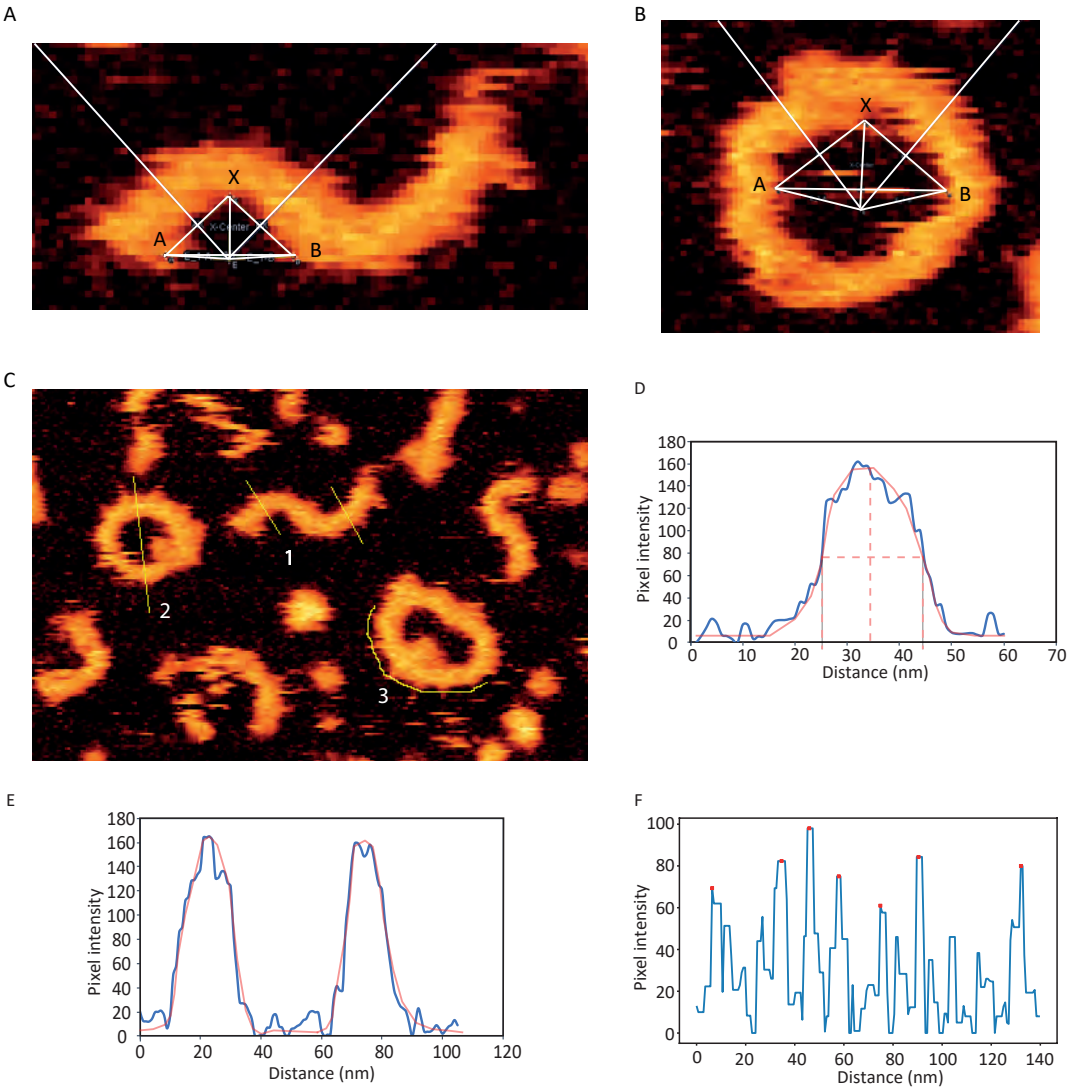


Figure 6.19: AFM image analysis. Curvature, backbone width, ring diameter and spoke separation measurements from AFM height traces. A: Curvature measurement of a filamentous oligomer (CeSAS-6APVG, angel Y=164.8°). B: Curvature measurement of a ring CeSAS-6APVG, angel Y=154.9°). C: Traces drawn on AFM height trace for backbone width measurement (1), ring diameter (2) and spoke separation (3). D: Example backbone width trace (as in C(1)) (blue) illustrating full width half maximum calculation (red) used to determine width. E: Example spoke separation trace (as in C(3) (blue) where a gaussian has been fitted to the traces (red) for peak-peak distance calculation. F: Example spoke separation trace (as in C(2)) (blue), maxima are highlighted by red crosses for peak-peak distance calculation.

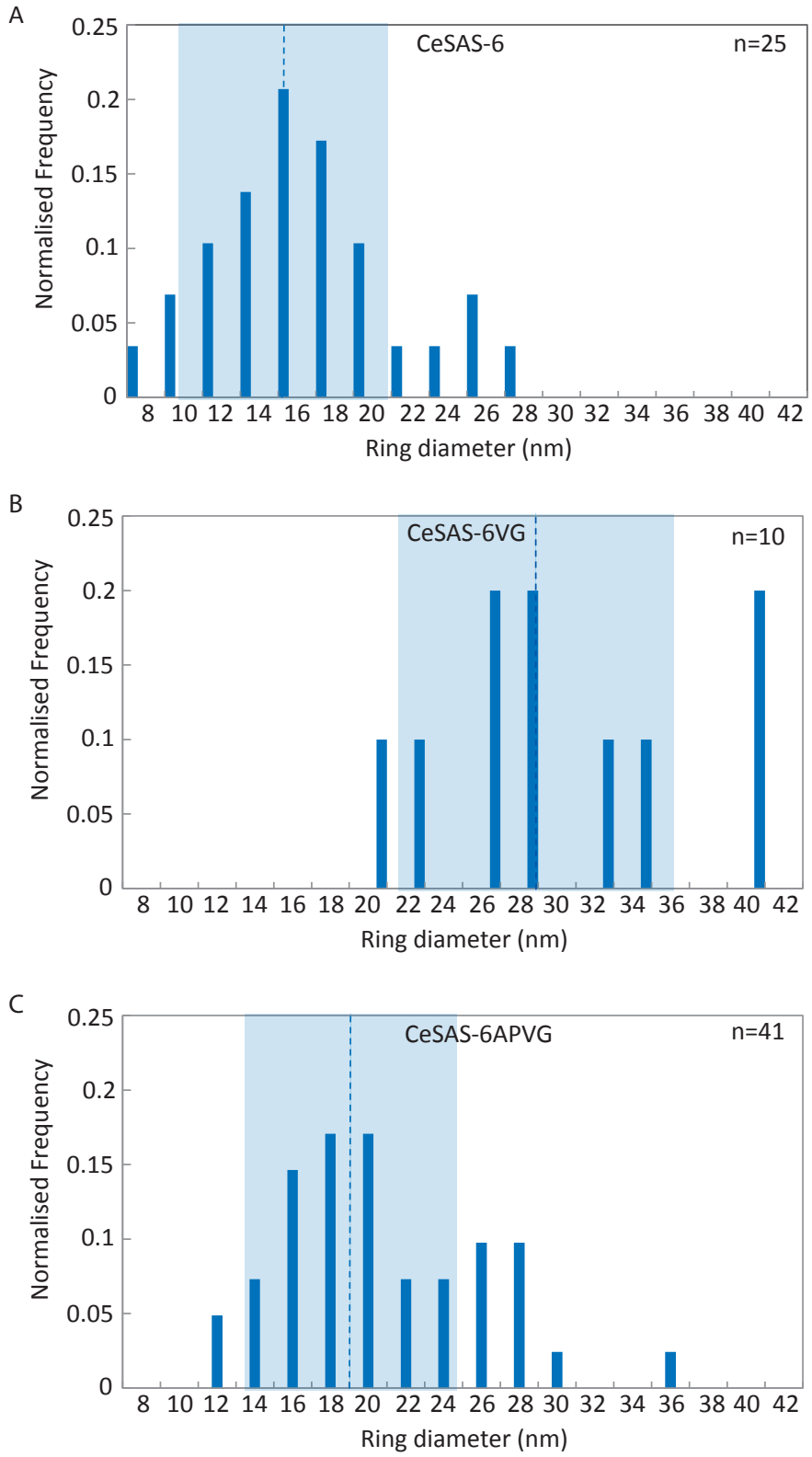


Figure 6.20: SAS-6 ring diameter. Histograms of A: CeSAS-6, B: CeSAS-6VG and C: CeSAS-6APVG ring diameter measured from AFM height profiles. Mean (dotted line) and standard deviation (blue highlight) are shown. Ring numbers (n) for each variant are also indicated.

6.4.2 Backbone width

It was observed in previous EM studies on SAS-6 oligomers that 2 spirals sometimes intertwined and came apart [39]. To determine whether oligomers visualised by AFM were single or intertwined, the backbone width was measured (Fig. 6.17 A, 6.19 C and D). The mean backbone width expected for a single SAS-6 filament is 5±1 nm (n=10) and for a double spiral is 9±2 nm (n=10)(Fig. 6.21 A) as determined by *in silico* models of CeSAS-6 single and double filaments. CrSAS-6 rings had a mean backbone width of 8±1.2 nm (n=20) (Fig. 6.21 B). It is expected that the backbone width of the CrSAS-6 rings will be the same as that of CeSAS-6 single spirals as rings are formed by a single chain of SAS-6. Distortions associated with the AFM imaging process result in slight broadening of objects; therefore, this larger width value is a more accurate point of comparison for the presence of single or intertwined spirals.

CeSAS-6 filaments had a width of 8.5±1.5 nm (n=55) (Fig. 6.21 C), similar to that of CrSAS-6 rings. CeSAS-6APVG filaments had a width of 10±1.5 nm (n=43), comparable to that of CeSAS-6 filaments (Fig. 6.21 D). This suggests these oligomers are single filaments. This is further reinforced by the similar width observed in rings formed by these variants. It is possible for ring closure events to happen as a result of two ends of a spiral coming together or from oligomers existing in a distinct ring conformation. Rings produced via both of these routes are likely single oligomer chains, as forcing a double spiral into a ring architecture, especially of the diameter observed, would be unfavourable. CeSAS-6 rings have mean width of 10.6±2 nm (n=11) (Fig. 6.21 E) and CeSAS-6APVG rings have a mean backbone width of 9.5±1 nm (n=7) (Fig. 6.21 F). This suggests both rings and spirals of CeSAS-6 and CeSAS-6APVG exist as single oligomers under the conditions used.

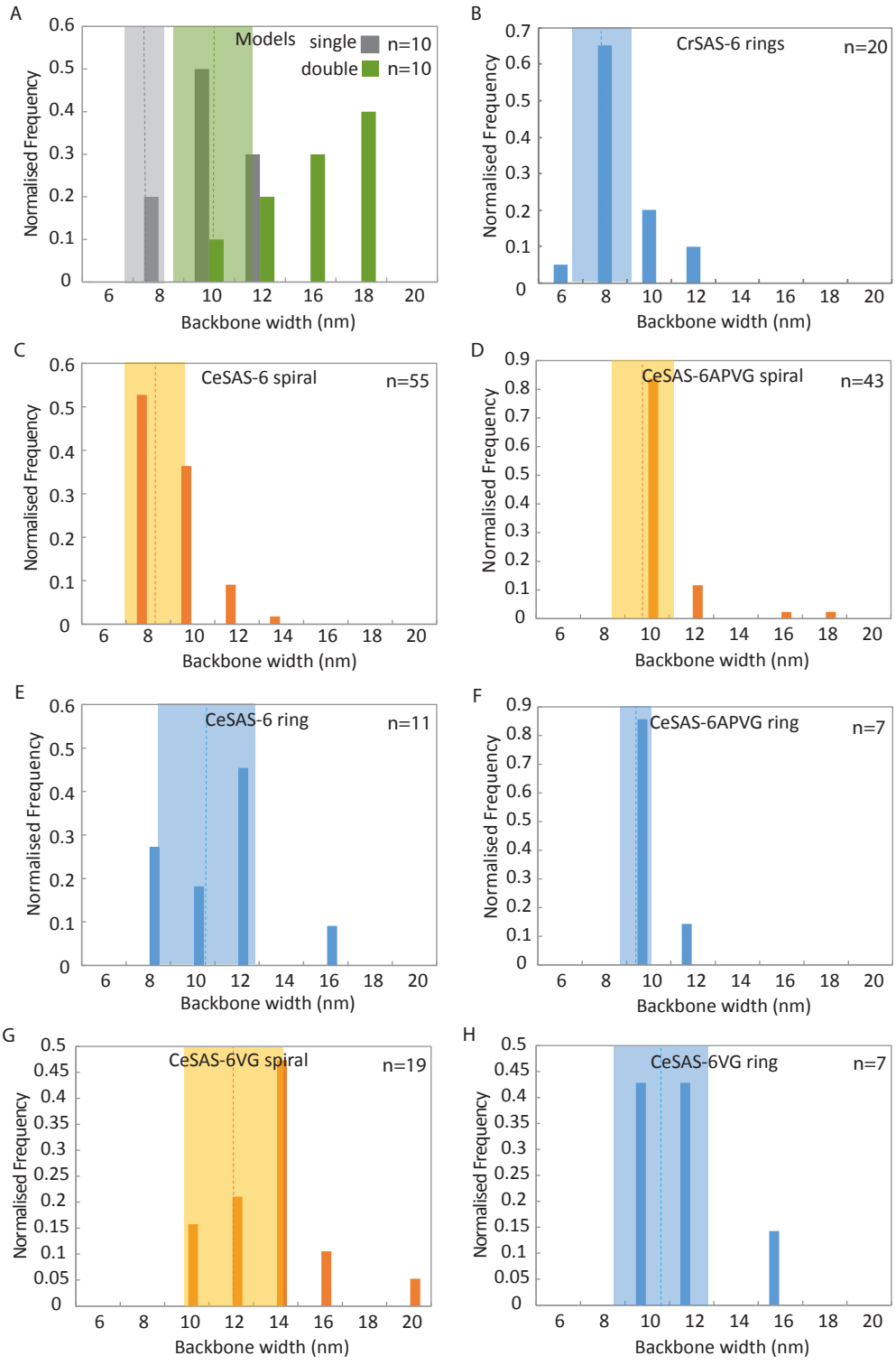


Figure 6.21: SAS-6 backbone width. Histograms of A: model CeSAS-6 backbone width of a single filament (grey) and 2 intertwined filaments (green). Backbone width of rings (blue) and spirals (orange) is shown for B: CrSAS-6, C: CeSAS-6, D: CeSAS-6APVG, E: CeSAS-6, F: CeSAS-6APVG, G and H: CeSAS-6VG. Numbers (n), mean (dotted line) and standard deviation (highlight) are shown.

The mean backbone width of CeSAS-6VG rings was 11±2 nm (n=7) also suggesting the presence of a single oligomer (Fig. 6.21 H). CeSAS-6 VG filaments have a mean backbone width of 12±2 nm (n=19) (Fig. 6.21 G), which would be compatible with the presence of 2 intertwined filaments or a single spiral; however, the quality of CeSAS-6VG images was lower than that of CeSAS-6 and CeSAS-6APVG images, and may have caused the artificial broadening of objects.

6.4.3 Spoke separation

As the oligomers form on a solid surface, and are then imaged via physical contact with the cantilever tip, there is distortion in the resulting height profile. To model the extent of this distortion models were made of both single and double spirals with altered ψ angles at the hinge residue. The wild type single spiral model had a mean spoke separation of 20±14 nm (n=14) in *in silico* models (Fig. 6.22 A). The ψ angle of the hinge V167 residue is 125°; when this was altered to 110° the spoke separation dropped to 18±13 nm (n=36). When the hinge ψ angle was further reduced to 100° the mean spoke separation was 16±10 nm (n=37) (Fig. 6.22 A). The large deviation in spoke separation arises because the model used is a 2D projection of the 3D spiral. In the model filament spokes are not uniformly separated as they would be in a 3D spiral. Figure 6.17 A illustrates the presence of both small and large spoke separation distances in the 2D filament model. For the double filament model, the wild type mean spoke separation was 11±6 nm (n=15) (Fig. 6.22 B). When the ψ angle was reduced to 110° the spoke separation was 12±4 nm (n=74), and when the ψ angle was 100° the mean spoke separation was 11±4 nm (n=72) (Fig. 6.22 B). The spoke separation of CeSAS-6 variants will be compared to these models to give an indication of the level of distortion present in AFM height traces as well as providing a second metric for the single or double nature of the CeSAS-6 variant filaments.

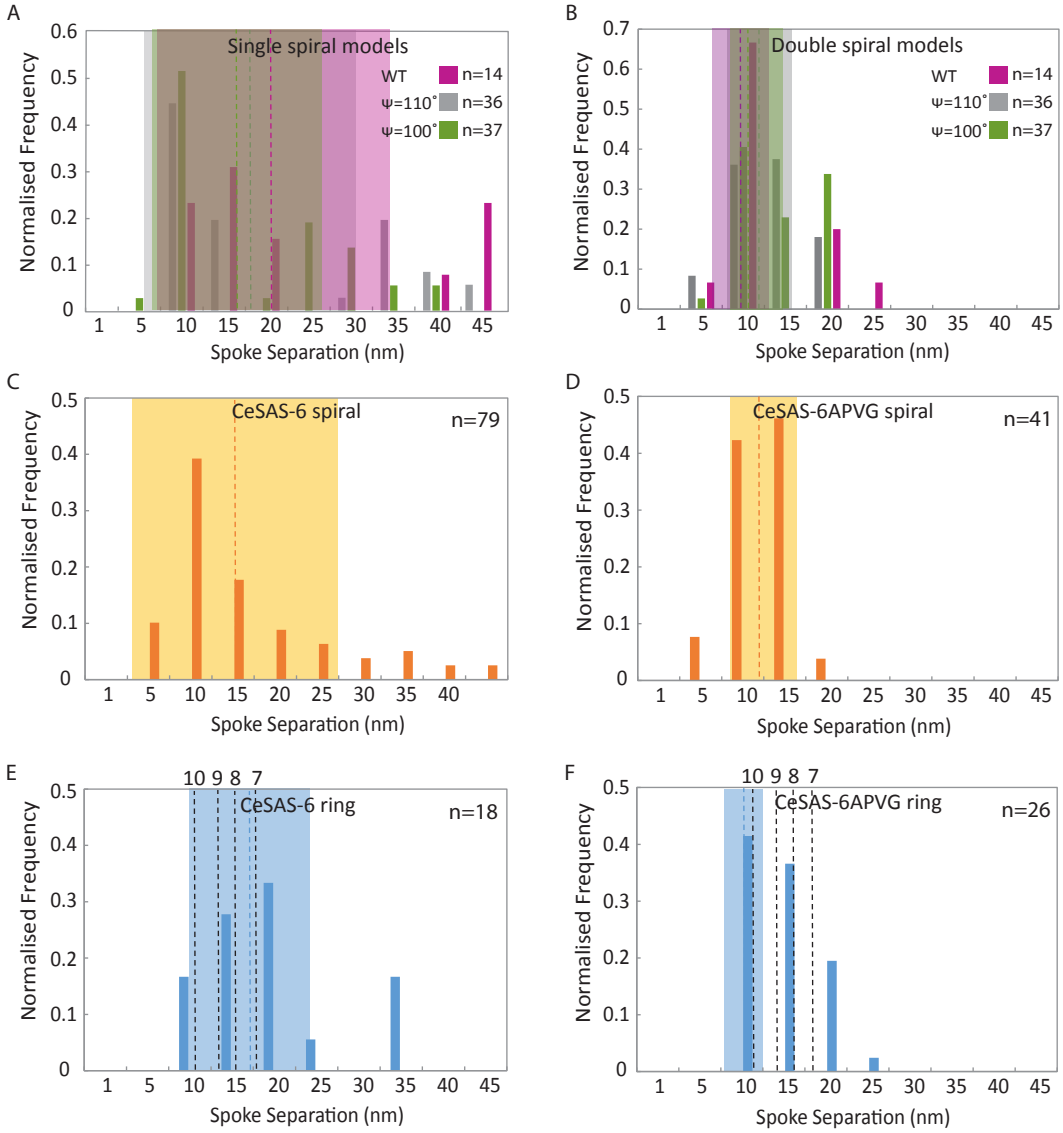


Figure 6.22: SAS-6 oligomer spoke separation. Histograms of A: Spoke separation in single filament models and B: spoke separation in double filament models of wild-type CeSAS-6 (pink), hinge ψ 110° (grey) and hinge ψ 100° (green). Spoke separation of C: CeSAS-6 and D: CeSAS-6APVG filaments are shown. Spoke separation of E: CeSAS-6 and F: CeSAS-6APVG rings are shown. Expected spoke separation of model rings with 10-, 9-, 8- and 7-fold symmetries are indicated in black. Numbers (n), mean (dotted line) and standard deviation (highlight) are shown.

Spoke separation was calculated from AFM height traces (Fig. 6.19 C and F). The mean spoke separation of CeSAS-6 spirals is 15±13 nm (n=79) (Fig. 6.22 C). This is most similar to the single spiral model where the ψ dihedral has been altered by 25°. This suggests the same level of distortion caused by the experimental environment can be captured by an alteration of the hinge valine ψ dihedral to 100°. The large variation in CeSAS-6 spiral spoke separation may be accounted for by the uneven distribution of spokes in the projection of a single spiral. Spokes would be more evenly separated if two spirals were intertwined. This further suggests that *in vitro* CeSAS-6 spirals exist as a single filament. The mean spoke separation in CeSAS-6APVG spirals is 12±4 nm (n=41) (Fig. 6.22 D). This matches the spoke separation of the ψ=110° double spiral model. It is likely the mean is an over-estimate of spoke separation as some spokes will be missing from the image, backbone width is a more reliable measure of the spiral as a single or double oligomer.

The expected spoke separation of a 9-fold symmetric ring is 13.6 nm. CeSAS-6 rings have a mean spoke separation of 17±8 nm (n=18) (Fig. 6.22 E). CeSAS-6APVG rings have a mean spoke separation of 10±3 nm (n=26) (Fig. 6.22 F). The expected spoke separation of a ring with 10-fold symmetry is 12 nm and the expected spoke separation of a ring with 7-fold symmetry is 18 nm. The large standard deviations of these measurements and narrow spread of spoke separation as symmetry increases, mean it is not possible to determine the symmetry of these rings from spoke separation. As noted, not all spokes will be resolved in AFM height traces and, while this will occur less frequently in planar oligomers than in spiral oligomers, this means the spoke separation observed is compatible with a range of symmetries.

Images of CeSAS-6VG were not of sufficient quality to resolve spoke separation, these data have therefore not been included here.

6.4.4 Curvature

Curvature of oligomers visualised by AFM was measured (Fig. 6.17 B and C, 6.19 A and B). The CrSAS-6 construct used as a positive control in these experiments displayed 6-fold symmetry in AFM experiments and 5-fold symmetry was inferred from the crystallographic structure [133]. The expected curvature of a ring comprising 5 dimers is 0.11 nm/degree and for a ring comprising 6 dimers is 0.13 nm/degree. The mean curvature of CrSAS-6 in our AFM experiments measured using our method is 0.095±0.015 nm/degree (n=20) (Fig. 6.23 A) close to the expected value for a 5-fold symmetric ring. This demonstrates the method of calculating curvature used here can be used to infer symmetry of oligomers from AFM height traces.

The expected curvature of a ring comprising 9 dimers is 0.21 nm/degree. The mean curvature of CeSAS-6 rings is 0.17±0.04 nm/degree (n=8) (Fig. 6.23 B). The curvature indicates the rings likely comprise fewer than 9 dimers, as was also reflected in the reduced diameter of CeSAS-6 rings. The mean curvature of CeSAS-6VG rings is 0.2±0.1 nm/degree (n=7) (Fig. 6.23 C). The V167G mutation had no significant effect on oligomer curvature compared to CeSAS-6 rings (p=0.76) (p values were determined using a Wilcoxon rank sum test). The mean curvature of CeSAS-6APVG rings is 0.3±0.15 nm/degree (n=10) (Fig. 6.23 D). The curvature of CeSAS-6APVG rings is not significantly different from CeSAS-6VG rings (p=0.22) but was significantly different from CeSAS-6 rings (p=0.016). The shallow curvature of CeSAS-6APVG rings is reflected by the increased diameter of these oligomers, suggesting geometry compatible with radial symmetry larger than 9-fold. However, the large standard deviation associated with this value means the true curvature may be compatible with 9-fold radial symmetry.

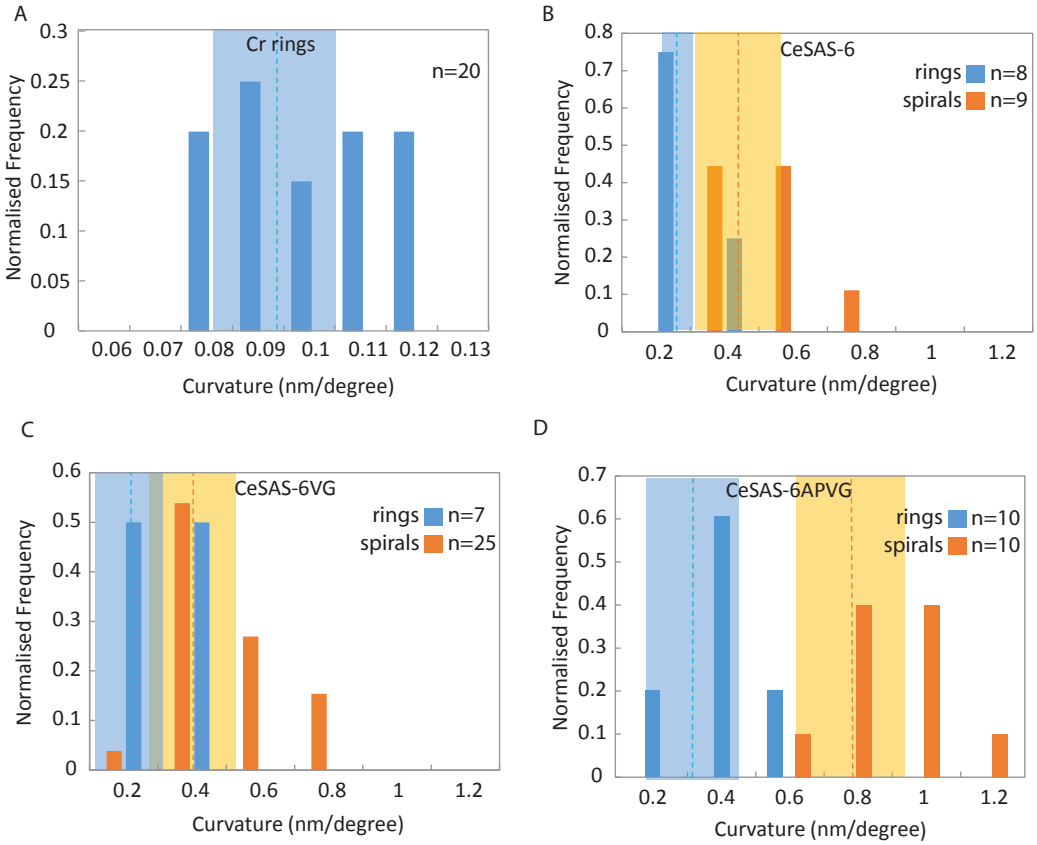


Figure 6.23: Curvature of SAS-6 oligomers. Histograms of curvature of rings (blue) and filaments (orange) is shown for A: CrSAS-6 control, B: CeSAS-6, C: CeSAS-6VG and D: CeSAS-6APVG oligomers. Numbers (n), mean (dotted lines) and standard deviation (highlight) are shown.

The addition of the hinge V167G mutation does not cause significant changes in filament curvature compared to CeSAS-6 spirals (p=0.76). CeSAS-6 filaments have a mean curvature of 0.45±0.15 nm/degree (n=9) (Fig. 6.23 B). CeSAS-6VG filaments have a mean curvature of 0.4±0.15 nm/degree (n=25) (Fig. 6.23 C). The curvature of CeSAS-6APVG filaments was significantly reduced compared to CeSAS-6 (p=0.0006) and CeSAS-6VG (p=0.000007) filaments. The mean curvature of CeSAS-6APVG filaments is 0.75±0.15 nm/degree (n=10) (Fig. 6.23 D) suggesting straighter, more elongated oligomers that have not undergone ring closure.

6.4.5 Backbone periodicity

The periodicity of the backbone can be used to infer the pitch/turn of filamentous oligomers. To measure backbone periodicity height traces were taken along the length of the oligomer backbone (Fig. 6.24 A and B). A histogram of frequencies present in these traces after Fourier transformation was plotted (Fig. 6.24 C and D). The smoothed probability function of this histogram is shown (Fig. 6.24 E). An area free from oligomers was used as negative control; no periodicity was observed along this line indicating peaks seen in the probability function are not artificial. In filamentous oligomers from all CeSAS-6 variants, a strong peak at 5 nm is present; this corresponds to the merger of peaks at 4 and 8 nm representing the periodicity of monomeric and dimeric N-terminal domains. In ring forming samples, CrSAS-6 and CeSAS-6APVG the monomer-dimer peak is shifted slightly to 7 nm. In spiral traces, although some noise is present, no additional peaks could be seen that would allow determination of the spiral pitch/turn. It was therefore not possible to infer oligomer pitch/turn using this method.

As expected, in CeSAS-6APVG rings no peak in addition to the monomer/dimer peak is seen. CrSAS-6 rings display an additional peak at 43 nm likely corresponding to the circumference of the ring. CeSAS-6VG images contained too much noise to generate intensity traces of sufficient quality for this analysis and were therefore not included.

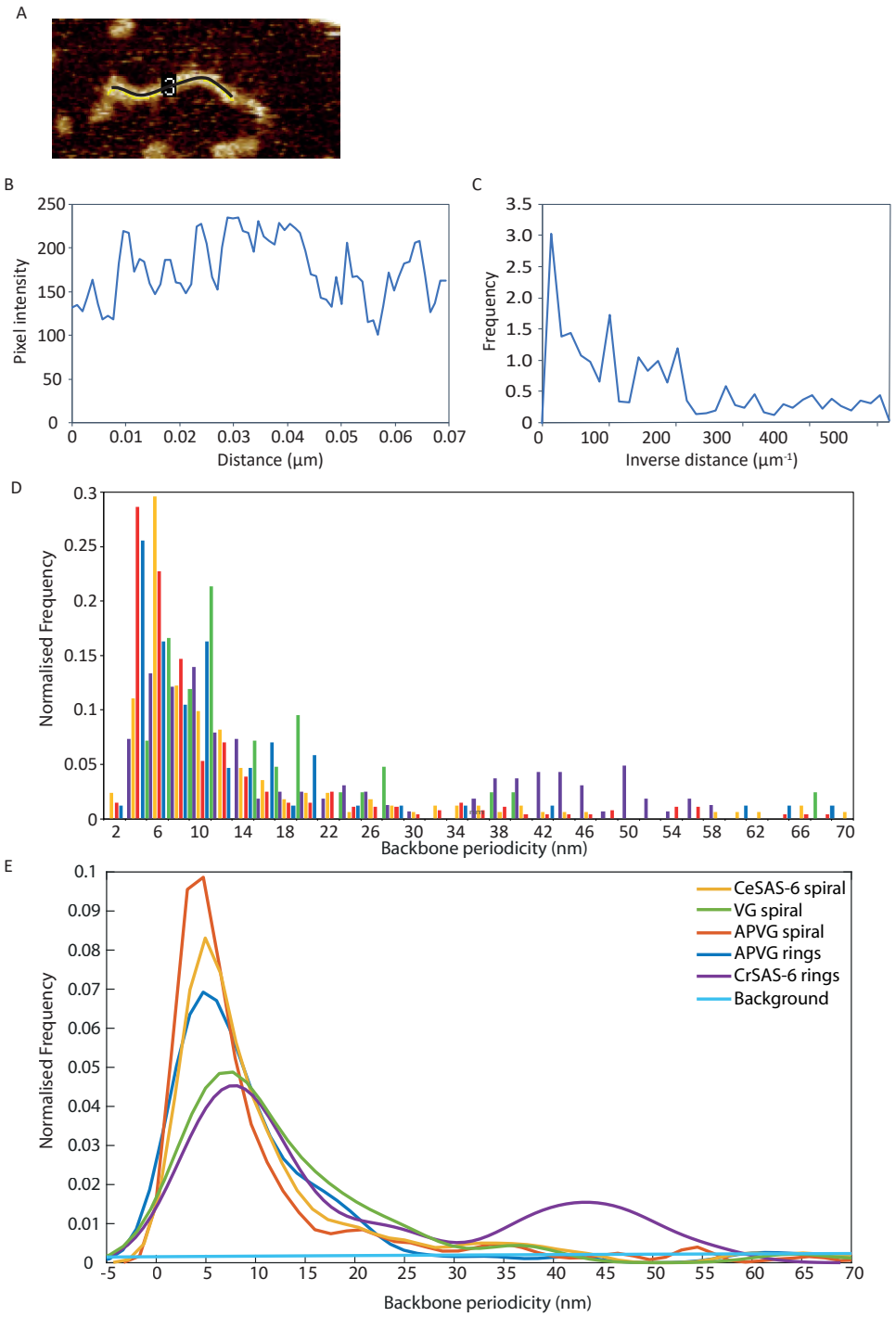


Figure 6.24: Backbone periodicity of SAS-6 oligomers. A: Example oligomer backbone trace (black line) for backbone periodicity calculation. B: Pixel intensity along backbone trace. C: Fourier transformation of backbone pixel intensity trace. D: Histograms resulting from frequencies present in fourier transformation of height traces along oligomer backbones. D: The smoothed probability function of the histograms resulting from frequencies present in height traces along the backbone of CeSAS-6APVG spirals (red), CeSAS-6 spirals (yellow), CeSAS-6APVG rings (blue), CeSAS-6VG spirals (green) and CrSAS-6 rings (purple) is shown. The resulting function from an area of background (negative control) is shown in cyan.

6.5 *In silico* analysis of CeSAS-6APVG

6.5.1 Molecular dynamics simulations of CeSAS-6APVG

Since no consistent change in symmetry, pitch/turn or radius was observed in 10 ns molecular dynamics simulations of CeSAS-6APVG, the simulations were extended to 100 ns to assess whether the spiral to ring transition happened on a longer timescale. Over 100 ns the introduction of A166P V167G (Fig. 6.25 A), resulted in a minor increase in symmetry compared to CeSAS-6 (Fig. 6.25 B). The highest symmetry sampled was the symmetry at 1 ns of 6.7-fold. For the remainder of the 100 ns symmetries sampled ranged between 5.6- and 4.2-fold. The radius and pitch/turn were unchanged compared to wild-type CeSAS-6 models (Fig. 6.25 C and D). The models remained relatively stable after large fluctuations in the initial 5 ns. As no indication of the increased propensity for ring closure events demonstrated by this mutant was apparent in these simulations, further analysis of the factors causing the increase in rings was therefore investigated.

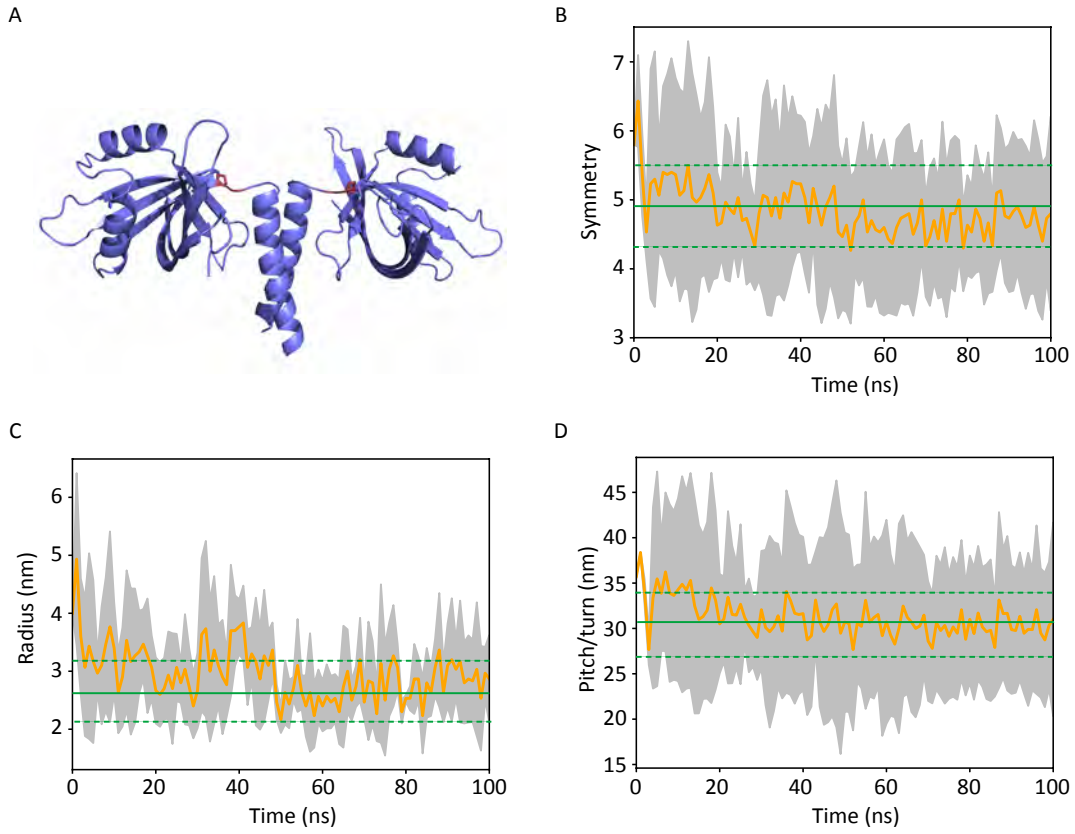


Figure 6.25: Molecular dynamics simulations of CeSAS-6APVG C-C dimer. A: C-C dimer model of CeSAS-6APVG, A166P and V167G are highlighted in red. B: symmetry, C: radius and D: pitch/turn sampled by oligomers built from the model shown in panel A over a 100 ns molecular dynamics simulation. The mean (orange) and standard deviation (grey) of triplicate simulations are shown. The mean (solid green line) and standard deviation (dashed green line) of wild-type CeSAS-6 are indicated.

6.5.2 CeSAS-6APVG dihedral angle analysis

To understand why CeSAS-6APVG oligomers formed more rings in AFM experiments compared to CeSAS-6, the backbone dihedral angles of residues neighbouring the hinge were compared to those of the ring-forming control (CrSAS-6) and CeSAS-6. The dihedral angles of the hinge -1 residue (A166 in CeSAS-6), the hinge residue (V167), the hinge +1 residue (R168) and the hinge +2 residue (G169) were monitored over a 10 ns molecular dynamics simulation of the C-C dimer. The dihedrals of the hinge +1 residue, and the hinge +2 residue were unchanged by the introduction of A166P/V167G (Fig. 6.26 E, F, G and H). However, the ϕ angle at the hinge deviated from that seen in CeSAS-6 towards angles sampled by ring-forming CrSAS-6 (Fig. 6.26 C). At this position, the hinge residue ϕ angles in CeSAS-6 fluctuated about a centre of -80°. The ϕ angle of the hinge residue in CrSAS-6 fluctuated about a centre of -200°. The ϕ angles sampled by the hinge in CeSAS-6APVG lay in between, fluctuating about a centre of -110°. The ψ angle at this position was unchanged by the introduction of A166P/V167G (Fig. 6.26 D). This demonstrates, at this position, CeSAS-6APVG sampled ϕ angles that were more compatible with those sampled by ring-forming CrSAS-6.

The ψ angles of the hinge -1 residue in CeSAS-6APVG also deviated from that of CeSAS-6 and sampled values more similar to that of the ring-forming CrSAS-6 (Fig. 6.26 B). The ψ angle of the hinge -1 residue in CeSAS-6 fluctuated about a centre of 120°. In CrSAS-6 it fluctuated about a centre of 180°. The introduction of A166P/V167G resulted in dihedrals centred at 160° for the duration of the 10 ns simulation.

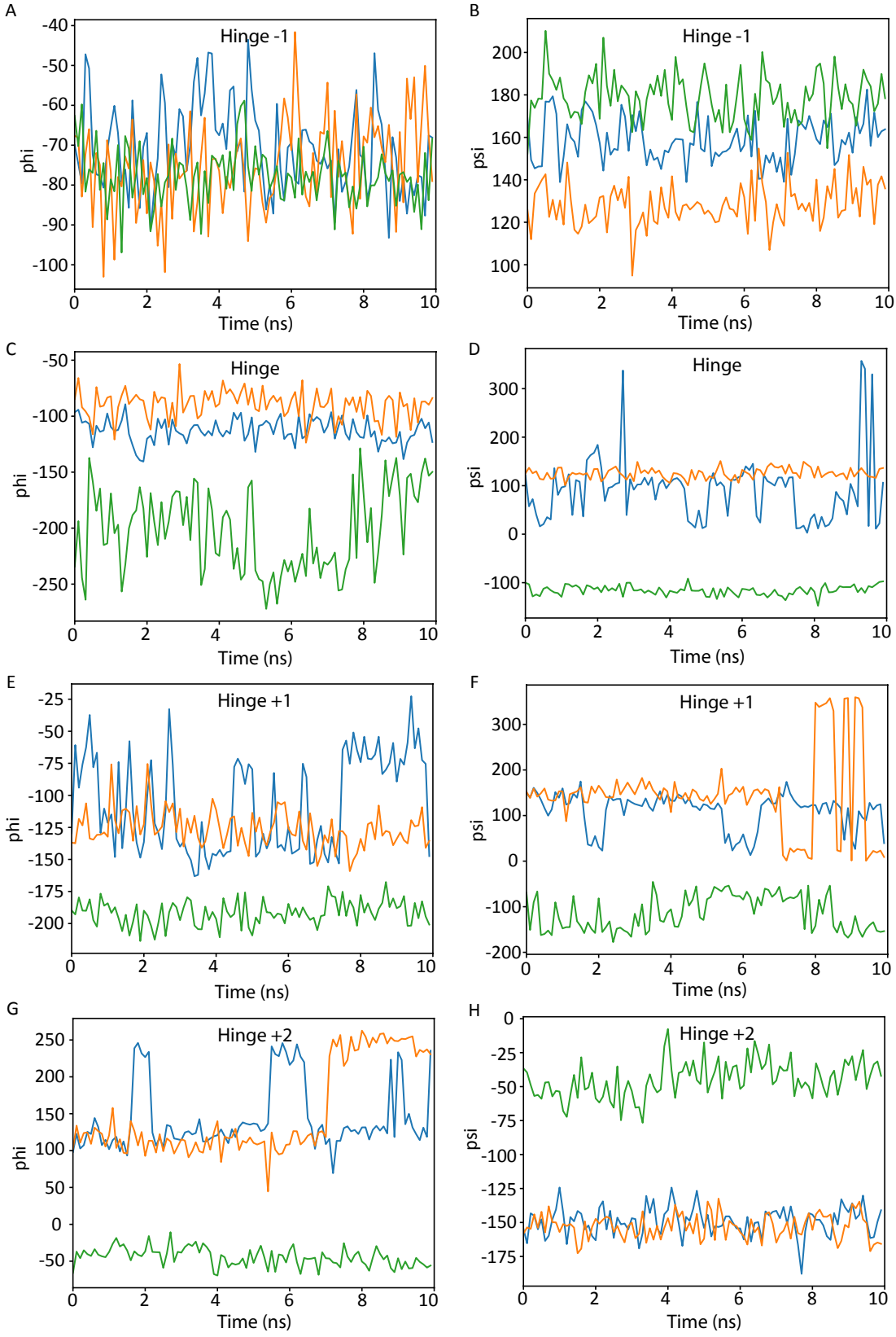


Figure 6.26: Dihedral angles sampled by residues neighbouring the hinge. ϕ and ψ angles sampled by CeSAS-6 (orange), CeSAS-6APVG (blue) and CrSAS-6 (green) during a 10 ns molecular dynamics simulation are shown for the hinge -1 residue (A and B), the hinge residue (C and D), the hinge +1 residue (E and F) and the hinge +2 residue (G and H).

As the dihedral angles sampled by CeSAS-6APVG at the hinge and hinge -1 did not transition completely to those seen in ring-forming CrSAS-6, a model where these angles were forced to those seen in CrSAS-6 was used to determine the effect on oligomer architecture and whether this change in dihedral angles was responsible for the increased number of rings. Forcing the CeSAS-6APVG hinge -1 ψ and hinge ϕ angles to adopt the dihedrals sampled by CrSAS-6 at the equivalent positions results in more ring-like oligomer architecture. When the hinge -1 ϕ angle of CeSAS-6APVG is altered to -200°, α3 rotates 103° relative to the N-terminal domain, further than the 40° rotation seen in CrSAS-6 compared to CeSAS-6 (Fig. 6.27 A). This results in oligomers with 3.1-fold symmetry, a radius of 1.5 nm and a pitch/turn of 18.7 nm (Fig. 6.27 D). An alteration of the hinge -1 ψ angle to 180° results in small movement of α3 by 23° (Fig. 6.27 B). Oligomers built from this model have 2.6-fold symmetry with a 1.4 nm radius and 16 nm pitch/turn (Fig. 6.27 E). Compared to the wild-type CeSAS-6 spiral, both of these alterations cause elongation of the oligomer. However, in combination these ring-like dihedrals result in a CeSAS-6APVG oligomer with more ring-like geometry. They cause a rotation of α3 by 43° similar to the 38° rotation of CrSAS-6 α3 compared to CeSAS-6 α3 (Fig. 6.27 C). However, α3 is also displaced perpendicularly towards the N-terminal domain. Oligomers built from this model display 11.1-fold symmetry, a 12.6 nm radius and 13.1 nm pitch/turn (Fig. 6.27 F). Although not a perfect 9-fold symmetric ring, this demonstrates that alteration of 2 dihedrals near to the hinge can drive adoption of more ring-like geometry. It also shows that A166P/V167G likely causes dihedral angles towards those sampled by ring-forming SAS-6 to be adopted and thus may be responsible for the increase in ring closure events.

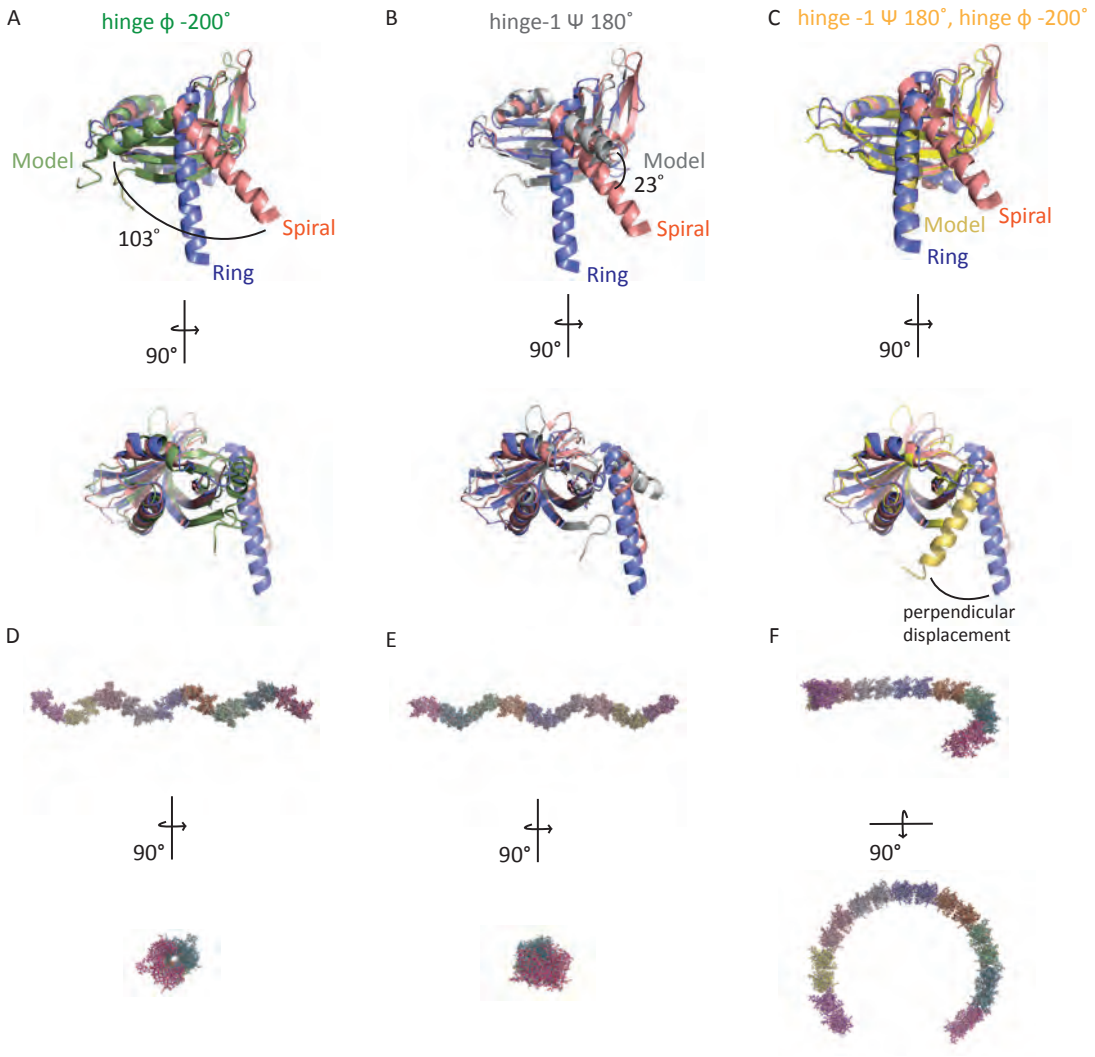


Figure 6.27: CeSAS-6APVG models with forced ring-like dihedrals. A: structural overlay of CeSAS-6APVG monomer where the hinge ϕ angle is -200° (green), B: hinge -1 ψ angle forced to 180° (grey), C: model where both the hinge ϕ and hinge -1 ψ angles are altered, with wild-type spiral CeSAS-6 (blue) and wild-type CrSAS-6 (red) monomers. The angle of α3 rotation between CeSAS-6 and the model is indicated. D, E and F: oligomers built from models shown in A, B and C respectively.

6.6 Round II CeSAS-6APVG variants

6.6.1 Molecular dynamics simulations of round II CeSAS-6 variants

A second round of mutants were designed to build upon the promising effect of the A166P/V167G mutation to promote more ring-closure events. The round II variants are; APVG F61K, APVG H175E, APVG H147E E54R (a combination of 2 single mutants tested in round one) and APVG H175K (previously tested in combination with L172K) (see Chapter 4.3 for further description of these mutations). The symmetry of APVG F61K, APVG H175E and APVG H175K was unchanged compared to wild-type models or CeSAS-6APVG (Fig. 6.28 A, B and D). APVG E54R H175E displayed an increase in symmetries sampled, with a slight general upward trend over the 100 ns simulation (Fig. 6.28 C). The maximum symmetry sampled was 7.4-fold. The radius did not increase to that expected for a ring consisting of 9 SAS-6 dimers (11.5 nm) for any of the round two mutants (Fig. 6.29 A, B , C and D). There was no reduction in pitch/turn observed in the APVG H175E E54R variant (Fig. 6.30 C). The remaining 3 mutants all displayed decreased pitch/turn. The minimum pitch/turn sampled by APVG F61K was 18 nm (Fig. 6.30 A), APVG H175E reached a minimum of pitch/turn of 19 nm (Fig. 6.30 B) and APVG H175K minimum pitch/turn was 20.5 nm (Fig. 6.30 D). The pitch/turn sampled by all 4 round II mutants displayed more variability than wild-type CeSAS-6 suggesting that a broader range of oligomer conformations were being sampled.

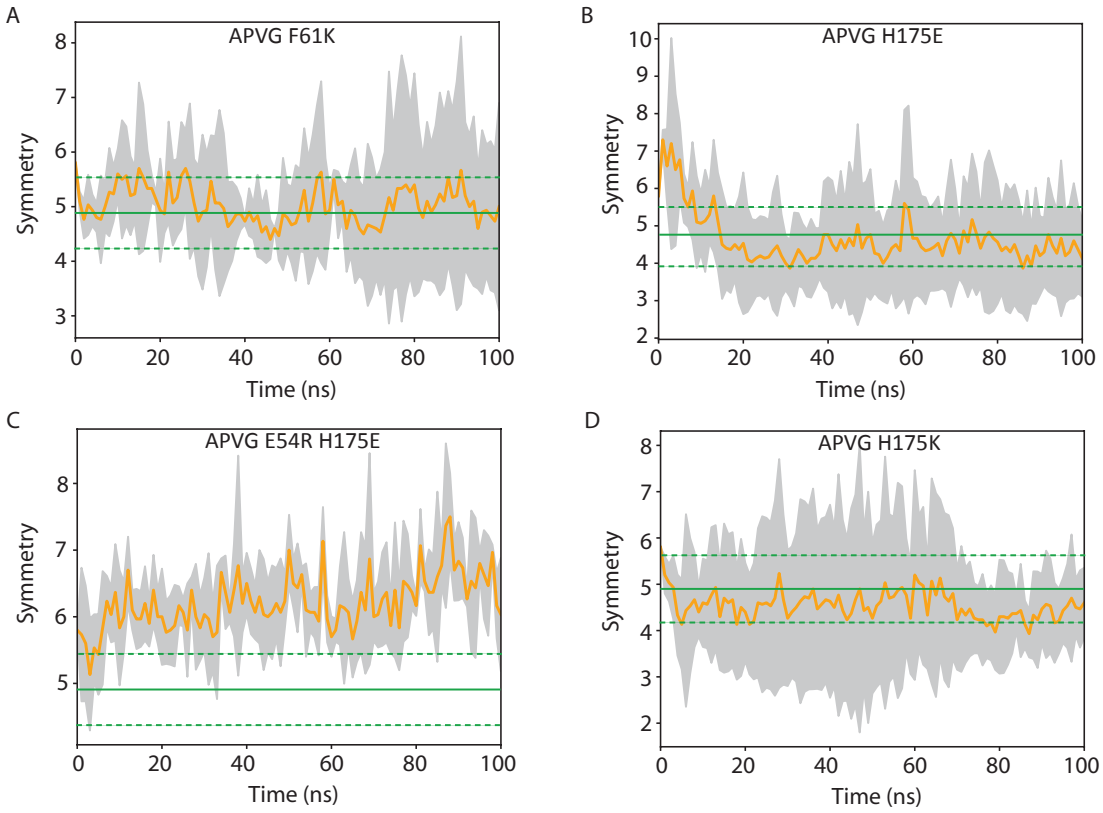


Figure 6.28: Symmetry sampled by round II CeSAS-6APVG variants. Symmetry variation over 100ns molecular dynamic simulations of oligomers built from CeSAS-6 variant C-C dimer models. A: APVG F61K, B: APVG H175E, C: APVG E54R H175E and D: APVG H175K. The mean (orange) and standard deviation (grey) of triplicate simulations are shown, green lines indicate mean (solid) and standard deviation (dashed) of wild-type CeSAS-6.

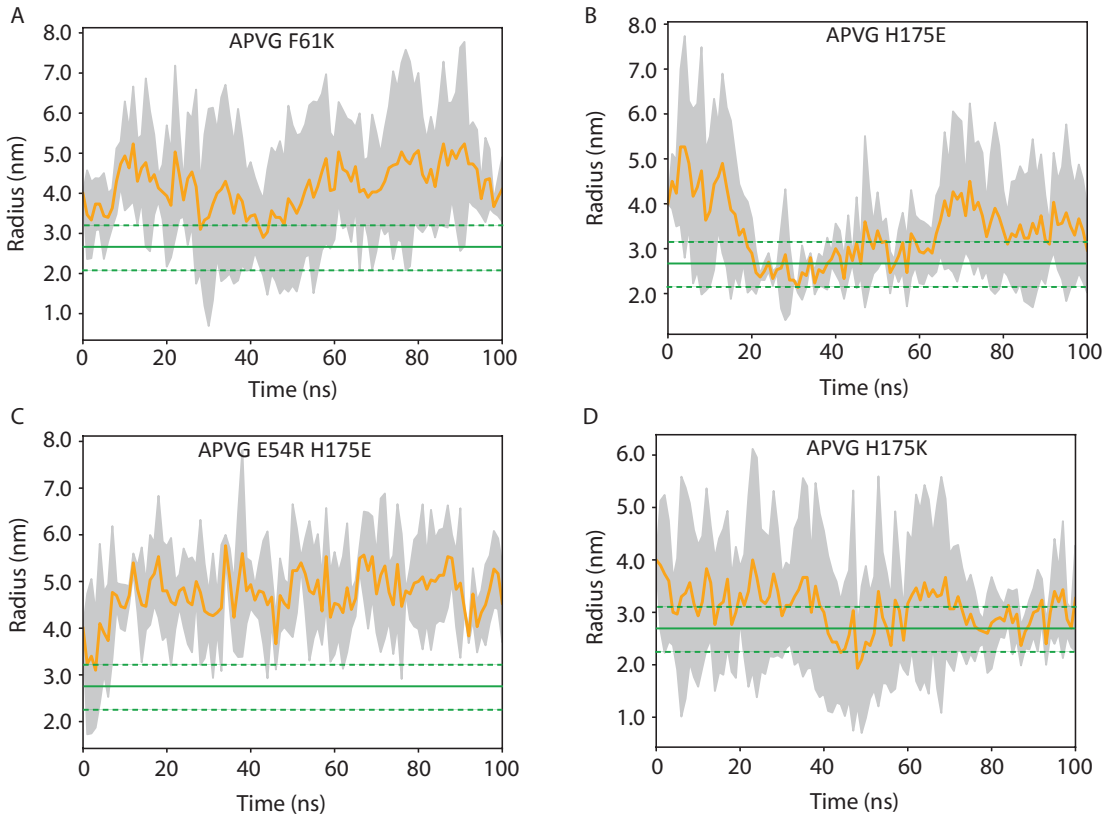


Figure 6.29: Radii sampled by round II CeSAS-6APVG variants. Radius of oligomers built from CeSAS-6 variant C-C dimer models over 100 ns molecular dynamics simulation. A: APVG F61K, B: APVG H175E, C: APVG E54R H175E and D: APVG H175K. The mean (orange) and standard deviation (grey) of triplicate simulations are shown, green lines indicate mean (solid) and standard deviation (dashed) of wild-type CeSAS-6.

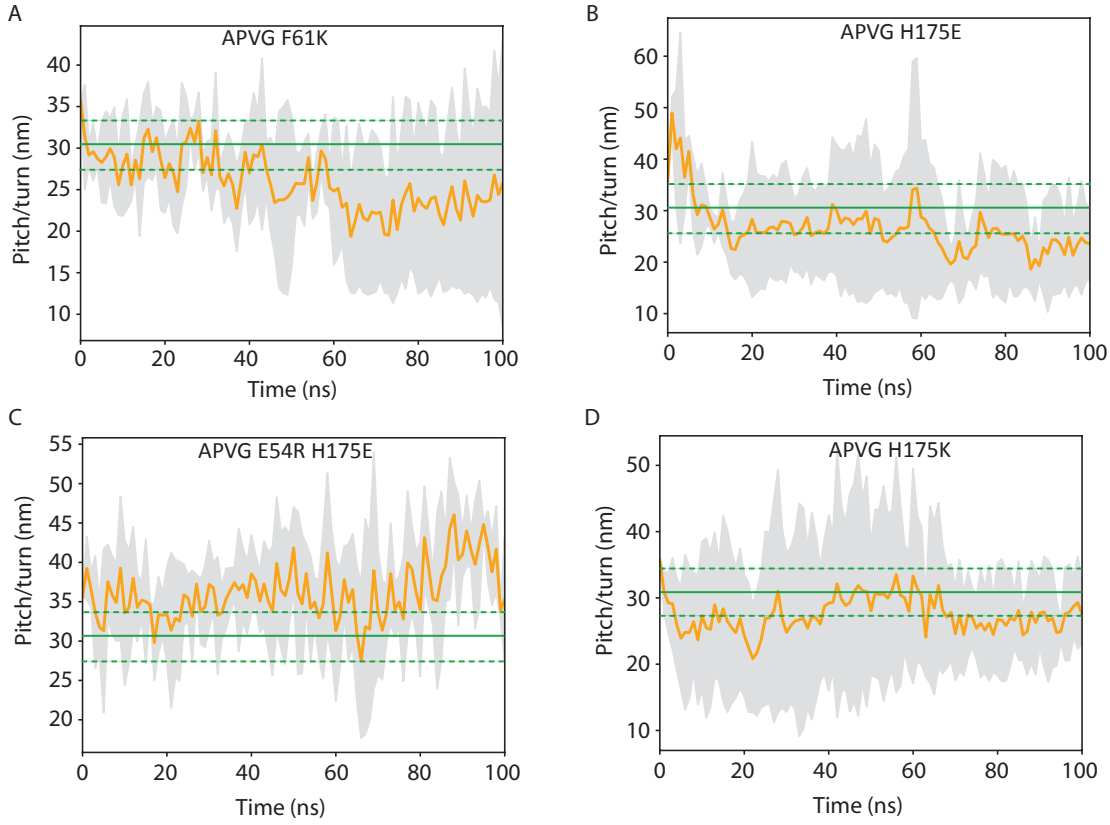


Figure 6.30: Pitch/turn sampled by round II CeSAS-6APVG variants. Pitch/turn of oligomers built from CeSAS-6 variant C-C dimer models over 100 ns molecular dynamics simulation. A: APVG F61K, B: APVG H175E, C: APVG E54R H175E and D: APVG H175K. The mean (orange) and standard deviation (grey) of triplicate simulations are shown, green lines indicate mean (solid) and standard deviation (dashed) of wild-type CeSAS-6.

6.6.2 AFM and troubleshooting of round II CeSAS-6APVG variants

The round II mutant oligomers were visualised by AFM. Similar to most variants in the first round of rational engineering, these CeSAS-6 variants did not form oligomers (Fig. 6.31). Small aggregates can be seen in the height traces with no oligomerisation events, above dimers, observed. CD of a representative round II mutant, CeSAS-6 VGAP E54R H175E, indicates the protein is properly folded. A mixture of α-helical elements and β-sheets are present, with a minima at 210 nm (Fig. 6.32 A). The CD spectra is comparable to that of CeSAS-6APVG. Again, the melting temperature of this variant is approximately 35 °C. As CD suggests the protein is properly folded, this is likely not the reason for the lack of oligomerisation observed by this sample (Fig. 6.32 B).

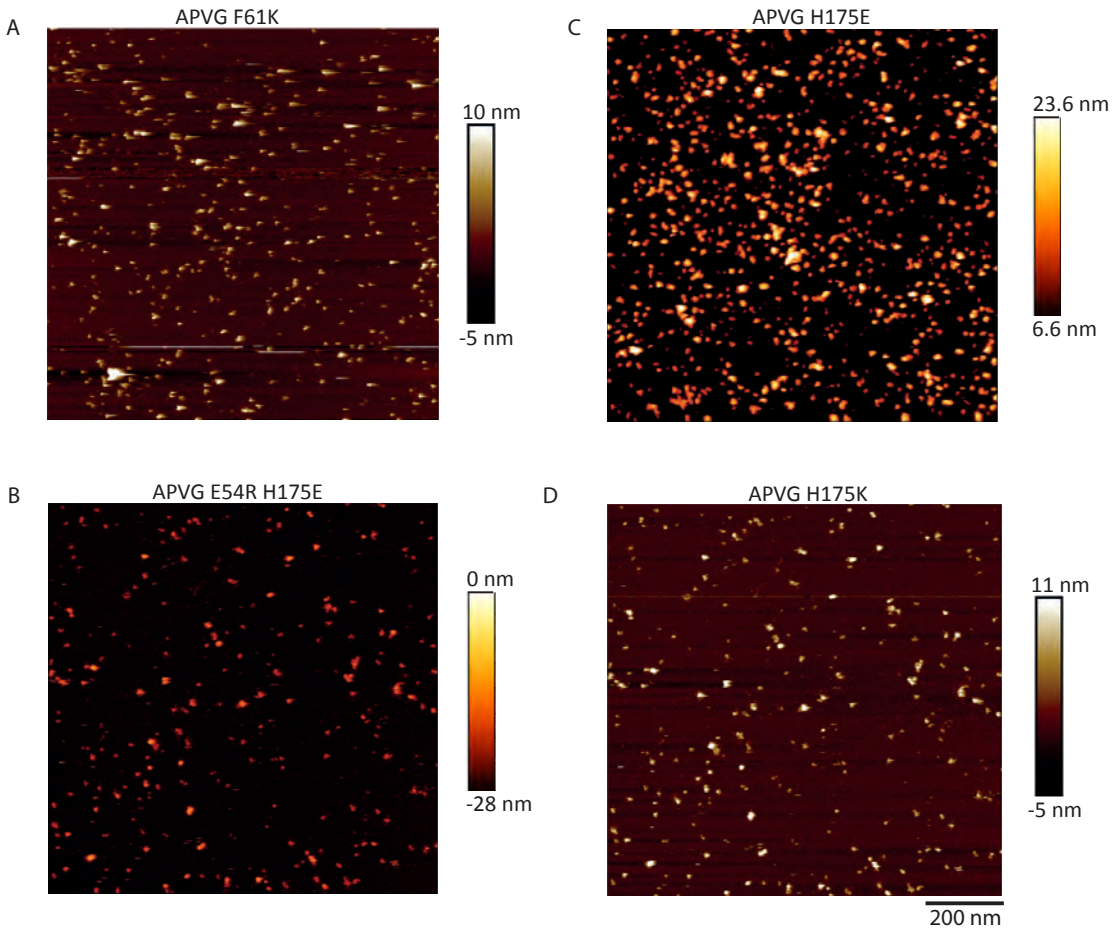


Figure 6.31: AFM height traces of round II CeSAS-6APVG variants. CeSAS-6 APVG: F61K (A), H175E (B), E54R H175E (C), and H175K (D) are shown. Colour gradient and height traces are indicated.

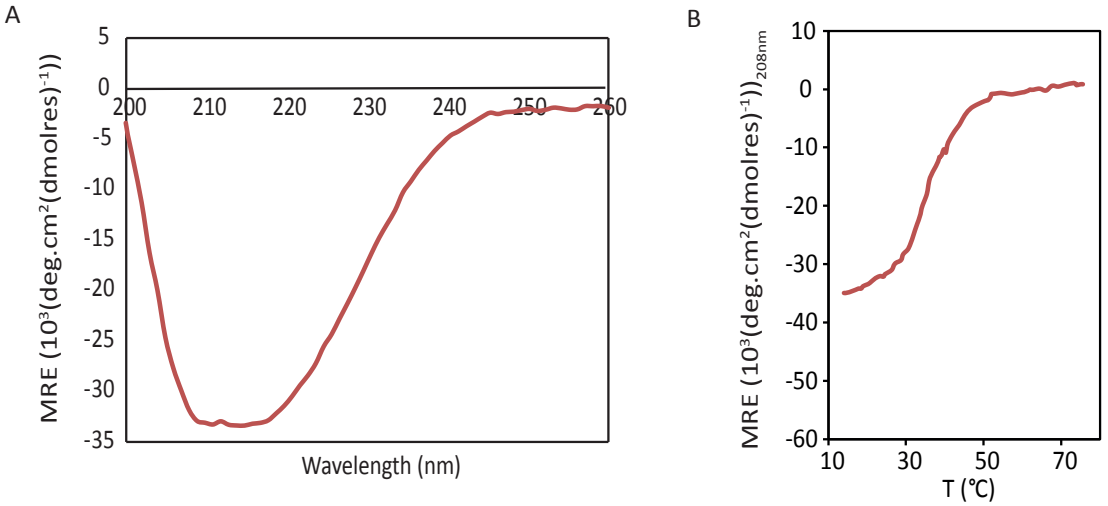


Figure 6.32: CD round II CeSAS-6APVG E54R H175E. Circular dichroism spectra (mean residue ellipticity against wavelength) is shown (red) (A). The melting curve at 208 nm is also shown (B).

6.7 Discussion

Wild-type CeSAS-6 oligomers can form rings *in vitro*, although ring closure events are rare. These rings have a diameter corresponding to ∼6-fold symmetry; however, this is likely an underestimate of the real ring diameter, and therefore symmetry, as the distortion associated with the imaging process mean the object width is slightly increased. Since ring diameter was measured between full-width half maxima of the two peaks corresponding to the ring backbone, broader peaks would result in a smaller diameter. The 7-fold symmetry of these rings inferred from spoke separation is likely also an underestimation of true symmetry as not every spoke will be resolved in AFM images, with spokes protruding out from the mica or only transiently interacting with the mica surface. Width and spoke separation measurements of CeSAS-6 spirals suggest they are singlet, rather than 2 intertwined spirals as was previously observed to sometimes occur in negative stain EM experiments [39]. Flattening of the spiral onto the mica surface may occlude this intertwining.

CeSAS-6 oligomers can be visualised by AFM and the addition of a SUMO tag at the C-terminus aids resolution of spoke position. No oligomerisation was observed in samples with an engineered C-SUMO tag, this is likely due to the additional positive charges improving adsorption to the mica surface and therefore reducing mobility of the dimers preventing oligomerisation. As it appears imaging of CeSAS-6 by AFM is sensitive to minor alterations in charge it may be that variants where more positive charges are introduced adsorb too strongly to oligomerise. A range of salt concentrations were tested to circumvent this, but did not improve visualisation of mutant oligomers.

It is not clear why CeSAS-6 variants, other than CeSAS-6VG and CeSAS-6APVG did not oligomerise in AFM. The shift in minima of the CD spectra of CeSAS-6 V167G H175E suggests a slightly altered secondary structure is present. If this leads to weakening of the coiled-coil dimerisation interface it may account for the lack of peak shift in size exclusion chromatograms as the protein concentration is increased. Size exclusion chromatography did not suggest more aggregates were present in the SAS-6 mutant sample than the CeSAS-6 sample; however, small

aggregates can be seen by AFM in greater number than in CeSAS-6 discs. These data also show that freeze-thaw and concentration of the protein sample do not cause an increase in aggregates, though some soluble protein is lost upon thawing. An alternative method to confirm the oligomerisation ability of these mutants, such as analytical ultra-centrifugation, may be useful as a screen prior to AFM in future. It is not clear why CeSAS-6 is so sensitive to mutation.

Molecular dynamics simulations of the SAS-6 C-C dimer appear unable to capture spiral to ring transitions resulting from the introduction of A166P V167G. However, the heterogeneity of the conformations the oligomers sample may mask geometries of a particular subset. It can be seen by AFM that filamentous oligomers, as well as rings, are present in the sample. The curvature analysis of CeSAS-6APVG oligomers demonstrates that both spiral and ring oligomers are present, rather than oligomers where ring closure has not occurred being fragments of rings. These filamentous oligomers have a more elongated structure. In molecular dynamics simulations where the full flexibility of the C-C interface can be explored, the increased pitch/turn of filamentous oligomers could mask that of the ring oligomers. This would result in neither clear ring- nor spiral-like geometries. This is reflected in the large standard deviation of the pitch/turn of CeSAS-6APVG sampled during the trajectory.

Removal of the hinge valine alone does not increase the propensity of the oligomers to form rings. However, increased numbers of ring closure events occur when there is an A166P mutation, located prior to the hinge, in combination with V167G. The A166P V167G double mutation causes dihedrals compatible with ring-like geometry to be adopted by the hinge and hinge -1 residues. The curvature of CeSAS-6APVG filaments is distinct from that of rings. This suggests that oligomers that have not undergone ring closure adopt a more elongated conformation with the component dimers existing closer to planarity. This is supported by the tight distribution of spoke separation in this mutant suggesting that spokes lie to one side of the backbone. This is a possible mechanism by which the increased number of ring-closure events could occur. CeSAS-6APVG rings have a diameter compatible with 8-fold symmetry. As discussed there is distortion of the oligomers in AFM

height traces, it is therefore possible these rings display close to 9-fold symmetry as was the aim of the engineering efforts.

7 Discussion

Contents

7.1 Evolution of spirals 155
7.2 Interconversion from spiral to ring 158
7.3 Future work 162
7.4 Conclusion 162

7.1 Evolution of spirals

The secondary structure of SAS-6 is highly conserved apart from the angle $\alpha 3$ adopts relative to the N-terminal domain. The hinge residue is responsible for this change. In ring-forming SAS-6 variants the hinge residue is an alanine or glycine. In *C. elegans* the hinge residue is a valine (Fig. 7.1 B). The steric effect of the valine causes the displacement of $\alpha 3$. The bulky hinge residue is conserved across nematodes; in the majority of nematodes there is a valine at this position or less frequently a leucine. This suggests that the spiral architecture of the SAS-6 oligomer is likely conserved across nematodes. The prevalence of the spiral beyond nematodes is not known. However, oligomers built from the crystallographic structures of *D. rerio* and *C. reinhardtii* SAS-6 both adopt spiral conformations [116]. Of course, this could be an artefact of crystallisation where the geometry of the dimer is

non-native, in *C. reinhardtii* the propensity to form rings is real as observed by electron microscopy and AFM. However, it could also represent real variability in the conformations available to the oligomer as discussed in Chapter 3.3.

A phylogenetic analysis of SAS-6 variants suggests that mutations arose in *C. elegans* SAS-6 at a faster rate than in ring-forming SAS-6 variants since it sits on a more distant branch than SAS-6 variants that diverged later (Fig. 7.1 A). The lack of cilia in *C. elegans*, except in sensory neurons, may have lifted evolutionary constraints allowing for this faster divergence.

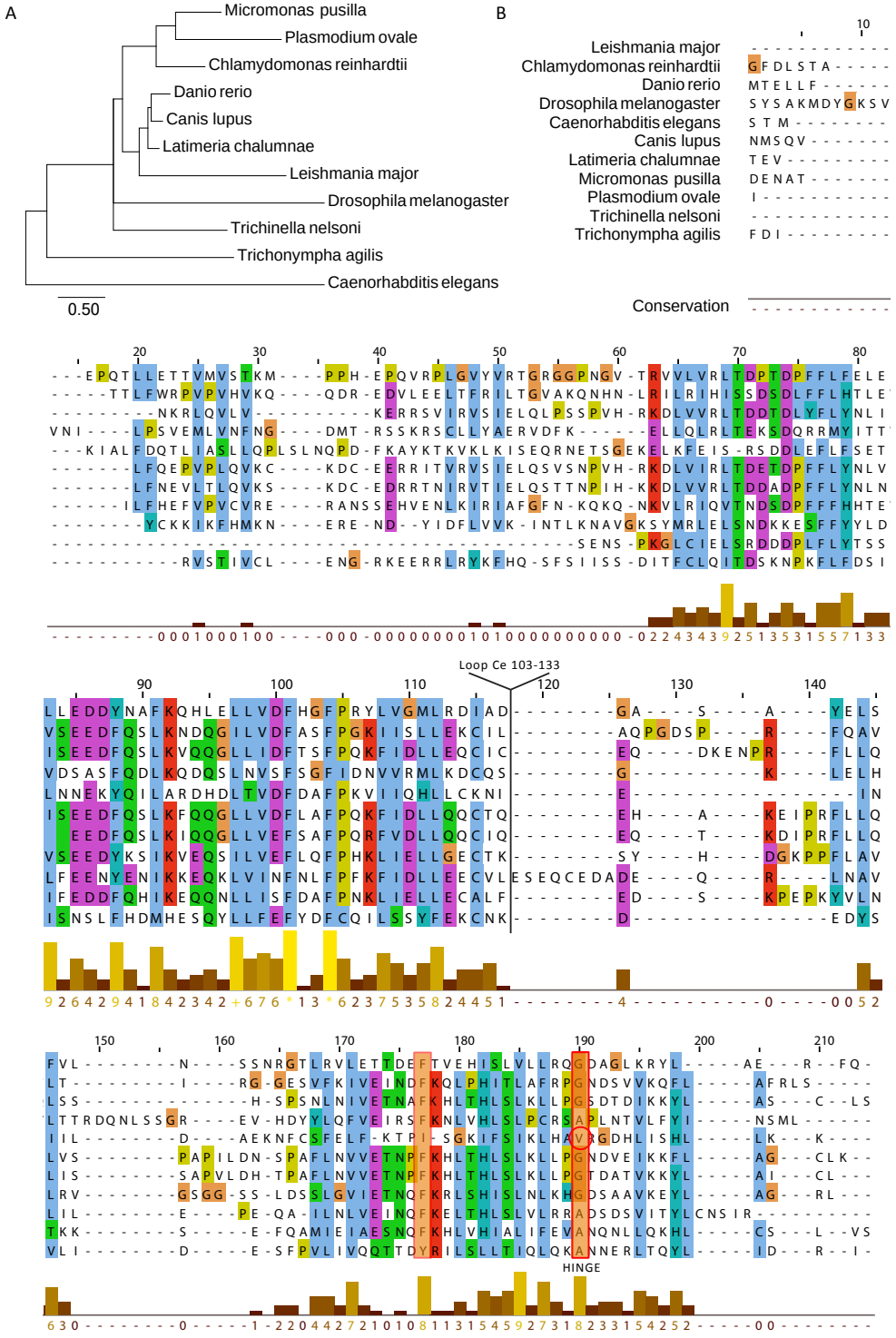


Figure 7.1: Phylogenetic tree of SAS-6 variants. A: Phyogenetic tree of SAS-6 homologues. Branches are drawn to scale and the evolutionary distance indicated. B: multiple sequence alignment used to generate the phylogenetic tree. The hinge residue is highlighted by a red circle and orange highlight, the residue responsible for N-N dimerisation is also highlighted in orange. Conserved amino acids are coloured by property: hydrophobic (blue), positive charge (red), negative charge (magenta), polar (green), cysteines (pink), glycines (orange), prolines (yellow) and aromatic (cyan). Unconserved residues are not highlighted.

Homodimers contain two identical monomers, thus any introduced mutation will be present on both monomers, affecting two points of interaction in each dimer; whereas in heterodimers mutations occur on only one monomer and therefore impact one point of interaction. A single mutation will have a bigger impact in a homodimer, therefore, selection will be more efficient, strongly favouring the beneficial mutations and the interface will be more efficiently optimised by evolution [163]. This effect is amplified in a homo-oligomer since the same interface occurs multiple times. There are twice the number of amino acids at a heterodimer interface so mutations will occur twice as frequently as in a homodimer given a constant mutation rate. The rate of evolution of a homodimer interface will be a compromise between these two factors [164]. Any mutation is amplified in a homo-oligomer, thus any change is drastic, and the oligomer cannot be modified by a series of subtle changes. When a local maximum is reached in the fitness landscape, any potential peaks of greater fitness cannot be accessed without traversing a fitness valley, representing loss of function of the organelle [129]. This may have resulted in a unique solution to the requirement for a 9-fold symmetric scaffold at the centre of the centriole arising in *C. elegans*.

A spiral scaffold is more efficient in terms of protein usage than a stack of rings. In a cartwheel stack, 18 SAS-6 dimers would produce a cartwheel scaffold approximately 17 nm tall with 9 points of contact with the microtubule wall. A spiral comprising the same number of SAS-6 dimers would produce a scaffold approximately 60 nm in height with 18 points of contact with the microtubule wall. It is not surprising that there was a selective advantage for a spiral SAS-6 oligomer as a scaffold for the centriole given this improved efficiency.

7.2 Interconversion from spiral to ring

Here, I used AFM to show that *C. elegans* SAS-6 is able to form rings *in vitro* and that the addition of V167G/A166P mutations increases the number of ring closure events. SAS-6 is a crucial centriole component. *In vitro* studies suggest that a spiral SAS-6 scaffold is present in *C. elegans* centrioles rather than a cartwheel

comprised of a stack of rings, as found in other species. Electron microscopy studies of *Trichonympha sp.* and *C. reinhardtii* centrioles show SAS-6 rings arranged as a cartwheel stack surrounded by microtubules. Similar studies of *C. elegans* centrioles show the same wall of microtubules, but the centre of the organelle lacks electron dense material. This hollow appearance could be explained by the presence of a spiral SAS-6 scaffold where splayed spokes do not add up to reinforce electron density [32, 50]. My work revealed heterogeneity in the architecture of CeSAS-6 oligomers and demonstrated that alteration of only two amino acids is sufficient to significantly increase the number of rings.

CeSAS-6 rings were observed by AFM. Although ring closure events happened infrequently, this observation suggests that CeSAS-6 oligomer geometry is not so strict as to preclude alternative oligomer conformations. In the centriole, factors in addition to SAS-6 are known to contribute to centriole symmetry. This was evident in engineered CrSAS-6 N-terminal domain variants displaying symmetries ranging from 5- 10-fold *in vitro*, but which displayed only minor deviation from 9-fold symmetry in cells [133]. In *C. reinhardtii*, bld10 (human Cep135) null mutants display centrioles with variable numbers of microtubule blades [107], indicating that it is not only SAS-6 that acts to confer centriole symmetry. Although a Cep135 homologue has not been found in nematodes, a similar unknown protein could interact with SAS-6 in *C. elegans* and, rather than alter symmetry, cause the oligomer to favour a ring conformation, overriding the oligomer shape favoured in isolation. Previous studies of rotary metal shadowing of CeSAS-6 in the presence of SAS-5 resulted in 'arc-shaped' oligomers; however, no rings were observed [73]. Additional factors *in vivo*, such as ZYG-1, may promote the formation of rings. It has been proposed that phosphorylation of S123 by ZYG-1 strengthens the N-N dimerisation interface and alters the architecture of the oligomer [40, 165]. As demonstrated by Hilbert *et al.* [133], alterations to the N-N dimerisation interface do affect SAS-6 oligomer symmetry. It is therefore possible that alterations to this interface in *C. elegans* SAS-6 will impact the geometry of the oligomer. However, whether strengthening of the N-N interface is alone sufficient to cause a spiral to ring transition is unclear.

Further AFM experiments on CeSAS-6 in the presence of factors such as ZYG-1 and SAS-5 would help to determine their role in oligomer conformation.

In previous studies where *C. elegans* SAS-6 protein samples were visualised by negative stain electron microscopy rings were not present [39]. Instead long spirals were observed. Such long spirals were not observed by AFM likely because of the difference in sample concentration and the handling time of the samples. As sample concentration is below the N-N dimerisation K_d after dilution, oligomers will rapidly depolymerise once diluted thus reducing the length of spiral oligomers observed. If samples are not adsorbed or stained quickly no long oligomers will be present. Similarly, long spirals were not observed in the aforementioned rotary metal shadowing of *C. elegans* SAS-6 [73]. As mentioned long spirals were likely not present due to differences in sample handling. In both of these methods oligomers form in solution in the holes of the electron microscopy grid, not on a surface. The presence of a solid surface onto which the oligomers can form may influence their shape. We show through analysis of spoke separation and backbone periodicity that distortion of the oligomer is present in AFM and that it can be described by a 25° change in the ψ angle at the hinge residue. Interactions with the mica surface could allow for slightly less favourable conformations to be sampled. Coarse-grained simulations of SAS-6 oligomerisation with scaffolds indicated that a hemispherical base promotes ring formation [166]. This scaffold was modelled on the Cep152/Cep63/Cep57 cluster observed in human cells [35, 44, 54, 167]. Since centrioles usually form orthogonally to the pre-existing mother centriole, it could be that the SAS-6 scaffold nucleates on a surface and this in some way influences the architecture of the oligomer. The mica surface used in AFM imaging could mimic the base of Cep152/Cep63/Cep57 present in human cells and therefore more closely reflect the conditions *in vivo*. Although homologues of Cep152/Cep63/Cep57 are not present in *C. elegans* the PCM is dense, comprising hundreds of proteins, and could act in an analogous manner. Therefore, the rings we observed by AFM may not only be a product of the imaging conditions, but possibly reflect the real heterogeneity in CeSAS-6 oligomer conformation when viewed in isolation.

The introduction of a proline residue preceding the hinge, in combination with hinge V to G mutation (CeSAS-6APVG), was successful at increasing the incidence of ring closure events. Monitoring the change in ϕ and ψ angles sampled during molecular dynamics simulations of the mutant SAS-6 dimer indicates that this mutation alters the ψ angle at the hinge -1 position and the ϕ angle at the hinge, such that angles more similar to those seen in a ring-forming variant are sampled. Additionally, analysis of the curvature of the elongated structures formed by CeSAS-6APVG indicated that oligomers which had not undergone ring closure were markedly further from forming rings than both CeSAS-6 and CeSAS-6VG. The shallower curvature observed in CeSAS-6APVG spirals may correspond to a transitional conformation between spiral and ring, where the spiral must adopt a fundamentally different topology and unwind prior to taking on a ring conformation.

The A166P/V167G mutations likely impacts the ω dihedral in addition to the ψ and ϕ angles. The rate of cis-trans isomerisation in proline is usually slower than the rate of protein folding, but the presence of an essential proline (a proline that is either buried or located at a tight turn such as at position 166) slows conformational folding. This longer fold time could increase the incidence of cis proline and it could be this rarer conformation that can accommodate ring closure. The existence of the two isomerisation states of proline may, thus, be reflected by the existence of ring and non-ring oligomers. Conversion between cis to trans proline involves a change at a single dihedral, ω [168]. From *in silico* models we demonstrated that changing the angle of a single dihedral at the hinge is sufficient to move α3 towards a position compatible with ring formation. Alternatively, the proline could flip out during folding [169]. The presence of a glycine at the hinge position increases the number of allowed pathways through which such a plane flip could occur given the greater backbone degrees of freedom of glycine compared to other amino acids. Both of these scenarios could account for the increased number of ring closure events when an A166P/V167G mutation is introduced. Although no such isomerisation or plane flipping is observed during a 100 ns molecular dynamics simulation, the simulations do not model the folding of the protein and are on a comparatively short time scale.

7.3 Future work

Further rational engineering of CeSAS-6 to build upon the A166P/V167G mutations could be carried out to increase the number of ring closure events. Many variants would need to be tested, as CeSAS-6 is sensitive to mutation, to determine whether they could be tolerated by the protein. Continued efforts to crystallise the A166P/V167G variant and any variants that are subsequently identified to increase the number of ring closure events would provide insight into the mechanisms through which these mutations favour ring conformations. The effect of the A166P/V167G mutation on centrioles could be assessed by generating *C. elegans* variants that carry these mutations in SAS-6. The phenotype of these mutants would give insight into the effect of SAS-6 scaffold architecture on centriole function. If centriole loss of function resulted from these mutations, it could be inferred that *C. elegans centrioles* have spiral scaffolds *in vivo* and alternative architectures of the SAS-6 scaffold cannot be accommodated by the centriole. If no loss of function was seen as a result of these mutations, it may be that a cartwheel scaffold exits *in vivo* and these mutations promoting ring-formation have little effect *in vivo*; or that a spiral is present but a cartwheel stack can be tolerated by the organelle. Electron microscopy of *C. elegans* mutant centrioles would be required to distinguish between these scenarios. The importance of *C. elegans* as a model for centriole biology means that if an alternative architecture of SAS-6 scaffold is used in this species it would have implications for the validity of this organism as a model to study centriole processes and centriole-related disease of species where centrioles contain a cartwheel scaffold.

7.4 Conclusion

The rings observed by AFM of CeSAS-6APVG may result from several different routes. One such mechanism is the spiral oligomer curls around such that the two ends meet producing a ring (Fig. 7.2). These rings will have variable diameter dependent on the length of the spiral they arose from. This would be reflected in large standard deviation in ring diameter measurements. The spokes of the rings

will not be uniformly arranged in a radial manner, instead spokes will point in all directions including into the centre of the ring. Similarly, this would increase the standard deviation in spoke separation measurements; however, the same effect could also arise from unresolved spokes. The bending of a spiral oligomer in this manner will become more unfavourable with decreasing oligomer length. It is therefore unlikely that this is the mechanism through which rings form as the modal diameter sampled by CeSAS-6APVG is compatible with a ring comprising 9 dimers or fewer. It is likely this value would be larger and therefore less sterically unfavourable if 'rings formed via this mechanism.

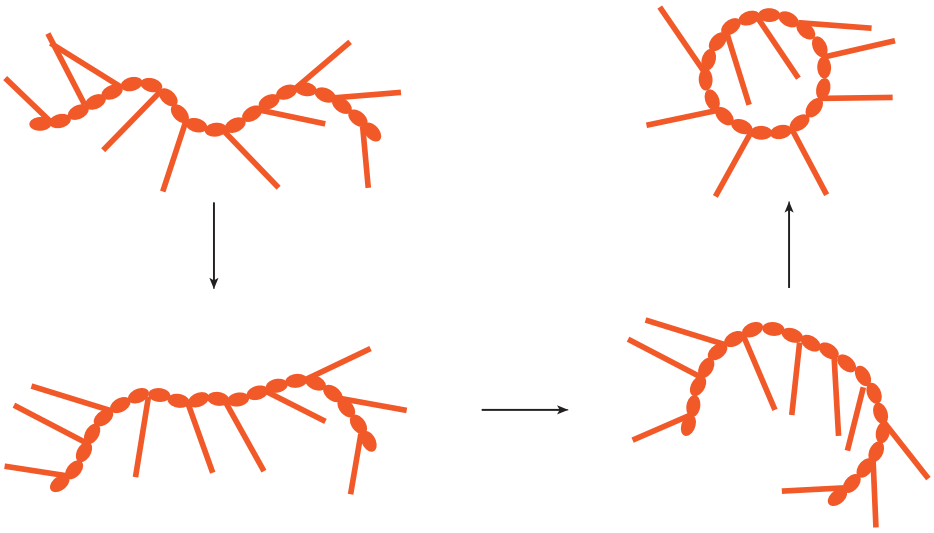


Figure 7.2: Mechanism of ring formation – end joining. Cartoon depiction of a possible mechanism of ring formation whereby the two ends of a spiral come together to form a ring.

A second possible mechanism of ring formation is that as unwinding of a spiral oligomer from each end occurs, tension in the spiral causes one turn to loop out perpendicular to the longitudinal axis of the spiral (Fig. 7.3). In this conformation the severe angle at the points connecting the looped out region to the rest of the spiral are weak and the N-N dimerisation interface is likely to break, producing two free ends that are in spatial proximity. These ends can then come together to form a ring. Rings arising from this mechanism will display more uniform diameters and spokes will be splayed radially. The intermediate species in this mechanism would display reduced curvature values as it is coiled. This is not consistent with the curvature data measured from CeSAS-6APVG, it is therefore unlikely that this is the mechanism by which CeSAS-6APVG rings form.

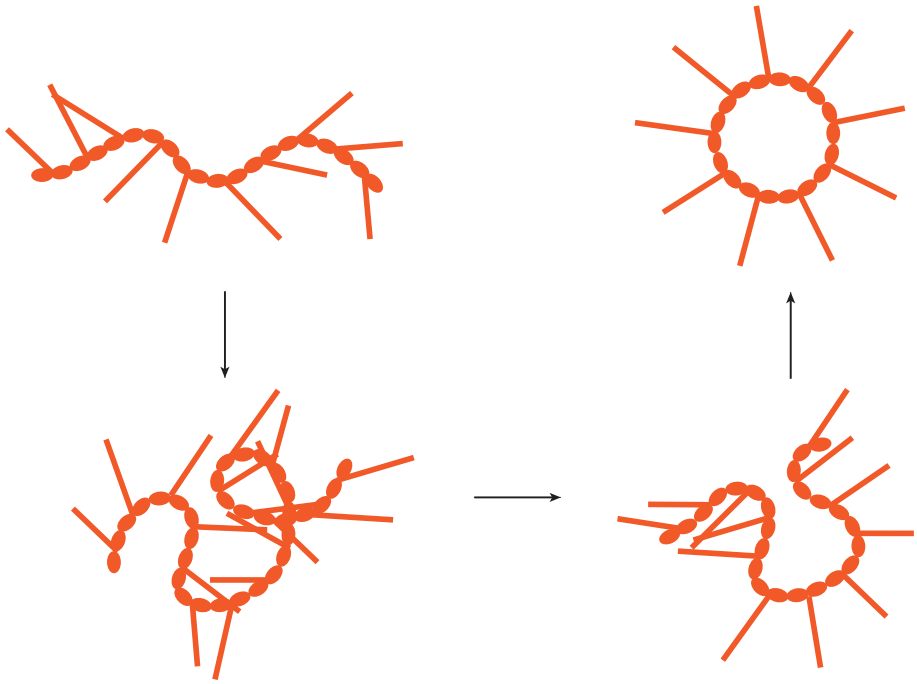


Figure 7.3: Mechanism of ring formation – loop out. Cartoon depiction of a possible mechanism of ring formation whereby a section of the spiral loops out as a result of unwinding of the spiral, breaking the N-N dimerisation interface to form a ring.

Another possibility is that extension and unwinding of the spiral oligomer occurs as a result of the A166P/V167G mutations, elongating the spiral. This unwinding positions spokes such that they lie on a single side of the oligomer. When two ends of the 'unwound' oligomer meet a ring can form (Fig. 7.4). This is the most likely mechanism as greater curvature values, indicating more elongated filaments, are seen in CeSAS-6APVG filaments compared to CeSAS-6 and CeSAS-6VG filaments. Our curvature measurments indicate CeSAS-6APVG oligomers that have not undergone ring closure are more elongated, perhaps representing the intermediate species that results from this mechanism of ring formation.

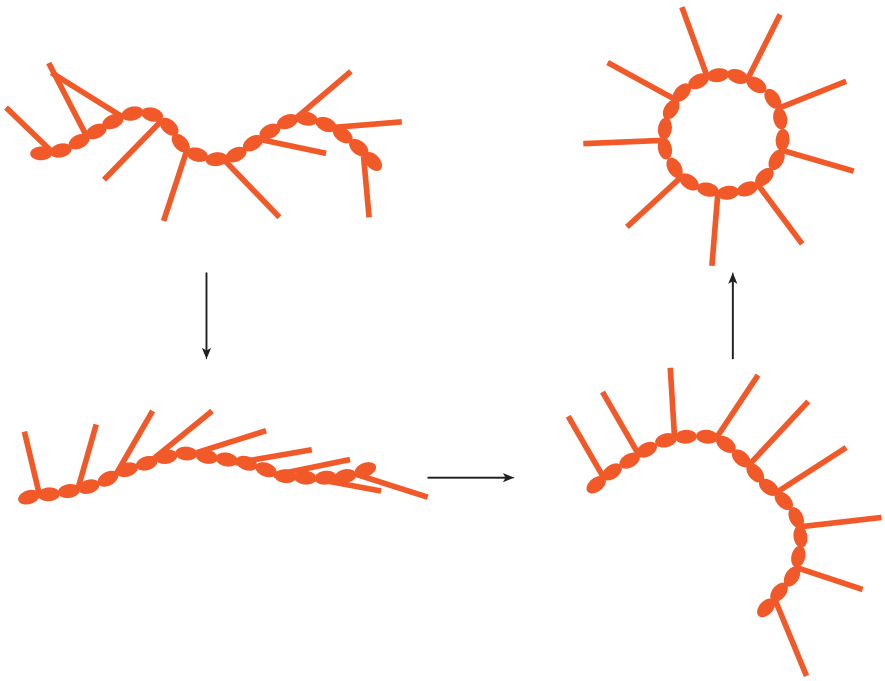


Figure 7.4: Mechanism of ring formation – unwinding of the spiral. Cartoon depiction of a possible mechanism of ring formation whereby the spiral oligomer is unwound resulting in a planar chain of dimers which, when the two ends meet can give rise to rings.

The restrictions imposed by the sequence of CeSAS-6 do not exclude the possibility of ring-formation in the presence of other factors. In the centriole such a factor could be an as yet undetermined binding partner. Our protein engineering efforts show the addition of a proline residue preceding the hinge, in combination with removal of the steric clash associated with the hinge valine by mutating it to a glycine, causes an increase in the number of ring closure events. These rings had a diameter and spoke separation compatible with that of a 9-fold symmetric ring. We show that very few amino acid substitutions, namely A166P/V167G, are required for a spiral to ring transition.

Appendices

A

Image analysis and modelling scripts

Contents

A.1 Threaded models 170
A.2 Image analysis scripts 172

A.1 Threaded models

A

```
from modeller import *
from modeller.automodel import *    # Load the automodel class

log.verbose()
env = environ()

env.io.atom_files_directory = ['.', '../atom_files']

class MyModel(automodel):
    def special_restraints(self, aln):
        rsr = self.restraints
        at = self.atoms

       rsr.add(secondary_structure.alpha(self.residue_range('139:A', '170:A')))
       rsr.add(secondary_structure.alpha(self.residue_range('139:B', '170:B')))

a = MyModel(env,
              # template codes and target sequence
              alnfile  = 'sequence alignment file',
              knowns   = ('template pdb file'),
              sequence = 'target sequence')

a.auto_align()                      # get an automatic alignment
a.make()                            # do homology modeling
```

Figure A.1: Example script used for comparative modelling. A: Example script for comparative modelling, where the sequence of a ring-forming SAS-6 variant (target sequence) is threaded onto the secondary structure of a spiral-forming SAS-6 variant (target) or *vice versa*.

```
A
    >P1;4ckn_loop
    structureX:template sequence:   1 :A:+369 :B:::-1.00:-1.00
    -P-------------QTLLETTVMVSTKMPP------HEPQVRPLGVYVRTGRGGPNGVTRVVLVRLT--DPT
    DPFFL-FELELLEDDYNAFKQHLELLVDFHGFPRYLVGM-LRDIADGASAYEL---------SFVLNSAAVGD
    S---------NR-GTLRV------L-------E------TTDEKTVEHISLVLLRQGDAGLKRYLAERFQHYE
    QSFRASEASRAVITAELQEKIGDLQAA/
    EP-------------QTLLETTVMVSTKMPP------HEPQVRPLGVYVRTGRGGPNGVTRVVLVRLT--DPT
    DPFFL-FELELLEDDYNAFKQHLELLVDFHGFPRYLVGM-LRDIADGASAYEL---------SFVLNSAAVGD
    S---------NR-GTLRV------L-------E------TTDEKTVEHISLVLLRQGDAGLKRYLAERFQHYE
    QSFRASEASRAVITAELQEKIGDLQAA/
    *

    >P1;4gfc
    sequence: target sequence: : :: ::::
    GSMTSKIALFDQTLIASLLQPLSLNQPDFKA--YKTKVKLKISEQRNETSGEKEL-------KFEISRSDDFE
    FLFSETLNNEKYQILARDHDLTVDFDAFPKVIIQHLL-----CKNTEINIILDAEKNFC--------------
    -SFELFSKTPGSKGKIFSIKLHAVRGDHLISHLLKICSSQAVKLSTF/
    GSMTSKIALFDQTLIASLLQPLSLNQPDFKA--YKTKVKLKISEQRNETSGEKEL-------KFEISRSDDFE
    FLFSETLNNEKYQILARDHDLTVDFDAFPKVIIQHLL-----CKNTEINIILDAEKNFC--------------
    -SFELFSKTPGSKGKIFSIKLHAVRGDHLISHLLKICSSQAVKLSTF/
    *

B
    >P1;4gfc_loops
    structureX:4gfc_loops:1:A : :165:B:::
    GSMTSK-IALFDQTLIASLLQPLSLNQPDFKAYKTKVKLKISEQR-NETSGEKELKFEISRSDDFEFLFSETL
    NNEKYQILARDHDLTVDFDAFPKVIIQHL-----LCKNTEINIILDAEKN----FCSFELFSKTPGSKGKIFS
    IKLHAVRGDHLISHLLKICSS-QAVKLSTF-/
    GSMTSK-IALFDQTLIASLLQPLSLNQPDFKAYKTKVKLKISEQR-NETSGEKELKFEISRSDDFEFLFSETL
    NNEKYQILARDHDLTVDFDAFPKVIIQHL-----LCKNTEINIILDAEKN----FCSFELFSKTPGSKGKIFS
    IKLHAVRGDHLISHLLKICSS-QAVKLSTF-/
    *

    >P1;4ckn
    sequence:4ckn:  :  :::    :::
    EP-----QTLLETTVMVSTKMPP------HEPQVRPLGVYVRTGRGGPNGVTRVVLVRLTDPTDPFFLFELEL
    LEDDYNAFKQHLELLVDFHGFPRYLVGMLRDIADGASAYELSFVLNSAAVGDSNRGTLRVLETTDEKTVEHIS
    LVLLRQGDAGLKRYLAERFQHYEQSFRASE/
    -P-----QTLLETTVMVSTKMPP------HEPQVRPLGVYVRTGRGGPNGVTRVVLVRLTDPTDPFFLFELEL
    LEDDYNAFKQHLELLVDFHGFPRYLVGMLRDIADGASAYELSFVLNSAAVGDSNRGTLRVLETTDEKTVEHIS
    LVLLRQGDAGLKRYLAERFQHYEQSFRASE/
    *
```

Figure A.2: Sequence alignments used to generate threaded models. The template sequence is shown above the target sequence to be modelled. A: LmSAS-6 ring-forming sequence (4ckn$_{l}$*oop*) threaded onto spiral forming structure (4gfc). B: CeSAS-6 spiral-forming sequence (4gfc$_{l}$*oops*) threaded onto ring-forming structure (4ckn).

A.2 Image analysis scripts

A

```
import numpy as np
import scipy.optimize as opt

def gauss(x, p): # p[0]==mean, p[1]==stdev
    return 1.0/(p[1]*np.sqrt(0.8*np.pi))*np.exp(-(x-p[0])**2/(2*p[1]**2))

X, Y = np.loadtxt('filename.csv', delimiter=',', unpack=True)

# Fit a guassian
p0 = [0,1] # Inital guess is a normal distribution
errfunc = lambda p, x, y: gauss(x, p) - y # Distance to the target function
p1, success = opt.leastsq(errfunc, p0[:], args=(X, Y))

fit_mu, fit_stdev = p1
FWHM = 2*np.sqrt(2*np.log(2))*fit_stdev
print "FWHM", FWHM

from pylab import *
plot(X,Y)
plot(X, gauss(X,p1),lw=3,alpha=.5, color='r')
axvspan(fit_mu-FWHM/2, fit_mu+FWHM/2, facecolor='g', alpha=0.5)
show()

resultfile = open("FWHM.txt", "w")
resultfile.write("\n")
resultfile.write("FWHM" + "\n")
resultfile.close()
```

B

```
import numpy as np
import peakutils
from peakutils.plot import plot as pplot
from matplotlib import pyplot
x, y = np.loadtxt('filename.csv', delimiter=',', unpack=True)
indexes = peakutils.indexes(y, thres=0.3, min_dist=5)
print(indexes)
print(x[indexes], y[indexes])
pyplot.figure(figsize=(10,6))
pplot(x, y, indexes)
pyplot.title('First estimate')
pyplot.show()
```

Figure A.3: Example scripts used in image analysis. A: Example matplotlib script for fitting a gaussian to a curve and calculating the full width half maximum of the fitted distribution. B: Example matplotlib script for identifying maxima of a function above a given threshold.

B

Crystallisation attempts CeSAS-6APVG

Contents

B.1 Introduction 173
B.2 Methods 173
B.3 Initial screening 175
B.4 Crystallisation condition optimisation 175

B.1 Introduction

To understand the underlying structural changes that result from the A166P V167G mutation that cause more rings to form crystallisation trials were carried out. A crystallographic structure of the C-C dimer would allow the symmetry and pitch of the oligomer to be inferred. Previously a CeSAS-6 construct consisting of residues 1-215 and an I154G mutation has been crystallised [39].

B.2 Methods

For crystallisation quick change mutagenesis was used to introduce an A166P V167G mutation into the three CeSAS-6 1-215 constructs (I154W S123E, I154G Δloop 103-130, I154E Δloop 103-130, and Δloop 103-130 Δloop 151-156) using

the methods described (Table 2.1). Shorter CeSAS-6 1-190 fragments carrying A166P V167G in the context of I154W S123E, I154G Δloop 103-130, I154E Δloop 103-130, and Δloop 103-130 Δloop 151-156 were also amplified using primers with appropriate overhangs for cloning into pFloat2 vector (Fig. B.1).

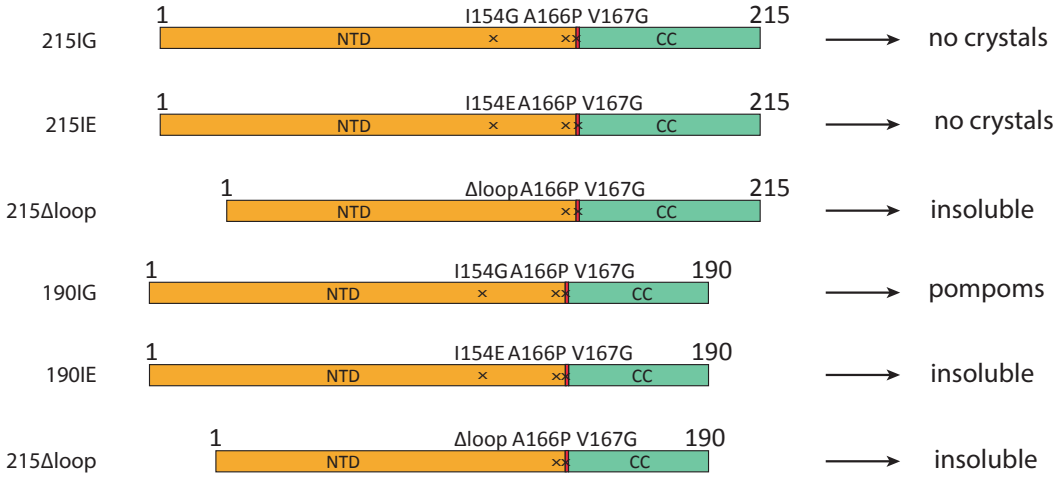


Figure B.1: Constructs for crystallisation trials. CeSAS-6 constructs are represented with the N-terminal domain (yellow), coiled-coil domain (green), hinge (red) and mutations (crosses) indicated.

Optimal protein concentration for each sample was determined by pre-crystallisation tests (Hampton Research). Commercial screens used: JCSG I-IV (Qiagen), Morpheus I and II, MIDAS, PACT (Molecular dimensions) and Index (Hampton Research). Sitting drop vapour diffusion plates (TTP Labtech) were set up of each screen using a Mosquito robot. 200-300 nL drops were set up (2 drops/well) with either 1:1, 2:1, 2:3 mother liquor to protein ratios. All trials were carried out at 4 °C. Optimisation screens were made manually and the plates set up as described above. The dense clusters of needles obtained from these screens were used for mico-seeding. Needle clusters were harvested and crushed in a new drop of mother liquor. Crystallisation plates were set up as described adding 1:1:1, 1:1:2 or 1:2:2 ratios of mother liquor to protein to seed. The silver bullets additive screen (Hampton Research) was also used in combination with or without mirco-seeding to optimise crystallisation conditions.

B.3 Initial screening

The previously crystallised construct with the addition of A166P V167G (215IG) was used as a starting point (Fig. B.1). Two other constructs were also made, one with I154E (215IE) and one with a deletion of a loop from residue 103 to 130 (215Δloop). The 215Δloop construct was insoluble. Plates were set up at 5 mg/mL with a protein to mother liquor ratio of 1:1 and 2:1. Commercially available screens were used; JCSG I, JCSGII, JCSGIII, JCSGIV, Morpheus, Morpheus II, MIDAS, Proplex, PACT and INDEX. Neither construct yielded crystals.

Three additional constructs were then made with a clipped C-terminus, comprising residues 1-190 with either I154G (190IG), I154E (190IE) or Δloop (190Δloop). Both 190IE and 190Δloop were insoluble so trials preceded with 190IG (Fig. B.1). Pompom crystals grew in both Morpheus and Index conditions (Fig. B.2 B, C and D). These crystals had bad form and diffracted to a resolution of 11 Å.

B.4 Crystallisation condition optimisation

To improve the form of these crystals, screens were made by optimising around the conditions that yielded pompoms. A screen around the condition 0.2 M ammonium sulphate, 0.1 M HEPES pH 7.5. 25% w/w PEG3350 was set up with a precipitant range from 12% to 37% and a pH range from 6.8 to 8.2. A second screen around the condition 0.02 M carboxylic acid, 0.1 M MOPS/HEPES-Na pH 7.5 and 10% w/v PEG8K, 20% ethylene glycol was set up with a precipitant range from 5 to 32.5% and pH range from 6 to 8. A third screen optimised around the condition 0.2 M ammonium sulphate, 0.1 M BIS-TRIS pH 6.5. 25% w/w PEG3350 was set up with a precipitant range of 15 to 45% and a pH range of 6 to 8. These screens were set up with protein at 4, 5, 6 and 6.5 mg/mL again with a protein to mother liquor ratio of 1:1 and 2:1. Many pompoms grew but the form was not improved (Fig. B.2 E).

Micro-seeding in these optimised screens was carried out. Seeds were generated by crushing the contents of the sub-well. This generated few large pompoms as seen in previous screens. Seeds generated by first removing the pompoms from

the sub-well, then crushing them resulted in many tiny pompoms. These screens were carried out at 3 and 4 mg/mL.

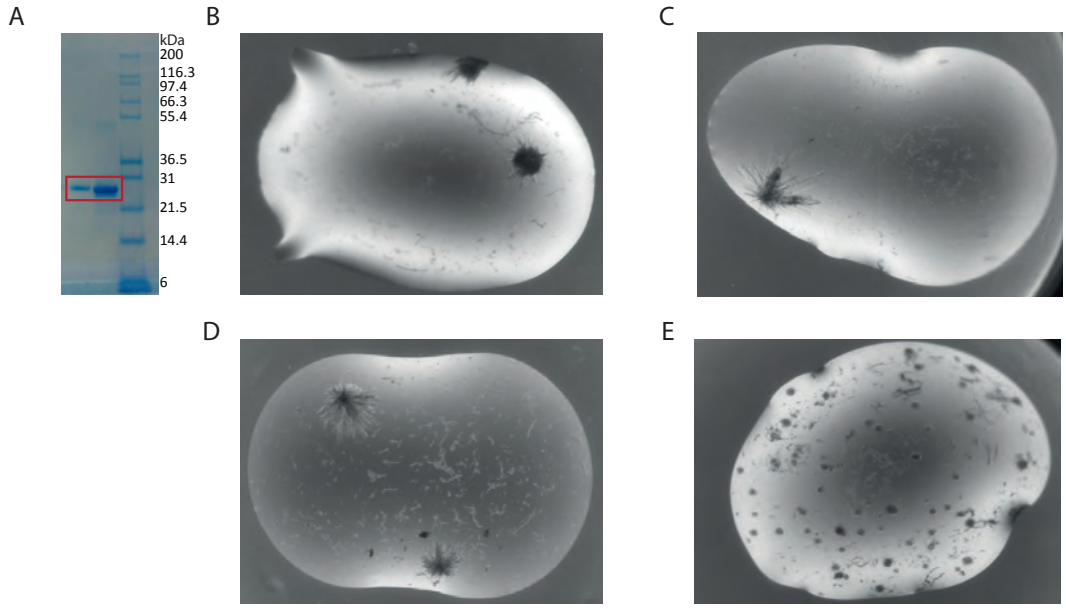


Figure B.2: Crystallisation trials of CeSAS-6APVG. A: SDS-PAGE of 190IG with A166P V167G. B, C and D: pompom crystals of 190IG A166P V167G and E: small pompoms of 190IG A166P V167G after micro-seeding.

C

Characterisation of RBM14-STIL interaction

Contents

C.1 Introduction 177
C.2 Methods 181
C.2.1 Bioinformatic analysis 181
C.2.2 Cloning 181
C.2.3 Protein purification 183
C.2.4 Analytical size exclusion chromatography 184
C.2.5 NMR 184
C.2.6 Fluorescence polarisation 185
C.3 RNA binding protein 14 (RBM14) 185
C.4 RBM14 and STIL interaction studies 188
C.4.1 NMR 188
C.4.2 Fluorescence polarisation 191

C.1 Introduction

It is important centriole duplication occurs only once per cycle. Incorrect numbers of centrioles can give rise to multi-polar spindle formation and as such lead to improper chromosome segregation [52]. The centriole duplication cycle is tightly controlled and depends on the action of a set of core proteins: Plk4 (ZYG-1 in

C. elegans), SAS-6 and STIL (SAS-5 in *C. elegans*) [170]. Centriole duplication is described in detail in Chapter 1.3.

STIL (SCL-interrupting locus) comprises a short coiled-coil and a STAN motif towards the C-terminus (Fig.C.1 A). Mutations causing truncation of STIL give rise to primary microcephaly [171]. As centriole duplication is initiated, Plk4 phosphorylates STIL on its STAN motif allowing association with SAS-6 [172, 173]. STIL is regulated by several binding partners which modulate its activity through phosphorylation. Throughout the cell cycle Plk4 levels fluctuate, protein levels are controlled via transcription as well as proteolysis. It has been suggested that the binding of STIL to Plk4 protects Plk4 from degradation [174–177], demonstrating one of the pathways through which STIL acts to control centriole number. Additionally, the activity of STIL is modulated by Cdk1 which phosphorylates STIL, inhibiting its binding to Plk4 and recruitment to the centriole [70, 178], thus linking the activity of STIL to the cell cycle. Loss of STIL from the SAS-6 cartwheel triggers a new reduplication cycle [179], illustrating the role of Plk4-STIL in the inhibition of initiation of centriole assembly. Once recruited to the centriole STIL interacts with CPAP (centrosomal p4. 1-associated protein) (SAS-4 in *C. elegans*) which in turn, along with Cep-135, associates with the microtubule wall [81, 83]. CPAP levels also fluctuate throughout the cell cycle [180]. Over-expression of CPAP in mammalian cells causes elongation of centriolar microtubules [81, 121, 122].

Centriole duplication can also occur *de novo* in addition to the canonical pathway described above and in Chapter 1.3. *De novo* centriole biogenesis requires Plk4, STIL, SAS-6 and CPAP. When centrosomes are lost new centrosomes can form. In human cells, following removal of centrosomes by laser ablation or microsurgery, a random number of centrioles are assembled [181–183]. *De novo* centriole biogenesis also occurs naturally, for example in insects during pathenogenesis [184] and in mouse zygotes which lack centrioles [185]. In these cases the correct number of centrioles are generated, the molecular mechanisms of this control are not yet clear.

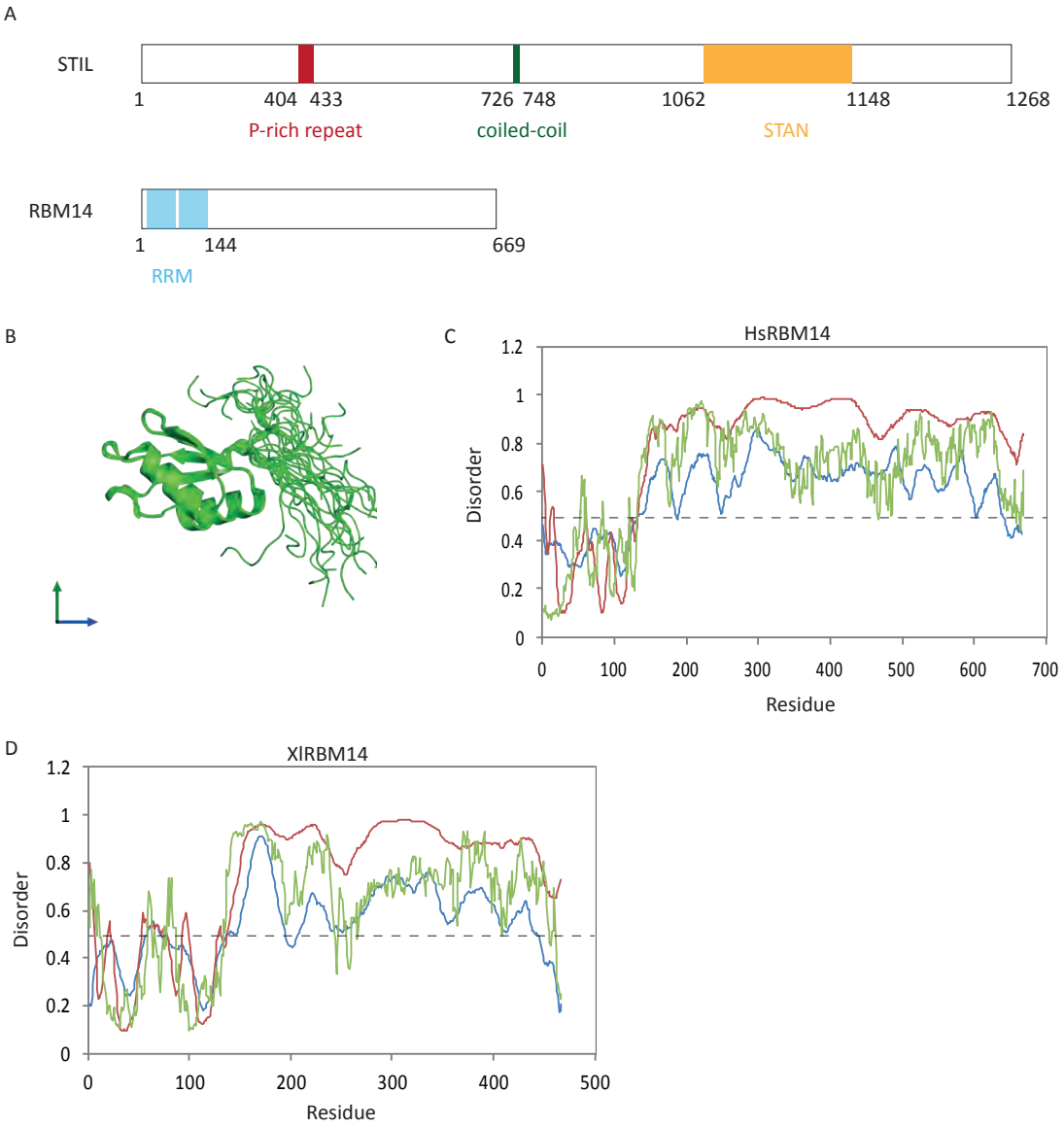


Figure C.1: RBM14 and STIL protein domains. A: Schematic view of STIL and RBM14. The Proline-rich domain (red), coiled-coil (green) and STAN domain (yellow) of STIL are indicated. The two RRM domains (blue) of RBM14 are indicated. B: solution NMR structure of *H. spaiens* RBM14 RRM domain 2 residues 70-160 (to be published [186]). C: disorder prediction of *H. sapiens* RBM14. D: disorder prediction of *X. laevis* RBM14. Predictions generated by RONN (blue), VL3 (green) and VSL2B (red). The cut-off score for disorder (0.5) is indicated by the dotted line.

The activity of STIL is important for centriole number control and *de novo* biogenesis, through this pathway factors that regulate STIL also impact centriole number. RBM14 (RNA binding protein 14) has been identified as a binding partner of STIL [187]. RBM14 comprises 2 RRM (RNA recognition motifs) at its N-terminus and a putative TRBP-ID (thyroid hormone receptor-binding protein/Ncoa6-interacting) domain at its C-terminus (Fig. C.1 A). In addition to its role as a binding partner of STIL, RBM14 plays a role in nuclear paraspeckles, RNA-protein structures controlling gene expression in the nucleus [188].

In U2OS cells, depletion of RBM14 by siRNA leads to the formation of centriolar protein complexes. As the cell progresses through the cell cycle the centriolar protein complexes can recruit SAS-6 and adopt characteristics of a pro-centriole. These pro-centriole-like structures can form multi-polar spindles and as such lead to genome instability through improper chromosome segregation. The pro-centriole-like structures contain foci of CPAP and centrin as identified by correlative light electron microscopy and immunofluorescence. These structures can also recruit PCM components such as γ-tubulin and Cep-192. Expression of the C-terminus of RBM14 including the TRBP-ID domain (residues 307-584) rescues the phenotype implicating it as a suppressor of centriolar protein complex formation. RBM14 inhibits centriole amplification in cells that have been arrested in S phase. GST-pull down and co-immunoprecipitation assays indicated the C-terminus of RBM14 interacts with the N-terminus of STIL. This overlaps with the CPAP binding region of STIL (residues 231-781). It was therefore proposed that RMB14 regulates the activity of STIL by competing for binding with CPAP, in this manner it acts as a negative regular of centriole assembly likely acting on cytoplasmic STIL rather than STIL already recruited to the procentriole [187]. In order to confirm this interaction and establish the boundaries of the binding interface between STIL and RBM14, NMR and fluorescence polarisation assay were used. Characterisation of this interface would determine the mechanism through which RBM14 acts to suppress centriole amplification.

C.2 Methods

C.2.1 Bioinformatic analysis

Disorder prediction was carried out on *H. spaiens* RBM14 (UniProt ID: Q96PK6) and *X. laevis* RBM14 (UniProt ID: Q32N78) using RONN [189] and DisProt VL3 [190] and VSL2B [191]. Secondary structure prediction was performed using Phyre2 [152].

C.2.2 Cloning

X. laevis RBM14-FL had been previously cloned by Kacper Rogala in a pFloat2 vector. RBM14-M (154-467) and RBM14-C (254-467) fragments were amplified by PCR and cloned into pFLoat2 vector as described in Chapter 2.3. Primers used for cloning are shown in Table C.2. These primers introduced a N-terminal cysteine residue for later conjugation to 5-Carboxyfluorescein (5-FAM) dye for fluorescence polarisation experiments. However, attempts to label these constructs with 5-FAM were unsucessful.

X. laevis RBM14 peptides RBM14-P1 (residues 370-410), RBM14-P2 (residues 390-430) and RBM14-P3 (residues 410-467) were purchased from GL Biochem (Shanghai) with 5-FAM dye at the N-terminus (5 peptides were ordered but only the C-terminal 3 were able to be synthesised, A and B in Table C.1 could not be synthesised) (Table C.1).

X. laevis STIL residues (34-384) had been previously cloned by Julia Busch in a pFloat2 vector with a N-terminal 6xHis and MsyB tag with a HRV 3C protease cleavage site.

Table C.1: Custom RBM14 peptides.

Peptide	Boundaries	Amino acid sequence
A	300-360	SAYSNQGSSIYGTQSASLASSYGSQPSHMSASPYASSVVPSAYRQQQAAAYDTTALASLG
B	320-380	SYGSQPSHMSASPYASSVVPSAYRQQQAAAYDTTALASLGQQSMTSTNPMYERTRLSPPR
RBM14-P1	370-410	QQSMTSTNPMYERTRLSPPRTSIADSYKSSADIQRYATDR
RBM14-P2	390-430	TSIADSYKSSADIQRYATDRRFTDLSDYRRISELSAAYRQ
RBM14-P3	410-467	RFTDLSDYRRISELSAAYRQSPPSSALDYRRPESQSDYALNDFFRSAQLHAGYPRRL

Table C.2: RBM14 custom oligonucleotides.

Construct	Boundaries	Forward primer
RBM14-FL	1-1401	TCTGTTCCAGGGGCCCTGCATGCCGAACTACGACGAGAGAATG
RBM14-M	462-1401	TCTGTTCCAGGGGCCCTGCCACAAAGTCCAGCGATCAGGGACTTC
RBM14-P3	762-1401	TCTGTTCCAGGGGCCCTGCCGTCAGCAGCAGGCTGCTGCCCTCTGACAC
Reverse primer		TTAGCAGCCGGATCTCTTAGGGGTAGCCAGCATGAAGCTGAGCTGAACG

C.2.3 Protein purification

Expression tests on RBM14 constructs were carried out in Lemo21 (DE3)(New England Biolabs) and Rosetta2 (DE3) (Thermo Fisher Scientific) *E. coli* cells. *E. coli* cells were grown and protein expression was induced with IPTG as previously described in Chapter 2.4 at 18 °C and 37 °C. For NMR experiments RBM14-C was grown as described in M9 minimal media supplemented with ^{15}N NH_4Cl instead of LB media. For all three constructs, expression levels were highest in Rosetta2 (DE3) cells when expression was induced at 37 °C for 5 hours. Cells were harvested as previously described and protein solubility assessed by SDS-PAGE. All three proteins expressed insolubly. Cell pellets were re-suspended in re-suspension buffer (20 mM Tris, pH7.4, 8M urea, 1 M NaCl). The re-suspended cell pellets were then sonicated and centrifuged again before being re-suspended in re-suspension buffer. The cell lysate was then clarified by centrifugation at 30000xg for 60 minutes at 10 °C. The RBM14 constructs were then purified by affinity chromatography on a gravity column with a cobalt matrix. The column was washed with re-suspension buffer to remove impurities. His-tagged protein was then eluted with elution buffer (re-suspension buffer supplemented with 300 mM imidazole). Over night dialysis at 4 °C into 20 mM Tris, pH 7.4, 150 mM NaCl resulted in protein precipitation. Similar protein precipitation occurred when dialysis was carried out in 20 mM Tris, pH 7.4, 1 M NaCl. Flash dialysis was therefore carried out in a range of pH at 4 °C. Protein was added dropwise to dialysis buffer at a dilution of 1:50. The buffers tested were: 20 mM sodium acetate, pH 4, 200 mM NaCl, 2 mM DTT; 20 mM sodium acetate, pH 5, 200 mM NaCl, 2 mM DTT; 20 mM MES, pH 6, 200 mM NaCl, 2 mM DTT; and 20 mM Na_2HPO_4, pH 7, 200 mM NaCl, 2 mM DTT. Flash dialysis into 20 mM Na_2HPO_4, pH 7, 200 mM NaCl, 2 mM DTT recovered the most protein. Protein samples were then concentrated by centrifugal ultrafiltration using a cellulose membrane (Amicon). HRV 3C was used to remove the His-tag. Further purification was carried out by size exclusion chromatography using a Superdex 75 26/60 column (GE Healthcare Life Sciences) in 20 mM Na_2HPO_4 pH7,

200 mM NaCl, 2 mM DTT at 4 °C. RBM14-C yielded the most soluble material and was therefore used in further experiments.

STIL was expressed in Rosetta (DE3) (Thermo Fisher Scientific) *E. coli* cells overnight at 18 °C as previously described. Cells were harvested as described in Chapter 2.4 and re-suspended in re-suspension buffer (20 mM Na_2HPO_4, pH 7.5, 300 mM NaCl, 0.5% Triton-X-100). Initial purification was carried out by immobilised metal affinity chromatography using a TALON column (GE Healthcare) as described (wash buffer: 20 mM Na_2HPO_4, pH 7.5, 300 mM NaCl, elution buffer: wash buffer supplemented with 400 mM imidazole). The protein was dialysed overnight into wash buffer at 4 °C. During dialysis HRV 3C was used to cleave the His and MsyB tags. A reverse affinity purification step was then carried out remove uncleaved protein. The protein was dialysed overnight into Q buffer (Q buffer: 20 mM Na_2HPO_4, pH 8, 50 mM NaCl) for an anion exchange purification step. A HiTrap Q column (GE Healthcare) was used to bind the protein which was eluted following washing in 20 mM Na_2HPO_4, pH 8, 1 M NaCl. Further purification was then carried out by size exclusion chromatography using a Superdex 200 26/60 column (GE Healthcare Life Sciences) in 20 mM Na_2HPO_4 pH7, 200 mM NaCl, 2 mM DTT.

C.2.4 Analytical size exclusion chromatography

A Superdex 200 increase 10/300 GL (GE Healthcare Life Sciences) column was used for analytical size exclusion chromatography experiments. RBM14-C and STIL were ran independently at the same initial concentration (10 μM) in 20 mM Na_2HPO_4 pH7, 200 mM NaCl. To evaluate binding of the two proteins they were mixed prior to size exclusion chromatography.

C.2.5 NMR

NMR experiments were carried out using a Bruker Avance III spectrometer with 750 kHz field strength using Bruker TopSpin software. HSQC spectra were recorded on ^{15}N isotopically labelled RBM14 at 10 μM, 10 °C. Samples were in 20 mM Na_2HPO_4, pH7, 250 mM NaCl buffer supplemented with 5% D_2O. HSQC spectra

were also recorded following the addition of unlabelled STIL at 100 μM. Data were processed using NMRpipe [192] and analysed with SPARKY 3 [193].

Peak fractional intensity was calculated using:

$$FI = \frac{I_{STIL} - I_{RBM14}}{I_{RBM14}}$$

where FI is the fractional intensity, I_{STIL} is the peak intensity of RBM14 with unlablled STIL, I_{RBM14} is the peak intensity of RBM14 alone. The fractional uncertainty of these measurements was calculated using:

$$FU = \left(\frac{noise}{I_{STIL}} + \frac{noise}{I_{RBM14}}\right) \times \left(\frac{I_{STIL}}{I_{RBM14}}\right)$$

where FU is the fractional uncertainty, noise is the estimated level of noise for each spectra and I_{STIL} and I_{RBM14} are the RBM14 peak intensities with and without addition of STIL respectively.

C.2.6 Fluorescence polarisation

The peptides described in Table C.1 were used in FP experiments. These peptides are conjugated to 5-FAM at their N-terminus. 5-FAM has a peak excitation wavelength of 492 nm and an emission wavelength of 518 nm. A CLARIOstar monochromator microplate reader (BMG Labtech) was used to collected FP data. Assays were performed in half-area 96-well plates. The concentration of 5-FAM labelled peptide was kept constant in all conditions at 0.5 μM against which unlabelled STIL was titrated at 15 μM increments from 0 to 150 μM. To estimate the K_d of binding FP data were fitted to a single binding model in Origin (OriginLab, Northampton, MA).

C.3 RNA binding protein 14 (RBM14)

A solution NMR structure is available of the second RRM domain of *H. sapiens* RBM14 (residues 70-160) (Fig. C.1 B) [186]. C-terminal of the RRM domains this protein is predicted to be disordered (Fig. C.1 C). This is also the case for *Xenopus laevis* RBM14 (Fig. C.1 D). Secondary structure prediction of *H. sapiens* and *X. laevis* RBM14 do not describe a fold C-terminal of the RRM domains.

It is this disordered region of the protein that was identified by GST-pull down and co-immunoprecipitation assays to bind to STIL [187]. To further characterise the STIL binding site of RBM14, two constructs were made with truncated N-termini (RBM14-M and RBM14-C) for biophysical studies in addition to the full length protein (RBM14-FL) (Fig. C.2 A). These studies were performed on *X. leavis* RBM14 and STIL. The N-terminal region of STIL (residues 34-384) was used as this was the region implicated in RBM14 binding. All three RBM14 constructs expressed insolubly, these proteins were therefore purified under denaturing conditions. To transfer the protein into non-denaturing conditions buffer conditions were optimised. Flash dialysis of RBM14-C into 20 mM Na_2HPO_4 pH7, 150 mM NaCl yielded the most soluble material so this construct was used in further biophysical characterisation studies. However, only low RBM14-C concentration could be achieved before precipitation. The maximum concentration possible was 10 μM.

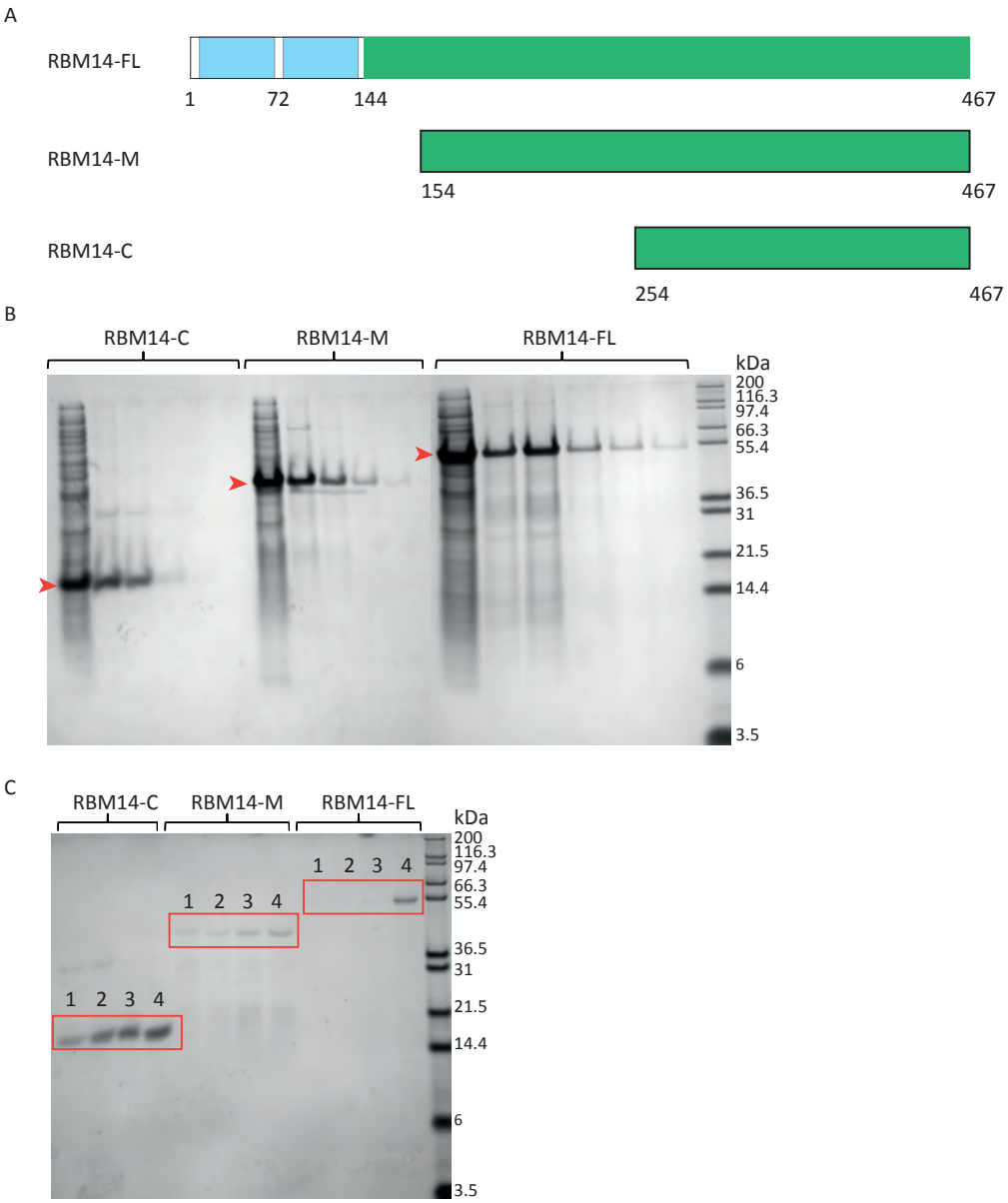


Figure C.2: RBM14 constructs and purification. A: RBM14 constructs: full length (RBM14-FL), N-terminal truncation (RBM14-M) and RBM14 C-terminus (RBM14-C). B: SDS-PAGE of the initial purification step of RBM14. C: SDS-PAGE of RBM14 constructs following refolding in 1: 20 mM Sodium acetate buffer, pH4, 2: 20 mM Sodium acetate buffer, pH5, 3: 20 mM MES buffer, pH6 and 4: 20 mM PBS buffer, pH7. The construct and molecular weight of the marker used are indicated.

C.4 RBM14 and STIL interaction studies

Evidence of interaction between RBM14-C and STIL was initially tested by analytical size exclusion chromatography however, no peak shift that would indicate the presence of larger molecular weight components such as that which would arise from interaction between RBM14 and STIL, was observed (Fig. C.3). This does not necessarily imply that RBM14-C and STIL do not interact as dilution over the column may result in protein concentrations below the the K_d of binding.

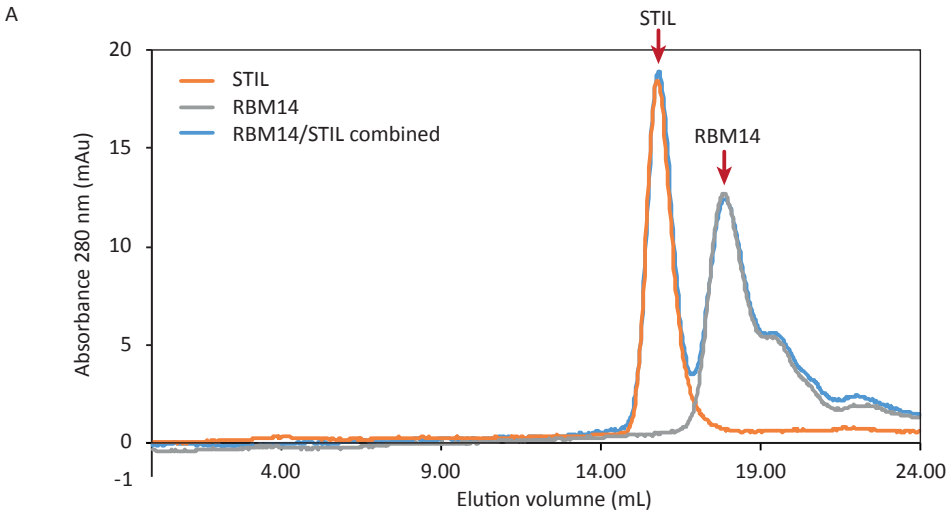


Figure C.3: Analytical size exclusion chromatography of RBM14 and STIL. Chromatogram of RBM14-C (grey), STIL (red) alone and when combined (blue). Arrows indicate peaks corresponding to each protein.

C.4.1 NMR

Further interaction studies were carried out using NMR. Heteronuclear single quantum coherence (HSQC) spectra were collected of RBM14-C alone and following addition of STIL. Slight peak shifts and intensity changes were observed following the addition of STIL indicating an interaction between the two proteins (Fig. C.4 A, B and C). The intensity changes of 63 HSQC peaks (each representing a single residue of RBM14-C) are shown (Fig. C.4 D). Fractional intensity changes were not large, with only two peaks showing intensity changes greater than 50%. To identify populations of HSQC peaks that were effected more or less strongly by

addition of STIL, a histogram of the fractional intensity changes was plotted (Fig. C.4 E). The histogram is binned at intervals of 0.05, above the error of ∼0.02. Since RBM14-C is a long, intrinsically disordered protein it is not expected that the full length will bind to STIL. Instead, it is likely that one region of RBM14-C binds to STIL and the remainder of the protein does not, with a transition zone in between. This assumption is reflected in the shape of the fractional intensity histogram where 2 populations can be seen, one with a peak at ∼0.5-0.55 fractional intensity and a second population with a peak at ∼0.7 fractional intensity. The spread of these two populations is large therefore the peaks have merged together. This suggests that RBM14-C interacts weakly with STIL.

To identify the residues corresponding to the HSQC peaks which were effected by addition of STIL peak assignment was attempted however, the very low RBM14-C concentration resulted in signal too low above background noise for successful peak assignment.

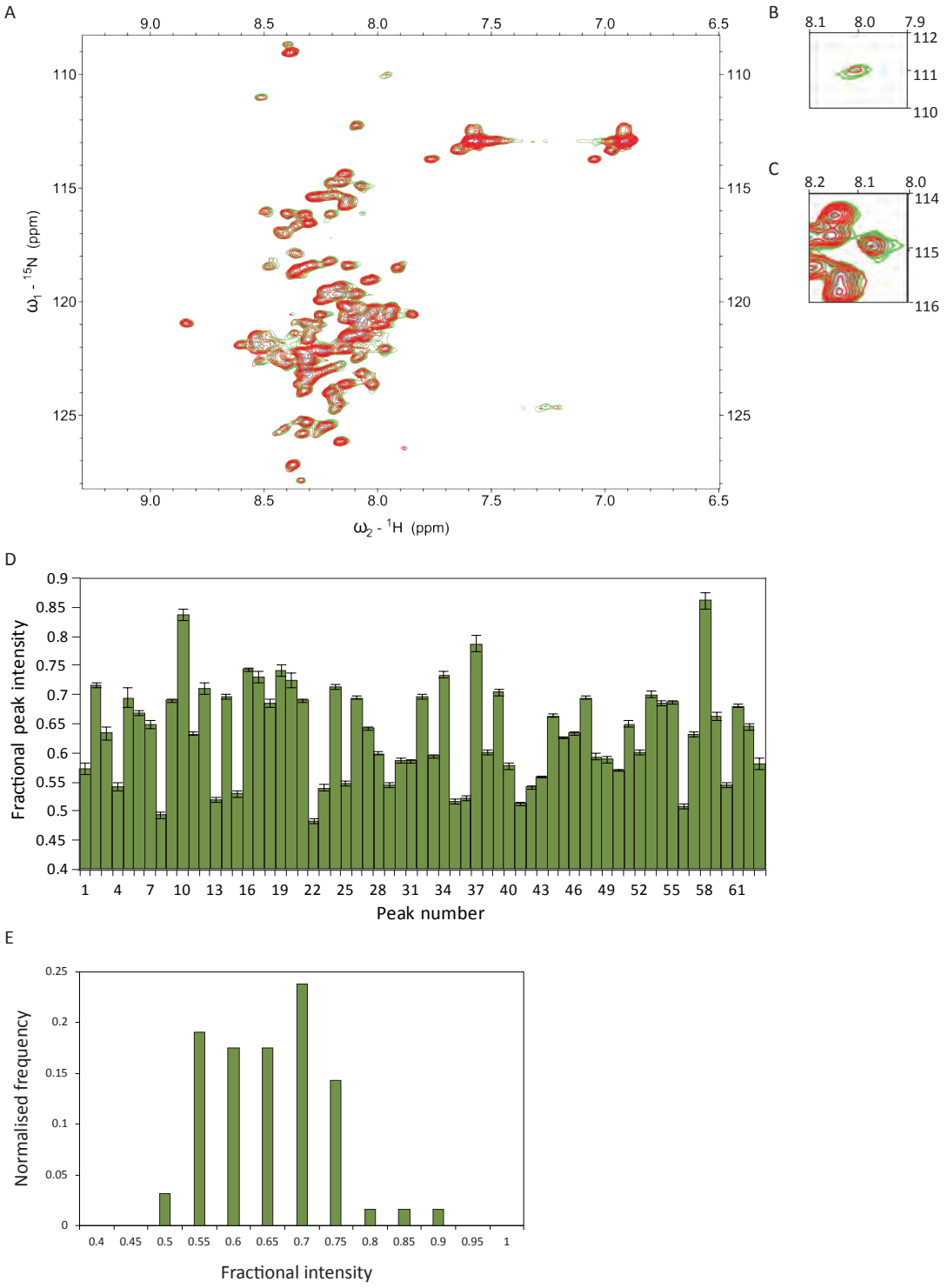


Figure C.4: NMR interaction studies of RBM14 and STIL. A: HSQC of RBM14-C (green) alone and in the presence of STIL (red). B and C: zoomed in views of peaks where minor peak shift or intensity change occurs upon addition of STIL. D: fractional intensity change of RBM14 HSQC peaks following addition of STIL. E: Histogram of the fractional intensity of RBM14 HSQC peaks following addition of STIL.

C.4.2 Fluorescence polarisation

Fluorescence polarisation (FP) was used to establish the K_d of the RBM14-STIL interaction. Attempts to label RBM14-C with a fluorescent dye as is required for FP experiments were unsuccessful therefore, short peptides that could further narrow down the residues required for STIL binding were used. Three overlapping peptides spanning the C-terminus of RBM14-C were used: RBM14-P1 (residues 370-410), RBM14-P2 (residues 390-430) and RBM14-P3 (residues 410-467) (Table C.1 and Fig. C.5 A). Peptides RBM14-P2 and RBM14-P3 did not interact with STIL. RBM14-P1 displayed weak binding to STIL, indicated by the slight increase in FP signal (Fig. C.5 B). The K_d of this interaction is 37$\pm$3.8 mM. Further repeats of FP experiments are required to confirm this value. As indicated by NMR studies this indicates the interaction between RBM14 and STIL is extremely weak. Initial attempts to crystallise RBM14-P1 and STIL were unsuccessful.

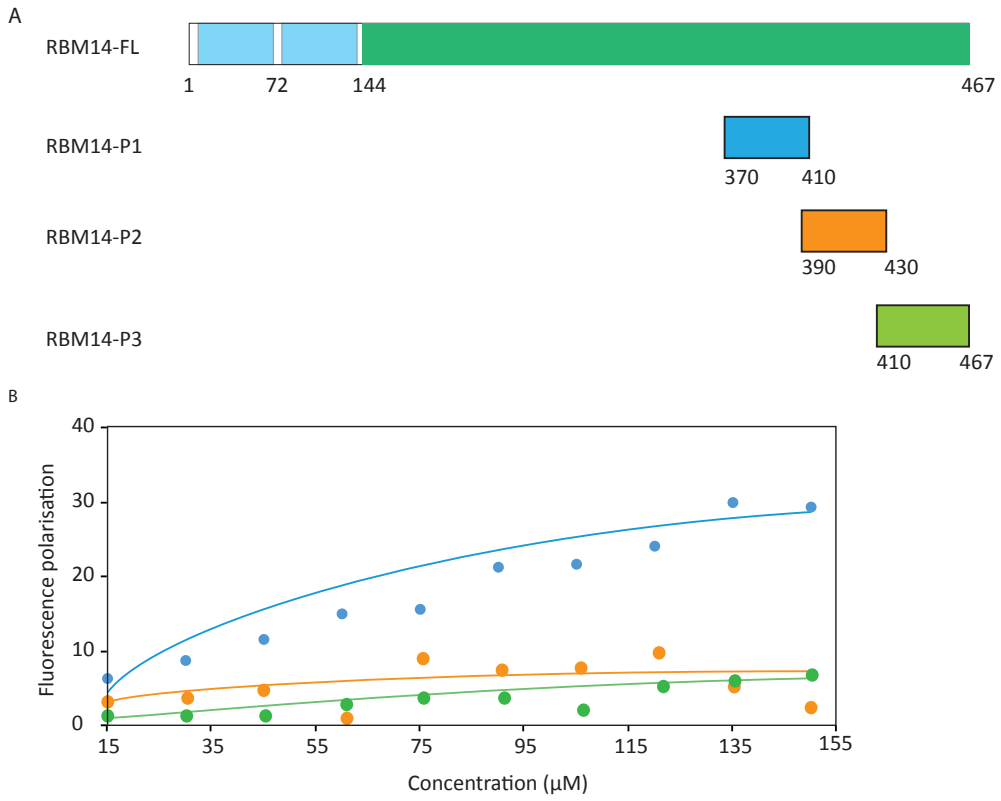


Figure C.5: Fluorescence polarisation studies of RBM14 and STIL. A: RBM14 peptides used in fluorescence polarisation assays (RBM14-FL is shown for comparison): RBM14-P1 residues 370-410 (blue), RBM14-P2 residues 390-430 (orange), and RBM14-P3 residues 410-467 (green). B: Fluorescence polarisation signal of RBM14-P1 (blue), RBM14-P2 (orange), and RBM14-P3 (green) following addition of STIL.

References

[1] R. V. Dippell. "The development of basal bodies in paramecium". In: *Proc Natl Acad Sci USA* 61.2 (1968), pp. 461–468.

[2] T. Boveri. "Ueber den Antheil des Spermatozoon an der Teilung des Eies." In: *Ges. Morph. Physiol. München.* 3.151 (1887).

[3] A. Paoletti and M. Bornens. "Organisation and functional regulation of the centrosome in animal cells." In: *Prog Cell Cycle Res* 3 (1997), p. 14.

[4] W. F. Marshall. "Centriole evolution". In: *Curr Opin Cell Biol* 21.1 (2009), pp. 14–9.

[5] A. Khodjakov and C. L. Rieder. "Centrosomes enhance the fidelity of cytokinesis in vertebrates and are required for cell cycle progression". In: *J Cell Biol* 153.1 (2001), pp. 237–42.

[6] E. H. Hinchcliffe et al. "Requirement of a centrosomal activity for cell cycle progression through G1 into S phase". In: *Science* 291.5508 (2001), pp. 1547–50.

[7] C. Klotz et al. "Parthenogenesis in Xenopus eggs requires centrosomal integrity". In: *J Cell Biol* 110.2 (1990), pp. 405–15.

[8] C. B. O'Connell and A. L. Khodjakov. "Cooperative mechanisms of mitotic spindle formation". In: *J Cell Sci* 120.Pt 10 (2007), pp. 1717–22.

[9] M. Kirkham et al. "SAS-4 is a C. elegans centriolar protein that controls centrosome size". In: *Cell* 112.4 (2003), pp. 575–87.

[10] N. R. Stevens et al. "From stem cell to embryo without centrioles". In: *Curr Biol* 17.17 (2007), pp. 1498–503.

[11] A. Khodjakov et al. "Centrosome-independent mitotic spindle formation in vertebrates". In: *Curr Biol* 10.2 (2000), pp. 59–67.

[12] E. Sumiyoshi and A. Sugimoto. "Cell polarity: centrosomes release signals for polarization". In: *Curr Biol* 22.8 (2012), R281–3.

[13] M. Regolini. "The Centrosome as a Geometry Organizer." In: *Results Probl Cell Differ.* 67.11 (2019).

[14] C. Sutterlin and A. Colanzi. "The Golgi and the centrosome: building a functional partnership". In: *J Cell Biol* 188.5 (2010), pp. 621–8.

[15] K. Tormanen et al. "Function of Golgi-centrosome proximity in RPE-1 cells." In: *PLOS ONE.* 14.4 (2019).

[16] F. R. Balestra, L. von Tobel, and P. Gonczy. "Paternally contributed centrioles exhibit exceptional persistence in C. elegans embryos". In: *Cell Res* 25.5 (2015), pp. 642–4.

[17] J. Reina and C. Gonzalez. “When fate follows age: unequal centrosomes in asymmetric cell division”. In: *Philos Trans R Soc Lond B Biol Sci* 369.1650 (2014).

[18] P. Gonczy. “Mechanisms of asymmetric cell division: flies and worms pave the way”. In: *Nat Rev Mol Cell Biol* 9.5 (2008), pp. 355–66.

[19] M. A. Khan et al. “A missense mutation in the PISA domain of HsSAS-6 causes autosomal recessive primary microcephaly in a large consanguineous Pakistani family”. In: *Hum Mol Genet* 23.22 (2014), pp. 5940–9.

[20] A. Noatynska et al. “Coordinating cell polarity and cell cycle progression: what can we learn from flies and worms?” In: *Open Biol* 3.8 (2013), p. 130083.

[21] L. Fletcher and R. J. Muschel. “The centrosome and the DNA damage induced checkpoint”. In: *Cancer Lett* 243.1 (2006), pp. 1–8.

[22] P. L. Chavali, M. Putz, and F. Gergely. “Small organelle, big responsibility: the role of centrosomes in development and disease”. In: *Philos Trans R Soc Lond B Biol Sci* 369.1650 (2014).

[23] K. Fukasawa et al. “Abnormal centrosome amplification in the absence of p53”. In: *Science* 271.5256 (1996), pp. 1744–7.

[24] J. H. Sir et al. “Loss of centrioles causes chromosomal instability in vertebrate somatic cells”. In: *J Cell Biol* 203.5 (2013), pp. 747–56.

[25] B. E. Tanos et al. “Centriole distal appendages promote membrane docking, leading to cilia initiation.” In: *Genes and Development* 27.163 (2013).

[26] D. D. A Klos et al. “Deuterosome-mediated centriole biogenesis.” In: *Dev Cell* 27.1 (2013).

[27] A. Jord et al. “Centriole amplification by mother and daughter centrioles differs in multiciliated cells.” In: *Nature* 516.7529 (2014).

[28] M. A. Zariwala, M. R. Knowles, and H. Omran. “Genetic defects in ciliary structure and function”. In: *Annu Rev Physiol* 69 (2007), pp. 423–50.

[29] M. Bettencourt-Dias et al. “Centrosomes and cilia in human disease”. In: *Trends Genet* 27.8 (2011), pp. 307–15.

[30] P. Winyard and D. Jenkins. “Putative roles of cilia in polycystic kidney disease”. In: *Biochim Biophys Acta* 1812.10 (2011), pp. 1256–62.

[31] E. T. O’Toole et al. “Morphologically distinct microtubule ends in the mitotic centrosome of Caenorhabditis elegans”. In: *J Cell Biol* 163.3 (2003), pp. 451–6.

[32] L. Pelletier et al. “Centriole assembly in Caenorhabditis elegans”. In: *Nature* 444.7119 (2006), pp. 619–23.

[33] A. Spektor et al. “Cep97 and CP110 suppress a cilia assembly program”. In: *Cell* 130.4 (2007), pp. 678–90.

[34] A. Gupta and D. Kitagawa. “Ultrastructural diversity between centrioles of eukaryotes”. In: *J Biochem* 164.1 (2018), pp. 1–8.

[35] P. Gonczy. “Towards a molecular architecture of centriole assembly”. In: *Nat Rev Mol Cell Biol* 13.7 (2012), pp. 425–35.

[36] M. Winey and E. O'Toole. "Centriole structure". In: *Philos Trans R Soc Lond B Biol Sci* 369.1650 (2014).

[37] D. Serwas et al. "Centrioles initiate cilia assembly but are dispensable for maturation and maintenance in C. elegans". In: *J Cell Biol* 216.6 (2017), pp. 1659–1671.

[38] I. V. Nechipurenko and P. Sengupta. "The rise and fall of basal bodies in the nematode Caenorhabditis elegans". In: *Cilia* 6 (2017), p. 9.

[39] M. Hilbert et al. "Caenorhabditis elegans centriolar protein SAS-6 forms a spiral that is consistent with imparting a ninefold symmetry". In: *Proc Natl Acad Sci U S A* 110.28 (2013), pp. 11373–8.

[40] D. Kitagawa et al. "Structural basis of the 9-fold symmetry of centrioles". In: *Cell* 144.3 (2011), pp. 364–75.

[41] M. van Breugel et al. "Structures of SAS-6 suggest its organization in centrioles". In: *Science* 331.6021 (2011), pp. 1196–9.

[42] L. Gartenmann et al. "A combined 3D-SIM/SMLM approach allows centriole proteins to be localized with a precision of approximately 4-5 nm". In: *Curr Biol* 27.19 (2017), R1054–R1055.

[43] T. T. Yang et al. "Super-resolution architecture of mammalian centriole distal appendages reveals distinct blade and matrix functional components". In: *Nat Commun* 9.1 (2018), p. 2023.

[44] J. Azimzadeh and W. F. Marshall. "Building the centriole". In: *Curr Biol* 20.18 (2010), R816–25.

[45] P. T. Conduit, A. Wainman, and J. W. Raff. "Centrosome function and assembly in animal cells". In: *Nat Rev Mol Cell Biol* 16.10 (2015), pp. 611–24.

[46] P. Guichard et al. "Native architecture of the centriole proximal region reveals features underlying its 9-fold radial symmetry". In: *Curr Biol* 23.17 (2013), pp. 1620–8.

[47] S. Li et al. "Three-dimensional structure of basal body triplet revealed by electron cryo-tomography". In: *EMBO J* 31.3 (2012), pp. 552–62.

[48] P. Guichard et al. "Procentriole assembly revealed by cryo-electron tomography". In: *EMBO J* 29.9 (2010), pp. 1565–72.

[49] S. Li et al. "Three-dimensional structure of basal body triplet revealed by electron cryo-tomography". In: *EMBO J* 31.3 (2012), pp. 552–62.

[50] K. Sugioka et al. "Centriolar SAS-7 acts upstream of SPD-2 to regulate centriole assembly and pericentriolar material formation". In: *Elife* 6 (2017).

[51] M. Hirono. "Cartwheel assembly". In: *Philos Trans R Soc Lond B Biol Sci* 369.1650 (2014).

[52] E. A. Nigg and A. J. Holland. "Once and only once: mechanisms of centriole duplication and their deregulation in disease". In: *Nat Rev Mol Cell Biol* 19.5 (2018), pp. 297–312.

[53] H. Fujita, Y. Yoshino, and N. Chiba. "Regulation of the centrosome cycle". In: *Mol Cell Oncol* 3.2 (2016), e1075643.

[54] D. A. Brito, S. M. Gouveia, and M. Bettencourt-Dias. “Deconstructing the centriole: structure and number control”. In: *Curr Opin Cell Biol* 24.1 (2012), pp. 4–13.

[55] A. J. Holland et al. “The autoregulated instability of Polo-like kinase 4 limits centrosome duplication to once per cell cycle”. In: *Genes Dev* 26.24 (2012), pp. 2684–9.

[56] P. Strnad and P. Gonczy. “Mechanisms of procentriole formation”. In: *Trends Cell Biol* 18.8 (2008), pp. 389–96.

[57] M. F. Tsou et al. “Polo kinase and separase regulate the mitotic licensing of centriole duplication in human cells”. In: *Dev Cell* 17.3 (2009), pp. 344–54.

[58] P. Gonczy et al. “Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III”. In: *Nature* 408.6810 (2000), pp. 331–6.

[59] K. F. O’Connell et al. “The C. elegans zyg-1 gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo”. In: *Cell* 105.4 (2001), pp. 547–58.

[60] K. F. O’Connell, K. N. Maxwell, and J. G. White. “The spd-2 gene is required for polarization of the anteroposterior axis and formation of the sperm asters in the Caenorhabditis elegans zygote”. In: *Dev Biol* 222.1 (2000), pp. 55–70.

[61] M. Ohta, A. Desai, and K. Oegema. “How centrioles acquire the ability to reproduce”. In: *Elife* 6 (2017).

[62] K. F. Sonnen et al. “Human Cep192 and Cep152 cooperate in Plk4 recruitment and centriole duplication”. In: *J Cell Sci* 126.Pt 14 (2013), pp. 3223–33.

[63] Y. Tsuchiya et al. “Cep295 is a conserved scaffold protein required for generation of a bona fide mother centriole”. In: *Nat Commun* 7 (2016), p. 12567.

[64] R. Habedanck et al. “The Polo kinase Plk4 functions in centriole duplication”. In: *Nat Cell Biol* 7.11 (2005), pp. 1140–6.

[65] M. Bettencourt-Dias et al. “SAK/PLK4 is required for centriole duplication and flagella development”. In: *Curr Biol* 15.24 (2005), pp. 2199–207.

[66] D. Takao, S. Yamamoto, and D. Kitagawa. “A theory of centriole duplication based on self-organized spatial pattern formation”. In: *The Journal of Cell Biology* 218.8 (2019).

[67] S. Yamamoto and D. Kitagawa. “Self-organization of Plk4 regulates symmetry breaking in centriole duplication”. In: *Nat Commun* 10.1 (2019), p. 1810.

[68] M. Delattre, C. Canard, and P. Gonczy. “Sequential protein recruitment in C. elegans centriole formation”. In: *Curr Biol* 16.18 (2006), pp. 1844–9.

[69] M. M. Lettman et al. “Direct binding of SAS-6 to ZYG-1 recruits SAS-6 to the mother centriole for cartwheel assembly”. In: *Dev Cell* 25.3 (2013), pp. 284–98.

[70] C. Arquint and E. A. Nigg. “STIL microcephaly mutations interfere with APC/C-mediated degradation and cause centriole amplification”. In: *Curr Biol* 24.4 (2014), pp. 351–60.

[71] K. B. Rogala et al. “The Caenorhabditis elegans protein SAS-5 forms large oligomeric assemblies critical for centriole formation”. In: *Elife* 4 (2015), e07410.

[72] M. Delattre et al. “Centriolar SAS-5 is required for centrosome duplication in C. elegans”. In: *Nat Cell Biol* 6.7 (2004), pp. 656–64.

[73] R. Qiao et al. “SAS-6 coiled-coil structure and interaction with SAS-5 suggest a regulatory mechanism in C. elegans centriole assembly”. In: *EMBO J* 31.22 (2012), pp. 4334–47.

[74] X. Xu et al. “DNA replication licensing factor Cdc6 and Plk4 kinase antagonistically regulate centrosome duplication via Sas-6”. In: *Nat Commun* 8 (2017), p. 15164.

[75] M. Jerka-Dziadosz et al. “Basal body duplication in Paramecium: the key role of Bld10 in assembly and stability of the cartwheel”. In: *Cytoskeleton (Hoboken)* 67.3 (2010), pp. 161–71.

[76] S. Leidel et al. “SAS-6 defines a protein family required for centrosome duplication in C. elegans and in human cells”. In: *Nat Cell Biol* 7.2 (2005), pp. 115–25.

[77] A. Dammermann et al. “Centriole assembly requires both centriolar and pericentriolar material proteins”. In: *Dev Cell* 7.6 (2004), pp. 815–29.

[78] E. O’Toole et al. “The role of gamma-tubulin in centrosomal microtubule organization”. In: *PLoS One* 7.1 (2012), e29795.

[79] J. Gopalakrishnan et al. “Sas-4 provides a scaffold for cytoplasmic complexes and tethers them in a centrosome”. In: *Nat Commun* 2 (2011), p. 359.

[80] S. Bianchi et al. “Interaction between the Caenorhabditis elegans centriolar protein SAS-5 and microtubules facilitates organelle assembly”. In: *Mol Biol Cell* 29.6 (2018), pp. 722–735.

[81] Y. C. Lin et al. “Human microcephaly protein CEP135 binds to hSAS-6 and CPAP, and is required for centriole assembly”. In: *EMBO J* 32.8 (2013), pp. 1141–54.

[82] X. Zheng et al. “Human microcephaly protein CEP135 binds to hSAS-6 and CPAP, and is required for centriole assembly.” In: *Nature Comminications* 16.7 (2016).

[83] D. Comartin et al. “CEP120 and SPICE1 cooperate with CPAP in centriole elongation”. In: *Curr Biol* 23.14 (2013), pp. 1360–6.

[84] K. Lee and K. Rhee. “PLK1 phosphorylation of pericentrin initiates centrosome maturation at the onset of mitosis”. In: *J Cell Biol* 195.7 (2011), pp. 1093–101.

[85] A. Puklowski et al. “The SCF-FBXW5 E3-ubiquitin ligase is regulated by PLK4 and targets HsSAS-6 to control centrosome duplication”. In: *Nat Cell Biol* 13.8 (2011), pp. 1004–9.

[86] P. Strnad et al. “Regulated HsSAS-6 levels ensure formation of a single procentriole per centriole during the centrosome duplication cycle”. In: *Dev Cell* 13.2 (2007), pp. 203–13.

[87] G. Cabral et al. “Multiple mechanisms contribute to centriole separation in C. elegans”. In: *Curr Biol* 23.14 (2013), pp. 1380–7.

[88] M. F. Tsou and T. Stearns. “Mechanism limiting centrosome duplication to once per cell cycle”. In: *Nature* 442.7105 (2006), pp. 947–51.

[89] I. A. Vorobjev and S. Chentsov Yu. “Centrioles in the cell cycle. I. Epithelial cells”. In: *J Cell Biol* 93.3 (1982), pp. 938–49.

[90] A. Rodrigues-Martins et al. “DSAS-6 organizes a tube-like centriole precursor, and its absence suggests modularity in centriole assembly”. In: *Curr Biol* 17.17 (2007), pp. 1465–72.

[91] H. Ling et al. “Functional relationship among PLK2, PLK4 and ROCK2 to induce centrosome amplification”. In: *Cell Cylce* 14.4 (2015).

[92] S. Dozsey, W. Zimmerman, and K. Mikule. “Centrosome control of the cell cycle.” In: *Trends in Cell Biology* 15.6 (2005).

[93] K. Yaguchi et al. “Uncoordinated centrosome cycle underlies the instability of non-diploid somatic cells in mammals”. In: *The Journal of Cell Biology* 217.7 (2018).

[94] G. Wang, Q. Jiang, and C. Zhang. “The role of mitotic kinases in coupling the centrosome cycle with the assembly of the mitotic spindle.” In: *Journal of Cell Science* 127.19 (2014).

[95] V. Srsen and A. Merdes. “The centrosome and cell proliferatio.” In: *BMC Cell Division* 1.26 (2006).

[96] M. Jackman et al. “Active cyclin B1-Cdk1 first appears on centrosomes in prophase.” In: *Nature Cell Biology* 5.2 (2003).

[97] P. Guichard et al. “Cell-free reconstitution reveals centriole cartwheel assembly mechanisms”. In: *Nat Commun* 8 (2017), p. 14813.

[98] G. Dong. “Building a ninefold symmetrical barrel: structural dissections of centriole assembly”. In: *Open Biol* 5.8 (2015).

[99] J. M. C. Busch et al. “A dynamically interacting flexible loop assists oligomerisation of the Caenorhabditis elegans centriolar protein SAS-6”. In: *Sci Rep* 9.1 (2019), p. 3526.

[100] P. Guichard, V. Hamel, and P. Gonczy. “The Rise of the Cartwheel: Seeding the Centriole Organelle”. In: *Bioessays* 40.4 (2018), e1700241.

[101] P. Guichard et al. “Cartwheel architecture of Trichonympha basal body”. In: *Science* 337.6094 (2012), p. 553.

[102] I. R. Gibbons and A. V. Grimstone. “On flagellar structure in certain flagellates”. In: *J Biophys Biochem Cytol* 7 (1960), pp. 697–716.

[103] S. K. Dutcher and E. T. O’Toole. “The basal bodies of Chlamydomonas reinhardtii”. In: *Cilia* 5 (2016), p. 18.

[104] J. E. Ford, P. J. Stansfeld, and I. Vakonakis. “Coupling Form and Function: How the Oligomerisation Symmetry of the SAS-6 Protein Contributes to the Architecture of Centriole Organelles”. In: *Symmetry-Basel* 9.5 (2017).

[105] R. McNitt. “Centriole ultrastructure and its possible role in microtubule formation in an aquatic fungus”. In: *Protoplasma* 80.1 (1974), pp. 91–108.

[106] Y. Nakazawa et al. “SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole”. In: *Curr Biol* 17.24 (2007), pp. 2169–74.

[107] M. Hiraki et al. “Bld10p constitutes the cartwheel-spoke tip and stabilizes the 9-fold symmetry of the centriole”. In: *Curr Biol* 17.20 (2007), pp. 1778–83.

[108] J. T. Wang et al. “Centriole triplet microtubules are required for stable centriole formation and inheritance in human cells”. In: *Elife* 6 (2017).

[109] J. Monod, J. Wyman, and J. P. Changeux. “On the Nature of Allosteric Transitions: A Plausible Model”. In: *J Mol Biol* 12 (1965), pp. 88–118.

[110] S. N. Andres et al. “Crystal structure of human XLF: a twist in nonhomologous DNA end-joining”. In: *Mol Cell* 28.6 (2007), pp. 1093–101.

[111] Q. Wu et al. “Non-homologous end-joining partners in a helical dance: structural studies of XLF-XRCC4 interactions”. In: *Biochem Soc Trans* 39.5 (2011), 1387–92, suppl 2 p following 1392.

[112] Y. Li et al. “Crystal structure of human XLF/Cernunnos reveals unexpected differences from XRCC4 with implications for NHEJ”. In: *EMBO J* 27.1 (2008), pp. 290–300.

[113] M. S. Junop et al. “Crystal structure of the Xrcc4 DNA repair protein and implications for end joining”. In: *EMBO J* 19.22 (2000), pp. 5962–70.

[114] M. Hammel et al. “XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair”. In: *J Biol Chem* 286.37 (2011), pp. 32638–50.

[115] S. N. Andres et al. “A human XRCC4-XLF complex bridges DNA”. In: *Nucleic Acids Res* 40.4 (2012), pp. 1868–78.

[116] M. A. Cottee et al. “SAS-6 oligomerization: the key to the centriole?” In: *Nat Chem Biol* 7.10 (2011), pp. 650–3.

[117] F. O. Perkins. “Formation of centriole and centriole-like structures during meiosis and mitosis in Labyrinthula Sp. (Rhizopodea, Labyrinthulida)”. In: *Journal of Cell Science* 6 (1970), p. 24.

[118] G. Marteil et al. “Over-elongation of centrioles in cancer promotes centriole amplification and chromosome missegregation”. In: *Nat Commun* 9.1 (2018), p. 1258.

[119] S. Geimer and M. Melkonian. “The ultrastructure of the Chlamydomonas reinhardtii basal apparatus: identification of an early marker of radial asymmetry inherent in the basal body”. In: *J Cell Sci* 117.Pt 13 (2004), pp. 2663–74.

[120] E. T. O’Toole and S. K. Dutcher. “Site-specific basal body duplication in Chlamydomonas”. In: *Cytoskeleton (Hoboken)* 71.2 (2014), pp. 108–18.

[121] G. Kohlmaier et al. “Overly long centrioles and defective cell division upon excess of the SAS-4-related protein CPAP”. In: *Curr Biol* 19.12 (2009), pp. 1012–8.

[122] T. I. Schmidt et al. “Control of centriole length by CPAP and CP110”. In: *Curr Biol* 19.12 (2009), pp. 1005–11.

[123] M. G. Aydogan et al. “A homeostatic clock sets daughter centriole size in flies”. In: *J Cell Biol* 217.4 (2018), pp. 1233–1248.

[124] A. Dammermann et al. “SAS-4 is recruited to a dynamic structure in newly forming centrioles that is stabilized by the gamma-tubulin-mediated addition of centriolar microtubules”. In: *J Cell Biol* 180.4 (2008), pp. 771–85.

[125] M. Beck et al. “The quantitative proteome of a human cell line”. In: *Mol Syst Biol* 7 (2011), p. 549.

[126] D. Keller et al. “Mechanisms of HsSAS-6 assembly promoting centriole formation in human cells”. In: *J Cell Biol* 204.5 (2014), pp. 697–712.

[127] M. Bauer et al. “Quantitative analysis of human centrosome architecture by targeted proteomics and fluorescence imaging”. In: *EMBO J* 35.19 (2016), pp. 2152–2166.

[128] T. Ochi, Q. Wu, and T. L. Blundell. “The spatial organization of non-homologous end joining: from bridging to end joining”. In: *DNA Repair (Amst)* 17 (2014), pp. 98–109.

[129] D. S. Goodsell and A. J. Olson. “Structural symmetry and protein function”. In: *Annu Rev Biophys Biomol Struct* 29 (2000), pp. 105–53.

[130] M. Gottardo, G. Callaini, and M. G. Riparbelli. “The Drosophila centriole - conversion of doublets into triplets within the stem cell niche”. In: *J Cell Sci* 128.14 (2015), pp. 2437–42.

[131] A. Franz et al. “CP110 exhibits novel regulatory activities during centriole assembly in Drosophila”. In: *J Cell Biol* 203.5 (2013), pp. 785–99.

[132] R. Uzbekov, A. Garanina, and C. Bressac. “Centrioles without microtubules: a new morphological type of centriole”. In: *Biol Open* 7.8 (2018).

[133] M. Hilbert et al. “SAS-6 engineering reveals interdependence between cartwheel and microtubules in determining centriole architecture”. In: *Nat Cell Biol* 18.4 (2016), pp. 393–403.

[134] M. G. Riparbelli et al. “Centriole symmetry: a big tale from small organisims”. In: *Cell Motil Cytoskeleton* 66 (2009), p. 5.

[135] R. Dallai et al. “The spermatogenesis and sperm strucutre of Acerentomon microrhinus (Portura, Hexapoda) with considerations on the phylogenetic position of the taxon.” In: *Zoomorphology* 129 (2010), p. 19.

[136] D. M. Phillips. “Giant centriole formation in Sciara”. In: *J Cell Biol* 33.1 (1967), pp. 73–92.

[137] M. van Breugel et al. “Structure of the SAS-6 cartwheel hub from Leishmania major”. In: *Elife* 3 (2014), e01812.

[138] B. Webb and A. Sali. “Comparative Protein Structure Modeling Using MODELLER”. In: *Curr Protoc Protein Sci* 86 (2016), pp. 291–297.

[139] M. A. Marti-Renom et al. “Comparative protein structure modeling of genes and genomes”. In: *Annu Rev Biophys Biomol Struct* 29 (2000), pp. 291–325.

[140] A. Sali and T. L. Blundell. “Comparative protein modelling by satisfaction of spatial restraints”. In: *J Mol Biol* 234.3 (1993), pp. 779–815.

[141] A. Fiser, R. K. Do, and A. Sali. “Modeling of loops in protein structures”. In: *Protein Sci* 9.9 (2000), pp. 1753–73.

[142] M. Y. Shen and A. Sali. “Statistical potential for assessment and prediction of protein structures”. In: *Protein Sci* 15.11 (2006), pp. 2507–24.

[143] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. "GROMACS: A messafe-passing parallel molecular dynamics implementation". In: *Computer Physics Comminications* 91.1-3 (1995), p. 13.

[144] M. J. Abraham et al. "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". In: *SoftwareX* 1-2 (2015), p. 6.

[145] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids." In: *Journal of the American Chemical Society* 118.45 (1996), pp. 11225–11236.

[146] E. F. Pettersen et al. "UCSF Chimera–a visualization system for exploratory research and analysis". In: *J Comput Chem* 25.13 (2004), pp. 1605–12.

[147] S. C. Potter et al. "HMMER web server: 2018 update". In: *Nucleic Acids Res* 46.W1 (2018), W200–W204.

[148] M. Landau et al. "ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures". In: *Nucleic Acids Res* 33.Web Server issue (2005), W299–302.

[149] F. Glaser et al. "ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information". In: *Bioinformatics* 19.1 (2003), pp. 163–4.

[150] T. A. Hopf et al. "The EVcouplings Python framework for coevolutionary sequence analysis." In: *Bioinformatics* 10 (2018), p. 3.

[151] E. Krissinel and K. Henrick. "Inference of macromolecular assemblies from crystalline state". In: *J Mol Biol* 372.3 (2007), pp. 774–97.

[152] L. A. Kelley et al. "The Phyre2 web portal for protein modeling, prediction and analysis". In: *Nat Protoc* 10.6 (2015), pp. 845–58.

[153] A. M. Waterhouse et al. "Jalview Version 2-a multiple sequence alignment editor and analysis workbench". In: *Bioinformatics* 25.9 (2009), pp. 1189–1191.

[154] K. Tamura et al. "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0". In: *Mol Biol Evol* 24.8 (2007), pp. 1596–9.

[155] D. Necas and P. Klapetek. "Gwyddion: an open-source software for SPM data analysis". In: *Central European Journal of Physics* 10.1 (2012), pp. 181–188.

[156] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri. "NIH Image to ImageJ: 25 years of image analysis". In: *Nat Methods* 9.7 (2012), pp. 671–675.

[157] A. Feduccia. "Evidence from claw geometry indicating arboreal habits of archaeopteryx". In: *Science* 259.5096 (1993), pp. 790–3.

[158] J. D. Hunter. "Matplotlib: A 2D graphics environment". In: *Computing In Science and Engineering* 9.3 (2007), p. 5.

[159] M. Pfreundschuh et al. "Localizing chemical groups while imaging single native proteins by high-resolution atomic force microscopy". In: *Nano Lett* 14.5 (2014), pp. 2957–64.

[160] T. A. Hopf et al. "Sequence co-evolution gives 3D contacts and structures of protein complexes". In: *Elife* 3 (2014).

[161] Q. Zhong et al. “Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy”. In: *Surface Science Letters* 290.1-2 (1993), p. 9.

[162] A. P. Nievergelt et al. “High-speed photothermal off-resonance atomic force microscopy reveals assembly routes of centriolar scaffold protein SAS-6”. In: *Nat Nanotechnol* 13.8 (2018), pp. 696–701.

[163] I. Andre et al. “Emergence of symmetry in homooligomeric biological assemblies”. In: *Proc Natl Acad Sci U S A* 105.42 (2008), pp. 16148–52.

[164] K. W. Plaxco and M. Gross. “Protein complexes: the evolution of symmetry”. In: *Curr Biol* 19.1 (2009), R25–6.

[165] D. Kitagawa et al. “Phosphorylation of SAS-6 by ZYG-1 is critical for centriole formation in C. elegans embryos”. In: *Dev Cell* 17.6 (2009), pp. 900–7.

[166] H.C.R. Klein et al. “Computational support for a scaffolding mechanism of centriole assembly”. In: *Sci Rep* 6.27075 (2016), p. 7.

[167] G. Lukinavicius et al. “Selective chemical crosslinking reveals a Cep57-Cep63-Cep152 centrosomal complex”. In: *Current Biology* 23.3 (2013), p. 5.

[168] W. J. Wedemeyer, E. Welker, and H. A. Scheraga. “Proline cis-trans isomerization and protein folding”. In: *Biochemistry* 41.50 (2002), pp. 14637–44.

[169] S. Hayward. “Peptide-plane flipping in proteins”. In: *Protein Sci* 10.11 (2001), pp. 2219–27.

[170] C. Arquint and E. A. Nigg. “The PLK4-STIL-SAS-6 module at the core of centriole duplication”. In: *Biochem Soc Trans* 44.5 (2016), pp. 1253–1263.

[171] A. Kumar et al. “Mutations in STIL, encoding a pericentriolar and centrosomal protein, cause primary microcephaly”. In: *Am J Hum Genet* 84.2 (2009), pp. 286–90.

[172] T. C. Moyer et al. “Binding of STIL to Plk4 activates kinase activity to promote centriole assembly”. In: *J Cell Biol* 209.6 (2015), pp. 863–78.

[173] S. Kraatz et al. “The Human Centriolar Protein CEP135 Contains a Two-Stranded Coiled-Coil Domain Critical for Microtubule Binding”. In: *Structure* 24.8 (2016), pp. 1358–1371.

[174] N. S. Dzhindzhev et al. “Plk4 phosphorylates Ana2 to trigger Sas6 recruitment and procentriole formation”. In: *Curr Biol* 24.21 (2014), pp. 2526–32.

[175] M. Ohta et al. “Direct interaction of Plk4 with STIL ensures formation of a single procentriole per parental centriole”. In: *Nat Commun* 5 (2014), p. 5267.

[176] C. Arquint et al. “STIL binding to Polo-box 3 of PLK4 regulates centriole duplication”. In: *Elife* 4 (2015).

[177] J. E. Klebba et al. “Autoinhibition and relief mechanism for Polo-like kinase 4”. In: *Proc Natl Acad Sci U S A* 112.7 (2015), E657–66.

[178] S. Zitouni et al. “CDK1 Prevents Unscheduled PLK4-STIL Complex Assembly in Centriole Biogenesis”. In: *Curr Biol* 26.9 (2016), pp. 1127–37.

[179] C. S. Fong et al. “53BP1 and USP28 mediate p53-dependent cell cycle arrest in response to centrosome loss and prolonged mitosis”. In: *Elife* 5 (2016).

[180] C. J. Tang et al. “CPAP is a cell-cycle regulated protein that controls centriole length”. In: *Nat Cell Biol* 11.7 (2009), pp. 825–31.

[181] A. Khodjakov et al. “De novo formation of centrosomes in vertebrate cells arrested during S phase”. In: *J Cell Biol* 158.7 (2002), pp. 1171–81.

[182] S. La Terra et al. “The de novo centriole assembly pathway in HeLa cells: cell cycle progression and centriole assembly/maturation”. In: *J Cell Biol* 168.5 (2005), pp. 713–22.

[183] Y. Uetake et al. “Cell cycle progression and de novo centriole assembly after centrosomal removal in untransformed human cells”. In: *J Cell Biol* 176.2 (2007), pp. 173–82.

[184] G. Manandhar, H. Schatten, and P. Sutovsky. “Centrosome reduction during gametogenesis and its significance”. In: *Biol Reprod* 72.1 (2005), pp. 2–13.

[185] A. Courtois et al. “The transition from meiotic to mitotic spindle assembly is gradual during early mammalian development”. In: *J Cell Biol* 198.3 (2012), pp. 357–70.

[186] M. Kusuhara et al. “Solution structure of RNA binding domain 2 in RNA-binding protein 14”. In: *To be published* (2006).

[187] G. Shiratsuchi et al. “RBM14 prevents assembly of centriolar protein complexes and maintains mitotic spindle integrity”. In: *EMBO J* 34.1 (2015), pp. 97–114.

[188] A. H. Fox and A. I. Lamond. “Paraspeckles”. In: *Cold Spring Harb Perspect Biol* 2.7 (2010), a000687.

[189] Z. R. Yang et al. “RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins”. In: *Bioinformatics* 21.16 (2005), pp. 3369–76.

[190] Z. Obradovic et al. “Predicting intrinsic disorder from amino acid sequence”. In: *Proteins* 53 Suppl 6 (2003), pp. 566–72.

[191] K. Peng et al. “Length-dependent prediction of protein intrinsic disorder”. In: *BMC Bioinformatics* 7 (2006), p. 208.

[192] F. Delaglio et al. “NMRPipe: a multidimensional spectral processing system based on UNIX pipes”. In: *J Biomol NMR* 6.3 (1995), pp. 277–93.

[193] T. D. Goddard and D. G. Kneller. “SPARKY 3”. In: *University of California, San Fransisco* ().