Site-Selective Chemical Phosphorylation and Activation of Kinases



A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy at the University of Oxford

Sébastien René Gabriel Galan

Department of Chemistry, University of Oxford

October, 2016

Keble College

AUTHOR'S DECLARATION

I declare the work presented in this thesis was carried out at the Chemistry Research Laboratory and Department of Biochemistry, University of Oxford, UK. All the work is my own, except where otherwise stated, and has not been submitted for any other degree at this or any other university.

Sébastien René Gabriel Galan

October 2016

ABSTRACT

Phosphorylation and dephosphorylation of proteins play an essential role in their cellular function. However, the study of protein phosphoforms in a molecular level is challenging, as the isolation and characterisation of homogeneous phosphoforms is not always possible. We propose to synthesise and study these phosphoforms using a tag-and-modify approach. This methodology allows the selective chemical phosphorylation of protein *via* the introduction of a phosphocysteine residue (pC).

The effect of individual phosphorylation of MEK1 kinase, which plays a central role in cell differentiation and proliferation, was examined. The selective formation of homogeneous MEK1pC218, MEK1pC222, as well as the doubly phosphorylated MEK1pC218pC222 mimics of the corresponding activate MEK1 phosphoforms, was successfully achieved. The engineering of new reactants and masking techniques was critical to the task. The development of a LCMS based procedure for kinetic studies unravelled the effect of phosphorylation on MEK1 binding to its substrates, ATP and ERK1. While initial crystallisation attempts showed that optimisation was still needed. Future work should stress the utility of such methodology towards the understanding of the structural variations created by phosphorylation.

The chemical mechanism of formation of dehydroalanine (Dha) was also investigated through the design and use of deuterated probes. The selective formation of α-deuterated cysteine residues in both amino-acid and protein models allowed the elucidation of an E1cB mechanism. Moreover, this study highlighted the particularity of 2,5-dibromohexane diamide (DBHDA) as a mild, water soluble and almost universal reagent for the modification of cysteine to Dha residues in protein.

ABBREVIATIONS

General Abbreviations

ADP adenosine diphosphate

AMP adenosine monophosphate

ATF2 Activating Transcription Factor 2

ATP adenosine triphosphate

BME β-mercaptoethanol

CAM S-carboxyamidomethylcysteine

CID collision-induced dissociation

CMC critical micelle concentration

COM S-carboxymethylcysteine

CD circular dichroism

CV column volume

Dha dehydroalanine

DMF dimethylformamide

DMSO dimethylsulphoxide

DNA deoxyribonucleic acid

DSF differential scanning fluorimetry

DTT DL-dithiothreitol

EGTA ethylene glycol-bis(2-aminoethylether)-N,N,N’,N’-tetraacetic acid

EM electromagnetic

ERA electrophoretic radioassay

ERK extracellular-signal regulated kinase

ES(I) electro-spray (ionisation)

ETD electron-transfer dissociation

FPLC Fast Protein Liquid Chromatography

GST glutathione S-transferase

h hour(s)

HCD high-energy collision dissociation

HEPES 4-(2-hydroxyethyl)-1-piperazine-1-ethanesulphonic acid

HePTP hematopoietic protein tyrosine phosphatase

HPLC high performance liquid chromatography

IPTG isopropyl _-D-1-thiogalactopyranoside

Kan kanamycin

(k)bp (kilo-)base pair

LB lysogeny broth

LC liquid chromatography

LC-MS liquid chromatography-mass spectrometry

MALDI matrix-assisted laser desorption/ionisation

MAPK mitogen-activated protein kinase

MEK MAP or ERK kinase

MKK MAP kinase kinase

MOPS 3-((N-morpholino)propanesulphonic acid

MS mass spectrometry

MWCO molecular weight cut-off

NAD nicotinamide adenine dinucleotide

NMR nuclear magnetic resonance

OD optical density

PCR polymerase chain reaction

pCys phosphocysteine

PEP phosphoenol pyruvate

ppm parts per million

pSer phosphoserine

pThr phosphothreonine

pTyr phosphotyrosine

rpm revolutions per minute

RT room temperature

sat. saturated

SB Super Broth

SDS sodium dodecyl sulphate

SGC Structural Genomics Consortium

SOC Super Optimal broth with Catabolite repression

TB Terrific Broth

TCEP tris(2-carboxyethyl)phosphine hydrochloride

TIC total ion current

UV ultra-violet

WT wild-type

Amino Acids

Ala, A alanine

Arg, R arginine

Asn, N asparagine

Asp, D aspartic acid

Cys, C cysteine

Gln, Q glutamine

Glu, E glutamic acid

Gly, G glycine

His, H histidine

Ile, I isoleucine

Leu, L leucine

Lys, K lysine

Met, M methionine

Phe, F phenylalanine

Pro, P proline

Ser, S serine

Thr, T threonine

Trp, W tryptophan

Tyr, Y tyrosine

Val, V valine

Nucleic Acid Bases

A adenine

C cytosine

G guanine

T thymine

ACKNOWLEDGMENTS

This thesis is dedicated to my wife, Bhairavi, for her patience, dedication and support. I know that she will be delighted with me coming back to the real world: the kitchen. As we say in French: “la patience a ses limites”.

I also want to dedicate this thesis to my friends and family. I have neglected you during the last few years. For the ones that were patient enough, I am now back.

Now the “thank-yous” as Maria, Reida and Amy would say.

Thank you, Ben (Professor Davis) for welcoming me into your lab, and letting me follow my scientific instincts and intuitions (not without some healthy arguments). They were not always right… Not always wrong… It’s how you learn! I guess I am a better scientist thank to that.

Thank you, Phin, for picking me. Mission “Phin Possible” was quite a ride. It will be remembered in the years to come… You initiated my thesis project. I hope you are proud to where I have taken it.

Jitka, thank you for finishing my work! No, no pressure! Thank you as well for your help during the write up and the corrections. I really appreciate it.

Thank you Ritu, for being you… Bored and tired... Joking apart, it was nice to, occasionally, talk science, and share great ideas. I wish you well for the future.

Of course all this work could not be done on my own. So I would like to thank, in no particular order, James, James, Tom, Tom, Tom, Elisabete, Ilaria, Maria, Rob, Sylvain, Bala, Regis, Billy, Lingzhi, Jens, and LG10 lab members, past and present for their moral and technical support. I apologise for the ones I forgot.

With no money, no PhD! So thank you, AstraZeneca, especially Alex, Jon and Gary, the BBSRC and their reminders, and Vertex through Bhairavi, for funding this adventure.

I would also like to thank my examiners Prof. Christian Hackenberger and Dr. Kilian Huber for their advice and giving me the opportunity to improve my manuscript greatly.

Finally, I stayed late at night, and sometimes came the weekend for you. You needed care and attention. I was sometime the only one able to reconnect you to the world. I guess it was worth it. It’s time to say good-bye. It’s not you, it’s me. I feel like I haven’t thank you enough. I’ll miss you, LCTA.

.

TABLE OF CONTENTS

AUTHOR'S DECLARATION i

ABSTRACT ii

ABBREVIATIONS iii

General Abbreviations iii

Amino Acids vi

Nucleic Acid Bases vii

ACKNOWLEDGMENTS viii

TABLE OF CONTENTS x

INTRODUCTION 0

Chapter 1: Introduction to MEK1 phosphorylation and activation 1

1.1. Activation of MEK1 kinase by phosphorylation 1

1.1.1. Protein phosphorylation 1

1.1.2. MEK, a dual specific Protein kinases 2

1.1.3. Phosphorylation of MEK1 kinase 3

1.1.4. Project synopsis 5

1.2. Biological methods towards the controlled activation and phosphorylation of MEK1 kinase 5

1.2.1. Preparation of constitutively active MEK1 by site directed mutagenesis 5

1.2.2. In situ and in vivo pMEK1 expression and purification 7

1.2.3. Insertion of pSer by amber stop codon suppression 8

1.3. Chemical phosphorylation of protein 13

1.3.1. (Semi)synthetic approaches to phosphorylated protein 13

1.3.2. Chemical phosphorylation of full-length protein 15

1.3.3. Dha insertion in protein 18

1.4. Site selective chemical phosphorylation of MEK1 26

1.4.1. Proof of concept 26

1.4.2. Challenges towards of site selective chemical phosphorylation of MEK1 via Dha 29

1.4.3. Thesis summary 30

RESULTS AND DISCUSSION 31

SCHEMATIC OVERVIEW OF THE ALKYLATING AGENTS USED TOWARDS REGIOSELECTIVE DHA FORMATION AND PHOSPHORYLATION 32

Chapter 2: Site-Selective Dehydroalanine Formation 33

2.1. Site selective formation of MEK1Dha222 with DBHDA 33

2.1.1. Reactivity of MEK1 native cysteine residues 33

2.1.3. Masking C207 with nucleotide and magnesium 37

2.2. Site-selective formation of MEK1Dha218 42

2.2.1. Over-alkylation of MEK1Dha218 with DBHDA in the presence of ADP-Mg complex 42

2.2.2. Alternative alkylating agents 43

2.3. A mild and selective formation of a dehydroalanine residue in a protein 47

2.4. An insight into the chemical mechanism of Dha formation 48

2.4.1. The different reactivity of bis-alkylating/elimination agents 48

2.4.2. Mechanistic insight into the formation of MEK1Dha222 50

Chapter 3: Mechanism of Elimination during Dehydroalanine Formation 53

3.1. The formation of a sulfonium ylid intermediate 53

3.1.1. Potential mechanisms of sulfonium ion elimination 53

3.1.2. The degradation of a stable sulfonium involved a sulfonium ylid intermediate 55

3.1.3. Visualisation of the sulfonium ylid in deuterated buffer 56

3.2. Site selective deuteration of cysteine residue in protein 58

3.2.1. Site selective protein deuteration strategy 58

3.2.2. Site-selective deuteration of SBLC156 60

3.2.3. Site-selective deuteration of histones H3C10 and H3C26 61

3.3. Sulfonium ylid as an intramolecular base 64

3.3.1. The fast solvent exchange hampered the mechanistic analysis 64

3.3.2. Site-selective deuteration of cysteine amino acid 65

3.3.3. Sulfonium ylid acted as a base in DMSO 67

3.3.4. DFT calculation and proposed E1cb mechanism 69

Chapter 4: Multiple Dha Formation on Protein 72

4.1. Crosslinking residues 218 and 222 in MEK1 72

4.1.1. DBHDA analogues crosslinking between cysteine residues 218 and 222 72

4.1.2. Lanthionine formation between position 218 and 222 75

4.2. Formation of double Dha using MDBP 77

4.2.1.Site-selective formation of MEK1Dha218Dha222 77

4.2.2. Versatility of MDBP: formation of p38αDha180Dha182 78

Chapter 5: Chemical phosphorylation and chemical activation of MEK1 84

5.1. MEK1pC222, mimics of MEK1pS222 84

5.1.1. Formation of MEK1pC222 84

5.1.2. Stability of the phosphocysteine residue in MEK1pC222 87

5.1.3. Assessment of the kinase activity of MEK1X222 constructs 87

5.1.4. Trial crystal structure MEK1pC222-ERK1-nucleotide Mg complex 92

5.2. Chemical activation of MEK1 at position 218, mimics of MEK1pS218 and MEKpS218pS222 95

5.2.1. Chemical phosphorylation of MEK1 at site 218 95

5.2.2. Activation of MEK1 at position 218 97

5.2.3. Crystallisation of MEK1 activated at position 218 99

5.3. Alternative approaches towards chemical phosphorylation of MEK1 101

5.3.1. Alternative thiamichael reagent 101

5.3.2. Stereo-and-site-selective phosphocysteine formation approach 102

5.3.3 Site-and-stereoselective and stable phosphocysteine mimics formation approach 104

5.3.4. Site-selective formation of stable phosphylated residues by radical reaction 106

CONCLUSION 108

Chapter 6: Conclusion 109

EXPERIMENTAL 113

Chapter 7: Procedures for the Site-Selective Chemical Activation of MEK1 114

7.1. Protein visualisation tools and general computational calculations 114

7.1.1. Molecular graphics 114

7.1.2. Protein properties calculation 114

7.2. Protein reaction monitoring and characterisation 115

7.2.1. Intact protein mass spectrometry 115

7.2.2. Analysis of digested proteins by LCMS-MS 119

7.2.3. Circular dichroism (CD) spectroscopy 123

7.3. Small molecule <500 Da reaction monitoring and characterisation 124

7.3.1. Nuclear magnetic resonance (NMR) spectroscopy 124

7.3.2. IR Spectroscopy 125

7.3.3. Mass spectrometry 125

7.4. Plasmids and protein expression and purification 126

7.4.1. General consideration for protein expression and purification 126

7.4.2. Plasmids and cloning 128

7.4.3. Protein expression and purification 128

7.5. Synthesis of small molecules as tools in the site-selective chemical phosphorylation of protein 133

7.5.1. General chemical considerations 133

7.5.2. Synthesis of bisalkylating/elimination agents 133

7.5.3. Synthesis of alkylating phosphorylated agents 140

7.6. Chemical protein modifications towards site-selective chemical phosphorylation of MEK1 141

7.6.1. General protein chemistry and reagents considerations 141

7.6.2. Treatment of MEKC222 with Ellman's reagent 141

7.6.3. Experiments towards the site-selective formation of MEK1Dha222 142

7.6.4. Experiments towards the site-selective formation of MEK1Dha218 155

7.6.5. Experiments towards the site-selective formation of MEK1Dha218Dha222 160

7.6.7. Alternative methods towards the formation of p38αDha180Dha182 167

7.6.8. Formation of MEK1pC222 168

7.6.9. Stability of MEK1pC222 172

7.6.10. Formation of MEK1pC218 and MEK1pC218pC222 173

7.6.11. Alternative methods of chemical phosphorylation of MEK1 175

7.6.12. Radical addition of $CH_2PO_3^{2-}$ and $CF_2PO_3^{2-}$ 177

7.7. Kinase activity of MEK1 constructs 179

7.7.1. Progress curve assay 179

7.7.2. Processing and results 180

7.7.3.Optimisation of the assay 181

7.8. Active MEK1 crystallisation 182

7.8.1. Attempt of crystallisation of MEK1pC222-ERK2-AMPNP 182

7.8.2. Crystallisation of truncated ΔMEK1 constructs 182

7.8.3. Proteolytic cleavage of MEK1 at Y300 183

Chapter 8: Procedures for the Sulfonium Elimination Mechanistic Elucidation on Protein 185

8.1. Protein reaction monitoring and characterisation 185

8.2. Chemical protein modification towards the elucidation of the mechanism of elimination in Dha formation 187

8.2.1. General protein and reagents considerations 187

8.2.2. Decomposition of GFPSul147 187

8.2.3. Formation of perdeuterated SBLDha156 188

8.2.4. Formation of SBLdAcc156 and attempt of deprotection to SBLdC156 189

8.2.5. Formation of SBLNbc156 190

8.2.7. Formation of H3dC10 191

8.2.8. Formation of H3dC26 194

8.2.9. Optimisation of the deuterium incorporation 195

8.3. Small molecule synthesis used in the site-selective chemical phosphorylation of protein 195

8.3.1. Synthesis of AcNHdCysOMe 195

8.3.2. Formation of AcNHLanOMe from the reaction of AcNHDhaOMe with Na_2S 198

8.3.3. Formation of AcNHdNbcOMe from AcNHDhaOMe with a photolabile thiol 198

8.3.4. Formation of THTCD 202

8.4. Mechanistic study by the detection of THTCD deuteration pattern with various substrate and solvent 203

8.4.1. Detection method of THTCD 203

8.4.2. Reaction SBLC156 and DBHDA in deuterated buffer 204

8.4.3. Reaction of H3C10 with DBHDA in deuterated buffer 206

8.4.4 .Reaction of SBLdC156 and DBHDA in non-deuterated buffer 207

8.4.5. H3dCX and DBHDA in water 208

8.4.6. AcNHCysOMe and DBHDA in D2O 209

8.4.7. AcNHdCysOMe and DBHDA in water 210

8.4.8. Control experiment with THTCD in deuterated buffer 210

8.4.9. Reaction of AcNHdCysOMe with DBHDA in DMSO-d6 211

8.4.10. Reaction of AcNHdCysOMe with DIB in DMSO-d6 211

8.4.11. DFT calculation of elimination mechanism 212

BIBLIOGRAPHY 215

Chapter 9: Bibliography 216

Appendix A: Naccess Analysis I

Appendix B: Protein Purification, Characterisation and Sequences L

Appendix C: Circular Dichroism of MEK1C218, MEK1C222 and MEK1C218C222 LX

Appendix D: NMR spectra for small molecules used in the Site-selective Activation of MEK1 LXI

Appendix E: NMR spectra for small molecules used in the mechanism elucidation LXVI

Appendix F: Progress curves of the phosphorylation of ERK1 by MEK1 constructs LXXVI

Appendix G: DFT calculations LXXVIII

Appendix H: Deuterium incorporation and optimisation LXXXI

INTRODUCTION

Chapter 1: Introduction to MEK1 phosphorylation and activation

1.1. Activation of MEK1 kinase by phosphorylation

1.1.1. Protein phosphorylation

Protein phosphorylation is a major post-translational modification (PTM), which affects dramatically the structure and function of proteins.[1-7] As such this PTM plays a pivotal role in signal transduction and in the regulation of cellular events.[8] The phosphorylation of serine, threonine and tyrosine[9] residues and its effect has been widely studied. Despite their inherent instability, the phosphorylation of other residues like cysteine, histidine, aspartic acid, lysine and arginine residues, has also been observed. For instance, the phosphorylation of cysteine residues, in *E. coli* and other bacteria, is essential in the mechanism of phosphorylation of glucose,[10] and also modulates bacterial virulence and antibiotic resistance.[11] Phosphocysteine intermediates are also formed during the dephosphorylation mechanism by eukaryotic protein tyrosine phosphatases.[12] It has also been demonstrated that the phosphorylation of histidine and aspartic acid was involved in the signal transduction in bacteria, as part of a two component systems phospho-relay.[13] Finally, few examples of the phosphorylation of lysine and arginine residues have been reported.[14-16]

The use of phosphocysteine residues (pC) to mimic phosphoserine residues (pS), which leads the phosphorylation of tyrosine and threonine residues, is described in this thesis.

1.1.2. MEK, a dual specific Protein kinases

Protein phosphorylation is a transient modification regulated by two families of proteins. While protein phosphatases catalyse the hydrolysis of the phosphate group attached to amino acid residues, protein kinases facilitates the transfer of the γ-phosphate of ATP to these same residues (Figure 1.1.A).

The human genome contains around 560 kinase genes divided into 134 families.[17] The dual specificity kinases (DSK) belong to a small family of kinases that phosphorylates serine/threonine residues as well as tyrosine residues.[18] This thesis focuses on the Mitogen-activated protein kinase kinase 1 (MAPKK), Extracellular signal-regulated Kinase (ERK) activator Kinase 1, MEK1. Once activated, MEK1 phosphorylates its only known substrate ERK1, at position T202 and Y204 (Figure 1.1.D & 1.2.B).[19] Active MEK1 structure has not been reported yet. However, a general MEK1 mechanism of phosphorylation can be described based on the conserved amino acid residues that compose the catalytic core of protein kinases.[20] Schematically, ATP binds in the ATP binding cleft of the protein, referred hereafter to “ATP pocket”, which contains an aspartic acid residue (D208) that stabilises the nucleotide *via* salt bridges with magnesium (Figure 1.1.B & C). The phosphorylation of residues in the activation loop is believed to affect the structure in 2 ways. The orientation of ATP is finely tuned by an additional salt bridge between a conserved lysine residue (K97) of the AXK-motif of the β3 strand and the glutamic acid (E114) belonging to the C-helix (Figure 1.1.B & C). The phosphate transfer is further catalysed by D190 residue of the conserved catalytic loop (β6/β7), which functions as a base acceptor for the proton transfer during the reaction (Figure 1.1.C).

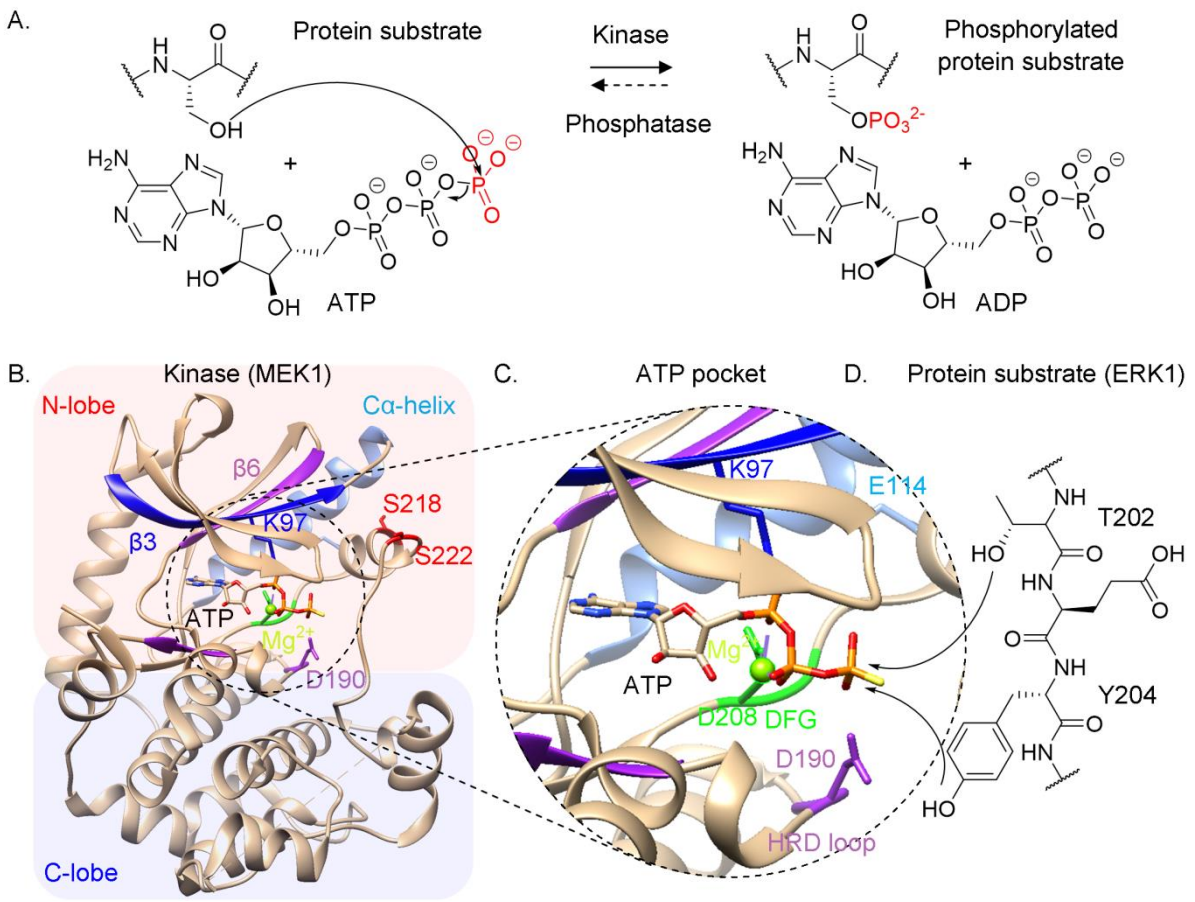


Figure 1.1. A. Phosphotransfer reaction between ATP γ-phosphate and a protein residue (here serine) catalysed by a kinase to form the phosphorylated protein and ADP, the dephosphorylation is catalysed by a phosphatase. B. General structure of a kinase bound to ATP (base on inactive MEK1-ATPγS-Mg, PDB 3EQD) showed two lobes, the N-lob in red and C-lobe in blue. The key residues and secondary structures involved in the binding of ATP and the activation of the phosphorylation are highlighted. C. The zoomed in "ATP pocket" shows the salt bridges between ATP-Mg and the protein. D. MEK1 phosphorylates ERK1 T202 and Y204.

1.1.3. Phosphorylation of MEK1 kinase

Phosphoproteomic studies revealed 12 "phosphorylatable" sites for MEK1.[21] However, the activation of MEK1 requires the phosphorylation of the two residues, S218 and S222, by upstream kinases like the Rapidly Accelerated Fibrosarcoma (Raf) kinases family in human (Figure 1.2.A & B),[22-25] MEKK1 in bacteria,[26] and Mos[27] in plants.

While the specific and detailed structural implications of MEK1 phosphorylation at S218 and S222 residues are not reported, the consequences of MEK1 activation have been identified. Considered as the "gate-keeper" of the RAF-MEK-ERK signalling pathway, involved in cell

differentiation and proliferation (Figure 1.2.A),[19, 28] MEK1 has logically attracted a considerable interest as a target towards cancer therapy, especially for cancer presenting mutations on BRAF, a member of the RAF family.[29] Trametinib (GSK1120212)[30] and cobimetinib (GDC-0973/XL518) (Figure 1.2.C),[31] allosteric MEK1 inhibitors, were approved for the treatment of BRAF-mutated (V600E) melanoma in combination with BRAF inhibitors. Interestingly, trametinib displays higher affinity for non-activated unphosphorylated MEK1,[32] whereas cobimetinib binds strongly to pMEK, making it highly potent in inhibiting BRAF(V600E) melanoma, since mutant BRAF constantly produces high levels of pMEK.[33] A methodology that enables the formation of the different phosphorylated state of MEK1 is therefore a particularly valuable tool towards the design of novel and potent MEK1 inhibitors.

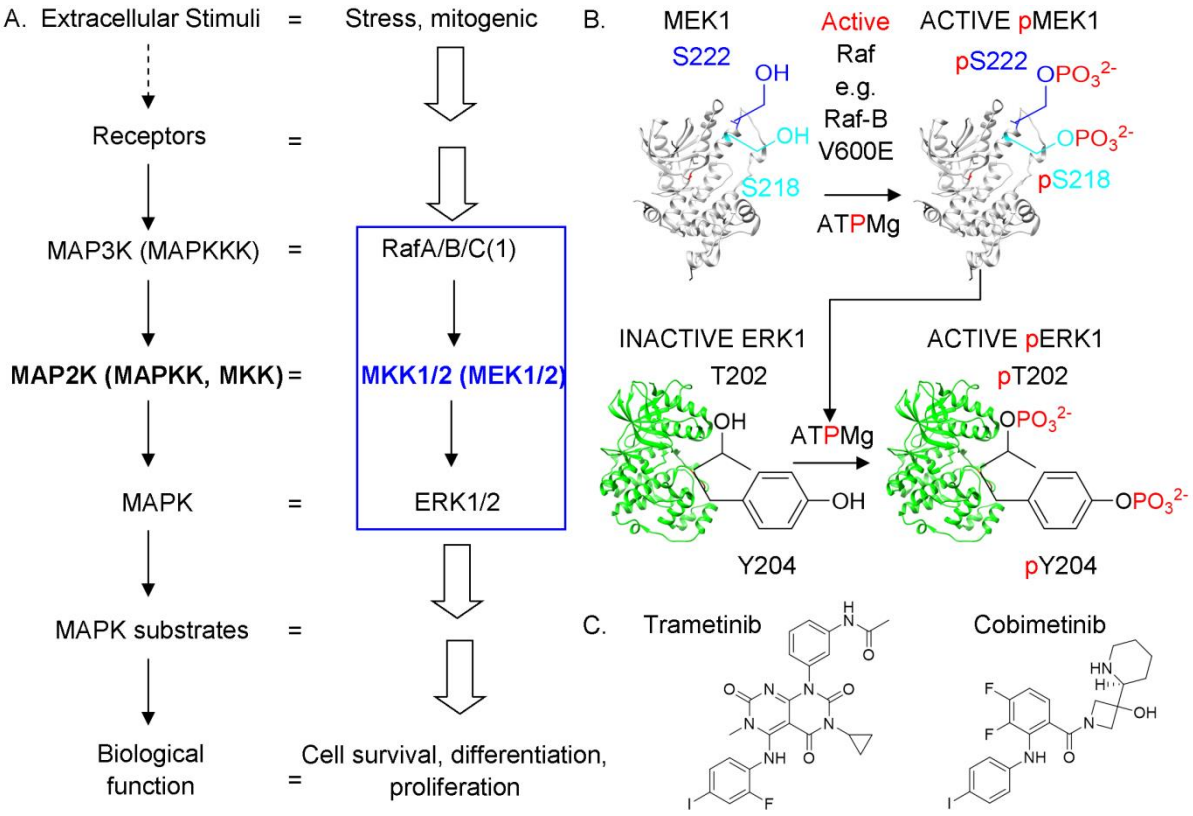


Figure 1.2.A. Schematic RAF-MEK-ERK signalling pathways. B. Phosphorylation of MEK1 at position S218 and S222 by Raf activates MEK1 to phosphorylate ERK1 at position T202 and Y204. C. Structures of 2 MEK1 allosteric inhibitors approved by the FDA to date.

1.1.4. Project synopsis

This thesis describes the development of a chemical methodology, which enables the site-and-regioselective phosphorylation of full length MEK1, *in vitro*. We believe that the use of this methodology is applicable to other kinases and proteins, and should facilitate the structural and dynamic study of protein phosphorylated and activated kinases by X-ray crystallography and NMR spectroscopy.

In the following part, the different biological methods to afford activated MEK1, phosphorylated or not, are described.

1.2. Biological methods towards the controlled activation and phosphorylation of MEK1 kinase

1.2.1. Preparation of constitutively active MEK1 by site directed mutagenesis

In the mid-nineties, the elucidation of MEK1 activation by the phosphorylation of S218 and S222 residues was performed by replacing these serine residues with alternative natural amino acid residues by site-directed mutagenesis (SDM).[22-24] Firstly, alanine scanning showed that the individual or double mutations S218A and S222A decreased MEK1 enzymatic activity, especially after subsequent MEK1 phosphorylation (Figure 1.3.A). It is worth noting that wild-type MEK1 possesses some basal activity. Interestingly, Zheng *et al.* concluded that the phosphorylation of both S218 and S222 was essential for MEK1 activation.[22] On the other hand, further S218T and S222T mutations studies combined with phosphorylation and dephosphorylation of these threonine residues demonstrated that the phosphorylation of only one residue was necessary for MEK1 activation.[23, 24] Finally, the introduction of a negative charge(s) created by S218D/E and S222D/E mutations was sufficient to activate MEK1 (Figure 1.3.A).[23-25] The exploitation of these so-called

constitutively active MEK1 constructs was crucial for the elucidation of MEK1 mechanism of action[34-36] and function.[28, 37-39]

However, the structure of activated MEK1 has yet to be resolved by X-Ray or NMR. The unique example of constitutively active MEK1 crystal structure, MEK1E218E222, in presence of diverse ligands (PDBs: 4AN2, 4AN3, 4AN9, 4ANB) exhibited no significant difference with inactive wild type MEK1 (MEK1wt) (Figure 1.3.B).[31] This result could be explained by the weak effect of a carboxylic acid compared to a phosphate moiety, which does not favour the stabilisation of the active conformation.

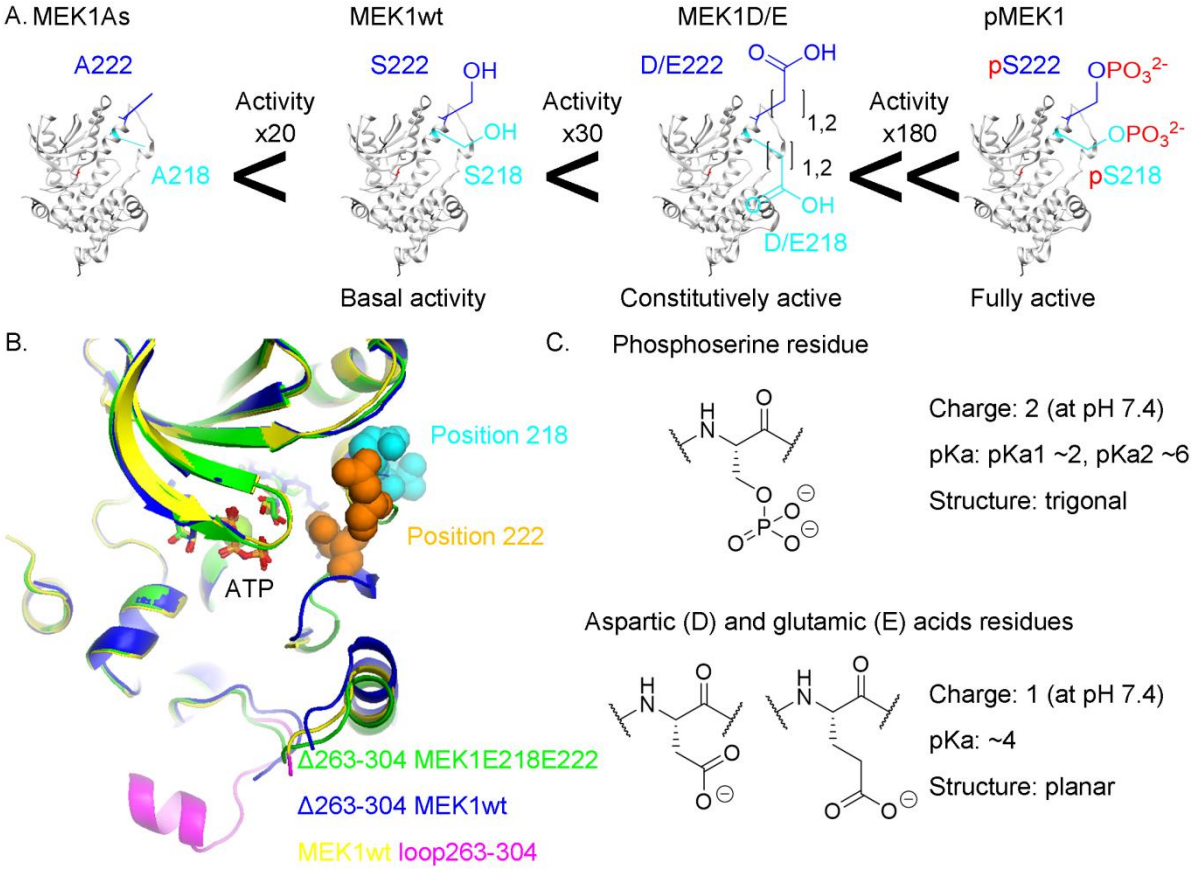


Figure 1.3.A. Comparative activity of MEK1 constructs. B. Overlay of truncated Δ263-304 MEK1E218E222 (green, PDB 4AN3[40]), truncated Δ263-304 MEK1 wt (blue, PDB 3WIG[41]) and MEK1 wt (yellow 3VVH) shows no significant difference between constitutively active and wild type MEK1. 218 and 222 are highlighted in cyan and orange respectively. The truncated loop (pink) is far away from the activation loop. C. Comparison of phosphoserine and carboxylic acid residues.

Despite their historical importance, constitutively active MEK1E/D does not appear as an ideal construct towards structural and dynamic studies of activated MEK1. Indeed, not only are constitutively active MEK1 constructs marginally more active that MEK1 wt,[23] but also carboxylic acids are imperfect bioisosteres of phosphates at physiological pH by their overall charge (-1 vs -2), pKa (4 vs ~2/~6) and geometry (planar vs trigonal) (Figure 1.3.C).[42, 43]

The expression and purification of pMEK1 has therefore been attempted.

1.2.2. In situ and in vivo pMEK1 expression and purification

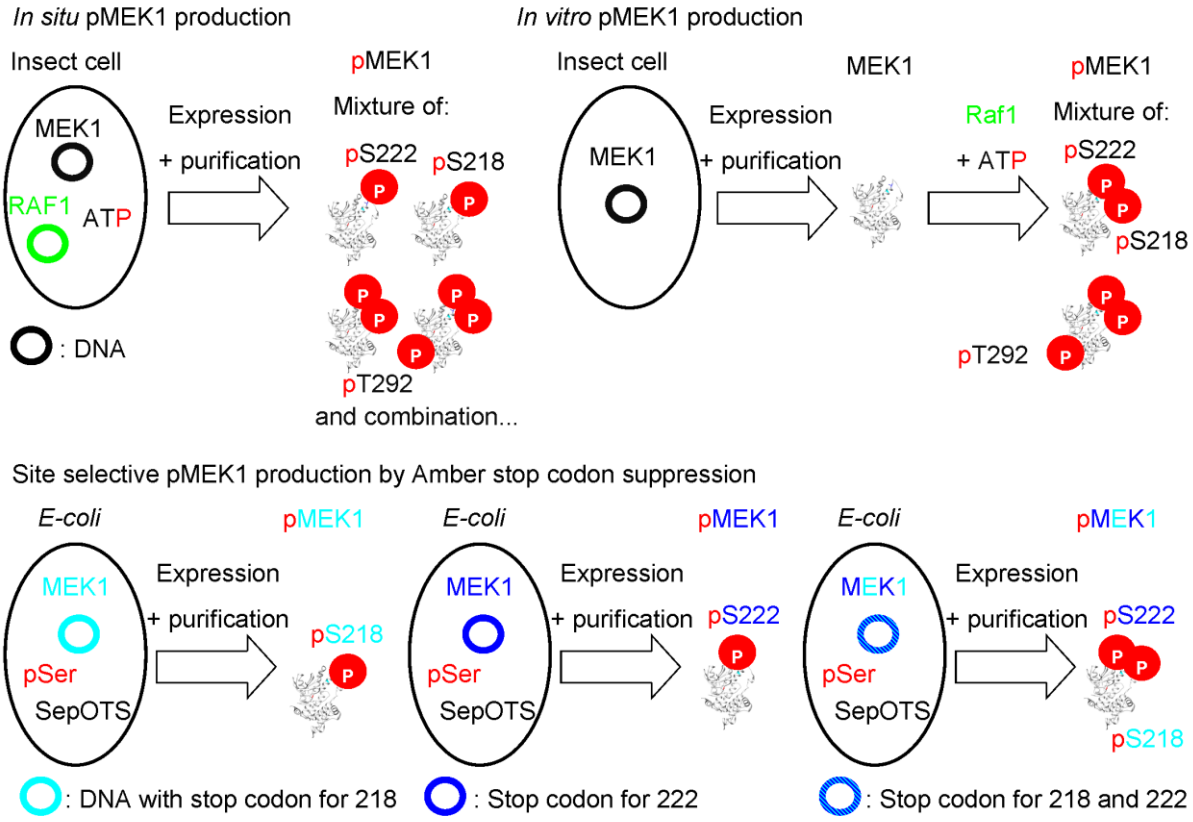
As previously described in paragraph 1.1.3, MEK1 is phosphorylated at position 218 and 222 by Raf *in vivo*. The *in situ* co-expression of MEK1 with baculo-virus Raf1 in insect cells, in presence of okadaic acid, a phosphatase inhibitor, generates phosphorylated MEK1 in good yield (5 mg/L, 27 mg after a 4 purification cycle).[44] LCMS and Western blot analysis confirmed the phosphorylation at position 218 and 222. However, LCMS clearly indicated a 1.8:2:1 mixture of singly, doubly and triply phosphorylated MEK1 (Figure 1.4).

Comparatively, the phosphorylation of purified MEK1 *in vitro* by incubation with constitutively active Raf1 Y340D/341D affords a mixture of doubly and triply phosphorylated MEK1: no single phosphorylation was observed.[45] While LCMS of tryptic peptides and Western blot analysis also confirmed the phosphorylation of S218 and S222, no indication of the phosphorylation ratio was mentioned in this case (Figure 1.4).

Both methods were used towards the understanding of binding of MEK1 and pMEK1 with inhibitors[45] or nucleotides.[46] Interestingly, the binding affinity of the 2 inhibitors studied, U0126 and PD032590, was different for *in vitro* pMEK1 and *in situ* pMEK1. This would indicate the importance of understanding the effect of each phosphorylation.

The lack of selectivity is a drawback towards X-ray structural studies, as the purity and homogeneity of the protein are crucial for the success of its crystallisation.[47] Amber stop codon technology appears as an attractive alternative to selectively phosphorylate MEK1.

The rapid improvement of this technology observed in the last decade, allowed the direct and selective incorporation of pSer residues into various proteins and MEK1 especially (Figure 1.4).



SepOTS: phosphoserine (Sep) orthogonal transcription system containing SeptRNA synthetase, EF-Sep, and tRNASep.

Figure 1.4. Schematic representation of biochemical phosphorylation of MEK1.

1.2.3. Insertion of pSer by amber stop codon suppression

In the genetic code, the 20 natural amino-acids that shape proteins are encoded by three based pair codons from the DNA (Figure 1.5.A). The expansion of this genetic code by amber stop codon suppression technology enabled the incorporation of unnatural amino-acids.[48] Schematically, a stop codon, typically the amber stop codon UAG, is used to encode for an unnatural amino acid (UAA) (Figure 1.5.A & B). More specifically, a readable orthogonal $tRNA_{anticodon}$, $tRNA_{CUA}$, is engineered and linked to the desired unnatural amino

acid either chemically or enzymatically with the use of aminoacyl-tRNA synthase and transferase (Figure 1.5.B). Using this methodology, phosphohomoalanine (Pha), a stable phosphoserine mimic, and phosphotyrosine and its mimic, phosphomethylenephenylalanine (Pmp) (Figure 1.5.C) were incorporated to the firefly luciferase protein.[49] These unnatural phosphoresidues are stable under the incorporation conditions. However, it is worth mentioning they are only mimics, and similarly to carboxylic acids, their structure and physicochemical properties differs from their phosphoresidues counterparts, e.g. pKa_2 ~6 for pSer and ~7.5 for Pha.[50] The formation and use of fluorinated derivatives like phosphodifluorohomoalanine (Pfa), and difluoromethylenephosphonophenylalanine (F_2Pmp) appear as a better alternative.[51] Another phosphotyrosine mimic, p-Carboxymethyl-L-phenylalanine (pCMF) (Figure 1.5.C) was successfully incorporated into the transcription factor, STAT1, in place of pY701 triggering its activation.[52] The limitations of using carboxylic acids as mimics of phosphates have been highlighted earlier (Figure 1.3.C). Therefore, the incorporation of caged-phosphoresidues (Figure 1.5.C), which possess a photocleavable-protecting group that masks the phosphate negative charge, and was nicely exemplified with the vasodilator-stimulated phosphoprotein, VASP.[53, 54] Interestingly, caged-phosphoresidues not only provide the desired phosphoresidues *in situ* after deprotection, but also act as a switch that enables the study of the effect of phosphorylation in real time. This study also claims successful incorporation of pSer, pThr and pTyr, also the incorporation yields were very low (Figure 1.5.C). Very recent advance in stop codon technology permitted the direct insertion of phosphotyrosine[55-57] and, more relevant in our case, phosphoserine (Sep) in high incorporation yield.

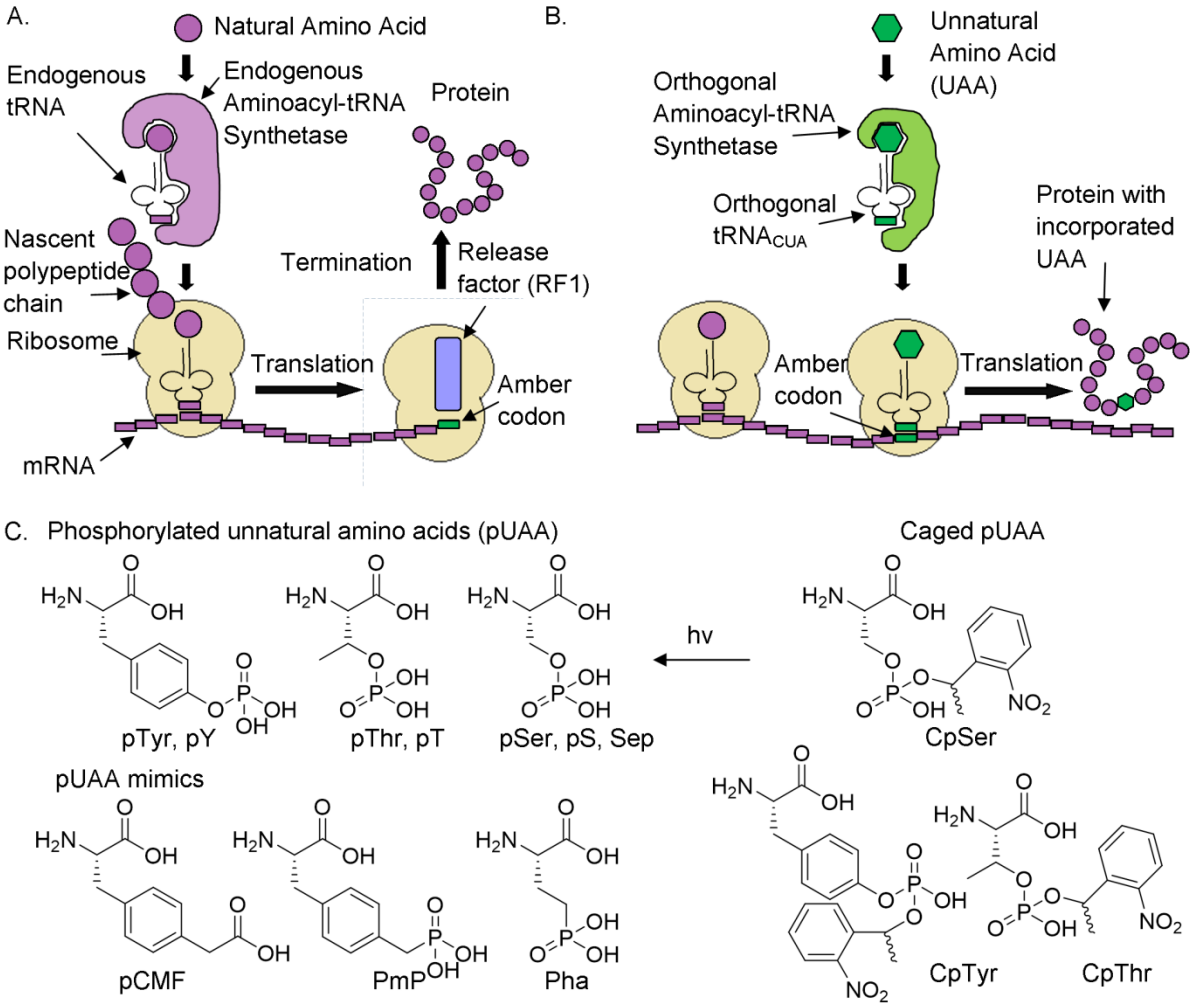


Figure 1.5.A. In normal translation for amber codon, the release factor 1 (RF1) recruited to terminate the translation of the polypeptide releasing the newly synthesized protein. B. Amber codon suppression, the amber codon is used as coding codon in translation to incorporate unnatural amino acids (UAA). The complementary amber $tRNA_{CUA}$ is acylated by an orthogonal aminoacyl-tRNA synthetase (aaRS) that is specifically designed to accept only unnatural amino acids to the incorporation of UAA. B. Examples of phosphorylated UAA, phospho-serine (pSer, pS, Sep), -threonine (pThr, pT) and –tyrosine (pTyr, pY), caged CpSer, CpThr, CpTyr and mimics p-carboxymethyl-L-phenylalanine (pCMF), phosphonomethylphenylalanine (Pmp), phosphohomoalanine (Pha).

The technology towards the insertion of uncaged phosphoserine residues (Sep) residues is based on the discovery that some methanogenic archaea incorporate cysteine residues into their proteins in two steps (Figure 1.6.A).[58] A phosphoserine residue (Sep) is first inserted to the $tRNA_{GCA}$ using the phosphoseryl-tRNA synthetase (SepRS) to afford $tRNA^{Sep}_{GCA}$. The residue is then converted by a second enzyme, SepCysS to cysteine residue (Cys) to afford $tRNA^{Cys}_{GCA}$. The incorporation of a cysteine over phosphoserine residue in protein is far more efficient as $tRNA^{Sep}_{GCA}$ seems a poor substrate for the elongation factor EF-Tu (Figure

1.6.A). In 2011, an *E-coli* strain containing an elongation factor EF-Sep and $tRNA^{Sep}_{UAG}$ was engineered to improve the incorporation of the phosphoserine residue. It led to the successful insertion of phosphoserine in myoglobin and more importantly on MEK1 kinase at position 222 (Figure 1.6.B).[59] However, the incorporation yields were very low due to the orthogonal transcription system (SepOTS) containing serine phosphatases and the competition with release factor RF-1. Therefore, an *E-coli* strain lacking the release factor RF-1, serine phosphatases and where 7 UAG codons in essential genes were converted to UAA, was created. This new *E-coli* strain slightly improved the pSer incorporation into GFP and WNP4, and was used successfully in histone H3.[60, 61] However, the yields of incorporation were still particularly low due to cell viability. Recent advance saw the incorporation of phosphoserine and phosphohomoalanine (Pha) (Figure 1.5.C), phosphoserine mimic, stable to phosphatase incorporated on NEK7 kinase (Figure 1.5).[62] Rinehart laboratory also developed a cell free protein synthesis (CFPS) system lacking the phosphatase pSer, and a SepOTS (orthogonal transcription system: SepRS, $tRNA^{Sep}$ and EF-Sep) with recoded *E-Coli* allowing the formation of MEK1pS218, pS222 and pS218pS222 (Figure 1.6.B).[63, 64] Indeed, the CFPS looked very promising as the yields for the 3 constructs were around 300 µg/mL for 15 µL-scale expressions. Scale up to 8x300 µL provided around 1 mg of MEK1pS218pS222. However, the product was not completely pure. Nevertheless, this technology looks very promising and the reagents are now available in Addgene as a kit for improved expression of recombinant phosphoproteins.

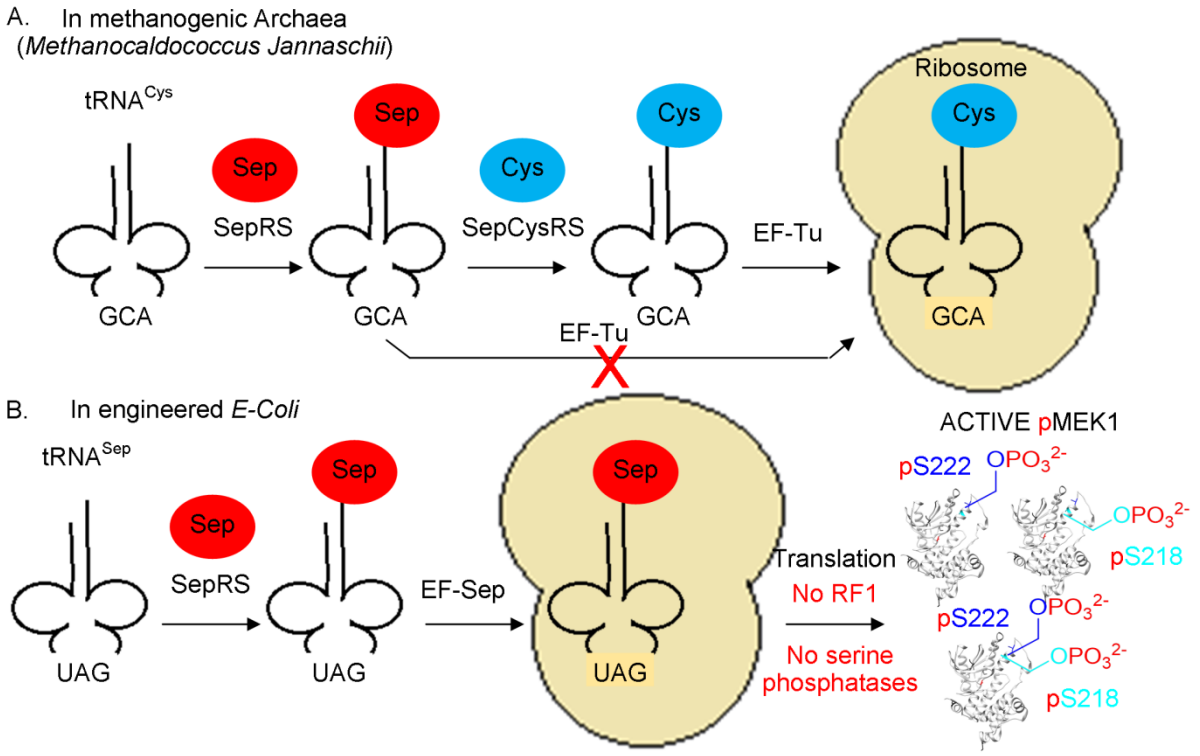


Figure 1.6.A. The incorporation of cysteine in some archaea involves first the use of phosphoseryl-tRNA synthetase (SepRS) to afford $tRNA^{Sep}$, which is converted to $tRNA^{Cys}$ by the SepCysRS enzyme. The elongation factor, EF-Tu, binds preferentially $tRNA^{Cys}$ to the ribosome. B. Engineered $tRNA^{Sep}$, elongation factor EF-Sep and *Ecoli* without the release factor, RF1 and serine phosphatases allowed the formation of MEK1pS218, MEK1pS222 and MEK1pS218pS222.

The efficacy and selectivity of the biological methods to activate MEK1 have recently grown exponentially from constitutively active MEK1 mimics to the formation of homogeneous MEK1 phosphoforms. However, the yields do not allow the production of 10 mg of pure phosphorylated MEK1 needed for NMR or X-ray structure studies yet. Moreover, obtaining pure and multiply phosphorylated MEK1 is still tricky.

Finally, the selective pSer incorporation was still a curiosity when my project started. The potential chemical methods towards the phosphorylation/activation of MEK1 were therefore assessed.[65]

1.3. Chemical phosphorylation of protein

1.3.1. (Semi)synthetic approaches to phosphorylated protein

The total synthesis of Human statherin, a low molecular weight phosphoprotein (5 kDa) containing phosphotyrosine residues, has been performed using solid-phase peptide synthesis (SPPS) (Figure 1.7).[66] However, due to MEK1 size (around 40 kDa), and the need to refold the protein afterwards, this technique was not applicable to our purpose.

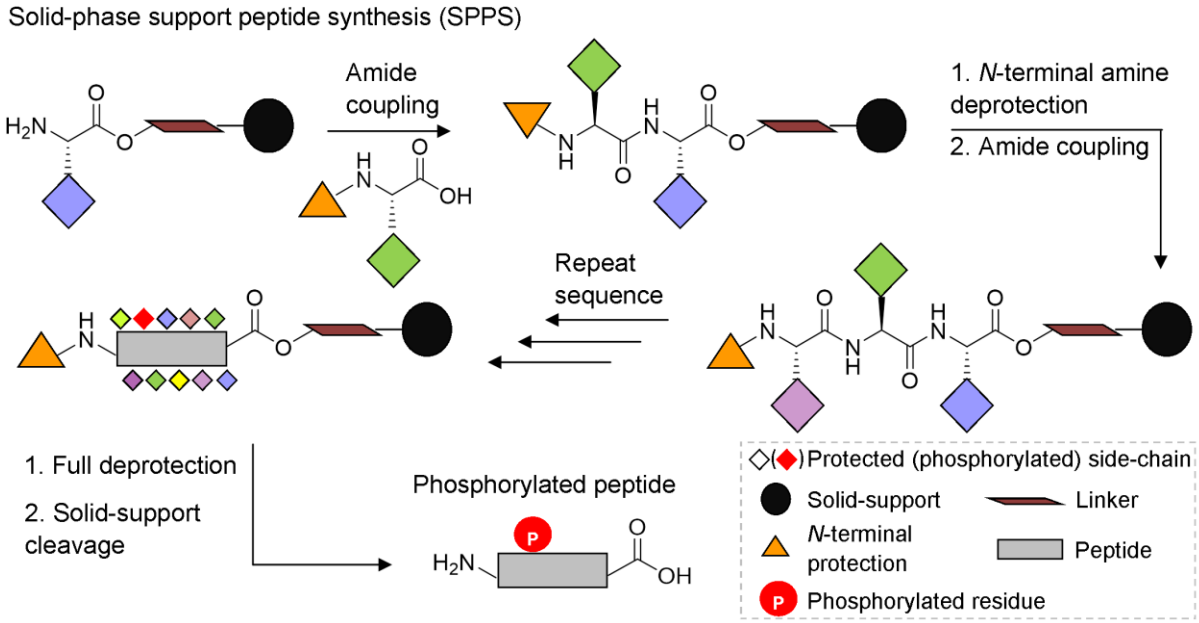


Figure 1.7. General principle of solid-phase support peptide synthesis for the synthesis of phosphorylated peptides.

The development of native chemical ligation (NCL), where a peptide, possessing a thioester moiety, reacts with another peptide containing a terminal cysteine residue, allowed the formation of bigger peptides and proteins (Figure 1.8.A). Some examples of phosphorylated proteins are histone H2B (14 kDa),[67] α-synuclein (14 kDa),[68] and phosphorylated ubiquitin analogues (17 kDa).[69] The size of the protein can be improved by expressed protein ligation (EPL). Exploiting intein-mediated protein splicing, large polypeptides containing a *C*-terminal thioester are expressed, and ligated to a synthetic peptide bearing an *N*-terminal cysteine residue in an NCL fashion (Figure 1.8.B).[70] Numerous phosphorylated proteins were synthesised using this technique: the phosphatases SHP-1 and SHP-2 (70 kDa),[51, 71, 72]

histones H3 (11 kDa),[73] and H2B (14 kDa),[74] SmadMH2 protein (26 kDa),[75] tau (46 kDa),[76] and C-terminal Src kinase (50 kDa).[77]

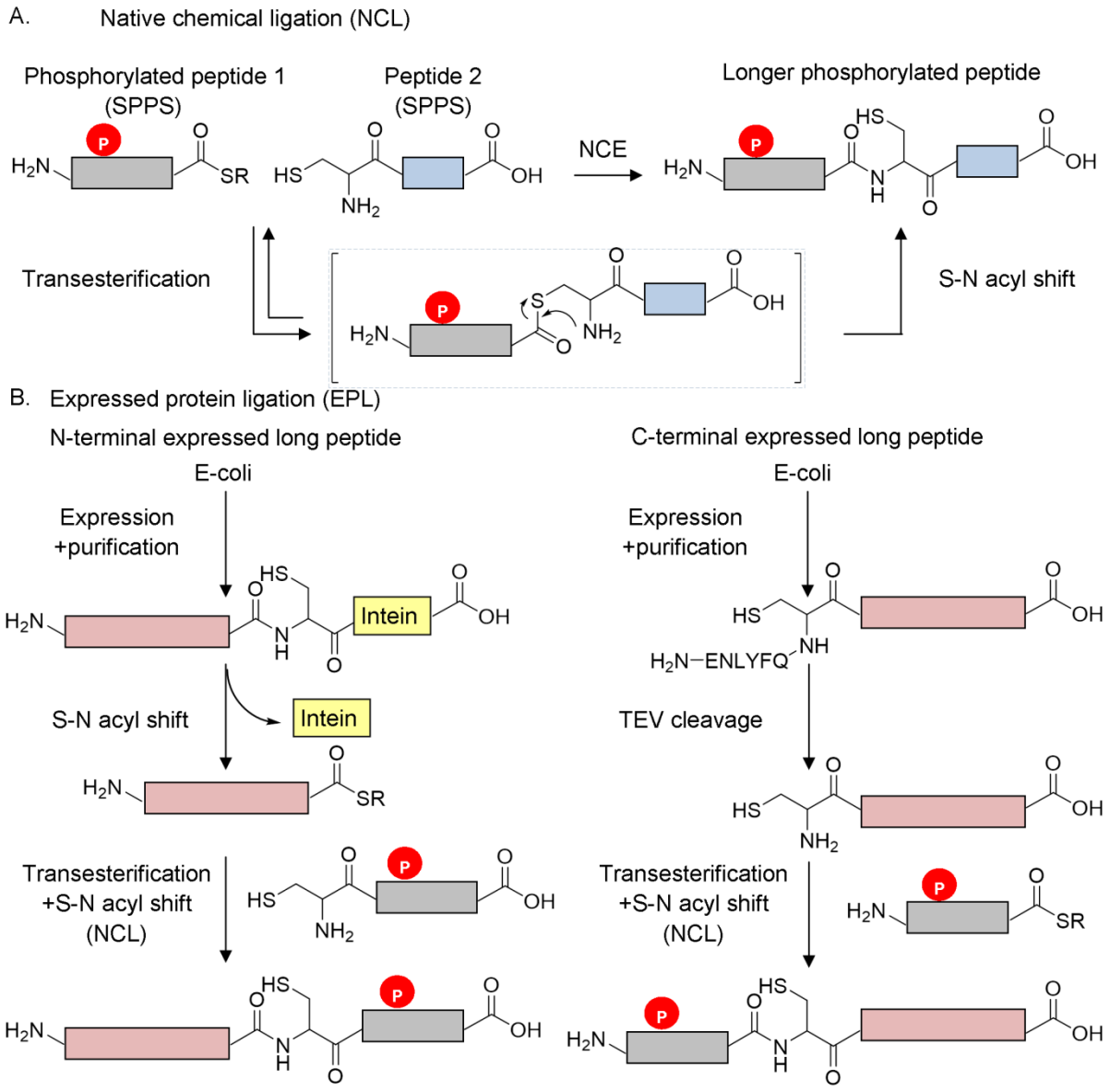


Figure 1.8.A. Principles of native chemical ligation (NCL) for the synthesis of phosphorylated peptides; B. Principles of express protein ligation (EPL) for the synthesis of phosphorylated peptides and proteins.

In all these examples, the phosphorylated residue is introduced in the C-terminal region. In MEK1, S218 and S222 are in the centre of the protein. The combination of both *N*-terminal and *C*-terminal ligation (Figure 1.8.B) could resolve this issue. A three-component synthesis can also be envisaged.[78] But more importantly, the need to refold the protein to its natural state, crucial for its activity, is a major limitation. Finally, both methodologies require the

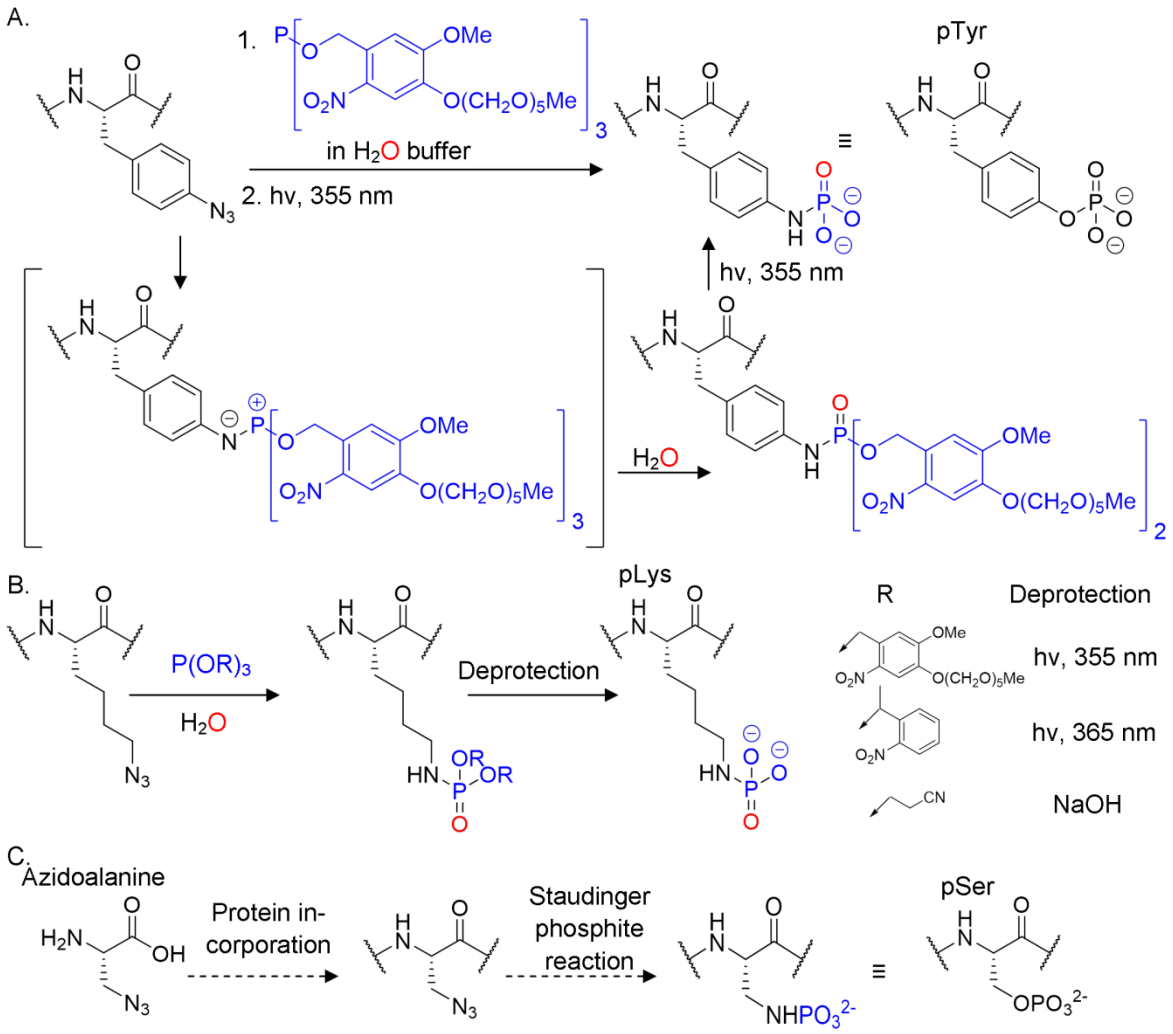
formation of (caged/protected) phosphopeptides and aminoacids.[79] Their synthesis is well-established.[49,53,80] Finally, the direct phosphorylation of the full peptide using diarylphosphorochloridate or phosphoramidite reagents on the serine, threonine or tyrosine residues is possible but needs a selectively orthogonal method for the removal of the protecting groups. Unwanted reactions were observed during deprotection in the synthesis of phosphotyrosine peptides. Indeed, the final deprotection of the protected phosphotyrosine residue led to an *N* to *O*-Acyl migration at a threonine residue contained in the peptide.[81]

The direct chemical phosphorylation of full length recombinant protein appears as a more amenable alternative.

1.3.2. Chemical phosphorylation of full-length protein

The direct (thio)phosphorylation of full length recombinant protein with (thio)phosphoramidite is not selective, and targets histidine residues.[82-84]

In order to phosphorylate a full length protein selectively, (bioorthogonally) reactive amino-acid residues were inserted. The insertion of a bioorthogonal azide moiety followed by a Staudinger-phosphite reaction afforded peptides and a protein containing a phosphotyrosine mimic (Scheme 1.1.A).[85] The same technology was applied with an alkyl azide group to form peptides bearing a phospholysine residue (pLys) (Scheme 1.1.B).[86, 87] The use of this technology, to lead to the corresponding phosphoserine mimic, necessitates 3-azido-L-alanine (Scheme 1.1.C). Despite its availability, the incorporation of azidoalanine into proteins is challenging. Cells auxotrophic to methionine have been used to that purpose with moderate success.[88, 89] As an alternative, the incorporation and use of a reactive natural amino acid has been envisaged.



Scheme 1.1.A. Staudinger-phosphite reaction allowed the formation of pTyr mimics and B. pLys mimics. C. Strategy towards the incorporation of a phosphoserine mimic from the Staudinger-phosphite reaction.

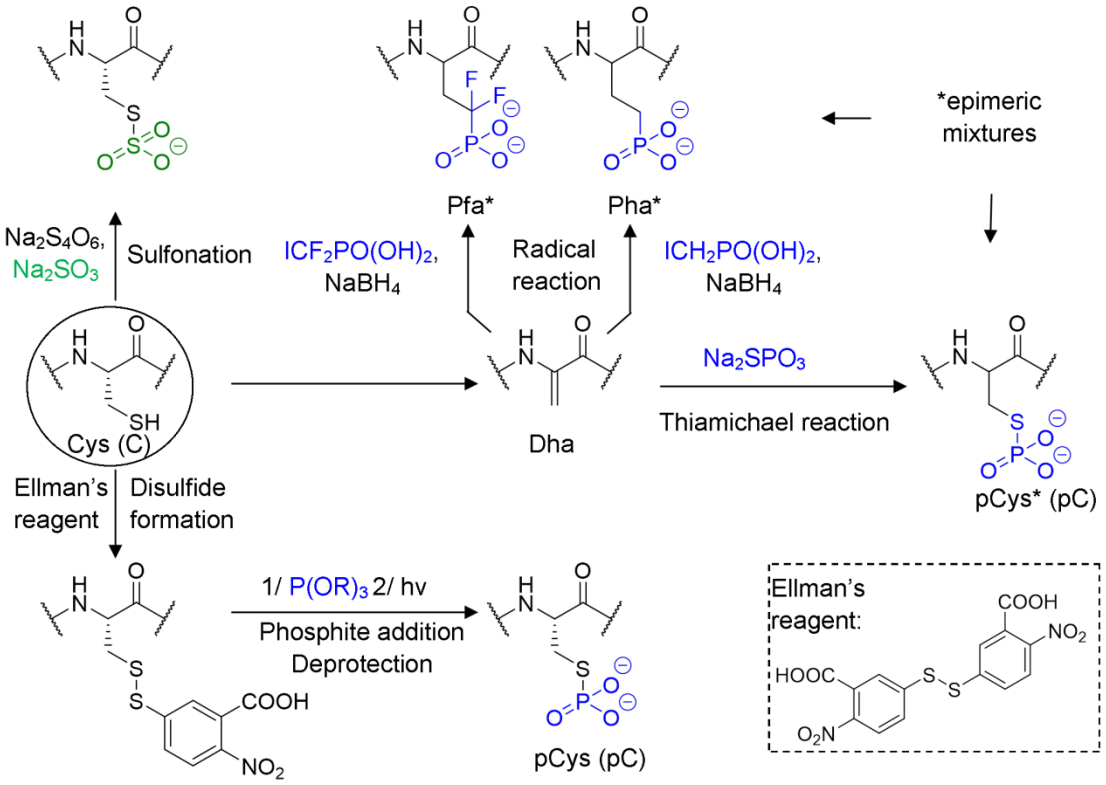
Naturally occurring cysteine residues (Cys) are relatively rare, characteristically around 2% mammalian proteins,[90] and reactive. The genetic incorporation of Cys by site-directed mutagenesis is well-established, even though the addition of an extra Cys can potentially scramble the folding in proteins containing multiple Cys. MEK1S218C was expressed. Its sulfonation provided a mimic of MEK1pS218 (Scheme 1.2).[91] However, the thiosulfate was not stable to reducing agent. In addition, this chemical modification was not investigated on the site 222 or on both sites.

Recently, the propensity of cysteine residues to form disulphide bonds was exploited. The treatment of a peptide, containing a cysteine residue, with Ellman's reagent affords an assymetrical disulphide bond, which leads to the formation of a phosphocysteine residue

when treated with a phosphite, in a Staudinger-phosphite fashion (Scheme 1.2).[92] The development and limitations of the method will be discussed Chapter 5.

Phosphocysteine residues can also be inserted from a thiamichael reaction between dehydroalanine residues (Dha) and sodium thiophosphate (Na_2SPO_3) (Scheme 1.2). The incorporation of Dha into protein is detailed next paragraph 1.3.3. Using this method, the incorporated phosphocysteine residue has been used as an on/off switch for antibody recognition.[93] More importantly for our purpose, it proved an acceptable mimic of phosphoserine and phosphothreonine residues in histone H3,[94] and AuroraA[95] and p38α kinases.[96] Contrary to the Staudinger-phosphite reaction, the thiamichael reaction is not stereoselective. The epimeric ratio is not known in these cases. However, the reaction of Dha-modified ubiquitin with thiols provided a 60:40 epimeric mixture based on HPLC.[97]

Dha can also react with alkyl halide in presence of sodium borohydride in a radical fashion. A recent report highlights the insertion of stable (difluoro)phosphonate ((Pfa)Pha) residues as mimic of pS in H3 (Scheme 1.2).[98]



Scheme 1.2. Phosphorylation or mimics by reaction of cysteine residues in full length protein or peptide.

Dehydroalanine appears as a versatile intermediate towards the phosphorylation of protein. Its formation is described below.

1.3.3. Dha insertion in protein

The formation of dehydroalanine (Dha) and dehydrobutyrine (Dhb) residues first requires the activation of serine, threonine, cysteine and selenocysteine residues. These activated residues can then undergo an α-elimination to provide the alkene moeity (Scheme 1.3).



X=O, Serine (Ser) or threonine (Thr), X=Se, Selenocysteine (Sec), X=C, cysteine

Scheme 1.3. General scheme of the formation of Dha from other residues

In nature, serine and threonine residues are directly or indirectly the providers of Dha and Dhb residues. The biosynthesis of lantibiotics, lanthionine-containing peptide antibiotics proceeds in two steps: the formation of Dha or Dhb and the intramolecular cyclisation with cysteine residues of the peptide (Figure 1.9.A & Scheme 1.4).[99] For instance, in the formation of Nisin (Figure 1.9.A & B), the serine and threonine residues are dehydrated to Dha and Dhb by the dehydratase NisB followed by the cyclisation with cysteine residues catalysed by NisC (Figure 1.9.A).[100] Interestingly, the mechanism of dehydratation involves the glutamylation of the serine and threonine residues with the co-factor glutamyl-tRNA^{Glu}.[101, 102] In another class of lantibiotics, Lacticin 481 synthetase, LctM dehydrates the serine and threonine residues of its substrate LctA to Dha and Dhb by a two-step phosphorylation/elimination process (Figure 1.9.A). This mechanism was exploited in a method to create peptide containing Dha and Dhb.[103] It required the attachment of a leader peptide, based on LctA, to the peptide of interest to be recognised by the enzyme.

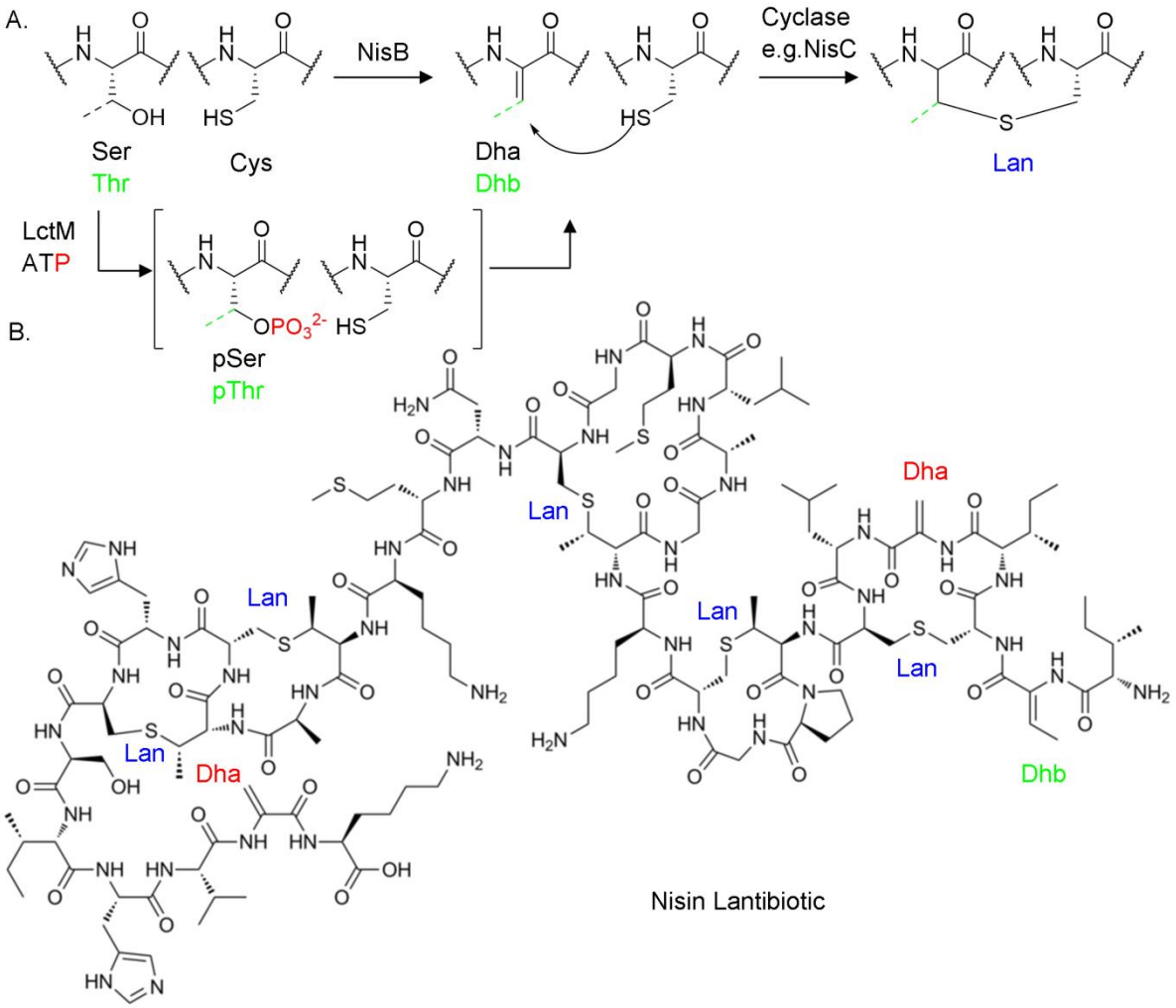
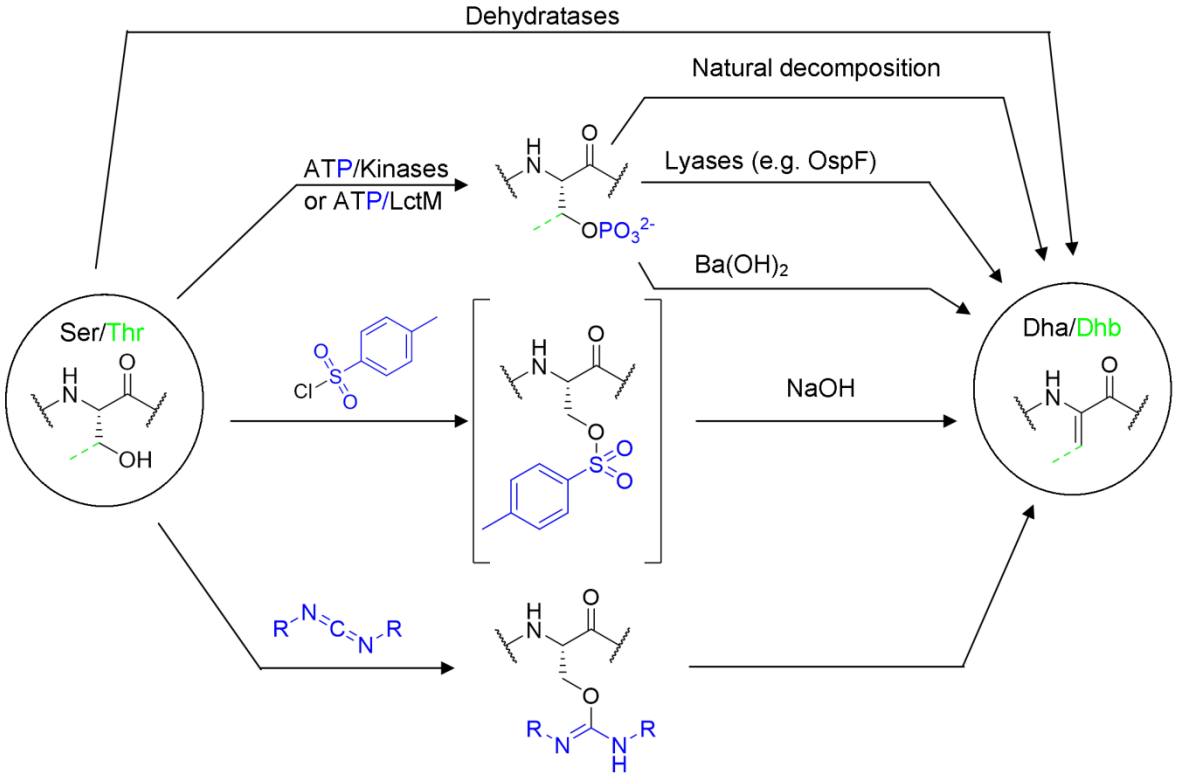


Figure 1.9.A. The creation of lanthionine bond (Lan) in lantibiotics involved the formation of dehydroalanine (Dha), the example of Nisin and Lacticin 481 are given. B. Nisin structure.

In some proteins, the phosphorylation of serine and threonine creates a phosphate leaving group that can eliminate. Indeed, phosphoserine residues decompose naturally to Dha in proteins like dentin and crystallin (Scheme 1.4). Similarly to the formation of lantibiotics, the Dha residues readily react with free thiols like glutathione to irreversibly inactivate this proteins (Scheme 1.4).[104, 105] In the prokaryotic world, lyase enzymes have the ability to deactivate MAPK by transforming pS and pT into the corresponding Dha and (Dhb) (Scheme 1.4).[106] Their mechanism has been studied by an ingenuous *in vivo* palladium mediated-deprotection of a caged-lysine residue, essential for their activity.[107] Chemically, the use of a dilute solution of barium hydroxide was enough to decompose pS to Dha (Scheme 1.4).[108]

This method has been successfully transferred to the formation of histone H3Dha, a chemical precursor towards the formation of methylated lysine residues.[109]

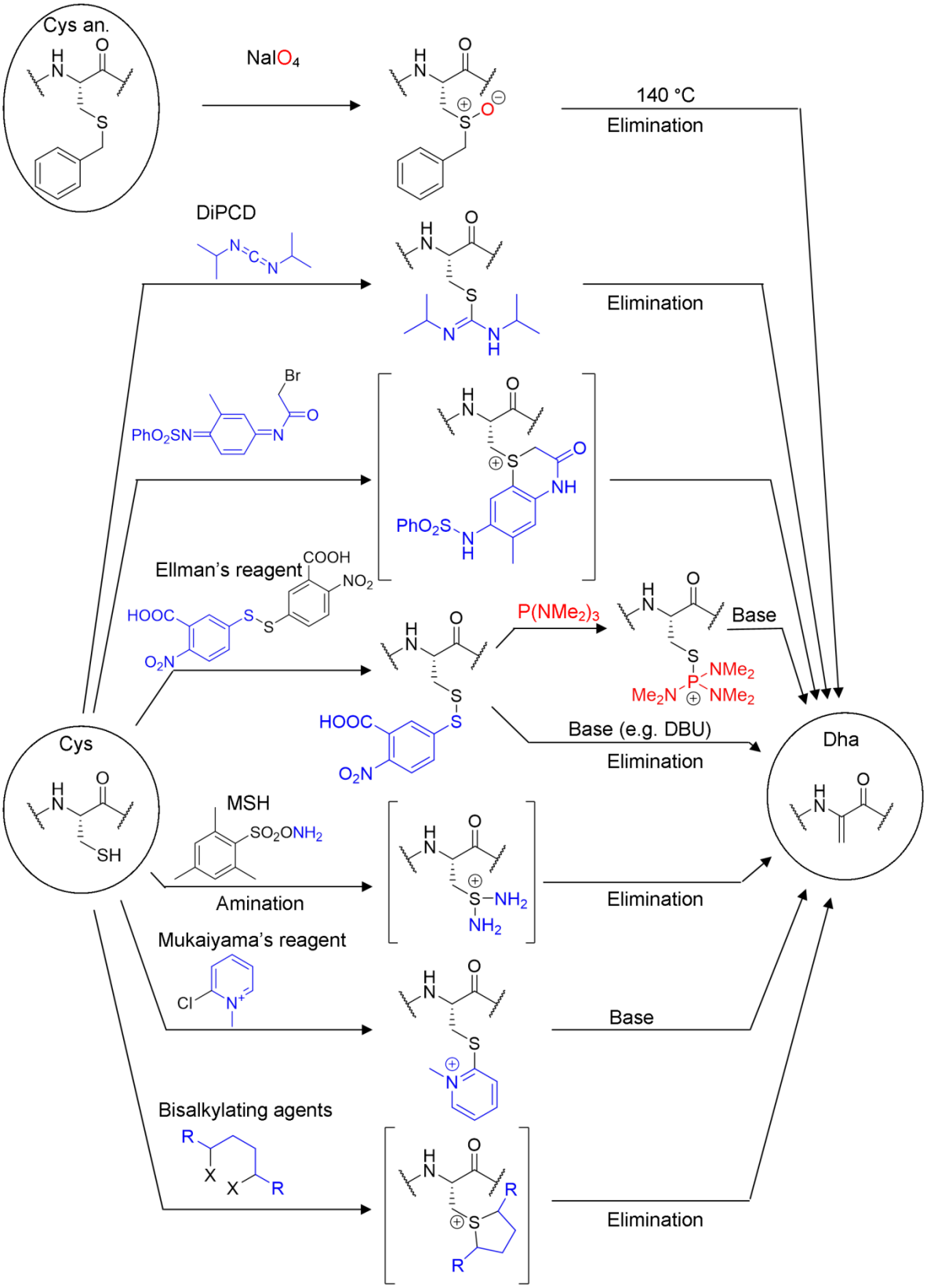
Finally, other chemical methods, using serine residues have been described towards the formation of Dha. In early work, the tosylation of serine residues, in the active site of chymotrypsin[110] and SBL[111] proteases, generated Dha after elimination in basic conditions (Scheme 1.4). The formation of isoureas followed by elimination was also investigated on aminoacids.[112] For our purpose, the formation of Dha from serine residues does not appear particularly selective, as proteins possess on average 6% of serine residues.[113, 114]



Scheme 1.4. A. Complementary modes of elimination of Ser and Tyr to Dha.

Therefore, rare, more reactive and/or unnatural amino acids were introduced to improve the selectivity of Dha formation. As described previously, cysteine residues (C) are relatively rare and reactive. In early work, the formation of Dha from cysteine residues was rarely

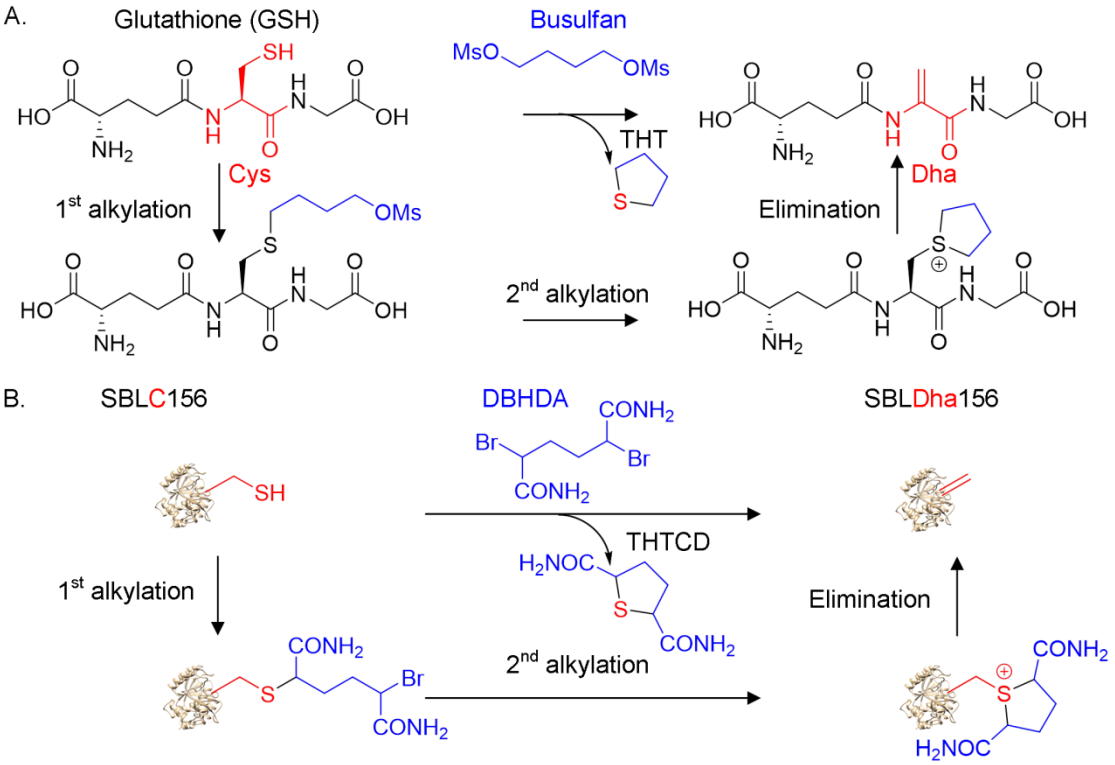
compatible with proteins. Oxidation/elimination of substituted cysteine residues was achieved at 140 °C (Scheme 1.5).[115] Similarly, the formation of isoureas from cysteine residues followed by elimination to Dha was performed in organic solvent (Scheme 1.5).[112] The formation of Dha, by addition of 2-methyl-*N'*-benzenesulfonyl-*N*,4-bromoacetylquinonediimide to ovalbumin and bovine pancreatic ribonuclease, was itself accompanied by the cleavage of the proteins (Scheme 1.5).[116] More recently, complementary methods towards the conversions of cysteine residues to Dha have been reported.[117] Cysteine residues promptly form unsymmetrical disulphide bond with Ellman's reagent, which eliminates either under basic conditions or after treatment with an electron rich phosphine (Scheme 1.5). Furthermore, O-mesitylenesulfonylhydroxylamine (MSH) converted SBLC156 to SBLDha156 by an amination/elimination process (Scheme 1.4).[118] Finally, the addition of Mukaiyama reagent to a cysteine residue followed by a treatment with base can promote the formation of Dha (Scheme 1.5). These methods are harsh, and not generalizable to all proteins. The formation of Dha can be achieved with treatment of cysteine residues with bisalkylating agents *via* a mild bisalkylation/elimination mechanism (Scheme 1.5).[117]



Scheme 1.5. Complementary modes of elimination of Cys and analogues to Dha.

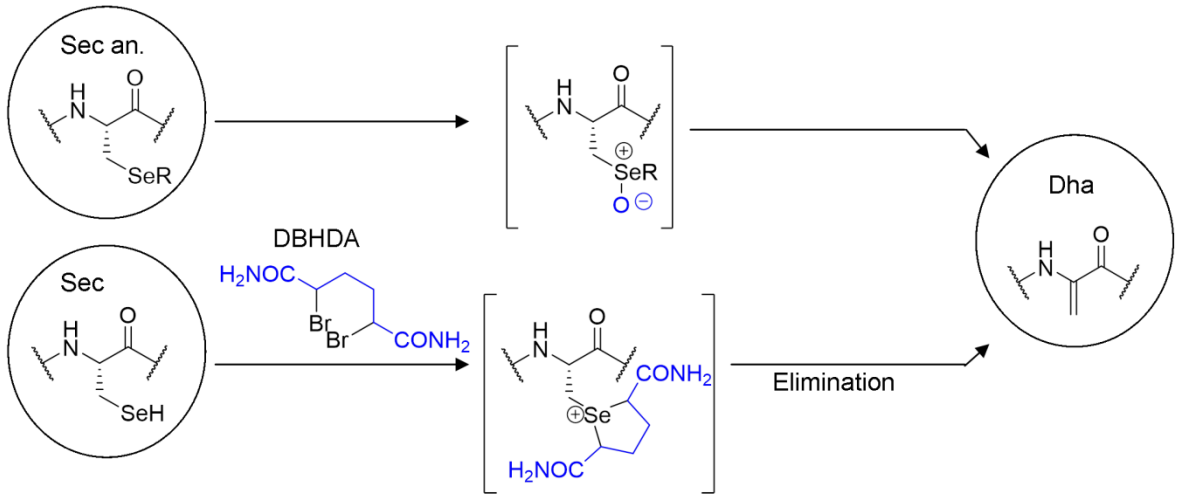
This milder method is based on the metabolic degradation of the cancer drug busulfan and selective protein cleavage.[116, 119, 120] Indeed, busulfan is a bis-alkylating agent, which reacts

with the cysteine residues of glutathione (GSH) to yield both Dha and a tetrahydrothiophene (THT) product (Scheme 1.6.A). This mechanism inspired the development of 2,5-dibromohexane diamide (DBHDA) as a water stable and soluble reagent towards the mild formation of Dha in proteins (Scheme 1.6.B).[117] This reagent has successfully been used with various proteins and applications: the protease SBL,[117] GFP (probes),[121] the kinases Aurora A,[95] and p38α (phosphate incorporation),[96] diubiquitin based probe[97, 122], the histones H2A/B (GlcNac incorporation),[123, 124] H3 (phosphate, modified lysine)[94] and H4 (modified lysine and arginine), cAb-Lys3 (on/off switch probe and glycosylation),[93, 98, 117] *N*-acetyl neuraminic acid lyase (modified lysine),[125] AcrA, annexin V and Npβ (protein model for radical reactions).[98, 117] Another bisalkylating agent, 2,5-methyldibromopentanoate has been developed to permit the incorporation of multiple Dha residues on peptides.[146] The use of other bisalkylating agents are detailed chapter 3 and 4.



Sheme 1.6.A. Busulfan metabolic degradation by glutathione generates a tetrahydrothiophene (THT) and the desulfurisation of the cysteine residue of GSH to Dha. B. DBHDA reacts with cysteine residues of protein to form Dha.

Finally, in order to improve selectivity, the incorporation of more reactive selenocysteine residues and analogues has also been explored. The oxidation/elimination of genetically incorporated unnatural substituted selenocysteine residues is described (Scheme 1.7).[126-130] The use of oxidative reagents is, however, a non-negligible limitation as some amino acids present in proteins like methionine (M) and cysteine (C) residues are sensitive to oxidation. Very recently, a study showed that selenocysteine residues were more reactive toward 2,5-dibromohexane diamide (DBHDA) than cysteine residues,. Moreover, the genetic incorporation of selenocysteine residues and unnatural substituted variants is not always efficient in proteins.[131]



Scheme 1.7. Complementary modes of elimination of Sec and analogues (Sec an.) to Dha.

For our purpose, we therefore propose a three-step chemical phosphorylation of MEK1 at residues 218 and/or 222. First, a cysteine residue is incorporated by site-mutagenesis at the desired sites. Then, the dehydroalanine residue is introduced with a bis/alkylating agent, 1,5-dibromohexanediamide DBHDA being our model reagent. Finally, a thiamichael reaction with sodium thiophosphate provides the phosphocysteine residue as mimics of phosphoserine S218 and S222.

1.4. Site selective chemical phosphorylation of MEK1

1.4.1. Proof of concept

This 3-step methodology was first reported for kinases by Blagg group to yield Aurora A pC288, mimic of pT288 (Figure 1.11).[95] However, thiamichael reaction with sodium thiophosphate (and other negatively charged thiols), to produce Aurora A pC288 proved to be particularly challenging and necessitated several rounds of purification.[132] Concurrently, the development of the method towards the chemical phosphorylation of p38α and MEK1 kinase was in progress in the Davis group. Our work was published soon after as "Synthetic Phosphorylation of p38α recapitulates Kinase Activity"[96] and is described in Dr. Phin Chooi's thesis citing my contribution. The work on p38α including the formation and characterisation of p38αpC172 and p38αpC180 as well as the analysis of their kinase activity (Figure 1.10) became the basis for my thesis. Fundamental findings are highlighted below.

The characterisation of the localisation of the p38α chemical phosphorylation by LCMS-MS of the tryptic peptide proved challenging.[133] The peptides containing the phosphocysteine residues could not be detected. First, the sample preparation was investigated. Excess dithiothreitol (DTT) or β-mercaptoethanol is commonly employed during tryptic digests, and is essential protein containing multiple cysteine residues.[134, 135] Our p38α constructs possessed 2 cysteine residues. Although, the potential instability of phosphocysteine residues towards DTT has been observed,[11] no cleavage of either p38αpC172 or p38αpC180 upon excess of DTT. Formic acid (FA) was also present to quench, purify tryptic digest and the corresponding peptides prior LCMS analysis. However, as phosphocysteine residues were labile under acidic conditions,[136] the use of FA was limited to the LCMS-MS run. We also decided to by-pass the C18 purification column prior injection due to the polarity of phosphopeptides. Finally, we first used Collision Induced Dissociation (CID) as a fragmentation method during LCMS-MS as it was previously reported for peptides containing phosphocysteine residues.[11] However, phosphorylated residues, especially serine and

threonine residues, tend to cleave under CID conditions,[137] the spectra are usually dominated by the neutral loss of the phosphate.[138] While it is not intrinsically an issue, our desired peptide could be not detected. Electron transfer dissociation (ETD) appeared as a milder option,[139] and was used successfully to validate the selectivity of the chemical phosphorylation. In a recent paper, Bertran *et al.* drew an identical conclusion.[92]

Mass spectrometry methods allowed the full characterisation of our chemically modified p38α constructs (Figure 1.10.A). A mass spectrometric method was also developed in order to assess the effect of the chemical phosphorylation on the p38α kinase activity (Figure 1.10.B). The phosphorylation of a full length transcription factor, ATF2, a natural substrate of p38α, was monitored by LCMS-ESI-TOF (Figure 1.10.B). Traditionally, radiolabelled ATPγ^{32}P, ATPγ^{33}P and even ATPγ^{35}S are used in kinase assays. The incorporation of radiolabelled phosphate into a protein is therefore directly measurable. However, due to the high cost and strict regulation associated with radioisotope assays, antibody-dependent and/or fluorescence/luminescence-based assays have been developed.[140, 141] These methods possessed their limitations. For instance, measuring ADP release is an indirect method that cannot dissociate ATPase activity from kinase activity. Furthermore, distinguishing between the single and the multiple phosphorylation of the protein substrate is difficult, although phos-tag gels have been recently developed for this purpose.[35, 142] On the other hand, LCMS is a direct method. The addition of +80 adduct is the indication of the phosphorylation. Activated p38α phosphorylates ATF2 at T69 and T71 (Figure 1.10.B). The single and double phosphorylations were directly detected by LCMS (Figure 1.10.C). The quantification, by measurement of the relative intensity of the protein mass, was performed without internal standard and relied on two hypotheses. Firstly, the mass difference of chemical phosphorylation was not significant when compared to the overall mass of the protein e.g. phosphorylation of ATF2 equated for less than 0.2% of its overall mass. Secondly, the addition of negatively charged phosphate groups to the protein was not dramatically affecting the charge state distribution of the protein, and therefore altering the

MS signal. These hypotheses were validated by comparison with a radiolabelled assay with ATPγ^{32}P (Figure 1.10.C).

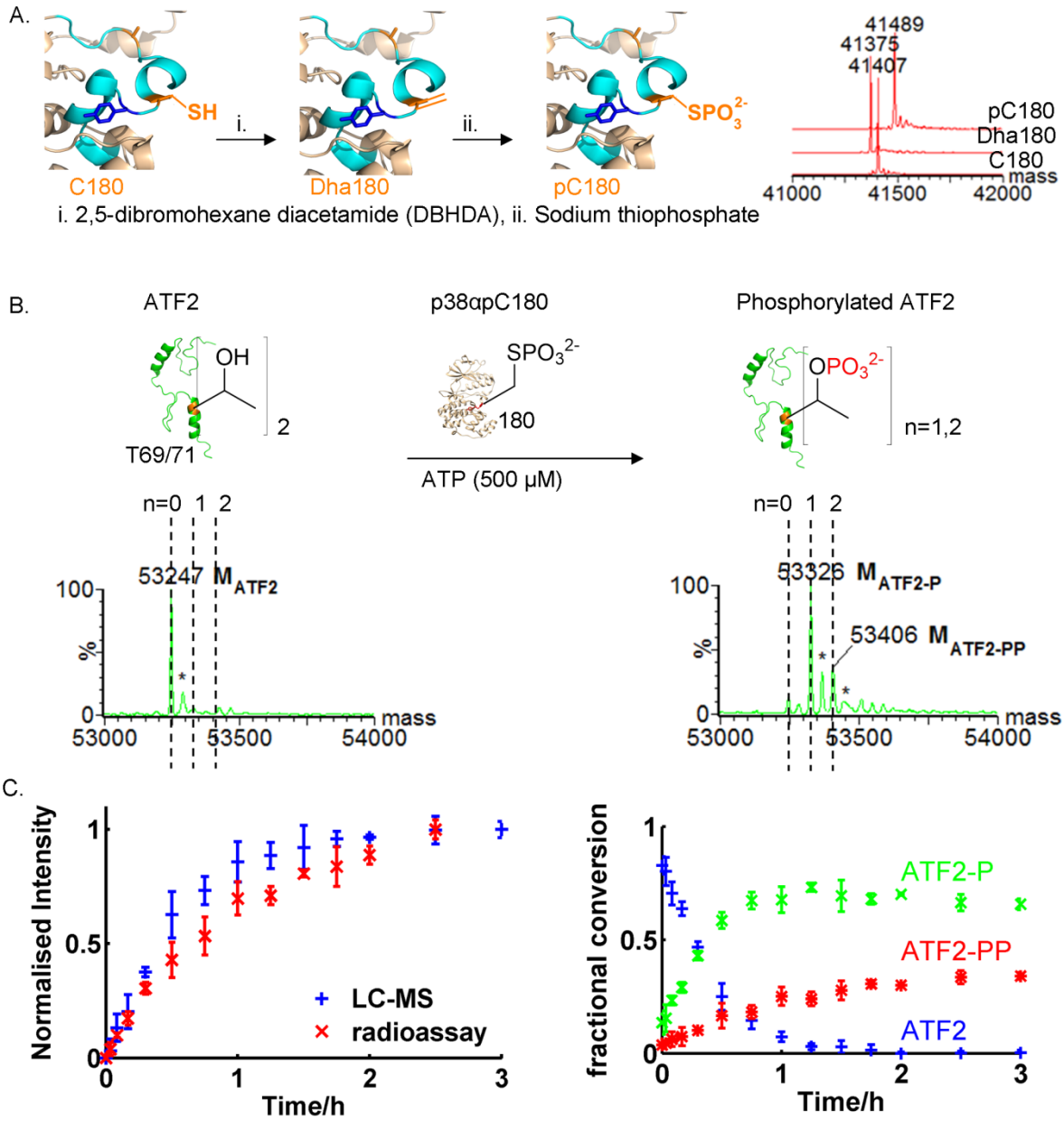


Figure 1.10.A.Formation of p38αpC180 from p38αC180 using Dha formation with DBHDA followed by thiamichael addition with sodium thiophosphate. B. LCMS based activity assay showed that p38αpC180 phosphorylated one of the natural substrate ATF2 in a similar way than p38αpT180. C. LCMS based activity assay and radioassay were comparable. The former differentiate mono and diphosphorylation.

Mastering these mass spectrometry techniques was a crucial part of my thesis. Mass spectrometry allows the monitoring of protein reactions, the full characterisation of chemical modifications, and the assessment of activity and chemical mechanism.

In this introduction, the methodologies towards the phosphorylation of MEK1 were discussed. While the recent development in amber stop codon suppression produced MEK1 singly phosphorylated in good yield and purity, no robust complementary chemical phosphorylation methodology has been described for MEK1. This thesis proposes to face the challenge. Some initial issues towards the site selective chemical phosphorylation of MEK1 at position 218 and 222 are presented next paragraph 1.4.2.

1.4.2. Challenges towards of site selective chemical phosphorylation of MEK1 via Dha

As discussed paragraph 1.3.2, the thiamichael reaction is not a stereochemical reaction. While affording an epimeric mixture could be detrimental for the crystallisation of our constructs MEK1pC, the effect of the protein environment around 218 and 222 is not known.[130] The determination of the epimeric ratio for MEK1pC222 is discussed chapter 5.

In most of the examples of Dha formation, proteins possess either a unique cysteine residue or several cysteine residues protected in intramolecular disulphide bonds. According to the crystal structure of MEK1, the protein contains 6 reduced cysteines that can potentially be alkylated. Methods to regioselectively target one cysteine residues amongst others in protein are rare. They usually require protein engineering[143, 144] or a privileged arrangement inside the protein.[145] We propose instead to modulate the reagents. The design and use of these reagents are illustrated chapter 2 and 4. The intimate mechanism of one of these reagents is exemplified chapter 3.

Finally, in order to make the doubly phosphorylated MEK1pC218pC222, the formation of multiple Dha is necessary. During my research, the use of a bisalkylating/elimination agent, 2,5-methyldibromopentanoate, has been reported for this purpose on peptides.[146] Chapter 4 shows how we came to the same conclusion.

1.4.3. Thesis summary

Chapter 2 describes the site and regioselective formation of Dha at position 218 and 222. It involves the use of nucleotides and the design of various bisalkylating agents.

Based on observations in Chapter 2, Chapter 3 describes the elucidation of the elimination mechanism in Dha formation using bis/alkylating agent. It involves the development of method to site selectively deuterated cysteine residues in proteins.

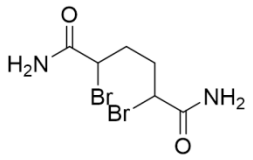
Chapter 4 reports the formation of multiple Dha on two proteins, MEK1 and p38α.

Chapter 5 finally discusses the challenges in the formation of MEK1pCs. The limitations of the method as well as alternative methods are discussed. The chapter also evokes the unexpected activity of MEK1Dha.

RESULTS AND DISCUSSION

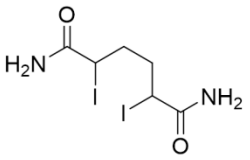
SCHEMATIC OVERVIEW OF THE ALKYLATING AGENTS USED TOWARDS REGIOSELECTIVE DHA FORMATION AND PHOSPHORYLATION

DBHDA



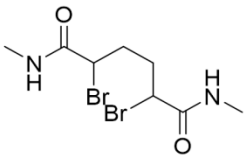
2,5-**di**bromo**h**exane **di**amide

DIHDA



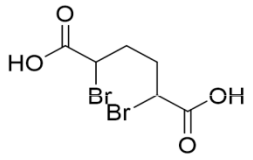
2,5-**di**iodo**h**exane **di**amide

DBDMet



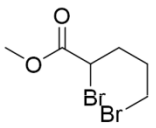
2,5-**di**bromo-*N*^{1},*N*^{6}-**di**m**et**hylhexanediamide

DBHDAc



2,5-**di**bromo**h**exane**di**oic **ac**id

MDBP



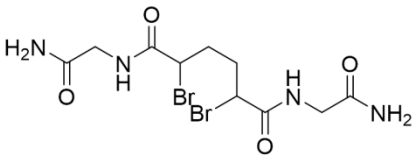
2,5-**m**ethyl**di**bromo**p**entanoate

DIB



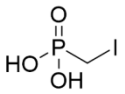
1,5-**di**iodo**b**utane

DBDGla



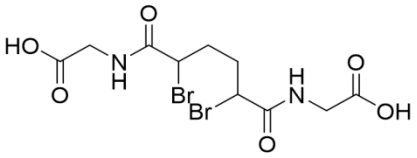
N1, *N6*-bis(2-amino-2-oxoethyl)-2,5-**di**bromohexanediamide
≡
Glycin**a**mide

IMPA



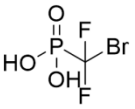
(**i**odo**m**ethyl)**p**hosphonic **a**cid

DBDGly



2,2'-((2,5-**di**bromohexanedioyl)bis(azanedyil) **di**acetic acid
≡
Glycine

BFPA



(**b**romodi**f**luoro**m**ethyl)**p**hosphonic **a**cid

Chapter 2: Site-Selective Dehydroalanine Formation

2.1. Site selective formation of MEK1Dha222 with DBHDA

2.1.1. Reactivity of MEK1 native cysteine residues

This chapter describes the site-selective formation of both MEK1Dha218 and MEK1Dha222, intermediates towards the synthesis of chemically phosphorylated MEK1pC218 and MEK1pC222. The native serine residues 218 and 222 were first mutated to cysteine residues. However, the potential reactivity of MEK1 native cysteine residues was also assessed in order to build our constructs.

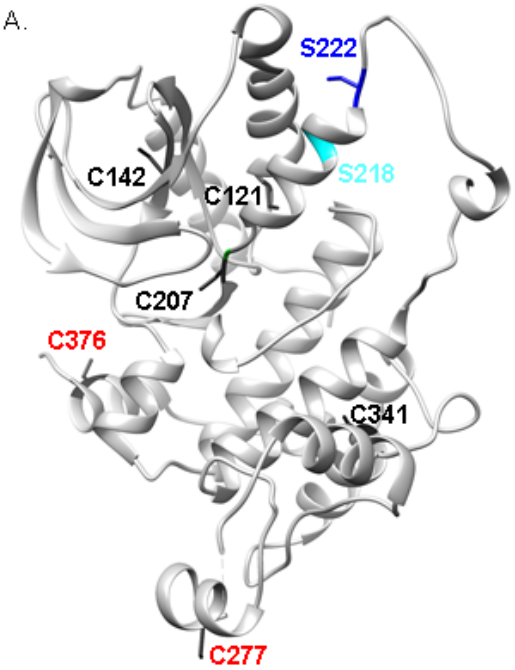
Full-length wild-type MEK1 contains 6 free cysteine residues at positions 121, 142, 207, 277, 341 and 376 (Figure 2.1.A). All appeared potentially reactive towards bis-alkylating agents used for Dha formation.

In order to discriminate their reactivity, their relative side-chain accessibility was calculated using Naccess (Figure 2.1.B.).[147] C207, C277 and C376 displayed relatively high side-chain accessibility, 22.4%, 93.5% and 21.9% respectively. Both C277 and C376 were mutated to serine residues to afford MEK1S277S376, referred hereafter to MEK1 (Figure 2.1.C). However, C207 was retained as the residue lies within the “ATP pocket”. Moreover, not only does the C207A mutant of the constitutively active MEK1D218D222 display reduced activity, but C207 has also been a target for MEK1 covalent inhibitors.[148, 149] Finally, C121, C142 and C341 were kept as they exhibited very poor accessibility and appeared crucial for MEK1

activity. For example, while C121S results in increased kinase activity and confers resistance to both RAF and MEK inhibition *in vitro*,[150] C341G inhibits ERK phosphorylation.[151]

In addition to the native cysteine residues, one or two additional cysteine residues were incorporated at the positions of interest 218 and 222 to afford MEK1C218, MEK1C222 and MEK1C218C222. As position 222 appeared the most accessible (Figure 2.1 B), the MEK1C222 construct, which contained five free cysteine residues at position 121, 142, 207, 222 and 341, was selected as our model for MEK1Dha formation.

A.



B.

Residues	Calculated side-chain accessibility
C121	5.5
C142	0
C207	22.4
C277	93.5
C341	0
C376	21.9
S218	8.9
S222	38.4

C.

Site-directed mutagenesis
C277S
C376S
S218C
S222C

Figure 2.1.A. Crystal structure of MEK1 (PDB 3W8Q) displaying its 6 free cysteine residues (black and red) and its 2 naturally phosphorylated serine residues of interest 218 (cyan) and 222 (blue). B. Calculated side-chain accessibility of the cysteine residues and residues of interest S218 and S222. C. Site-directed mutagenesis (SDM) of the most accessible cysteine residues, C277S and C376S (red) as well as the sites of interest S218C (cyan) and S222C (blue).

In a preliminary experiment, the treatment of MEK1C222 with Ellman's reagent or 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB), under denaturing conditions, and monitored by

LCMS,* confirmed all five cysteine residues were not oxidised, and forming intramolecular disulphide bonds under our reaction conditions (Figure 2.2.B). The reaction under non-denaturing conditions revealed that one cysteine residue was not accessible to Ellman's reagent (Figure 2.2.A).

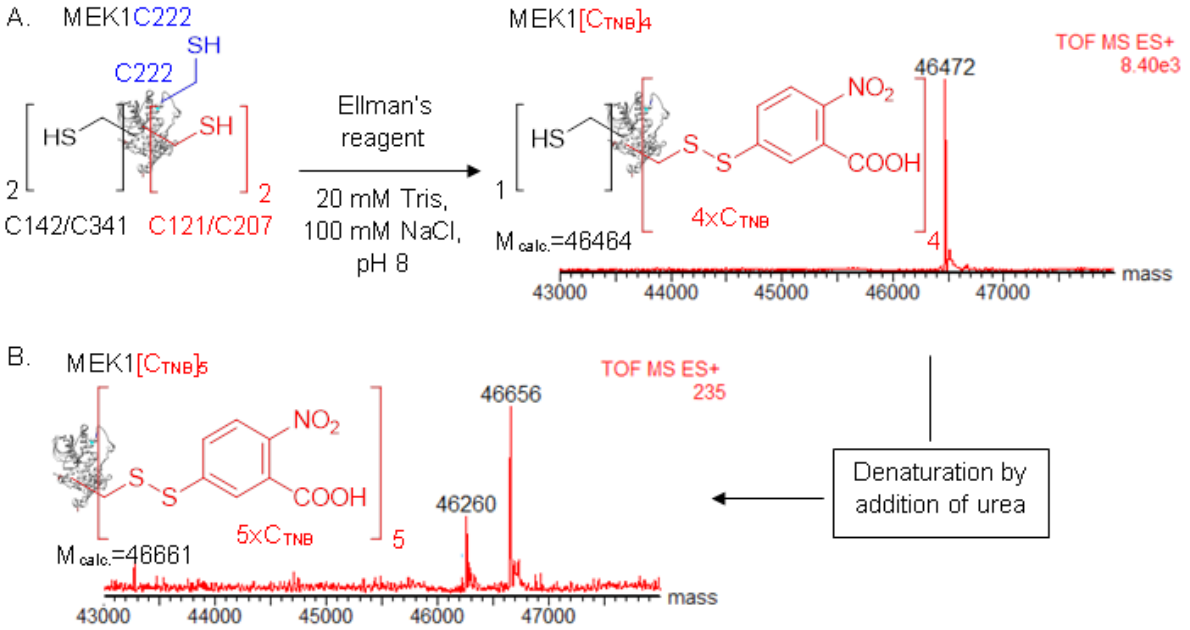


Figure 2.2. A. LCMS analysis of the reaction of MEK1C222 with Ellman's reagent showed the formation of a unique compound $M_{obs.}$=46472 corresponding to MEK1 with four cysteine residues bearing Ellman's adduct (C_{NTB}), MEK1[C_{NTB}]$_4$, $M_{calc.}$=46464. B. Under denaturing conditions using urea, the reaction afforded mainly $M_{obs.}$=46656 corresponding to MEK1[C_{NTB}]$_5$, $M_{calc.}$=46661. The minor product $M_{obs.}$=46260 corresponded to MEK1[C_{NTB}]$_3$, $M_{calc.}$=46267 see chapter 7 for explanation.

2.1.2. C207 reaction with DBHDA

Treatment with DBHDA fully converted MEK1C222 to over-alkylated MEK1Dha after 4 hours at pH 8 and 37 °C (Figure 2.3.A). The over-alkylation occurred on a constrained cysteine residue (C_{BHDA}). Indeed, under denaturing conditions, over-alkylated MEK1Dha converted to MEK1[Dha]$_2$. LCMS-MS of the tryptic peptides mixture revealed more precisely that DBHDA reacted with both C207 and C222 to form MEK1C_{BHDA}207Dha222 (Figure 3.3.B). This result was consistent with the calculation as 207 and 222 were the most

* Typically, the number of free cysteine residues was estimated by measuring the absorbance at 412 nm of NTB^{2-} in solution by spectrophotometry.[152-154]

accessible cysteine residues (Figure 3.1.B). Varying the temperature, the pH and the concentration of both protein and alkylating agent did not prevent the over-alkylation of MEK1Dha222.

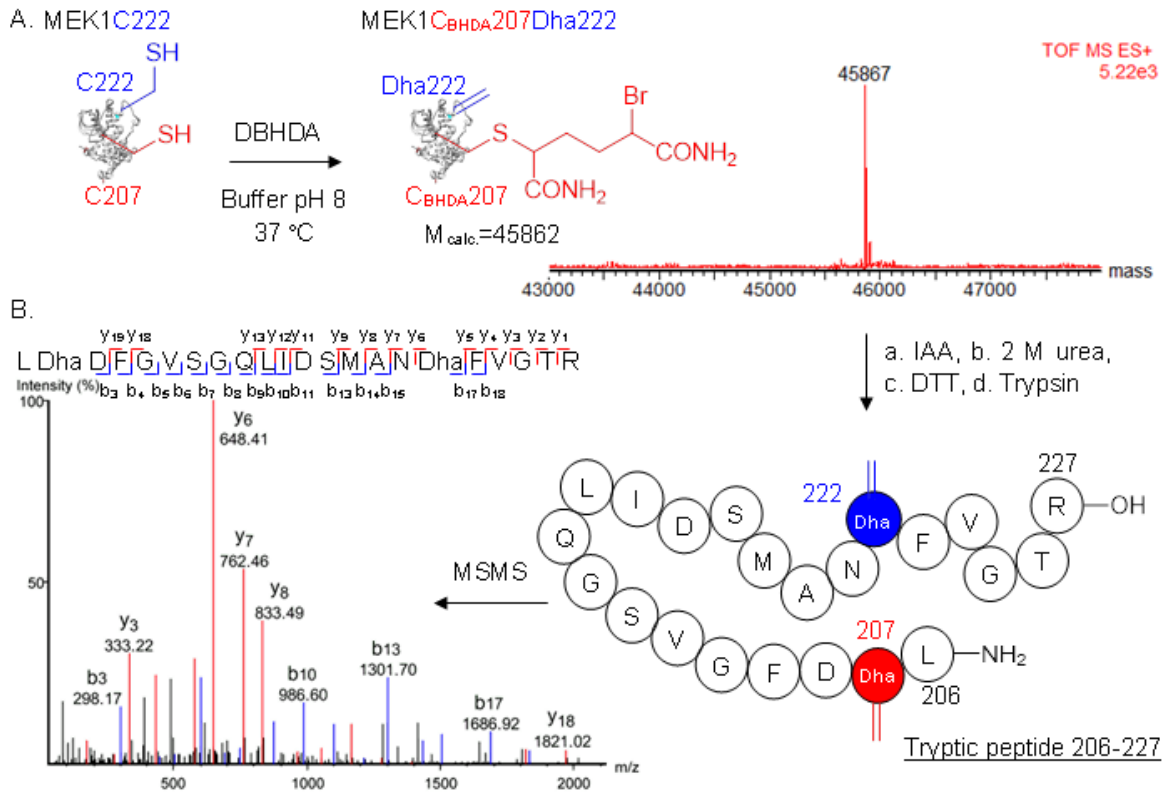


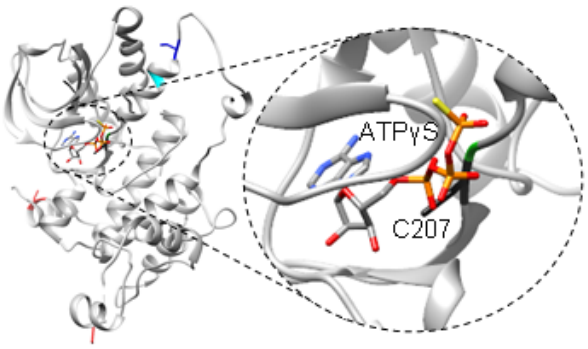
Figure 2.3.A. The treatment of MEK1C222 with DBHDA at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=45867 was consistent with overalkylation of MEK1Dha, $M_{calc.}$=45862. B. Tryptic peptides were prepared from MEK1C_{BHDA}207Dha222 denatured with urea, reduced with DTT and with the free cysteine residues capped with iodoacetamide (IAA), diluted and treated with trypsin at 37 °C. LCMS-MS of tryptic peptide 206-227 confirmed that both position 207 and 222 reacted with DBHDA.

In another study, the reaction of MEK1C218 with a mixture of sodium sulfite and sodium tetrathionate was performed in presence of ATP and magnesium.[91] According to the authors, ATP "protected vital cysteine residues". However, it was not mentioned which ones. We suggested that ATP would protect C207 from alkylating as it is located within the "ATP pocket".

2.1.3. Masking C207 with nucleotide and magnesium

Crystal structure analysis and side-chain accessibility calculations, based on the binary MEK1-ATPγS crystal structure (PDB 3W8Q), showed that ATPγS† indeed masked C207. Its presence significantly decreased C207 accessibility, from 22.4% to 0% (Figure 2.4.A and 2.4.B). In contrast, the accessibility of C222 in these calculations was not affected. However, when the reaction of MEK1C222 with DBHDA was performed in the presence of excess ATP, the nucleotide only partially masked C207. Indeed, 15% of over-alkylated MEK1Dha222 was detected by LCMS (Figure 2.4.C).

A.



B.

Residues	Calculated relative side-chain accessibility	
	ApoMEK1	MEK1-ATPγS
121	5.5	5.5
142	0	0
207	22.4	0
222	38.4	39.4
341	0	0

C.

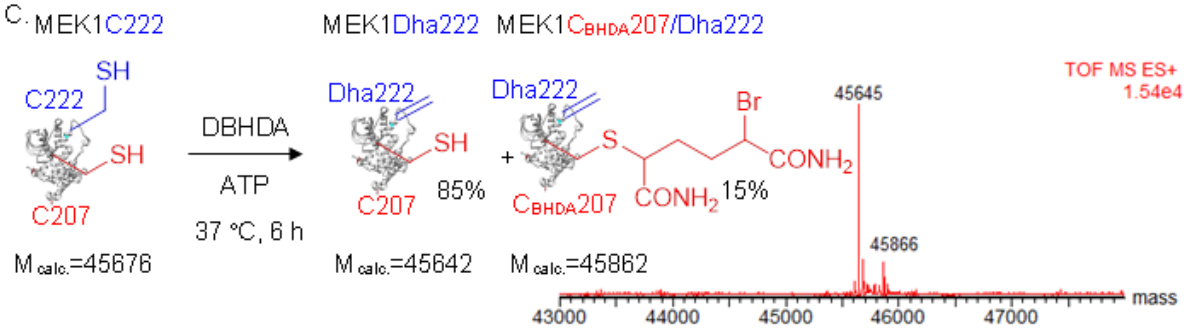


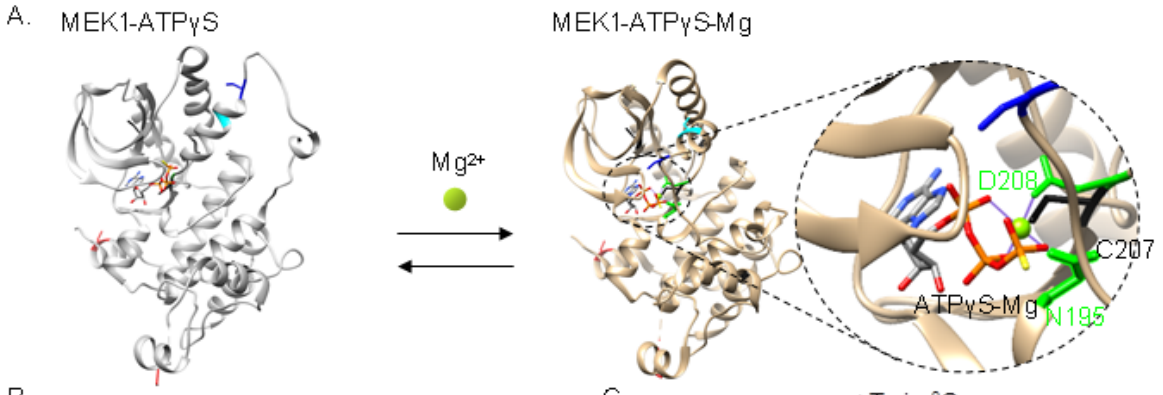
Figure 2.4.A. In the binary MEK1-ATPγS crystal structure (PDB 3W8Q), ATPγS masked C207. B. Comparison of calculated side-chain accessibility of cysteine residues (black and red) and site of interest 222 (blue) in the presence or absence of ATPγS based on PDB 3W8Q. C. LCMS analysis of the reaction of MEK1C222 with DBHDA in the presence of ATP showed the formation of two new masses $M_{obs.}$=45645 corresponding to MEK1Dha222, $M_{calc.}$=45642, and $M_{obs.}$=45866 corresponding to over-alkylated MEK1Dha222, $M_{calc.}$=45862.

† ATPγS is an analogue of ATP. It contains one sulfur atom in place of one oxygen atom on the γ-phosphate, which makes it more stable to hydrolysis than ATP. It is considered structurally equivalent to ATP.

As reported K_d values of ATP for MEK1 varied significantly in the presence or absence of magnesium,[46, 155] respectively ~2 µM and >100 µM, addition of magnesium to the reaction mixture was envisaged to stabilise the nucleotide-MEK1 complex. The crystal structure of MEK1-ATPγS-Mg (PDB 3EQD)[156] showed that magnesium created salt bridges between the D208 of the DFG loop, N195, and the nucleotide (Figure 2.5.A). The determination of the melting temperatures of MEK1-ATP and MEK1-ATPγS complexes with or without magnesium using a thermal shift assay by differential scanning fluorimetry[157] confirmed the stabilizing effect of the metal. Indeed, a 5 °C increase of the melting temperature of the complexes was observed in the presence of magnesium (Figure 2.5.C).

Furthermore, the presence of the metal also induced a conformational change of the MEK1 activation loop (Figure 2.5 A), which notably enhanced the accessibility of position 222 from 38.4% to 93.4% (Figure 2.5.B). More importantly for our purpose, the ATP-Mg complex lowered C207 accessibility. In the presence of a stoechiometric excess of ATP-Mg, DBHDA did not alkylate C207. However, 20% of phosphorylated MEK1Dha222 (pMEK1Dha222) was observed by LCMS analysis (Figure 2.5.D) along with the formation of ADP, by-product of the phosphorylation reaction (Figure 2.5.E). It thus appeared that MEK1 autophosphorylated during the reaction. Proteomic studies have revealed the presence of several autophosphorylation sites, e.g. T23, S298 and Y300, for MEK1.[27, 158] In addition, MEK1 ATPase activity has also been reported.[159] In our hands, LCMS-MS of the tryptically digested sample of the reaction mixture for the determination of the autophosphorylation sites was not conclusive.

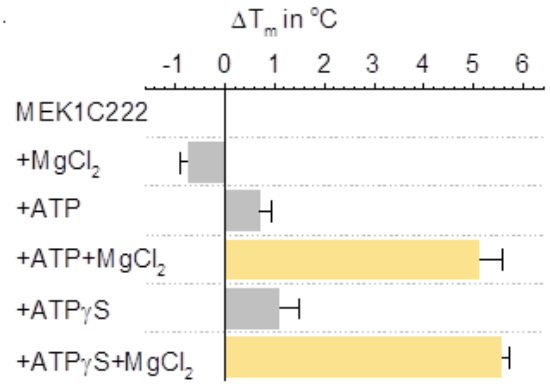
A.



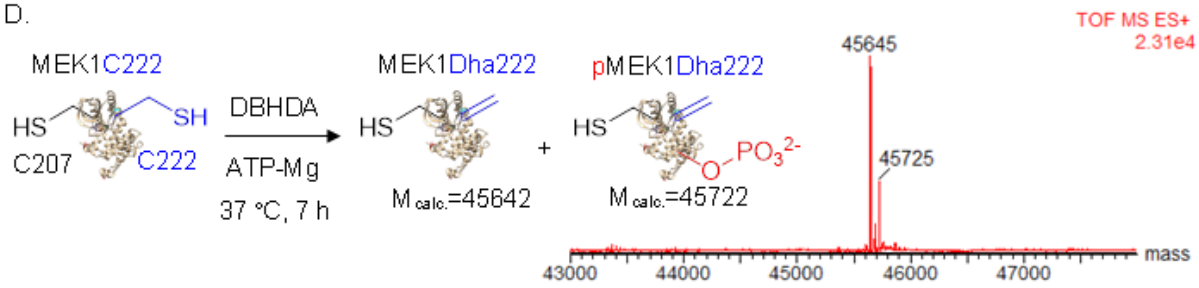
B.

Residues	Calculated relative side-chain accessibility	
	ApoMEK1	MEK1-ATPγS-Mg
121	5.5	0.7
142	0	0
207	22.4	4.4
222	38.4	99.6
341	0	0

C.



D.



E.

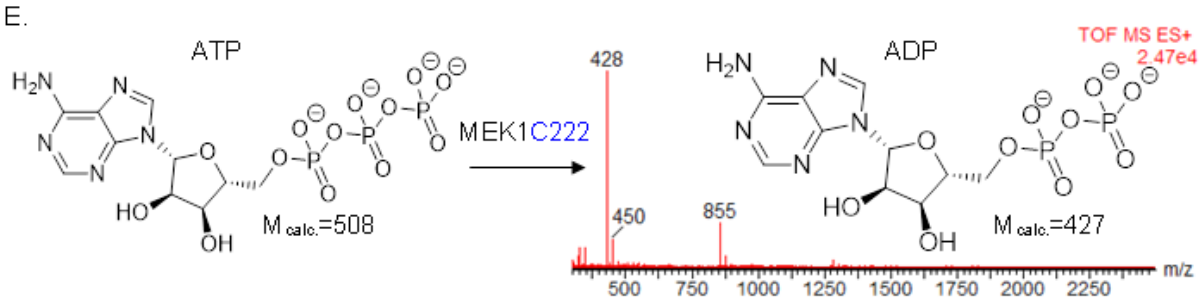


Figure 2.5.A. Difference in conformation of MEK1-ATPγS complexes in the presence of magnesium. Magnesium links the residues N195 and D208 to the nucleotide by salt bridge (PDB 3EQD). B. Calculated side chain accessibility of MEK1 in the presence of ATPγS-Mg (PDB 3EQD). C. Thermal shifts, ΔT_m, determined by differential scanning fluorimetry (DSF) of MEK1C222 in the presence or absence of ATP, ATPγS and $MgCl_2$. D. LCMS analysis of the reaction of MEK1C222 with DBHDA in the presence of ATP-Mg showed the formation of two new masses $M_{obs.}$=45645 corresponding to MEK1Dha222, $M_{calc.}$=45642, and $M_{obs.}$=45725 corresponding to phosphorylated MEK1Dha222, $M_{calc.}$=45722. E. Formation of ADP, 428 [M+H]+, was detected by LCMS.

A non-hydrolysable nucleotide was therefore used. ADP was readily available, and despite being the byproduct of phosphorylation, displayed similar affinity to MEK1 with near identical

K_d to that of ATP.[46] In addition, side chain accessibility calculations and thermal shift assay showed that ADP-Mg behaved similarly to ATP-Mg (Figure 2.6.A and 2.6.B). Pleasingly, the reaction of MEK1C222 with DBHDA performed in the presence of ADP-Mg lead exclusively to the selective formation of MEK1Dha222 without the autophosphorylation side reaction (Figure 2.6.C and 2.6 D).

A.

Residues	Calculated relative side-chain accessibility	
	ApoMEK1	MEK1-ADP-Mg
121	5.5	1.1
142	0	0
207	22.4	3.5
222	38.4	93.4
341	0	0

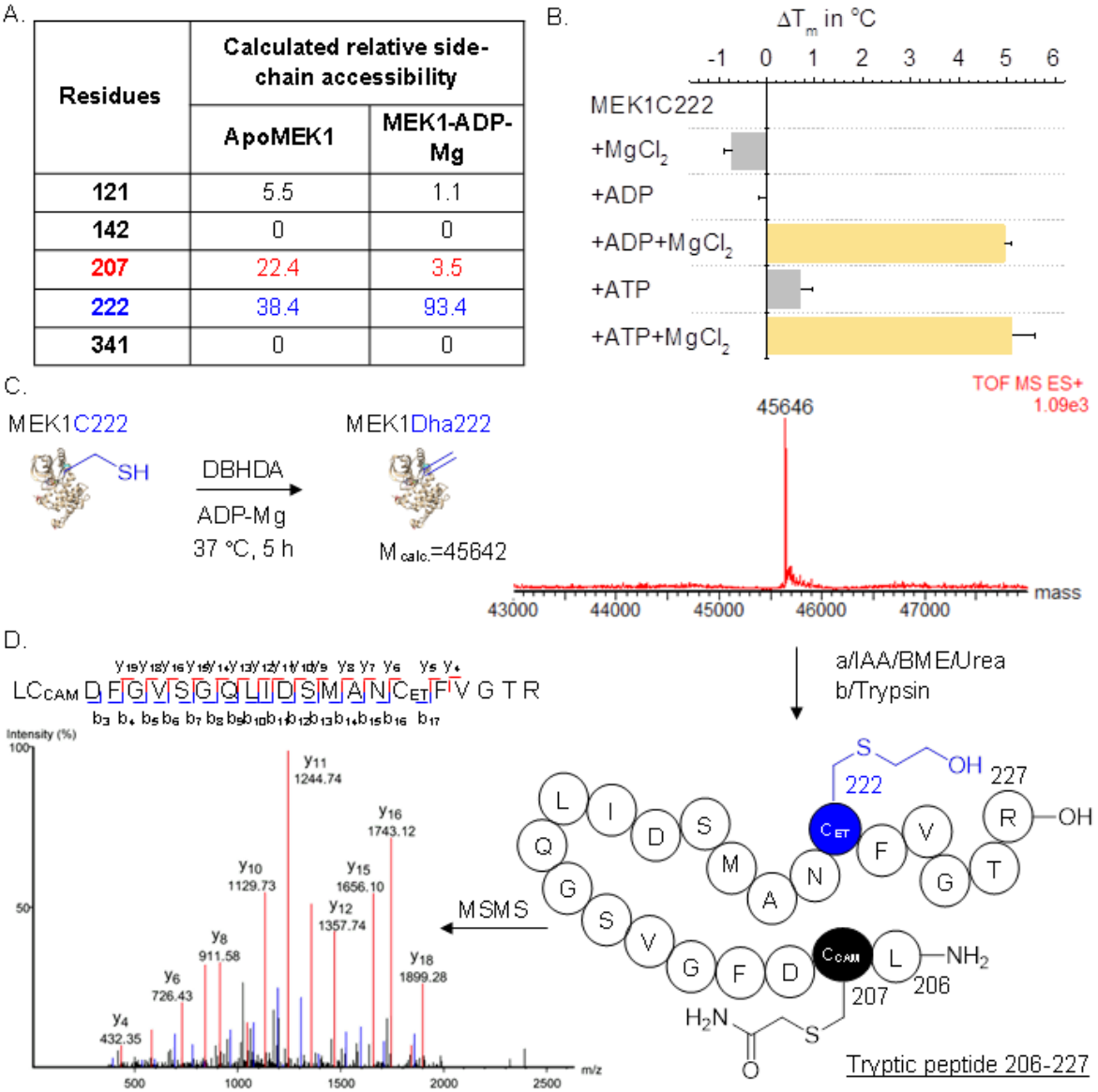


Figure 2.6. A. Relative side chain accessibility was calculated for apoMEK1 (PDB 3W8Q) and MEK1-ADP-Mg (PDB 3EQI). B. Thermal shifts, ΔT_m, determined by differential scanning fluorimetry (DSF) of MEK1C222 in the presence or absence of ADP, ATP and $MgCl_2$. C. The treatment of MEK1C222 with DBHDA in the presence of ADP-Mg at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=45646 was consistent with MEK1Dha222, $M_{calc.}$=45642. D. Tryptic peptides were prepared from MEK1Dha222 denatured with urea, reduced with β-mercaptoethanol (BME) and with the free unreacted cysteine residues capped with iodoacetamide (IAA), diluted and treated with trypsin at 37 °C. LCMS-MS of tryptic peptide 206-227 confirmed that position 207 was a capped unreacted cysteine residues and 222 was a dehydroalanine, which reacted with BME during the preparation for the digest.

To confirm site-selectivity, full mapping of MEK1 cysteine residues was performed. LCMS-MS of peptides derived from a trypsin/AspN digest of MEK1Dha222, with unreacted cysteine residues capped with iodoacetamide confirmed the selectivity of the reaction (Figure 2.7).

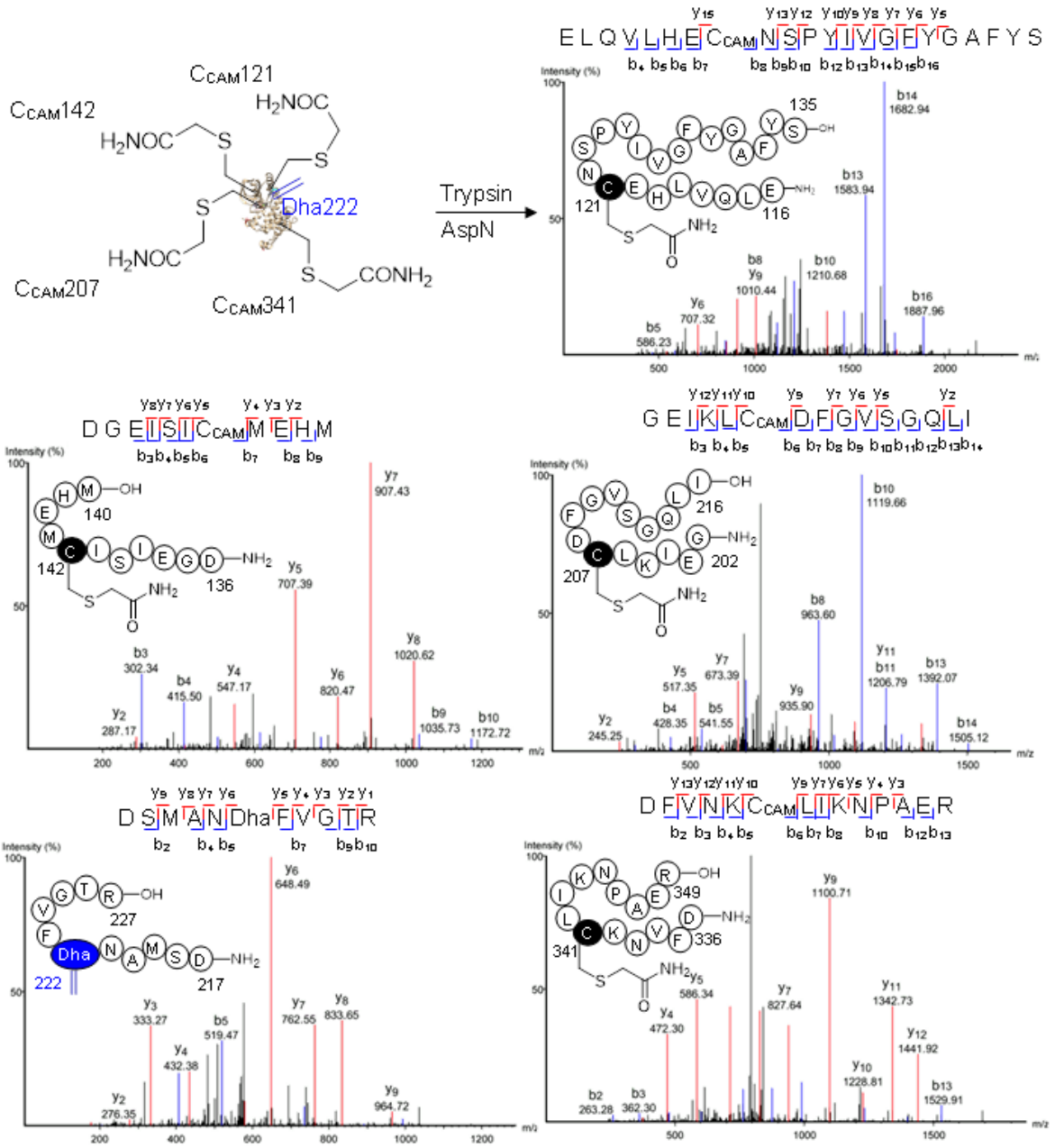


Figure 2.7. Peptides from the tryptic/AspN digestion of MEK1Dha222 with the unreacted cysteine residues capped with iodoacetamide was analysed by LCMS-MS. C121, C142, C207 and C341 were capped with iodoacetamide. Position 222 was confirmed to be a dehydroalanine residue.

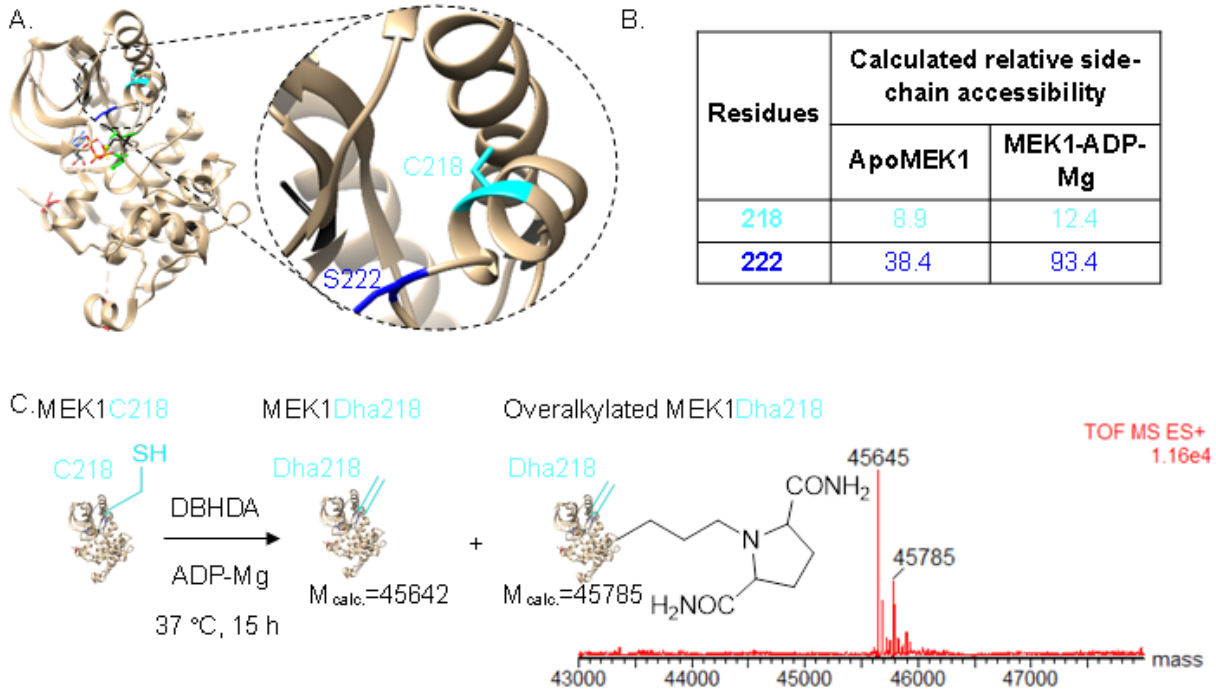
Using ADP-Mg as masking agent, the unwanted alkylation of C207 was prevented. True "region/chemo"site-selective formation of Dha was achieved on a protein.

The same methodology was subsequently applied towards the formation of MEK1Dha218.

2.2. Site-selective formation of MEK1Dha218

2.2.1. Over-alkylation of MEK1Dha218 with DBHDA in the presence of ADP-Mg complex

Similar to the residue 180 in p38α,[96] MEK1's C218 residue is part of an α-helix (Figure 2.8.A). Its accessibility was moderately enhanced by the addition of ADP-Mg but still remained relatively low at 12.4% compared to 93.4% for the position 222 (Figure 2.8.B). As a result, the reaction of DBHDA with MEK1C218 in the presence of ADP-Mg was slower in comparison with MEK1C222. MEK1C218 was fully converted to MEK1Dha218 in 15 hours against the 5 hours observed for the formation of MEK1Dha222 under identical conditions. A side product was also detected with a mass $M_{obs.}$=45785 (Figure 2.8.C). This mass corresponded to the addition of didebrominated DBHDA adduct on MEK1Dha218. But, in contrast to the previously observed C207 over-alkylation, this alkylation did not lead to the formation of Dha. The double alkylation of lysine residues, leading to the formation pyrrolidines, was suspected. Longer reaction times led to multiple over-alkylation adducts. Several lysine residues of MEK1 were potentially reactive as post-translationally modified in cells, either *via* ubiquitinilation or acetylation.[160, 161] However, it was not possible to identify the lysine residues involved by LCMS-MS of the tryptic digest.



Residues	Calculated relative side-chain accessibility	
	ApoMEK1	MEK1-ADP-Mg
218	8.9	12.4
222	38.4	93.4

Figure 2.8.A. Crystal structure of MEK1 (PDB 3EQD) showed that position 218 is part of an α-helix, while 222 is in a loop. B. Calculated side chain accessibility for position 218 with or without ADP-Mg in was compared the corresponding value for position 222. C. LCMS analysis of the reaction of MEK1C218 with DBHDA in the presence of ADP-Mg showed the formation of two new masses $M_{obs.}$=45645 corresponding to MEK1Dha218, $M_{calc.}$=45642, and $M_{obs.}$=45785 corresponding to over-alkylated MEK1Dha218, $M_{calc.}$=45785.

An alternative strategy, involving the modulation of the properties of the bis-alkylating/elimination agent, was therefore investigated to prevent over-alkylation.

2.2.2. Alternative alkylating agents

A number of DBHDA analogues were first designed and synthesised to present different physico-chemical properties for improving reactivity and accessibility towards the desired 218 (Figure 2.9.A): a more reactive 2,5-diiodohexane diamide (DIHDA), bulkier and more lipophilic 2,5-dibromo-N^1,N^6-dimethylhexanediamide (DBDMet) and N^1, N^6-bis(2-amino-2-oxoethyl)-2,5-dibromohexanediamide (DBDGla) to prevent their access to the "ATP pocket" and to the potential accessible lysine residues as well as the more soluble carboxylates, 2,5-dibromohexanedioic acid (DBHDAc) and 2,2'-((2,5-dibromohexanedioyl)bis(azanediyl) diacetic acid (DBDGly), to increase the rate of alkylation.

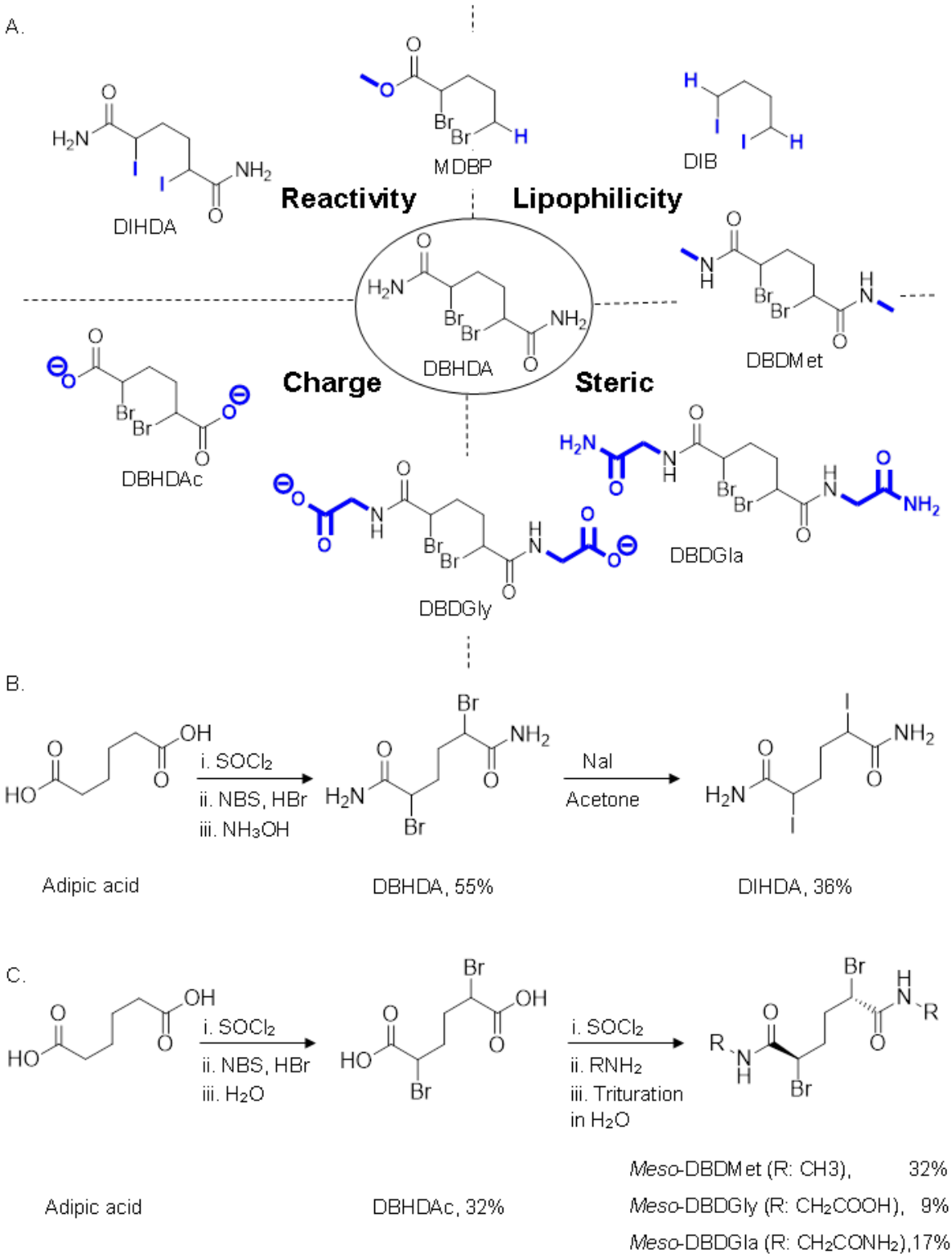


Figure 2.9. A. Alternative bis-alkylating agents towards the selective formation of Dha at position 218, arranged according to their physicochemical properties, B. Synthesis of DIHDA from DBHDA. C. Synthesis of the intermediate DBHDAc followed by its use in the formation of *meso*-DBHDA analogues, DBDMEt, DBDGly and DBDGla.

As they were reported in the formation of Dha moiety from cysteine residues at least on amino acids and peptides commercially available 1,5-diiodobutane (DIB) and 2,5-methyldibromopentonate (MDBP) were also considered.[117, 146] They are also less sterically hindered and potentially more reactive.

The synthesis of the intermediate 2,5-dibromohexanedioic acid (DBHDAc) followed a modified Hell-Vohlard-Zelinski reaction, where adipic acid was treated with thionyl chloride and NBS in a similar fashion to that reported with DBHDA.[162-165] DBHDAc could in turn be functionalised using amide coupling conditions to yield the DBHDA analogues as a mixture of diastereoisomers.[166] Subsequent trituration in water afforded pure *meso*-forms, albeit in very low recovery (Figure 2.9.C). DIHDA was made from DBHDA by reaction with sodium iodide in acetone (Figure 2.9.B).

Unfortunately, the use of these bis-alkylating/elimination agents without addition of ADP-Mg did not provide pure MEK1Dha218.

In addition to DBDMet and DIHDA, the water soluble DBHDAc and DBDGly did not react with MEK1C218. The reaction of MEK1C218 with DIB in the presence of ADP-Mg was very slow and only 15% conversion was observed to MEK1Dha218 after 26 hours at 37 °C (Figure 2.10).

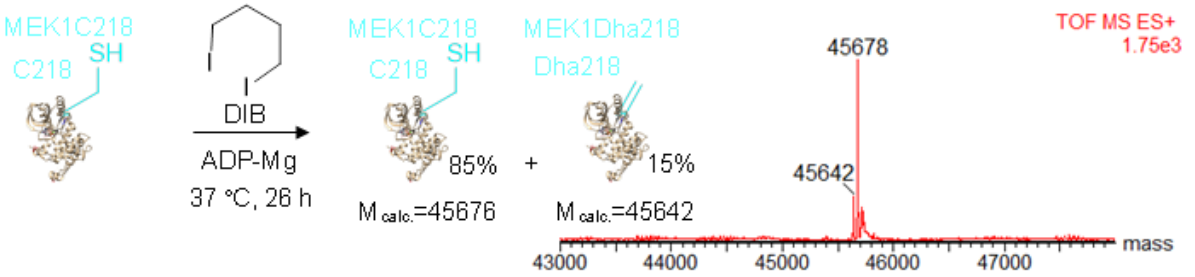


Figure 2.10. DIB reaction with MEK1C218, $M_{obs.}$=45678, $M_{calc.}$=45676 afforded 15% of MEK1Dha218, $M_{obs.}$=45642, $M_{calc.}$=45642.

In contrast, the reaction with DBHDGla afforded MEK1Dha218, site-selectively, after 20 hours at 37 °C (Figure 2.11).

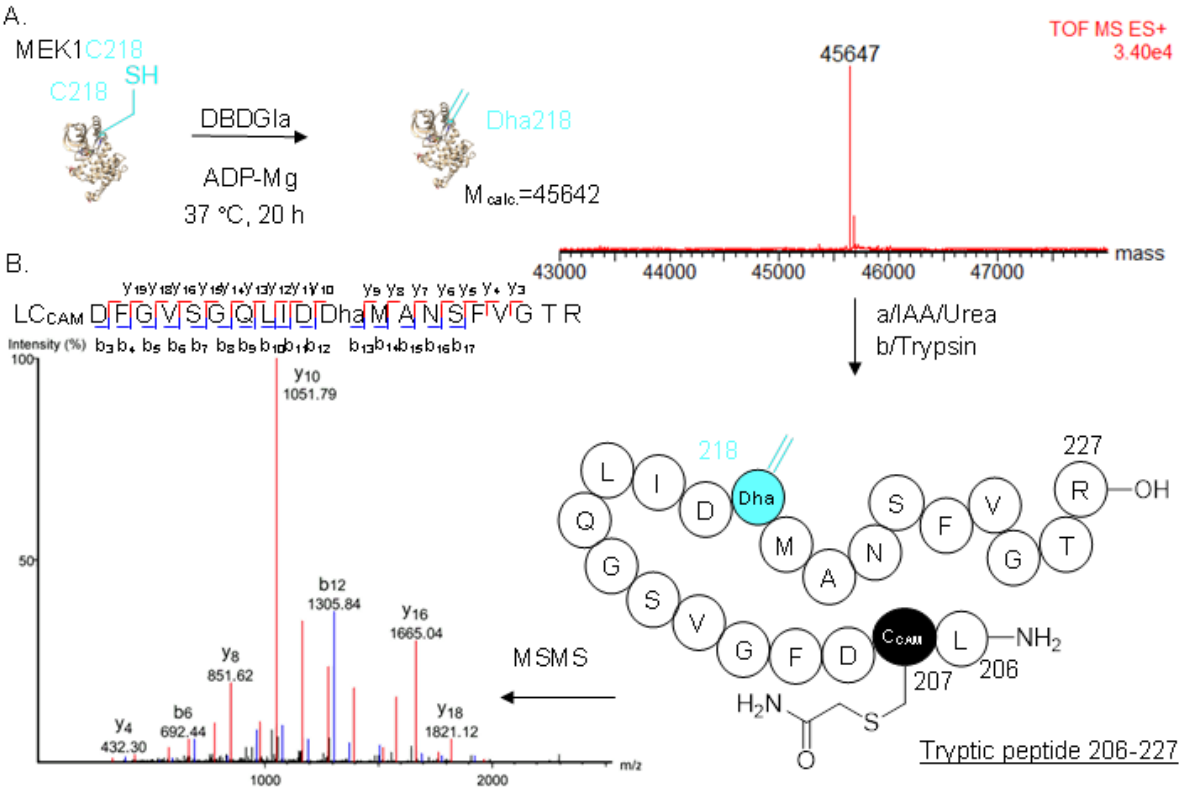


Figure 2.11. A. The treatment of MEK1C218 with DBDGla in the presence of ADP-Mg at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=45647 was consistent with MEK1Dha218, $M_{calc.}$=45642. B. Tryptic peptides were prepared from MEK1Dha218 denatured with urea and with the free unreacted cysteine residues capped with iodoacetamide (IAA), diluted and treated with trypsin at 37 °C. LCMS-MS of tryptic peptide 206-227 confirmed that position 207 was a capped unreacted cysteine residues and 218 was a dehydroalanine (Dha).

In order to accelerate the reaction rate, the more reactive MDBP was used. It allowed full conversion of MEK1C218 to MEK1Dha218 in just 3 hours at 37 °C (Figure 2.12).

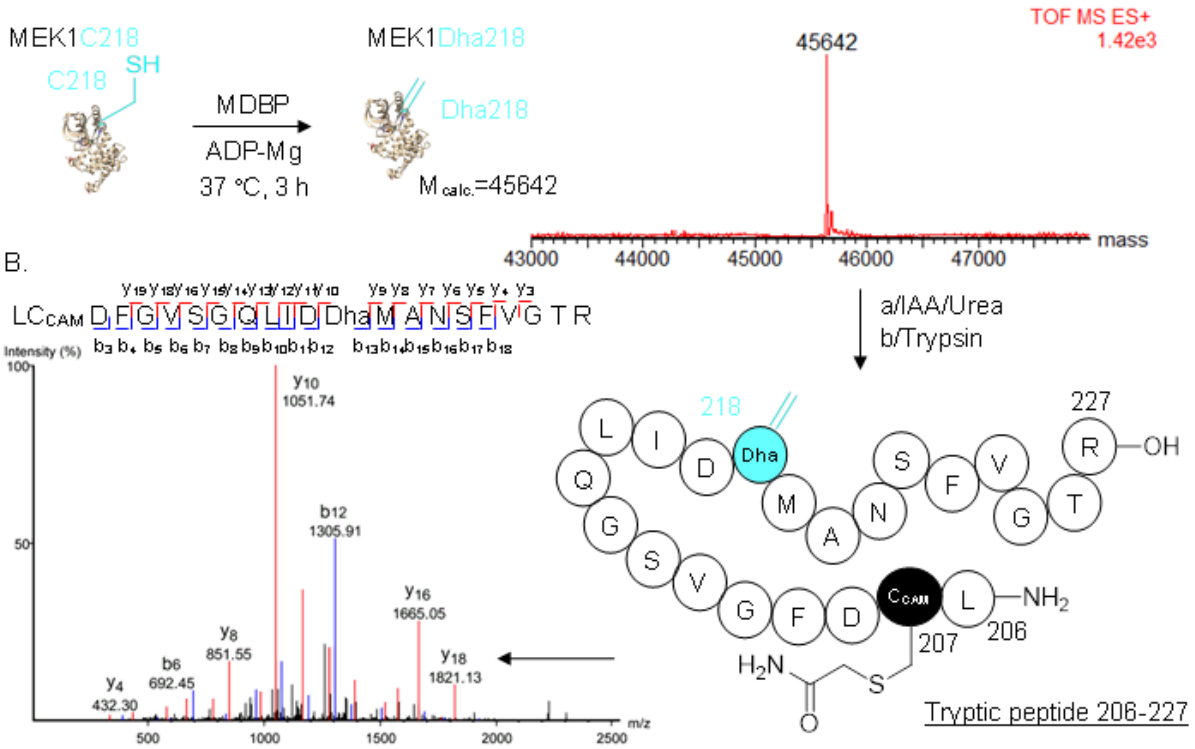


Figure 2.12. A. The treatment of MEK1C218 with MDBP in the presence of ADP-Mg at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=45642 was consistent with MEK1Dha218, $M_{calc.}$=45642. B. Tryptic peptides were prepared from MEK1Dha218 denatured with urea and with the free unreacted cysteine residues capped with iodoacetamide (IAA), diluted and treated with trypsin at 37 °C. LCMS-MS of tryptic peptide 206-227 confirmed that position 207 was a capped unreacted cysteine residues and 218 was a dehydroalanine (Dha).

2.3. A mild and selective formation of a dehydroalanine residue in a protein

Site-selective Dha formation has been achieved on MEK1, which contains numerous, accessible and reactive free cysteine residues.

The formation of Dha allowed the introduction of a unique electrophilic handle at the key positions 218 and 222 of MEK1 (Chapter 5).

Importantly, circular dichroism studies revealed that the methodology did not disrupt the protein secondary structure, necessary to retain its activity (Figure 2.13). The use of ADP-Mg as stabilising agent and a mild non-oxidative cysteine-specific bis-alkylating agents explains it. While the conformational flexibility of the loop around the position 222 would accommodate the transformation of a sp^3 to sp^2 centre during Dha formation, one could

argue that variations in backbone geometry at the position 218 laying in an α-helix could have structural consequences.[167] Any such changes, however, might be too small to be detected by CD.

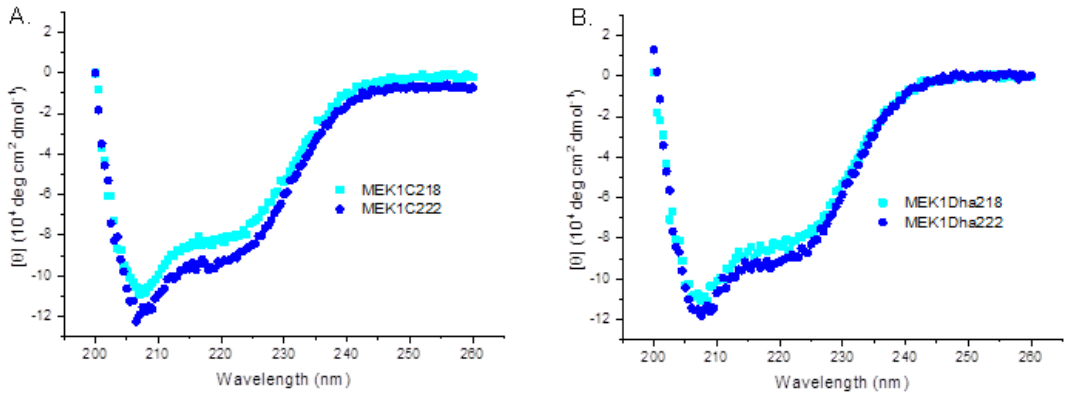


Figure 2.13. A. Comparative circular dichroism (CD) spectra of MEK1C218 and MEK1C222. B. Comparative CD spectra of MEK1Dha218 and MEK1Dha222.

2.4. An insight into the chemical mechanism of Dha formation

2.4.1. The different reactivity of bis-alkylating/elimination agents

During the optimisation of MEK1Dha218 formation, the various bis-alkylation/elimination agents were also tested with MEK1C222 (Table 2.1).

DBDGly did not react with this construct (Entry 6). Promising selectivity towards C222 against C207 was observed when DBDMet was used without the presence of nucleotide-magnesium complex (Entry 2 and 3). However, as the alkylating agent did not react with MEK1C218, and possessed very poor aqueous solubility, DBDMet was abandoned. The other alkylating agents, DIB, DBDGla, MDBP and DBHDAc, displayed selectivity only in the presence of ADP-Mg. The reaction of MEK1C222 with DIB was slow, but full conversion to MEK1Dha222 was observed after 18 hours at 37 °C (Entry 7). The rate of formation of MEK1Dha222 with DBDGla was similar to the one with DBHDA, around 5 hours (Entry 7).

Unexpectedly, the formation of MEK1Dha222 with the highly reactive MDBP was slower than the formation of MEK1Dha218 (Entry 8).

Table 2.1. MEK1C222 treated with different bis-alkylating/elimination agents

Entry	Alkylating agent	Additive	Time	$M_{obs.}$ in Da	Product ($M_{calc.}$ in Da)	Conversion
1	DBHDA	ADP-Mg	5 h[a]	45645	MEK1Dha222 (45642)	100%
2	DBDMet	None	16 h[a]	45646 45891	MEK1Dha222 (45642) oaMEK1Dha222[c] (45891)	90% 10%
3	DBDMet	ADP-Mg	16 h[a]	45645	MEK1Dha222 (45642)	100%
4	DBDGla	None	5 h[a]	45981	oaMEK1Dha222[c] (45977)	100%
5	DBDGla	ADP-Mg	5 h[a]	45645	MEK1Dha222 (45642)	100%
6	DBDGly	ADP-Mg	16 h[a]	45676	No reaction	0%
7	DIB	ADP-Mg	20 h[a]	45645	MEK1Dha222 (45642)	100%
8	MDBP	ADP-Mg	5 h[a]	45645	MEK1Dha222 (45642)	100%
9	DBHDAc	ADP-Mg	4 h[b]	45840	(45836)	100%

[a]Reaction performed at 37 °C. [b]Reaction performed at room temperature. [c]oa: C207 over-alkylation.

Finally, DBHDAc reacted with MEK1C222 (Entry 9). LCMS analysis showed the full conversion to a MEK1 construct with $M_{obs.}$=45840. We hypothesised that the sulfonium ion did not eliminate but opened under the nucleophilic attack of hydroxide ions, $M_{calc.}$=45836 (Figure 2.14.A). The nucleophilic attack on a stable sulfonium ion generated by the reaction of DBHDA on the green fluorescent protein construct, GFPS147C has been previously described (Figure 2.14.B).[168] In this case, the sulfonium ion was used as bioorthogonal handle, which allowed the insertion of azide, phtalimide and thiols.[121]

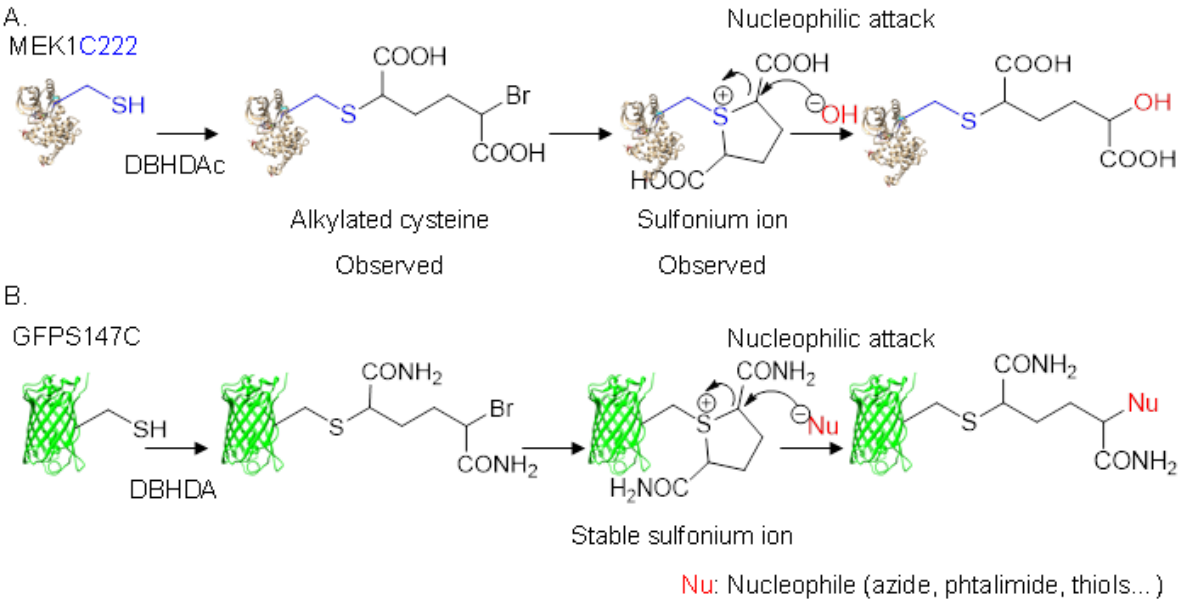


Figure 2.14. A. LCMS monitoring of the reaction of DBHDAc with MEK1C222 revealed the nucleophilic attack of hydroxide with the sulfonium ion. B. The treatment of GFPS147C construct with DBHDA afforded a stable sulfonium ion, which was used as a bioorthogonal handle, proned to open with a variety of nucleophiles.[121, 168]

The different reactivity displayed by the bis-alkylation/elimination agents was intriguing. The reactivity and accessibility of the reagents could not explain on their own the slow reaction of C222 with MDBP and the product of the reaction with DBHDAc. Therefore, a deeper analysis of the reaction profile was performed.

2.4.2. Mechanistic insight into the formation of MEK1Dha222

The proposed mechanism of the formation of Dha using bis-alkylating agents follows a three-step process (Figure 2.15).[117,121] The cysteine residue is first alkylated. Then, a sulfonium ion is formed by the intramolecular nucleophilic substitution of the second halide by the alkylated cysteine sulfur. The sulfonium ion then readily eliminates, *via* the abstraction of the α-proton of the back bone amide, to yield the dehydroalanine residue.

The different intermediates were detectable by LCMS monitoring of the reaction of MEK1C222 with various bis-alkylating/elimination agents (Figure 2.15). The sulfonium ion was not observed when DBHDA and its analogues were used (Figure 2.15.A). It implied that the rate of elimination was very fast. In contrast, the first alkylation product was not

detectable when DIB was used, although the sulfonium ion was (Figure 2.15.B). These observations highlighted two aspects of the mechanism of Dha formation with DIB. Firstly, the second alkylation was particularly fast, which was probably due to fact that DIB was sterically smaller and could cyclise more easily. Secondly, the stability of the sulfonium ion may indicate a different mechanism for its elimination compared to the elimination of the sulfonium ion generated by DBHDA and its analogues. Finally, the relatively slow formation of MEK1Dha222 with MDBP was explained by the additional hydrolysis of the sulfonium ion ester. The carboxylic acid sulfonium was particularly stable (Figure 2.15.C), which slowed down its elimination towards the formation of Dha. However, this hydrolysis was not observed in the formation of MEK1Dha218.

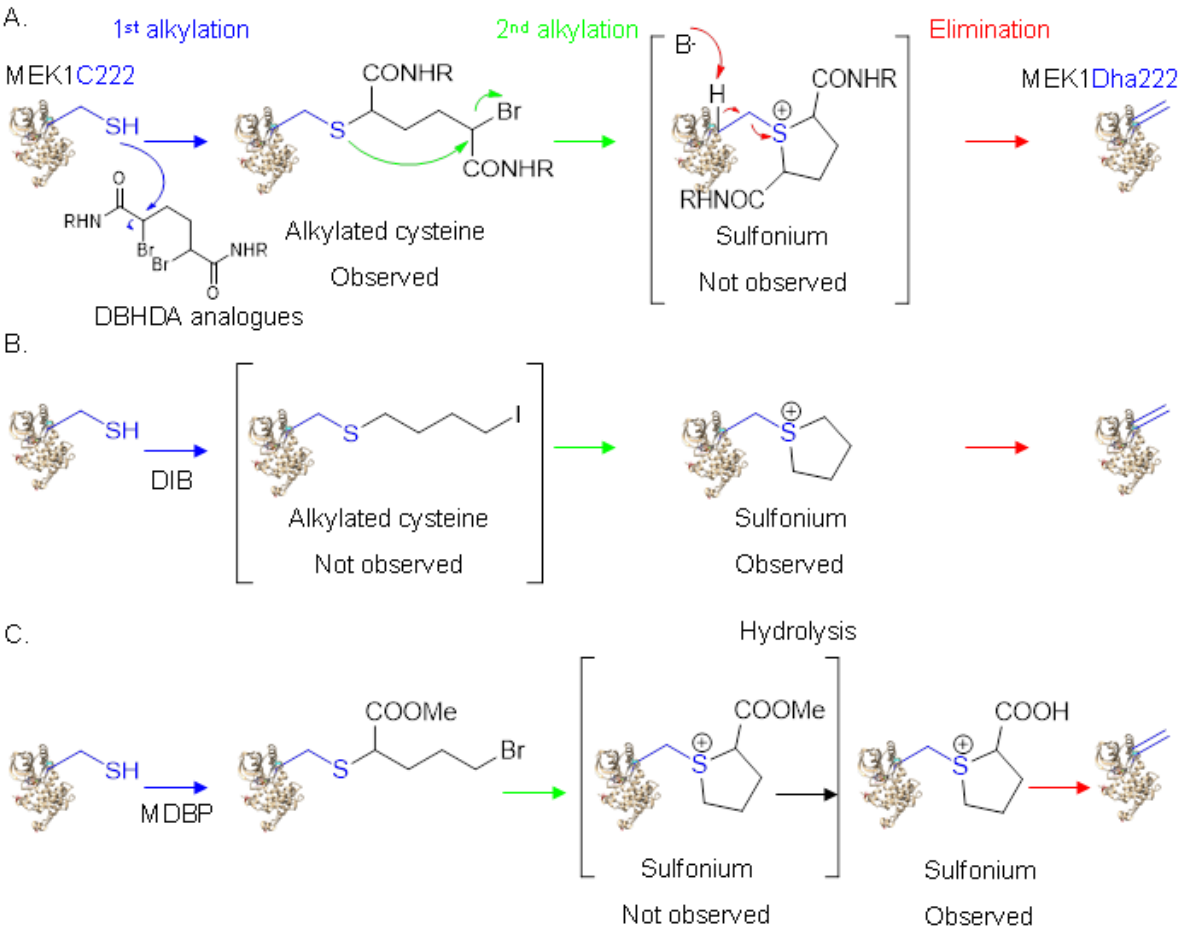


Figure 2.15. A. The formation of Dha followed a three step process: first alkylation (blue), second alkylation (green) and elimination (red). The sulfonium ion, product of the second alkylation was not observed on LCMS monitoring when DBHDA analogues were used. B. With DIB, the product of the first alkylation was not detected by LCMS. C. LCMS monitoring highlighted the hydrolysis of the ester moiety, which stabilised the sulfonium ion intermediate.

Further understanding the difference of the mechanism of elimination between DIB and DBHDA would lead to the design of more efficient alkylating agents. The elucidation of the mechanism of elimination is described in next chapter (Chapter 3).

Chapter 3: Mechanism of Elimination during Dehydroalanine Formation

3.1. The formation of a sulfonium ylid intermediate

3.1.1. Potential mechanisms of sulfonium ion elimination

In chapter 2, various alkylating agents were used to form a dehydroalanine from a cysteine residue in MEK1. LCMS analysis of the reactions allowed the direct observation of the three-step process: the first alkylation was followed by a second alkylation and the formation of a sulfonium ion, which eliminated to protein Dha. The cyclic tetrahydrothiophene (THT) was liberated as the evidence of the formation of a sulfonium ion (Figure 3.1.A).

The sulfonium ion elimination was investigated further. In principle, under basic conditions once the suflonium was formed, three positions could be deprotonated: Hα on the backbone amide, and Hβ and Hγ next to the positively charged sulfur atom (Figure 3.1.A). Therefore, variations on four mechanisms of elimination of a sulfonium ion could be considered.[169, 170] In the concerted E2 mechanism, the base, hydroxide ions in aqueous media, attacks Hα, which initiates a concerted anti-periplanar elimination (Figure 3.1.B). The deprotonation of Hγ leads to the formation of a sulfonium ylid, which readily eliminates in a syn fashion by the extraction of Hα (Figure 3.1.C). In both cases, the created αC carbanion can be stable and promote a non-concerted E1cb elimination (Figure 3.1.D). Finally, Hβ can also be deprotonated to form a carbene (Figure 3.1.E). In all cases, the result is the elimination of the sulphonium and formation of dehydroalanine. Early mechanistic studies favoured the

E2-elimination and syn-elimination as the elimination mechanisms of small molecule sulfonium ion.[171-175]

In this chapter, the elimination of the sulfonium generated by DIB and DBHDA is compared. Firstly, the formation of the stable sulfonium formed with DBHDA on GFPC147 (Figure 2.14) was examined.

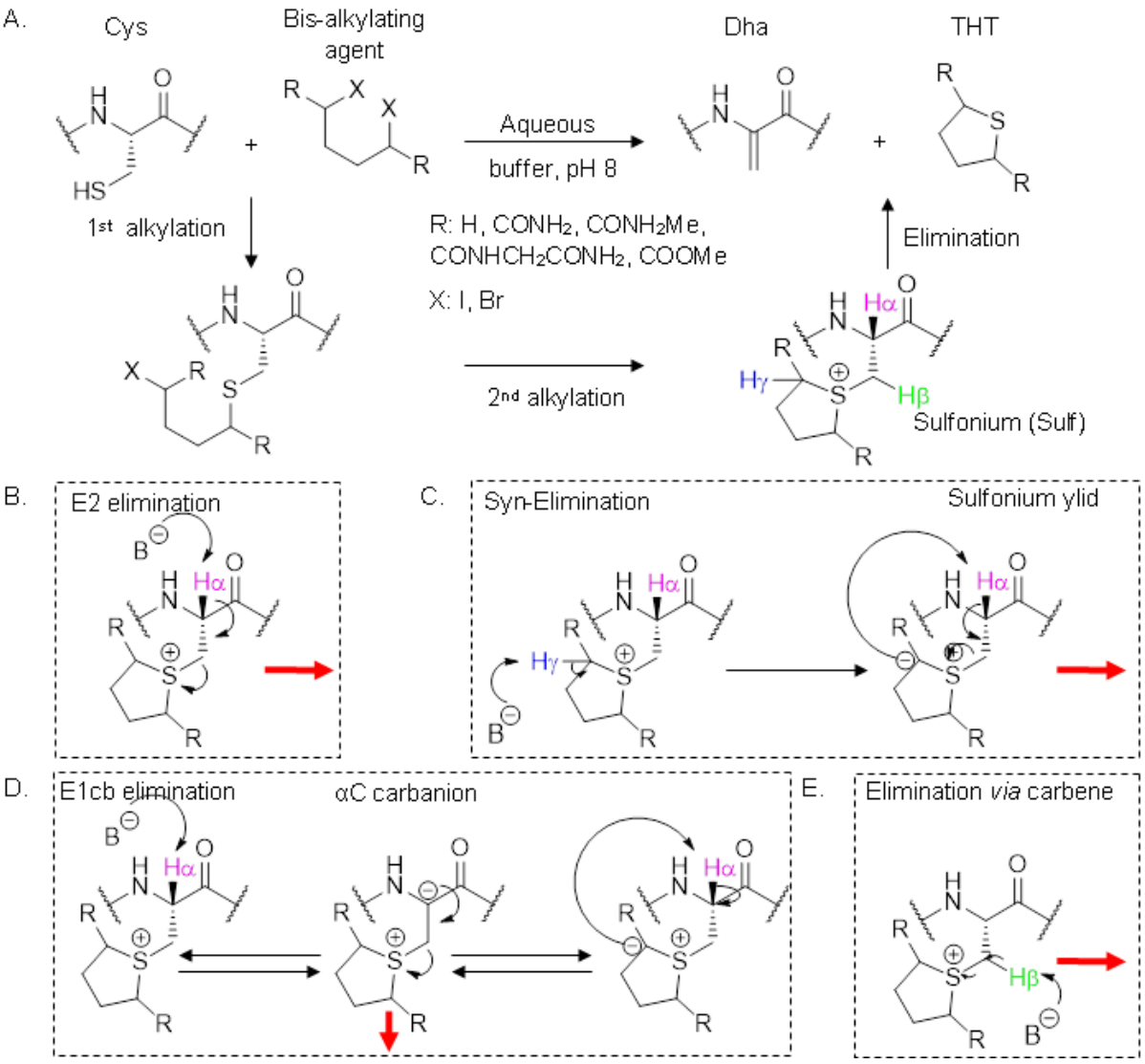


Figure 3.1. A. The treatment of cysteine residues with bis-alkylating agents led to the formation of dehydroalanine residue and tetrahydrothiophene by-product (THT). B. Hα deprotonation triggers the antiperiplanar elimination (red arrow) of the sulfonium in an E2 fashion. C. Hγ deprotonation forms a sulfonium ylid, which abstracts Hα, and initiates a syn elimination. D. The direct Hα or indirect Hγ deprotonation lead to the formation of a αC carbanion, which promotes the elimination in an E1cb manner. E. Hβ deprotonation produces a carbene that causes the formation of Dha.

3.1.2. The degradation of a stable sulfonium involved a sulfonium ylid intermediate[‡]

As described in Chapter 2, the treatment of GFPC147 with DBHDA did not afford GFPDha147 but the stable sulfonium GFPSul147 (Figure 3.2). The stability of the sulfonium ion was explained by the inaccessibility of the α-proton of the backbone protein as estimated by Naccess calculation.[121] However, the sulfonium ion stability was relative, and it degraded over time at 37 °C. LCMS analysis was consistent with the direct fragmentation of the sulfonium by [3+2] cycloreversion to form the corresponding protein vinyl sulfide with concomitant elimination of acrylamide, which has been observed on small molecules.[173, 176] Such cycloreversion reaction implied the formation of a sulfonium ylid intermediate (Figure 3.2).

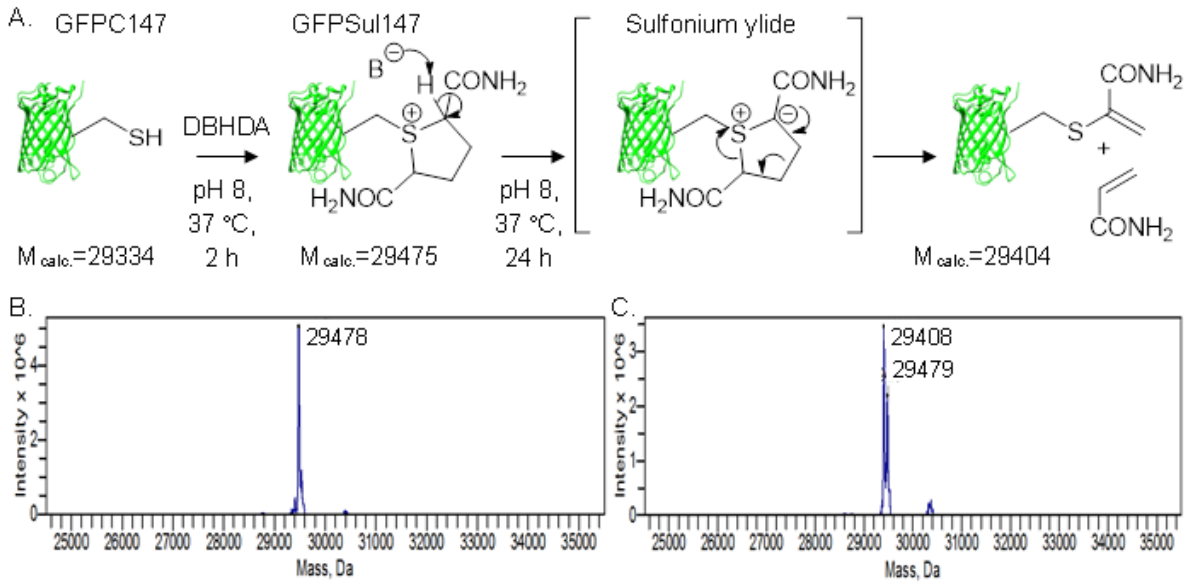


Figure 3.2. A. LCMS analysis of the reaction between GFPC147 construct and DBHDA revealed the formation of a stable sulfonium, $M_{obs.}$=29478, $M_{calc.}$=29475, which degraded by a [3+2] cycloreversion triggered by a sulfonium ylid intermediate to form the corresponding protein vinyl sulfide $M_{obs.}$=29408, $M_{calc.}$=29404 and the acrylamide. B. & C. Corresponding MS of sulfonium and vinyl sulfide

[‡] In collaboration with Dr. Vijay Chudasama and Prof. Stephen Caddick, UCL, UK.

3.1.3. Visualisation of the sulfonium ylid in deuterated buffer

In order to visualise the formation of the sulfonium ylid, the reaction was performed in deuterated buffer, which enabled the deuterium labelling of any deprotonated sites, since these sites would be quenched by the buffer (Figure 3.3.A). Two proteins were chosen for this study: the protease subtilisin bacillus lentus, SBL, and the histone H3. These proteins have previously been used as model proteins in the optimisation of the formation of Dha by bis-alkylation/elimination methodology.[94] In particular, SBLS156C (SBLC156) and H3S10C (H3C10) constructs have been described to readily form the corresponding SBLDha156 and H3Dha10 with DBHDA.[94, 117] LCMS analysis of the reactions in deuterated buffer with both proteins revealed that the tetrahydrothiophene generated by DBHDA, referred hereafter to THTCD, was deuterated on the *γ*-C (Figure 3.3.B and 3.3.C).

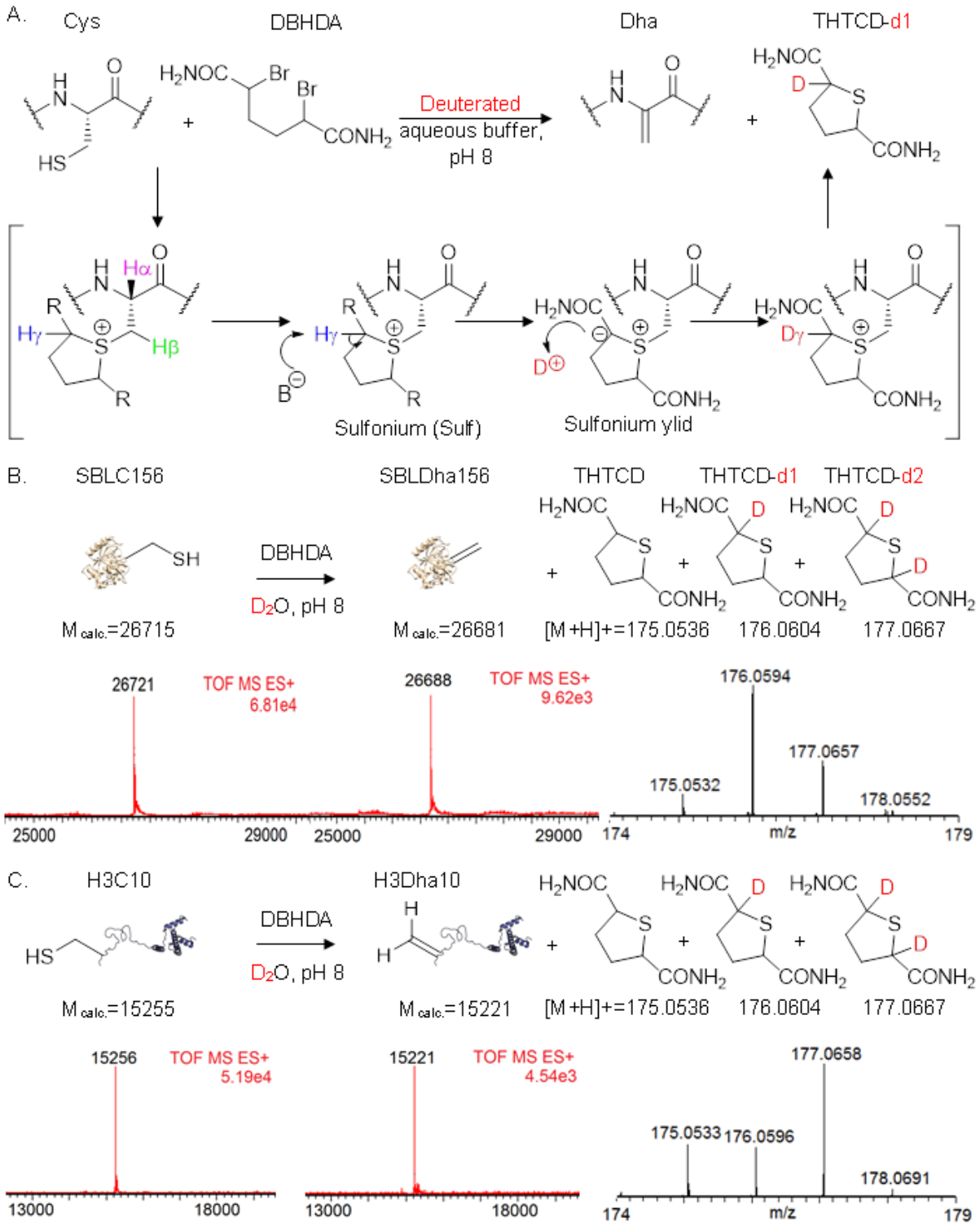


Figure 3.3. A. The mechanism of deuteration of tetrahydrothiophene THTCD involved the formation of a sulfonium ylid. B. LCMS analysis of the reaction between SBLC156 and DBHDA in deuterated buffer confirmed the formation of deuterated THTCD. C. LCMS analysis of the reaction of H3S10 with DBHDA in deuterated buffer confirmed the formation of deuterated THTCD.

As control experiments indicated that neither DBHDA nor THTCD were deuterated at γ-C under the reaction conditions, the deuteration could only occur on the sulfonium ion, which

confirmed the formation of a sulfonium ylid. The formation of doubly deuterated THTCD (THTCD-d2) indicated a fast protonation/deprotonation equilibrium with the buffer. In addition, as the dehydroalanine residue was not deuterated, the mechanism involving a carbene formation was therefore excluded.

Once the existence of the sulfonium ylid was confirmed, its potential role as a base was explored. The strategy involved the use of α-deuterated cysteine residues. Indeed, the syn-elimination initiated by the sulfonium ylid would be detectable by the LCMS analysis of the THTCD by-product. The transfer of the deuterium would confirm that the sulfonium ylid could act as a base (Figure 3.4).

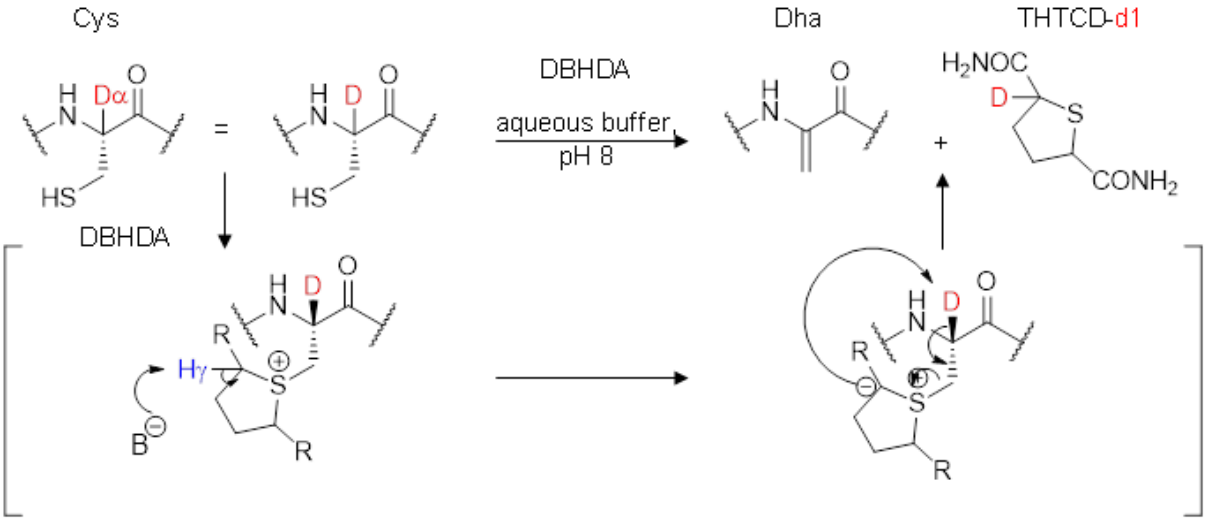


Figure 3.4. In a syn-elimination triggered by the sulfonium ylid, the α-deuterium is transferred to the by-product THTCD.

T

he site-selective α-deuteration of a cysteine residue on a protein was therefore attempted.

3.2. Site selective deuteration of cysteine residue in protein

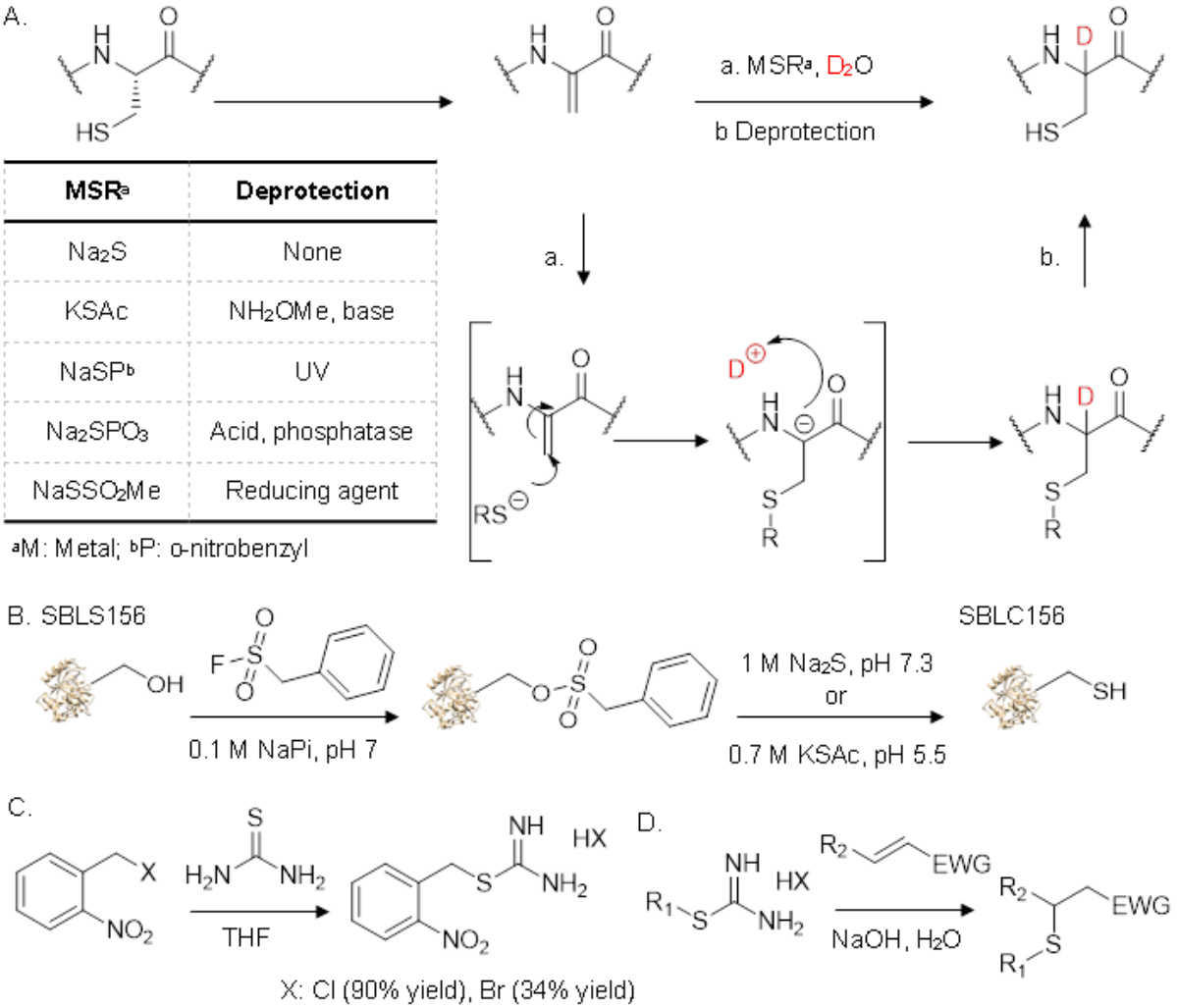
3.2.1. Site selective protein deuteration strategy

Methods for site-selective incorporation of deuterium into proteins are limited. One approach is the segmental isotopic labelling strategy.[177,178] This approach involves

expressed protein ligation (EPL) and *trans*-splicing techniques using synthetically prepared peptides containing isotopically labelled amino acids to afford full-length selectively labelled protein.[179-181] Alternatively, the synthesis of deuterated amino acids followed by biosynthetic incorporation leads to a global, and therefore a non-site-selective labelling of any given amino acid.[182-186] The incorporation of unnatural deuterated amino acids by amber-stop codon technology resolved the selectivity issue.[187] However, the selective α-deuteration of cysteine amino acid is not common and often convoluted.[188-190]

Inspired by early reports about the partial exchange of the C2 hydrogen of histidine residues with deuteroxide (D_2O) in bovine pancreatic ribonuclease, we postulated that site-selective α-C labelling might be achieved *via* Dha itself through conjugate addition of an appropriate thiolate (MSR) in deuterated buffer (Figure 3.5.A).[191] The addition would be followed by a cleavage step to remove the R group when necessary, leaving the desired cysteine deuterated at α-C. Five different thiolate sources were considered: sodium sulphide (Na_2S), potassium thioacetate (KSAc), *o*-nitrobenzyl thiol derivative, sodium thiophosphate (Na_3SPO_3) and sodium methanethiosulfonate ($NaSSO_2Me$). Both Na_2S and KSAc have successfully employed on protein substrate for the chemical modification of SBLS156 to SBLC156 (Figure 3.5.B).[111,192] In these cases, the acetate was hydrolysed under the reaction conditions. However, other studies, on other proteins, have described methoxylamine as a thioacetate cleaving agent.[193] The light sensitive *o*-nitrobenzyl cysteine (Nbc)[194-198] have been introduced using readily available odourless *o*-nitrobenzyl thiol derivatives[197] (Figure 3.5.C) to afford the corresponding thiolate ion under basic conditions in small molecule models (Figure 3.5.D).[199] The nitrobenzyl moiety could then be removed using UV light.[200, 201] Na_3SPO_3 was also envisaged as a suitable candidate for conjugate addition/deprotection, as its reaction with SBLDha156 and H3Dha10 yielded the corresponding phosphocysteine residues, SBLpC156 and H3pC10. Phosphocysteine residues hydrolysed back to cysteine residues under acidic conditions,[136] and some phosphoserine phosphatases have been shown to dephosphorylate them.[93] Finally, $NaSSO_2Me$ appeared as an attractive option for

conjugate addition, since common reducing agents, such as dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP), would subsequently afford the desired cysteine residue.



MSR[a]	Deprotection
Na_2S	None
KSAc	NH_2OMe, base
NaSP[b]	UV
Na_2SPO_3	Acid, phosphatase
$NaSSO_2Me$	Reducing agent

[a]M: Metal; [b]P: o-nitrobenzyl

Figure 3.5. A. Strategy of site-selective deuterium labelling of a cysteine residue involved the site-selective formation of dehydroalanine followed by the thiamichael addition of protected thiolate, MSR (M: Metal) in deuterated buffer. Various thiolates were considered: sodium sulphide, Na_2S, potassium thioacetate, KSAc, with their corresponding deprotection conditions. B. SBLS156C chemical mutagenesis used Na_2S and KSAc.[111, 192] C. Alkylation of thiourea afforded odourless thiolate in good yields. D. The thiamichael reaction of Michael acceptor with odourless thiolate in aqueous basic conditions has been reported.[197, 199]

3.2.2. Site-selective deuteration of SBLC156

Initial attempts in non-deuterated buffer showed that the reaction of SBLDha156 with sodium sulphide, potassium thioacetate and *o*-nitrobenzyl thiourea provided the corresponding SBLC156, SBLAcc156 and SBLNbc156 (Figure 3.6.A, B and C). However,

the reaction with potassium thioacetate was not pursued as the acetate moiety did not hydrolyse under acidic conditions or cleave with the addition of methoxylamine. The reaction with sodium sulphide was optimised, and transferred to deuterated buffer to afford SBLdC156 in good conversion (Figure 3.6.C). However, the full characterisation using LCMS-MS of the tryptic peptides was not successful.

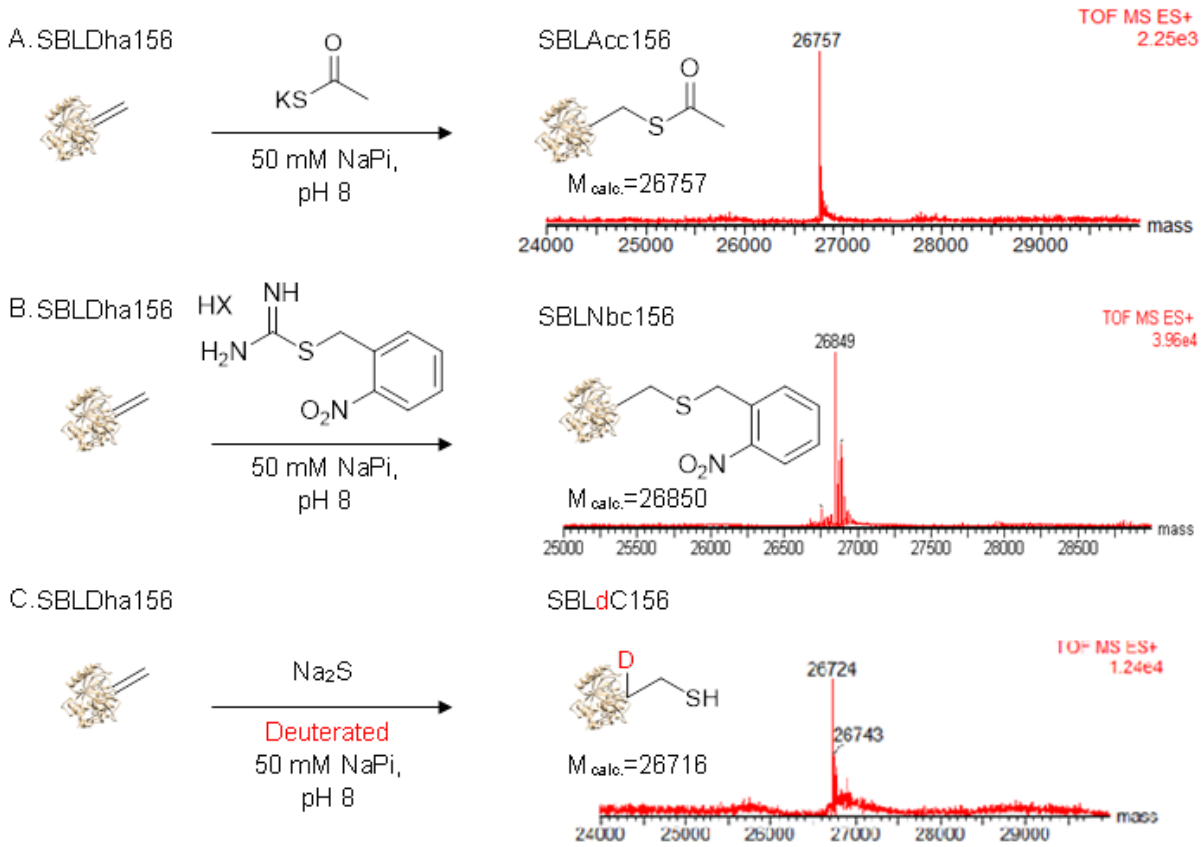


Figure 3.6. A. The treatment of SBLC156 with KSAc at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=26757 was consistent with SBLAcc156, $M_{calc.}$=26757. B. The treatment of SBLC156 with Nbc precursor at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=26849 was consistent with SBLNbc156, $M_{calc.}$=26850. C. The treatment of SBLC156 with Na_2S at 15 °C and pH 7 in deuterated buffer was monitored by LCMS. $M_{obs.}$=26724 was consistent with SBLdC156, $M_{calc.}$=26716.

Deuteration of histone H3, as our protein model, was then investigated.

3.2.3. Site-selective deuteration of histones H3C10 and H3C26

The formation of deuterated cysteine residues with Na_2S under the same conditions optimised for SBL (Figure 3.6.C) was not transferable to histone H3. However, H3Dha10 reacted with sodium thiophosphate in denaturing deuterated buffer to afford H3dpC10

(Figure 3.7.A). The deprotection of the phosphocysteine to the cysteine residue was achieved using two distinct methods. First, the addition of free glutathione, to lower the pH of the buffer to 4, readily hydrolysed H3dpC10 to H3dC10 (Figure 3.7.A). The phosphocysteine residue could also be cleaved by PP1, the protein phosphatase regulating the phosphorylation of H3S10 *in vivo* (Figure 3.7.A).[202] LCMS-MS of the tryptic peptides not only confirmed the site of deuteration at position 10 (Figure 3.7.B) but also allowed the estimation of the level of deuteration (Figure 3.7.C). The deuterium incorporation level depended greatly on the deuteration of all the protein exchangeable protons, and the use of anhydrous sodium thiophosphate. Prior to the reactions, the protein was therefore lyophilised several times in deuteroxide, and the commercially available sodium thiophosphate hydrate was dried in dry methanol using reported conditions.[203] To assess the generality of the method, another site of H3 was investigated. The site R26, known to methylate,[204] was chosen as it did not phosphorylate/dephosphorylate *in vivo*. The formation of H3dpC26 was successful, and its deprotection under acidic conditions afforded H3dC26. In a slower rate than H3dpC10, H3dpC26 was also dephosphorylated by PP1.

Site-selective deuteration of cysteine residues was achieved on SBLC156, H3C10 and H3C26 using the corresponding constructs containing dehydroalanine.

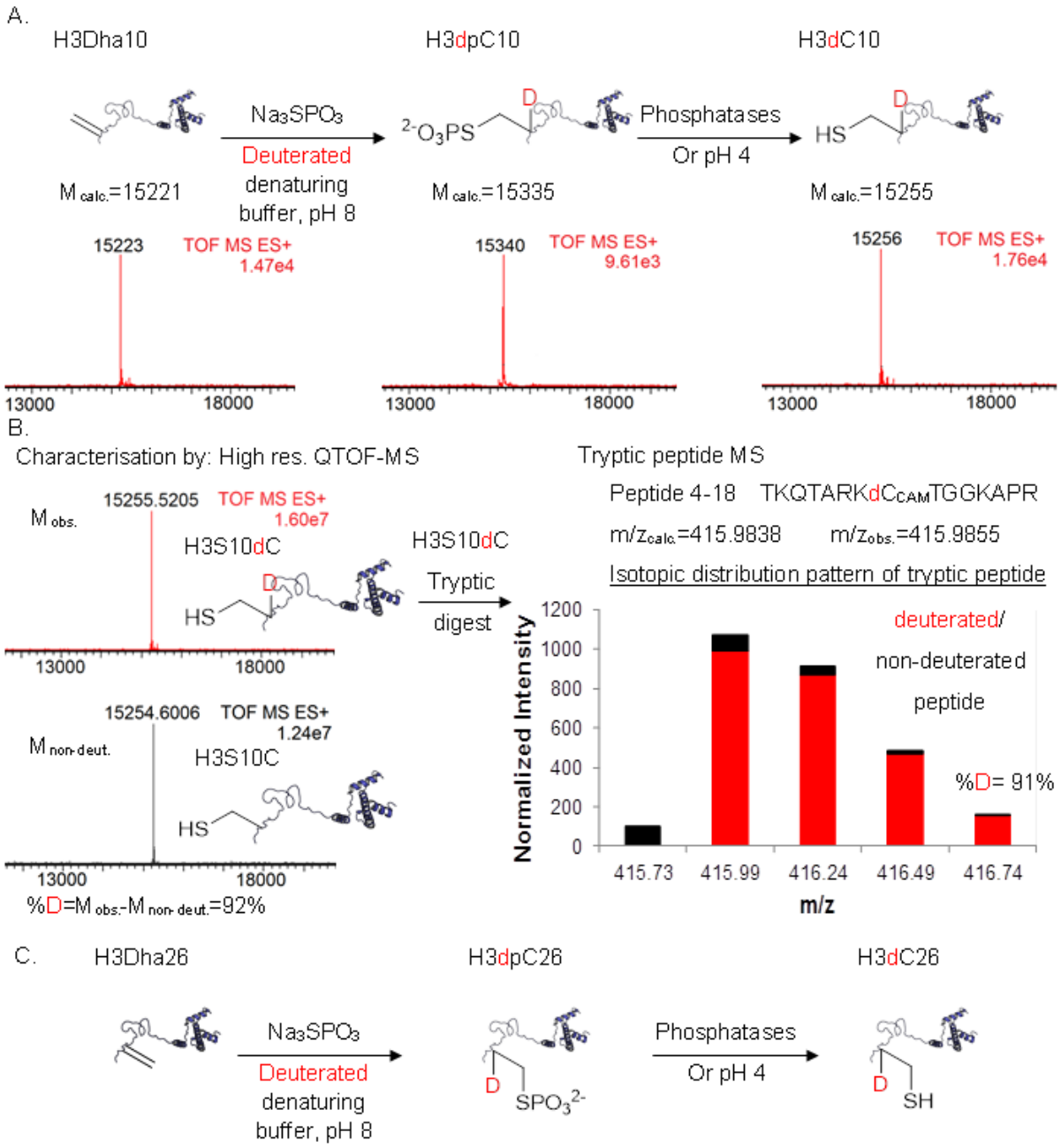


Figure 3.7. A. Formation of H3dC10 from H3Dha10 was performed in two steps: the thiamichael with sodium thiophosphate in deuterated buffer followed by hydrolysis in acidic conditions or by dephosphorylation with PP1 phosphatase. B. High resolution QTOF-MS and subsequent tryptic digest-MS analyses (as acetamide dC_{CAM}) confirmed D incorporation>90%; shown for H3dC10. C. The reaction was also performed on another position 26.

3.3. Sulfonium ylid as an intramolecular base

3.3.1. The fast solvent exchange hampered the mechanistic analysis

When SBLdC156, H3dC10 and H3dC26 were reacted with DBHDA in aqueous buffer, only non-deuterated THTCD was observed by LCMS (Figure 3.8.A, B and C). Although an E2 mechanism could not be discarded, this result was likely due to the fast exchange between the hydrogen Hγ of the sulfonium ion in the buffer upon ylid formation. Any deuterium insertion from the Dα onto the sulfonium ion was therefore washed out before elimination.

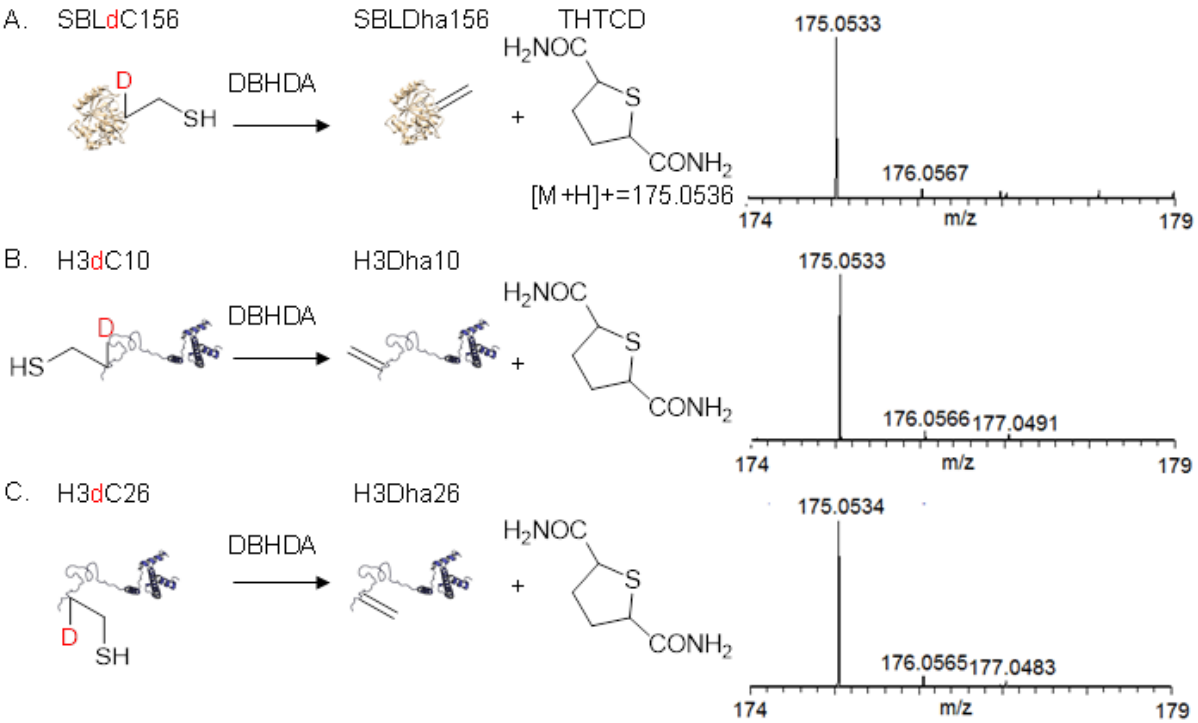


Figure. 3.8. A. The reaction of SBLdC156 with DBHDA afforded SBLDha156 and THTCD, $[M+H]+_{obs.}$=175.0533, $[M+H]+_{calc.}$=175.0536. B. The reaction of H3dC10 with DBHDA afforded H3Dha10 and THTCD $[M+H]+_{obs.}$=175.0533, $[M+H]+_{calc.}$=175.0536. C. The reaction of H3dC26 with DBHDA afforded H3Dha26 and THTCD, $[M+H]+_{obs.}$=175.0534, $[M+H]+_{calc.}$=175.0536.

To probe the role of the sulfonium ylid as a base and limit the influence of the buffer, the reactions were performed in an aprotic solvent.

3.3.2. Site-selective deuteration of cysteine amino acid

Dry DMSO was chosen as the aprotic solvent. However, this solvent was not compatible with protein chemistry. Therefore, an amino acid model, methyl acetyl-L-cysteinate, AcNHCysOMe, was selected. AcNHCysOMe treated with DBHDA in deuterated buffer led to the formation of AcNHDhaOMe, which was readily attacked by the remaining AcNHCysOMe to form the deuterated lanthionine, AcNHdLanOMe (Figure 3.9.A). More significantly, LCMS also indicated the deuteration of THTCD (Figure 3.9.B), which validated the use of AcNHCysOMe as a model.

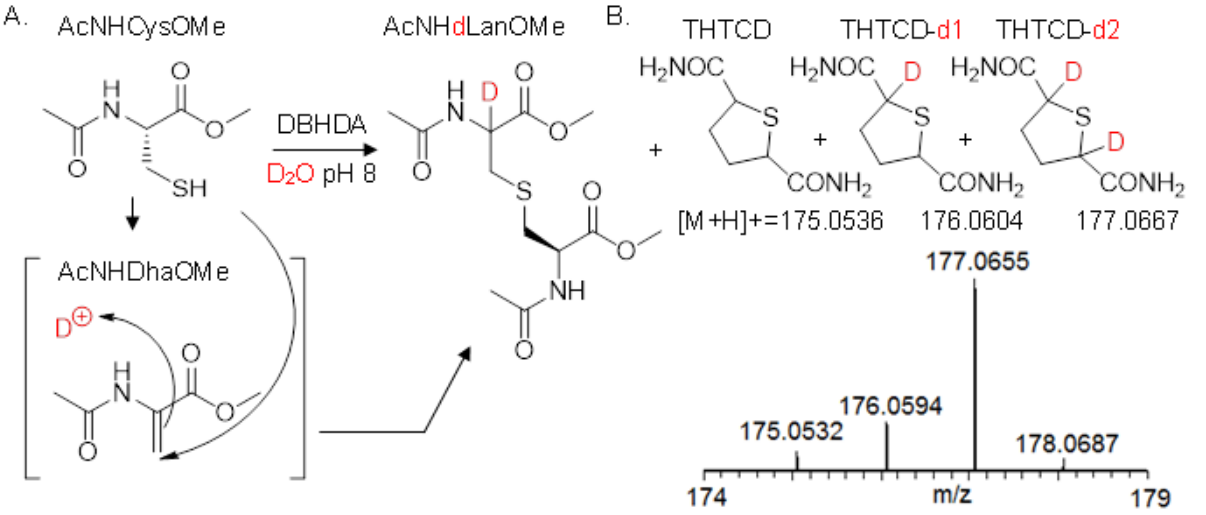


Figure 3.9. A. The reaction of AcNHCysOMe with DBHDA in deuterated buffer afforded the deuterated thioether AcNHdLanOMe, product of the reaction between the newly form AcNHDhaOMe and the remaining starting material AcNHCysOMe. B. The by-product THTCD was also deuterated, THTCD $[M+H]+_{obs.}$=175.0532, $[M+H]+_{calc.}$=175.0536, $[M+H]+_{obs.}$=176.0594, $[M+H]+_{calc.}$=176.0604 and $[M+H]+_{obs.}$=177.0655, $[M+H]+_{calc.}$=177.0667.

The same strategy of deuteration described for protein (Section 3.2.3) was used to form deuterated AcNHdCysOMe. Sodium methanethiosulfonate (Na_2SSO_2Me) did not react with AcNHDhaOMe (Figure 3.10.A). The treatment with sodium sulphide afforded AcNHLanOMe as described above (Figure 3.9.A and 3.10.B). The thiamichael reaction with o-nitrobenzyl thiourea deprotected with NaOD was successful and yielded the deuterated AcNHdNbcOMe (Figure 3.10.C). ^{1}H NMR and LCMS indicated that the deuteration also occurred at the

benzylic position. However, despite full deuteration of the αC position, initial deprotection attempts using UV light at 366 nm were not successful.[200] The use of additives like ascorbic acid and methoxylamine to quench both the oxidation of the cysteine residue and the formation of the aldehyde by-product necessitated extensive purification, not compatible with an efficient synthesis of AcNHdCysOMe.

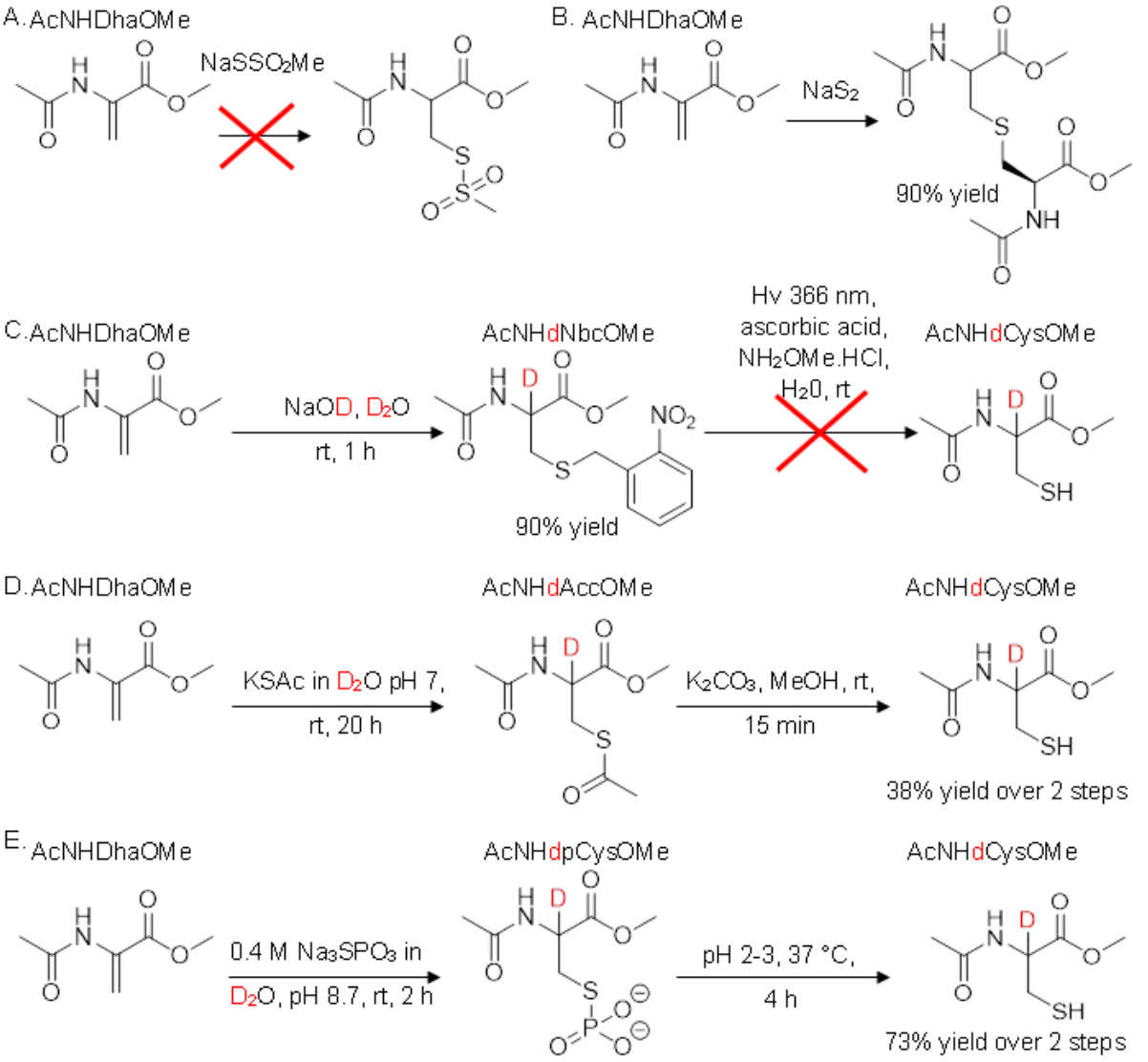


Figure 3.10. A. $NaSSO_2Me$ did not react with AcNHDhaOMe. B. The treatment of AcNHDhaOMe with Na_2S in water afforded AcNHLanOMe. C. The treatment of AcNHDhaOMe with o-nitrobenzyl thiourea in basic aqueous solution afforded deuterated AcNHdNbcOMe. The isolation of AcNHdCysOMe was not possible. D. The treatment of AcNHDhaOMe with KSAc afforded the desired AcNHdCysOMe after deacetylation with K_2CO_3 in 38% overall yield. E. The treatment of AcNHDhaOMe with Na_3SPO_3 afforded the desired AcNHdCysOMe after hydrolysis in acid in 73% overall yield.

The use of potassium thioacetate and sodium thiosulphate for synthesising AcNHdCysOMe was met with more success. The potassium thioacetate addition in D_2O was poor yielding (Figure 3.10.D). However, AcNHdAccOMe was readily hydrolysed to the desired AcNHdCysOMe with potassium carbonate in methanol in 38% overall yield. Finally, the treatment of AcNHDhaOMe with a deuterated solution of 0.4 M thiophosphate pH 8.7 followed by dephosphorylation at pH 2-3 gave the desired AcNHdCysOMe in a good 73% overall yield (Figure 3.10.E).

3.3.3. Sulfonium ylid acted as a base in DMSO

The successful synthesis of *α*-C deuterated AcNHdCysOMe allowed us to investigate the mechanism of elimination further. Indeed, the reaction with this amino acid model could be carried out in aprotic solvents.

The reaction of AcNHdCysOMe with DBHDA was done in both aqueous buffer and DMSO. As expected, the reaction in water afforded exclusively non deuterated THTCD, which confirmed the observations on the protein templates, and validated the amino acid model (Figure 3.11.A and Figure 3.8).

In contrast, the reaction in DMSO-d_6 followed by ^{1}H NMR indicated the formation of monodeuterated THTCD-d_1 (Figure 3.11.B and C). Indeed, the integrated intensity of Hγ hydrogen (Hc) was equivalent to one hydrogen, when compared with the integrated intensities of Dha protons (Ha and Ha'). The incorporation of one deuterium in THTCD implied the deprotonation of the only deuterium atom available in the α-position of AcNHdCysOMe by the sulfonium ylid (Figure 3.11.C).

In comparison, the same conditions were used with DIB. No deuteration was detected by ^{1}H NMR (Figure 4.12.A and B). The absence of deuterium insertion into THT suggested that the

elimination did not proceed *via* ylid formation. As a matter of fact, direct deprotonation of Hα leading to trans-E2 or E1cb elimination appeared more likely to proceed (Figure 3.12.C). In DMSO, the pK_a of the α-protons of THT sulfonium was calculated to be at around 17, while the α-proton of the amino acid residue was around 15.[205,206]

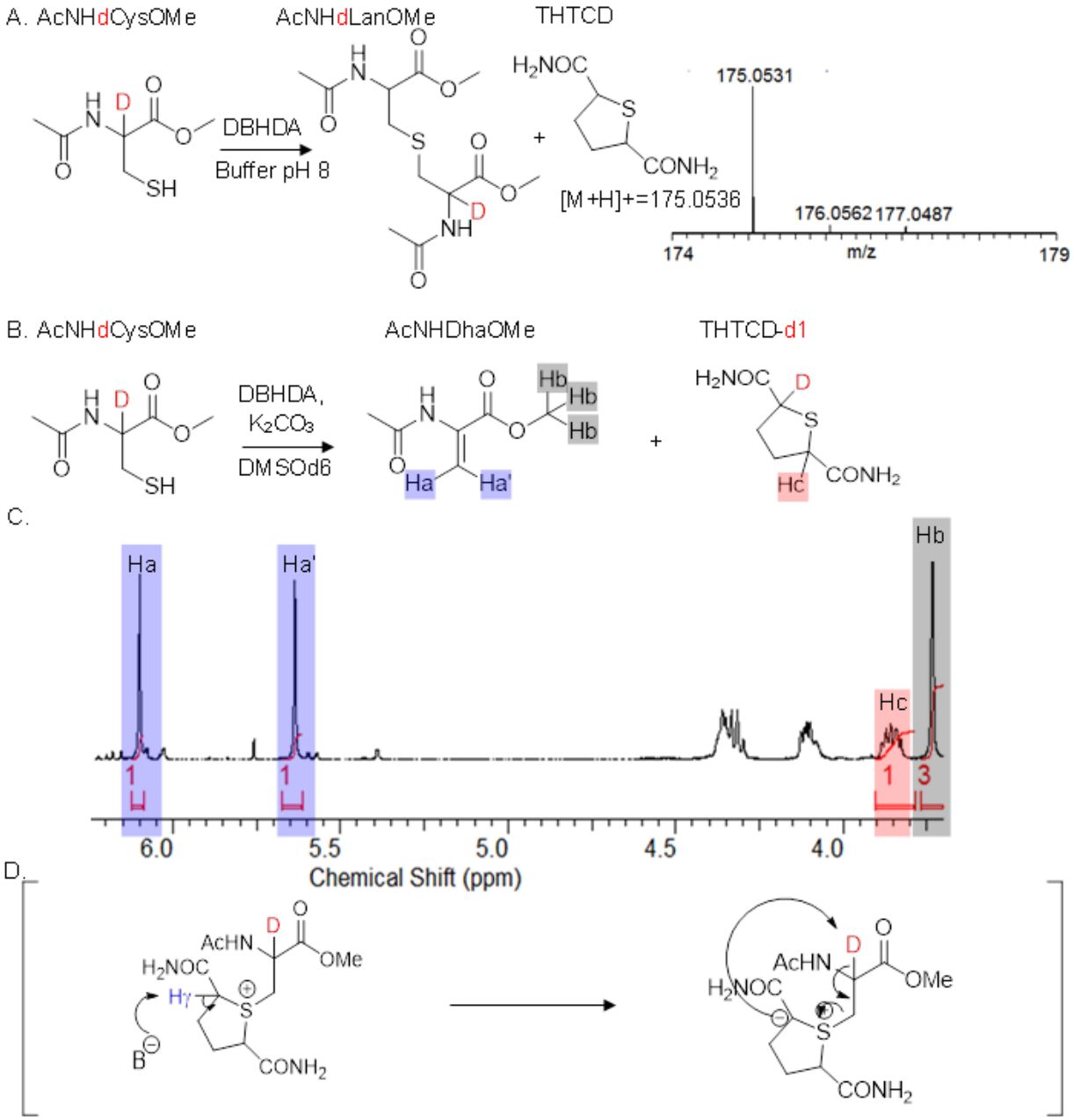


Figure 3.11. A. The reaction of AcNHdCysOMe with DBHDA in deuterated buffer afforded AcNHdLanOMe and THTCD, [M+H]+obs.=175.0531, [M+H]+calc.=175.0536. B. The reaction of AcNHdCysOMe with DBHDA in DMSO-d6 afforded AcNHDhaOMe and THTCD-d1. C. The reaction was analysed by ^{1}H NMR. integrated intensity of THTCD Hc was one, indication of the deuteration of THTCD. D. The proposed mechanism of deuteration of THTCD involved a sulfonium ylid, which acted as an intramolecular base.

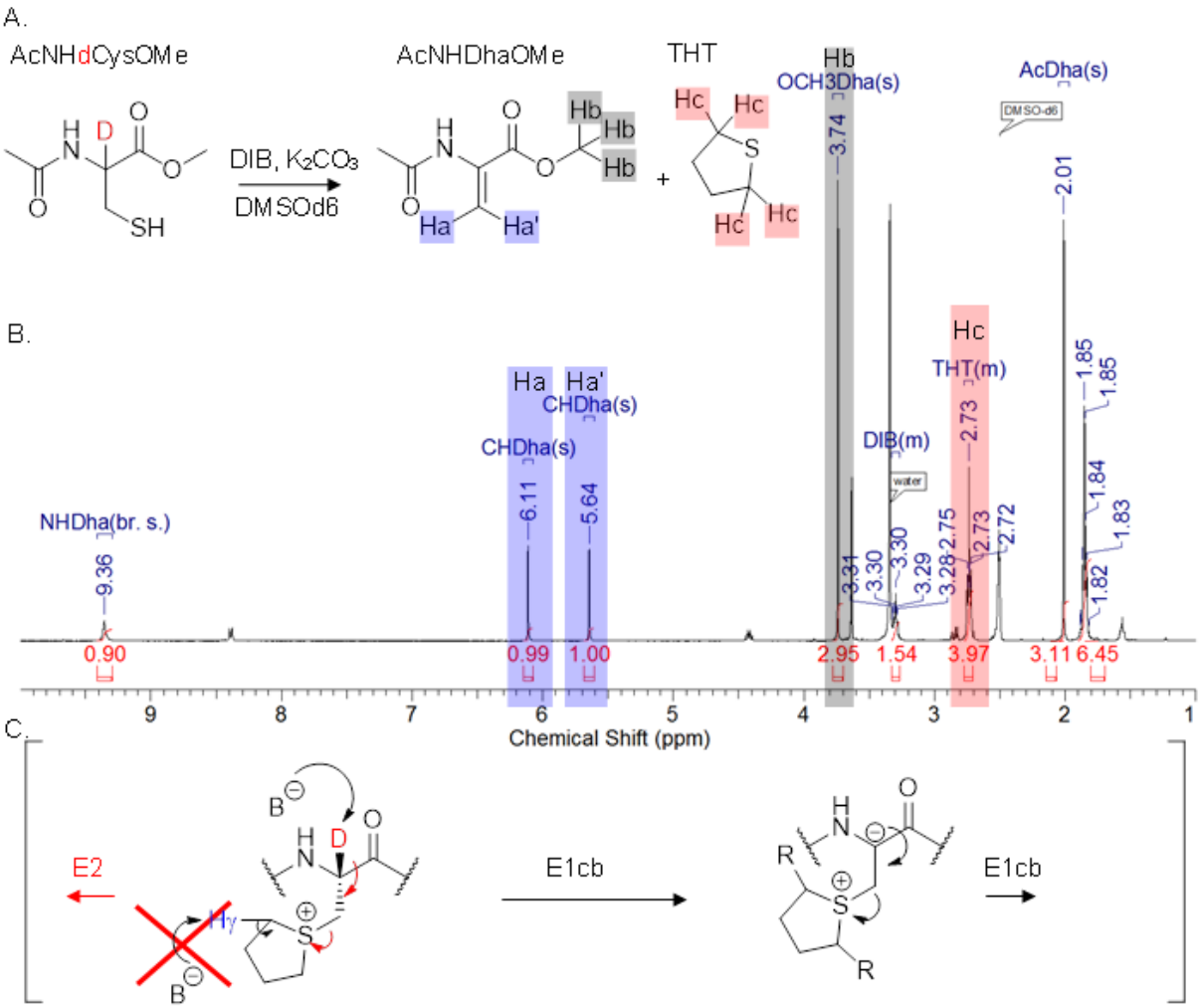


Figure 3.12 A. The reaction of AcNHdCysOMe with DIB in DMSO-d6 afforded AcNHDhaOMe and non-deuterated THT. B. The reaction was analysed by ^{1}H NMR. integrated intensity of THT Hc was four, indication that THT was not deuterated. C. The proposed mechanism involved the deprotonation of α-C of the amide, which underwent a concerted E2 or not E1cb elimination.

3.3.4. DFT calculation and proposed E1cb mechanism

Density functional theory (DFT) calculations were performed for the elimination reaction of AcNHCysNH$_2$ treated with DBHDA in water and hydroxide as a base (Figure 3.13.A).[§] Sulfonium ylid formation appeared relatively facile and marginally endergonic (i.e. uphill), making it reversible. The subsequent step, intramolecular deprotonation (formally a retro-ene reaction) is thermally feasible and was the overall rate-determining step. E1cb-like kinetics were expected for this process. In contrast, the E2-elimination barrier was

[§]Computational studies were performed in collaboration with Dr. Robert Simion and Prof. Rob Paton.

*ca.*10 kcal/mol less favorable than that computed for the intramolecular process, making this pathway substantially kinetically disfavoured (Figure 3.13).

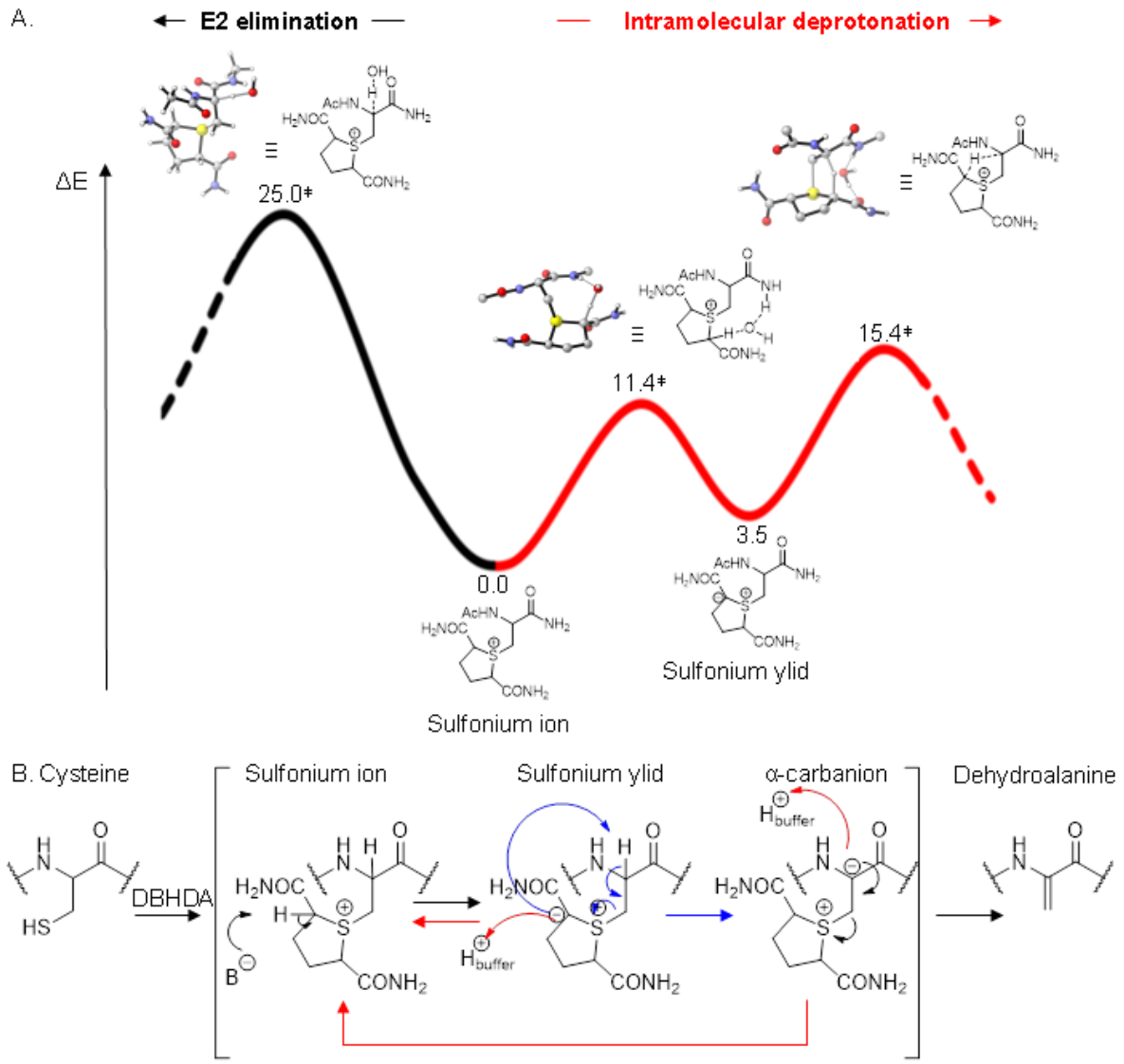


Figure 3.13. A. Computational calculations showing G_{rel} (kcal/mol) energies of transition states and intermediates of an amino acid model, $AcNHCysNH_2$ treated with DBHDA at pH 8. B. Proposed mechanism of Dha formation with DBHDA as alkylating/elimination agents.

Both steps of Dha formation *via* bis-alkylation/elimination have been mechanistically investigated in Chapters 2 and 3. The mechanism of elimination proposed in this chapter was also consistent with some observations made in previous chapter. Indeed, the stabilisation of both sulfonium ions bearing a carboxylic acid observed in the reactions of MEK1C222 with DBHDAc or MDBP could now be explained (Figure 3.14), not only did the

carboxylate anion stabilise the sulfonium cation, but the sulfonium ylid formation, which was essential for elimination to occur, was also disrupted by the close proximity of the carboxylate negative charge.

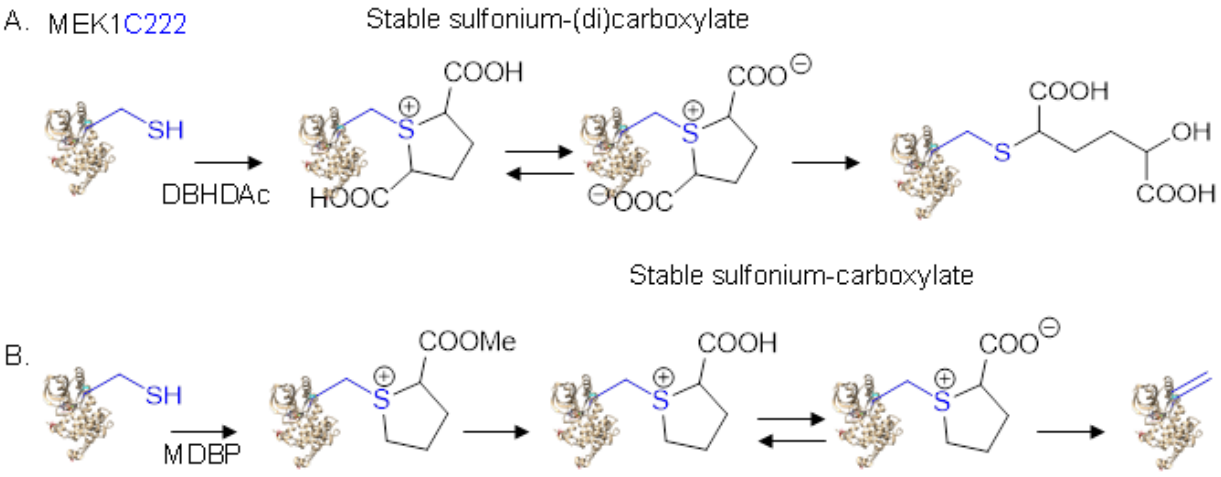


Figure 3.14. A. Treatment of MEK1C222 with DBHDAc afforded a stable sulfonium carboxylate intermediate. B. Treatment of MEK1C222 with MDBP afforded a stable sulfonium intermediate.

Chapter 4: Multiple Dha Formation on Protein

4.1. Crosslinking residues 218 and 222 in MEK1

4.1.1. DBHDA analogues crosslinking between cysteine residues 218 and 222

Following the successful site-selective formation of MEK1Dha218 and MEK1Dha222, the introduction of multiple Dha on MEK1C218C222 to afford MEK1Dha218Dha222 presented a different challenge. The use of DBHDA without any nucleotide-magnesium additive produced a unique compound with $M_{obs.}$=46057. This mass matched the mass of an over-alkylated MEK1 with C218 and C222 stapled by the alkylating agent, $M_{calc.}$=46052 (Figure 4.1.A). The over-alkylation corresponding to the alkylation at cysteine 207 was described in chapter 2. The relatively short distance between C218 and C222, around 7 Å, favoured the stapling over the sulfonium formation followed by elimination to form Dha (Figure 4.1.B and 4.2.B).

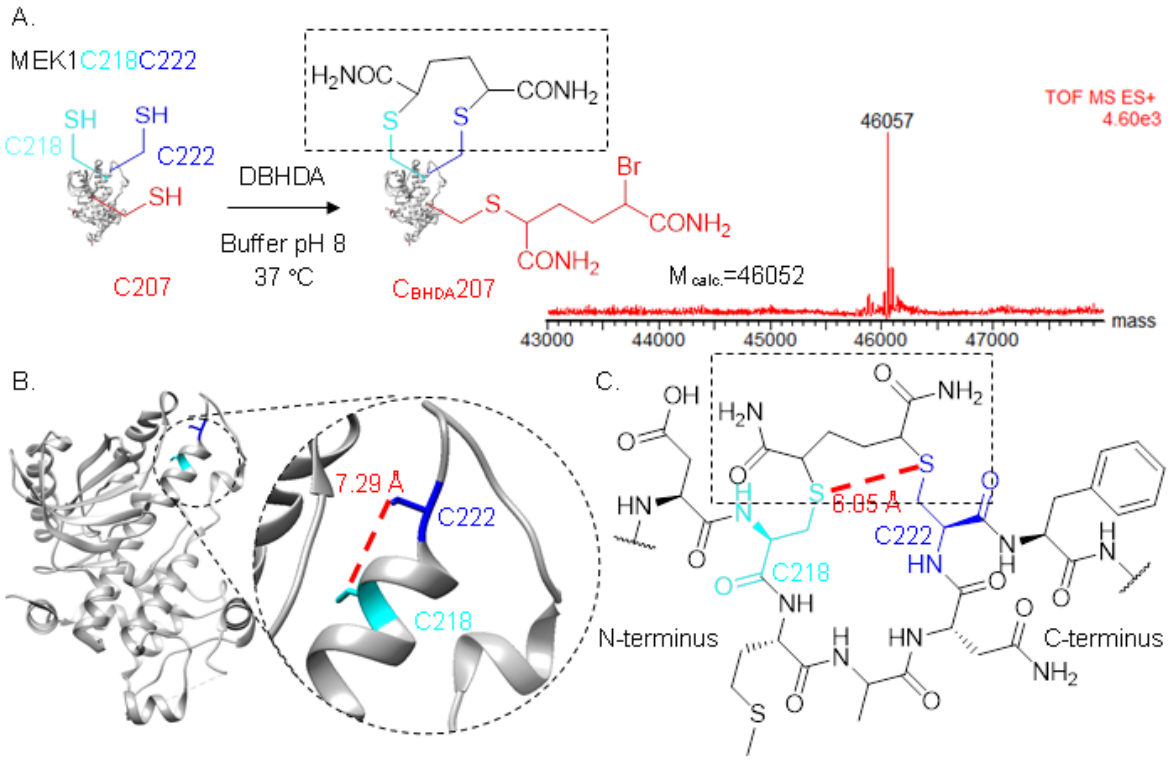


Figure 4.1. A. The treatment of MEK1C218C222 with DBHDA at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=46057 was consistent with alkylation of C207 and alkylative stapling of C218 and C222, $M_{calc.}$=46052. B. Distance between position 218 and 222 (PDB 3W8Q). C. Distance between stapled C218 and C222 in a model peptide representing residues 217-223.

Crystal structures have previously shown that the addition of nucleotide-magnesium to MEK1 induced conformational changes. These not only varied the side-chain accessibility of residue and the overall stability of MEK1 (Chapter 2), but also increased the distance between the residues 218 and 222 from 7 Å to 13 Å with ATP-Mg, and 11 Å with ADP-Mg (Figure 4.2.A). The increase in distance between the two residues favoured the formation of the sulfonium ion and therefore the formation of Dha over stapling (Figure 4.2.B). However, the addition of nucleotide-magnesium complex was insufficient to produce exclusively MEK1Dha218Dha222. Indeed, although MEK1C218C222 was converted to 80% of MEK1Dha with 20% of stapled C218 and C222 when ATP-Mg was used, these species were mostly phosphorylated (Figure 4.2.C). Alternative reaction conditions were explored. DBDGla was also used, having the added benefit of avoiding the alkylation of lysine residues as observed with MEKC218 construct after 14 hours reaction with DBHDA (Chapter 2). Nevertheless, ADP-Mg complex could not prevent the stapling from occurring, and only a

maximum of 23% conversion to MEK1Dha218Dha222 was obtained (Figure 4.2.D). Other bis-alkylating/elimination agents were subsequently evaluated.

A.

Ligand	Distance between C218 and C222 residues (Å)
None	7
ATP-Mg	13
ADP-Mg	11

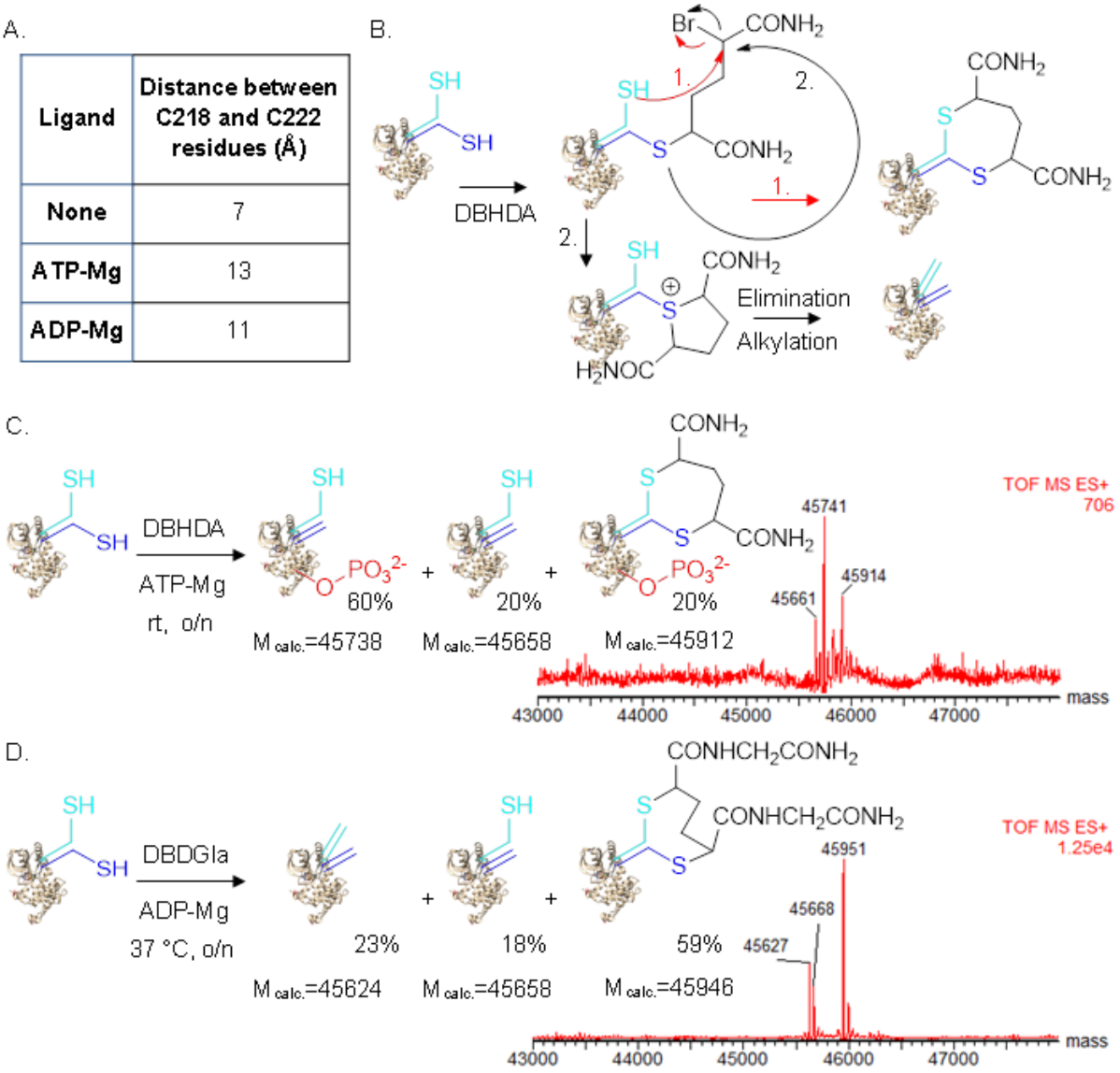


Figure 4.2.A. Distance between C218 and C222 in the presence or not of nucleotide-Mg complex. B. Two mechanisms described the formation of crosslinked residue 1 (red) and sulfonium ion 2 (black) leading to Dha formation. C. LCMS analysis of the reaction of MEK1C218C222 with DBHDA in the presence of ATP-Mg showed the formation of a mixture of 60% $M_{obs.}$=45741(phosphorylated MEK1Dha222), 20% $M_{obs.}$=45661 (MEK1Dha222) and 20% $M_{obs.}$=45914 (phosphorylated MEK1 with stapled C218 and C222). D. LCMS analysis of the reaction of MEK1C218C222 with DBDGla in the presence of ADP-Mg showed the formation of a mixture of 23% $M_{obs.}$=45627 (MEK1Dha218Dha222), 18% $M_{obs.}$=45668 (MEK1Dha222) and 60% $M_{obs.}$=45951 (MEK1 with stapled C218 and C222).

4.1.2. Lanthionine formation between position 218 and 222

When used with MEK1C218 and C222, DIB presented a different alkylating profile compared to DBHDA and its amide analogues (Chapter 2 and 3). Indeed, the first alkylation with DIB was very slow but selective towards C222. However, the second alkylation was extremely fast as the sulfonium ion was formed readily. Therefore, the treatment of MEK1C218C222 was postulated to provide MEK1C218Dha222 (Figure 4.3). The addition of second alkylating agents like DBDGla, which was used successfully towards the formation of MEK1Dha218, would then afford MEK1Dha218Dha222 (Figure 4.3).

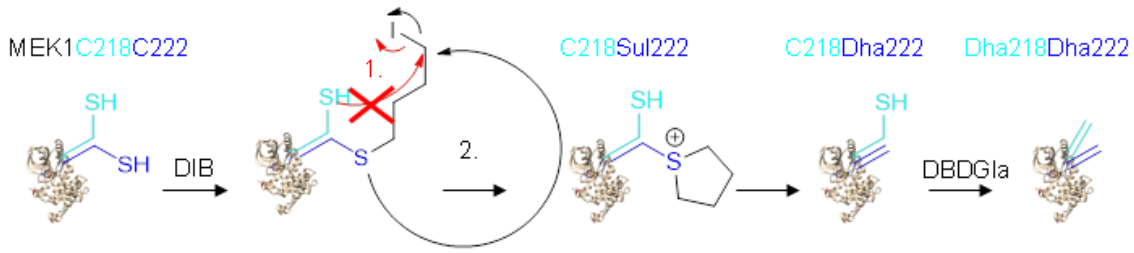


Figure 4.3. Proposed strategy and mechanism of formation of MEK1Dha218Dha222 with two alkylating agents in sequence. The poor reactivity of DIB to C218 does not allow the stapling of C218 and C222 with DIB (1.), and favours the formation of the sulfonium ion on C222, MEK1C218Sul222 (2.).

LCMS analysis of the reaction of MEK1C218C222 with DIB led to the formation of compound which mass corresponded to MEK1C218Dha222 after 15 hours at 37 °C (Figure 4.4.A). Unexpectedly however, further reaction with DBDGla did not afford MEK1Dha218Dha222 (Figure 4.4.B). Furthermore, the thiamichael addition of β-mercaptoethanol (BME) to Dha222 did not occur (Figure 4.4.B). It was therefore hypothesised that a nucleophilic attack of C218 towards Dha222 led to the formation of a thioether bridge called lanthionine (Figure 4.4 A), which would block C218 towards alkylation by DBDGla and the control reaction by the thiamichael addition of BME with Dha222. The lanthionine formation was confirmed by LCMS-MS analysis of the tryptic peptides (Figure 4.4.C). Further analysis of the LCMS of the tryptic peptide containing the lanthionine revealed an epimeric mixture (Figure 4.4.D), likely LL and LD. This result was consistent with a cyclisation carried out from the N- to the C-terminus,[207, 208] and confirmed the Dha

formation at position 222. If Dha218 had formed, the cyclisation to lanthionine would have given only one stereoisomer.[127, 209] Indeed, the cyclisation from C- to N-terminus produces an endo-enolate, which protonation leads to only one stereoisomer.[208]

As the reaction did not afford effectively the desired MEK1Dha218Dha222 or its intermediate MEK1C218Dha222, it was abandoned. However, the selectivity observed for C222 was encouraging.

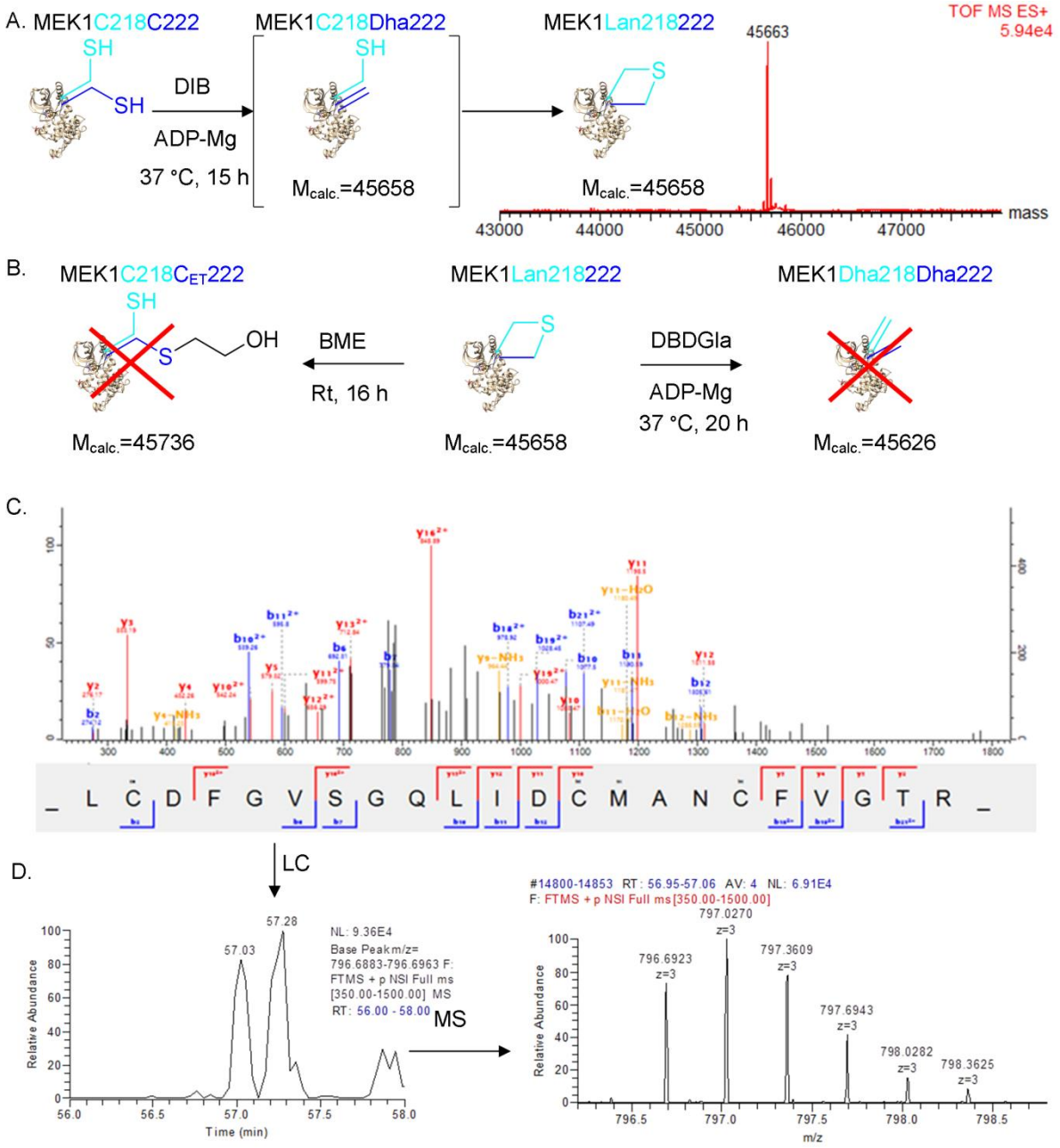


Figure 4.4.A. Formation of lanthionine MEK1Lan218222 from MEK1C218C222 with DIB $M_{calc.}$=45658, $M_{obs.}$=45663. B. The control experiment by treatment of the product with β-mercaptoethanol (BME) and the addition of DBDGla showed that neither Dha222 nor C218 were reactive. C. MSMS of tryptic

digest peptides confirmed the formation of lanthionine. D. MS1 chromatogram highlighted the epimeric mixture resulting from the intramolecular thiamichael addition.

4.2. Formation of double Dha using MDBP

4.2.1.Site-selective formation of MEK1Dha218Dha222

As described in Chapter 2, MDBP combined the relatively fast first alkylation of DBHDA on both C218 and C222 residues and the fast sulfonium ion formation of DIB. These properties would prevent the stapling of C218 and C222 and allow the formation of MEK1Dha218Dha222. Concurrently, Morrison et al. also observed that DBHDA was stapling cysteine residues separated by six amino acids on peptides, and used MDBP in DMF to yield multiple Dha.[146] The transfer of this chemistry to protein in aqueous buffer was successful. Treatment of MEK1C218C222 with MDBP in aqueous buffer led to the full and selective conversion to MEK1Dha218Dha222 (Figure 4.4 A and 4.4.B). The reaction did not disrupt the secondary structure of the protein as confirmed by CD spectroscopy (Figure 4.4.C).

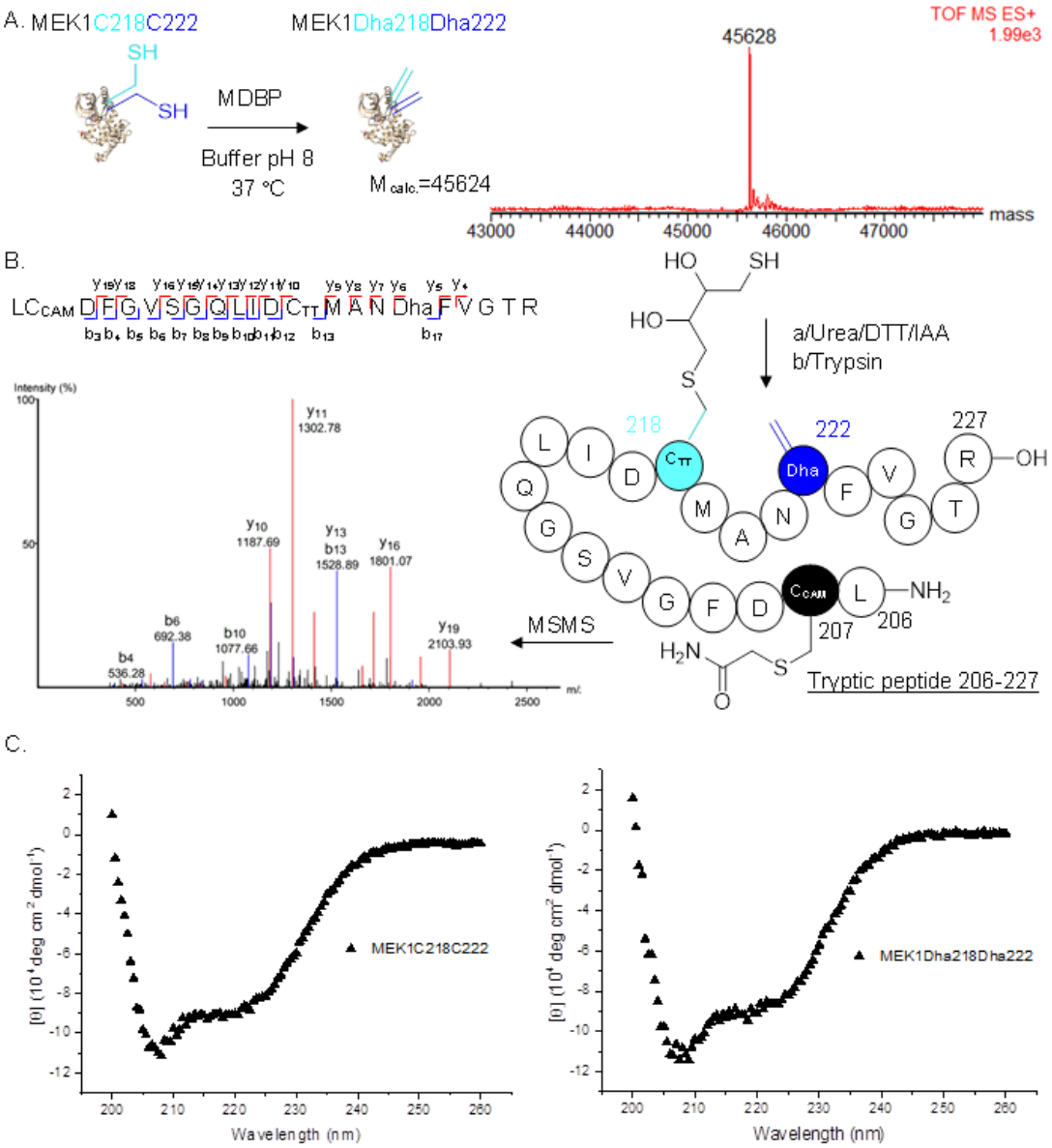


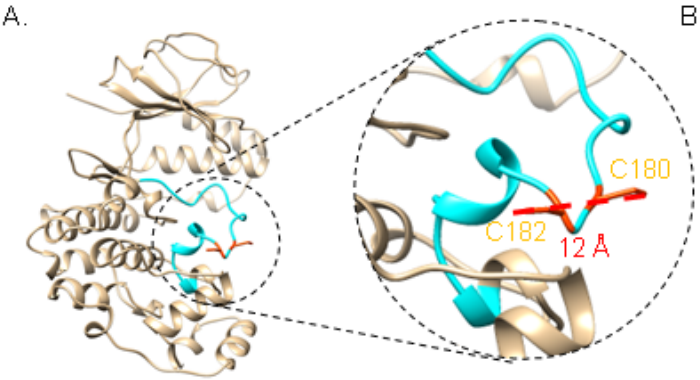
Figure 4.4. A. The treatment of MEK1C218C222 with MDBP in the presence of ADP-Mg at 37 °C and pH 8 was monitored by LCMS. $M_{obs.}$=45628 was consistent with MEK1Dha218Dha222, $M_{calc.}$=45624. B. Tryptic peptides were prepared from MEK1Dha218Dha222 denatured with urea, treated with DTT and with the free unreacted cysteine residues capped with iodoacetamide (IAA), diluted and treated with trypsin at 37 °C. LCMS-MS of tryptic peptide 206-227 confirmed that position 207 was a capped unreacted cysteine, position 218 was a dehydroalanine, which reacted with DTT, and position 222 was a dehydroalanine. C. Comparative circular dichroism of MEK1C218C222 and MEK1Dha218Dha222 indicated a similar profile.

4.2.2. Versatility of MDBP: formation of p38αDha180Dha182

During the study of the chemical phosphorylation of p38α, the formation of p38αpC180pC182 was envisaged. Indeed, p38αpC180pC182 was thought to be a relevant

model to investigate the effect of inhibitors in comparison with p38αpT180pY182. Furthermore, it would also allow the investigation of pC as bioisostere of pY. The formation of p38αDha180Dha182 was therefore necessary. Contrary to MEK1C218C222, the cysteine residues were only separated by one glycine residue in p38αC180C182. In a p38αC180C182 homology model structure, the cysteine residues pointed in opposite directions, and the distance between them was calculated as 12 Å (Figure 4.5.A). According to the relative side chain accessibility, the position 182 was calculated to be as accessible as position 180 (Figure 4.5.B).

A.



B.

Residues	Calculated relative side-chain accessibility
	p38αC180C182
180	71.8
182	63.7

Figure 4.5. A. The distance calculated between the two sulfur atoms of C180 and C182 in a modelled structure of p38αC180C182 was 12 Å. B. The calculated relative side chain accessibility of both positions was very similar and estimated at 71.8% and 63.7% respectively.

Similarly to the reaction with MEK1C218C222, DBHDA was not reactive enough to avoid the stapling of both cysteine residues, C180 and C182 (Figure 4.6.A and 4.6.B). After 24 hours at 37 °C, DIB did not provide p38αDha180Dha182. Instead, the reaction led to the formation of a mixture of p38αDha and monoalkylated $p38\alpha C_{IB}$ (Figure 4.6.C). After 31 hours reaction, no material was detected by MS. Only the use of MDBP afforded the desired p38αDha180Dha182 (Figure 4.6.D). However, more optimisation was necessary as the cysteine residues 180 and 182 oxidised over time to form an intramolecular disulphide bond, which prevented the alkylation to occur.

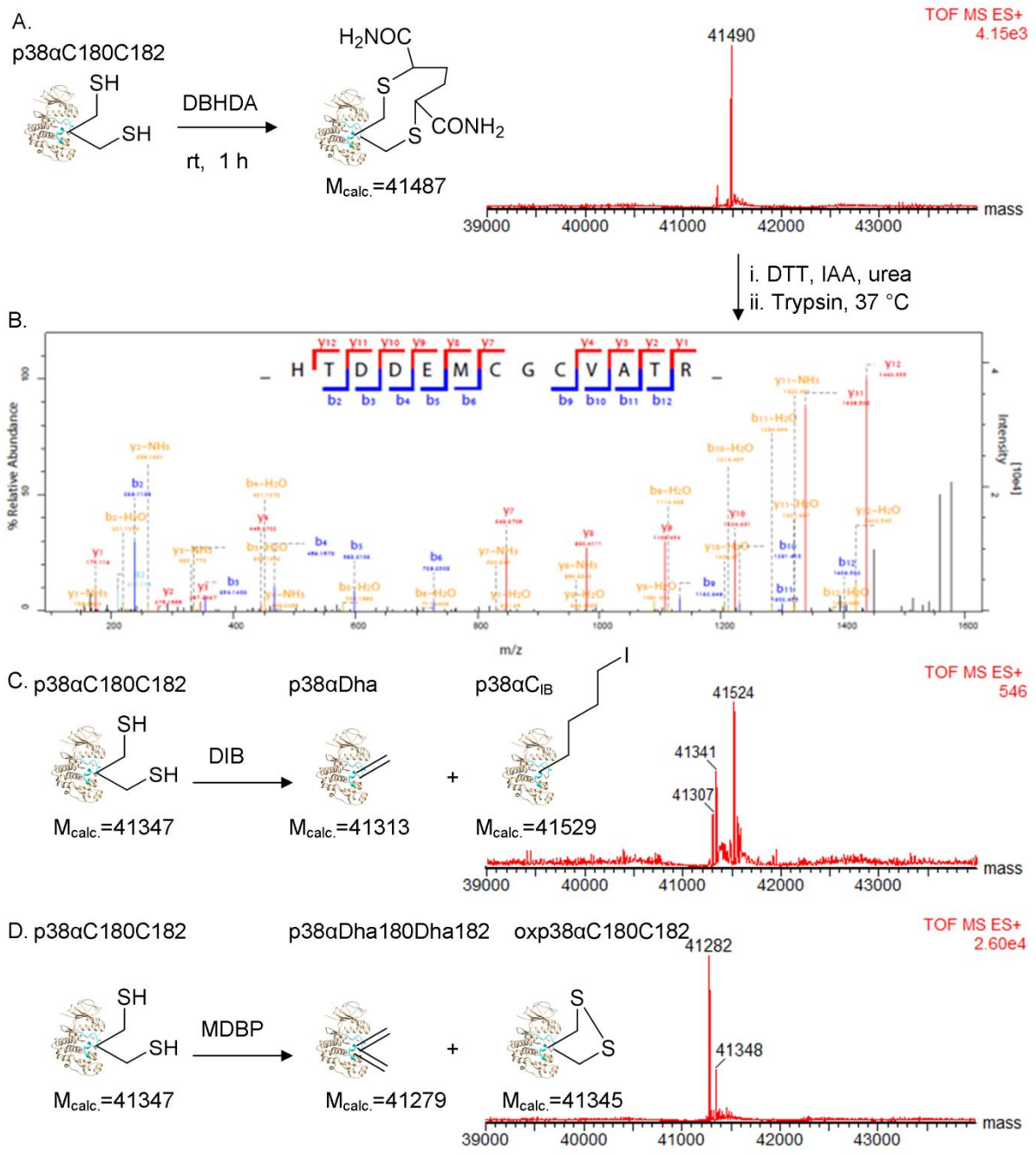


Figure 4.6. A. The reaction of p38αC180C182 with DBHDA afforded stapled p38αC180C182, $M_{calc.}$=41487, $M_{obs.}$=41490. B. LCMS-MS of tryptic peptide 174-186 confirmed the stapling of C180 with C182. C. The reaction of p38αC180C182 with DIB was not complete and afforded a mixture of 30% p38αC180C182, $M_{calc.}$=41347, $M_{obs.}$=41341, 15% p38αDha, $M_{calc.}$=41313, $M_{obs.}$=41307 and 55% monoalkylated $p38αC_{IB}$, $M_{calc.}$=41529, $M_{obs.}$=41524. C. The reaction of p38αC180C182 with MDBP afforded a mixture of 30% oxidised oxp38αC180C182, $M_{calc.}$=41345, $M_{obs.}$=41351, and 70% p38αDha180Dha182, $M_{calc.}$=41279, $M_{obs.}$=41283.

Several non-alkylating strategies based on work developed in our group on SBL model protein were also tested towards the formation of p38αDha180Dha182.

First, the reaction of SBLC156 with Ellman's reagent led to the formation of a mixed disulphide, which when treated with the electron rich phosphine, hexamethylphosphorus triamide (HMPT) followed by base, or directly by base, afforded the desired SBLDha156.[117] The treatment of p38αC172 with Ellman's reagent yielded a mixture p38αC_{NTB} and p38αC_{NTB}x2, which indicated that Ellman's reagent was not selective. However, the addition of 100 equivalents to p38αC180C182 gave mainly p38αC_{NTB}180C_{NTB}182. Nevertheless, the addition of HMPT or base did not provide the desired p38αDha180Dha182, only noise was detected by LCMS (Figure 4.7.A).

Similarly to SBLC156,[117] the Mukaiyama reagent was selectively installed on p38α (Figure 4.4.B). However, the elimination did not occur with addition of base. As described with Ellman's adduct, the protein was no longer observed by LCMS.

Finally, o-mesitylenesulfonylhydroxylamine (MSH), which was described in the synthesis of SBL156Dha,[118] was also envisaged (Figure 4.4.C). However, in our case, p38αDha180Dha182 was not observed by LCMS.

These conditions were harsher than the mild bis-alkylating/elimination methodology as they mostly resulted in the degradation of the protein. Moreover, site-selectivity of Ellman's and Mukaiyama reagent proved to be a challenge even on p38α, which contains buried and poorly accessible native cysteine residues.

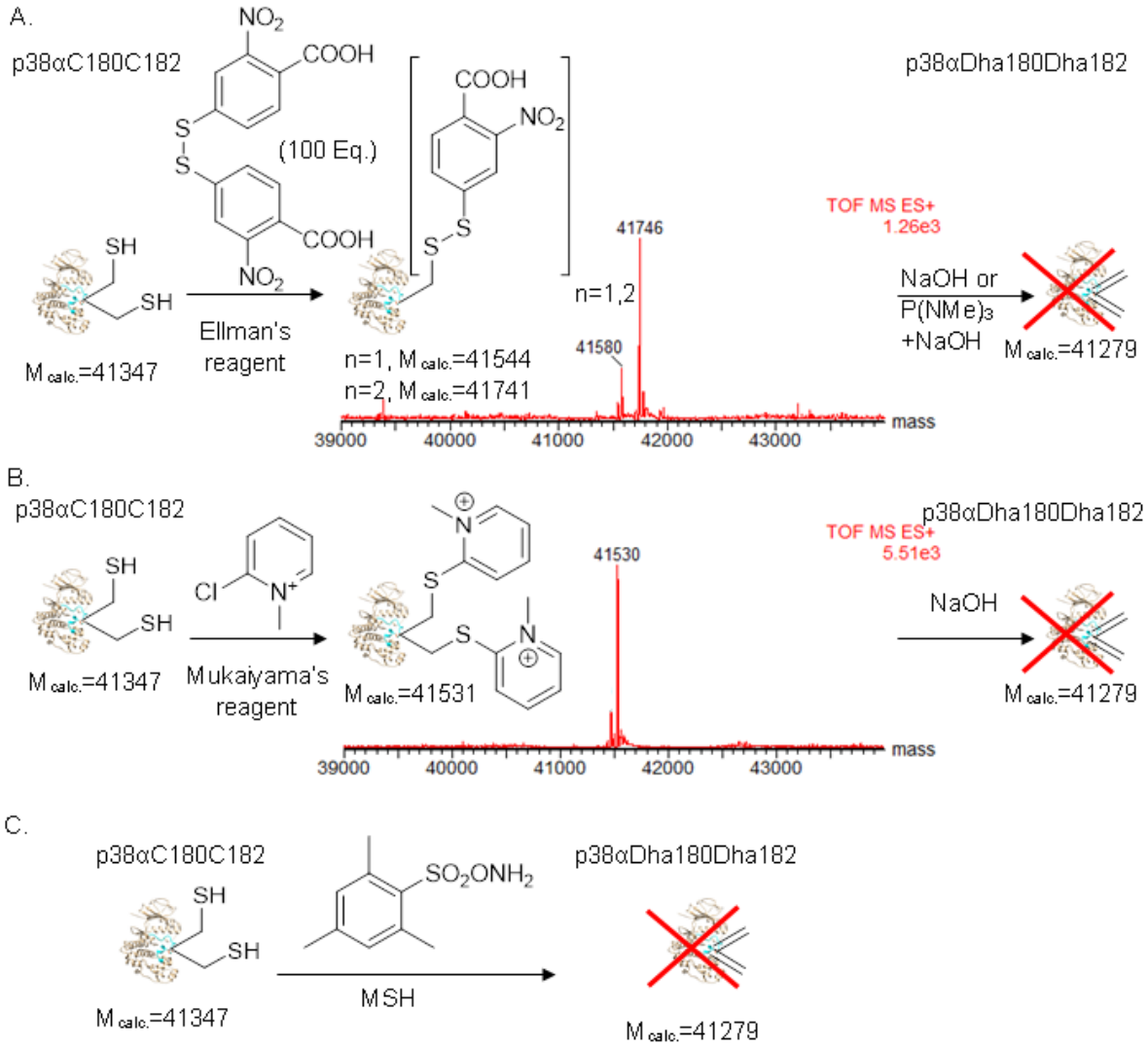


Figure 4.7. A. The reaction of p38αC180C182 with Ellman's reagent afforded a mixture of 20% monosubstituted p38αC$_{NTB}$, $M_{calc.}$=41544, $M_{obs.}$=41580, and 80% disubstituted p38αC$_{NTB}$x2, $M_{calc.}$=41741, $M_{obs.}$=41746. The treatment of the mixture with either NaOH or HMPT+NaOH did not lead to the formation of p38αDha180Dha182. B. The reaction of p38αC180C182 with Mukaiyama reagent afforded a 85% p38αC$_{MP}$x2, $M_{calc.}$=41531, $M_{obs.}$=41530. C. The reaction of p38αC180C182 with MSH was not detectable by LCMS.

Finally, the use of non-alkylating strategies did not improve on the 70% conversion to p38αDha180Dha182 produced using MDBP. The optimisation of p38αDha180Dha182 formation was however not further pursued. The work in progress was left in favour of the work on MEK1Dhas and their subsequent chemical phosphorylation, as demonstrated in next chapter.

In summary, thanks to the careful analysis of the mechanism of action of the different bis-alkylating agents (Chapter 2 and 3), the discovery of the asymmetric bis-alkylating agent

MDBP enable the introduction of multiple Dhas to various proteins in aqueous buffer. It was therefore showed that bis-alkylation/elimination was a mild and efficient method towards the selective introduction of single and multiple Dhas in complex protein containing several reactive cysteine residues.

The formation of MEK1Dha218, MEK1Dha222 and MEK1Dha218Dha222 was taken further to investigate the formation of MEK1pC218, MEK1pC222 and MEK1pC218pC222 mimics of the active phosphorylated MEK1pS218, MEK1pS222 and MEK1pS218pS222.

Chapter 5: Chemical phosphorylation and chemical activation of MEK1

5.1. MEK1pC222, mimics of MEK1pS222

5.1.1. Formation of MEK1pC222

In principle, the electrophilic handle installed on MEK1 phosphorylation and activation sites, 218 and 222, allowed the formation of MEK1pC218, MEK1pC222 and MEK1pC218pC222 by thiamichael addition of sodium thiophosphate. This chapter describes the formation of these MEK1pCs as mimics of the natural MEK1pSs.

The synthesis of MEK1pC222 was first undertaken as the position 222 was the most accessible. In an effort to create a straightforward chemical phosphorylation method, MEK1Dha222 was first made in a buffer containing ADP-Mg, and treated without isolation with a large excess of sodium thiophosphate in a “one-pot” fashion. The presence of ADP-Mg in the thiamichael additions not only stabilised the protein at 37 °C but also protected C207 from potential oxidative addition of sodium thiophosphate leading to the formation of S-SPO_3^{2-} bond.[210] Although the full conversion to MEK1pC222 was observed, the construct could not be isolated in an apo form necessary for assessing its activity (Figure 5.1).

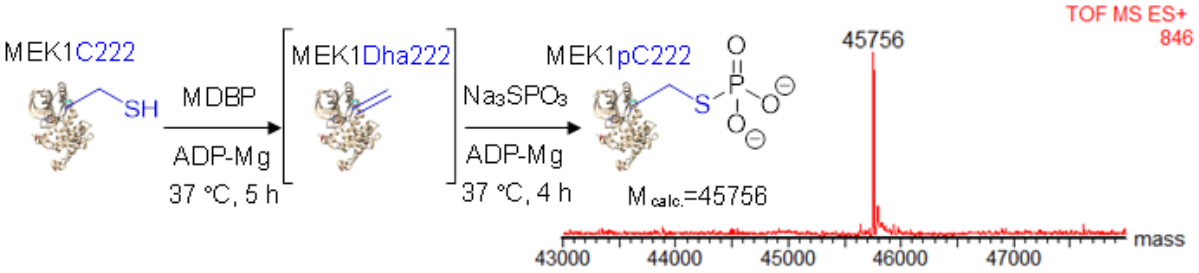


Figure 5.1. Treatment of MEK1C222 in buffer containing ADP-Mg with MDBP followed by sodium thiophosphate afforded MEK1pC222, $M_{calc.}$=45756, $M_{obs.}$=45756.

MEK1Dha222 was, therefore, first isolated into a buffer free of ADP-Mg, and then sodium thiophosphate was added. The addition of ATP into the buffer improved the profile of the reaction. The thiamichael addition afforded the desired MEK1pC222 after 20 hours at room temperature (Figure 5.2.A and 5.2.B). CD confirmed that the thiamichael reaction was mild, and did not affect the overall folding of the protein (Figure 5.2.D). The confirmation of C222 phosphorylation by LCMS-MS was more challenging. The desired tryptic peptide was not detected, even by LCMS. Precautions in the sample preparation were taken as described in the introduction (paragraph 1.4.1), and stability tests over acid and reducing agents are described in the next paragraph 5.2.1. The tryptic peptide is typically 22 amino acids long, with one arginine at the C-terminus and 2 aspartic acid residues. Considering the potential addition of the negative charges of a phosphate moiety, this peptide does not appear as an ideal candidate for LCMS-MS. However, the presence of 2 aspartic acid residues was used to our advantage. A combined tryptic/AspN digest was performed. The shorter peptide was not detected by LCMS-MS using Electron Transfer Dissociation (ETD), likely due to the lack of positive charged residues. However, the desired peptide containing the phosphocysteine residue was observed and characterised using Collision Induced Dissociation (CID) with a neutral loss. LCMS analysis of this tryptic/AspN peptide revealed two distinct peaks corresponding to the two epimers (Figure 5.2.C). The epimeric ratio was around 40:60, which was consistent with reported values.[97] As indicated in the introduction, this epimeric ratio can vary with the environment, and we should expect differences between the addition at 222 and 218.

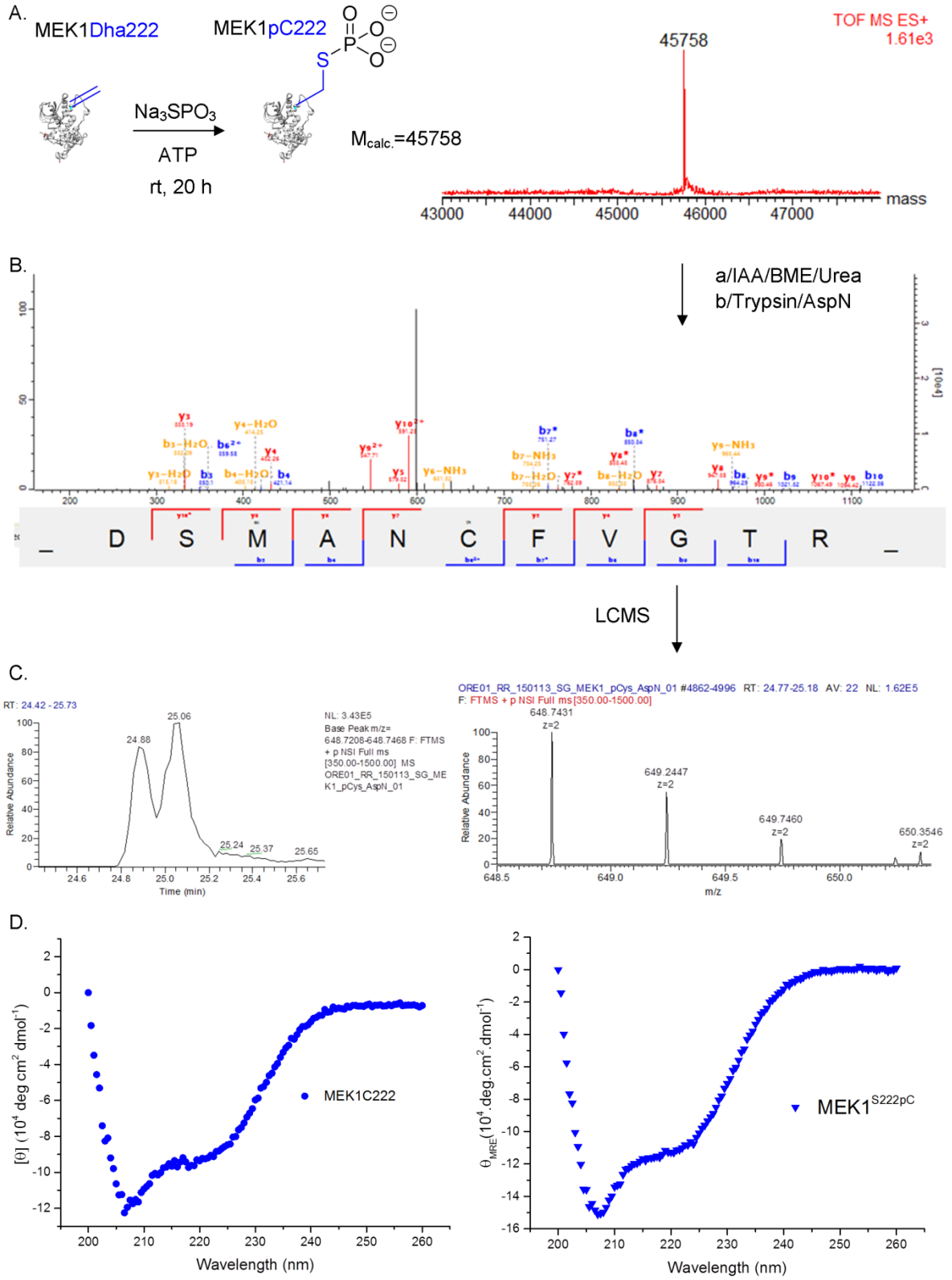


Figure 5.2. The treatment of MEK1Dha222 with sodium thiophosphate in the presence of ATP at room temperature was monitored by LCMS. $M_{obs.}$=45758 was consistent with MEK1pC222, $M_{calc.}$=45756. B. LCMS-MS of tryptic/AspN peptide 219-227 confirmed the selective phosphorylation site. C. LCMS of tryptic/AspN peptide showed a ~1:1 epimeric mixture. D. Comparative circular dichroism (CD) spectra of MEK1C222 and MEK1pC222.

5.1.2. Stability of the phosphocysteine residue in MEK1pC222

As p38αpC172 and p38αpC180,[96] the stability of the phosphocysteine residue was assessed. The phosphocysteine residue was stable in a buffer containing 10 mM reducing agents such as DTT, BME and glutathione (Figure 5.3.A). However, as described in Chapter 3 with the phosphorylated histone, H3pC10, MEK1pC222 was hydrolysed to MEK1C222 under acidic condition (Figure 5.3.B), and by the protein phosphatase involved in the diphosphorylation of MEK1pS218pS222 *in vivo*, PP2A[211-214] (Figure 5.3.C).

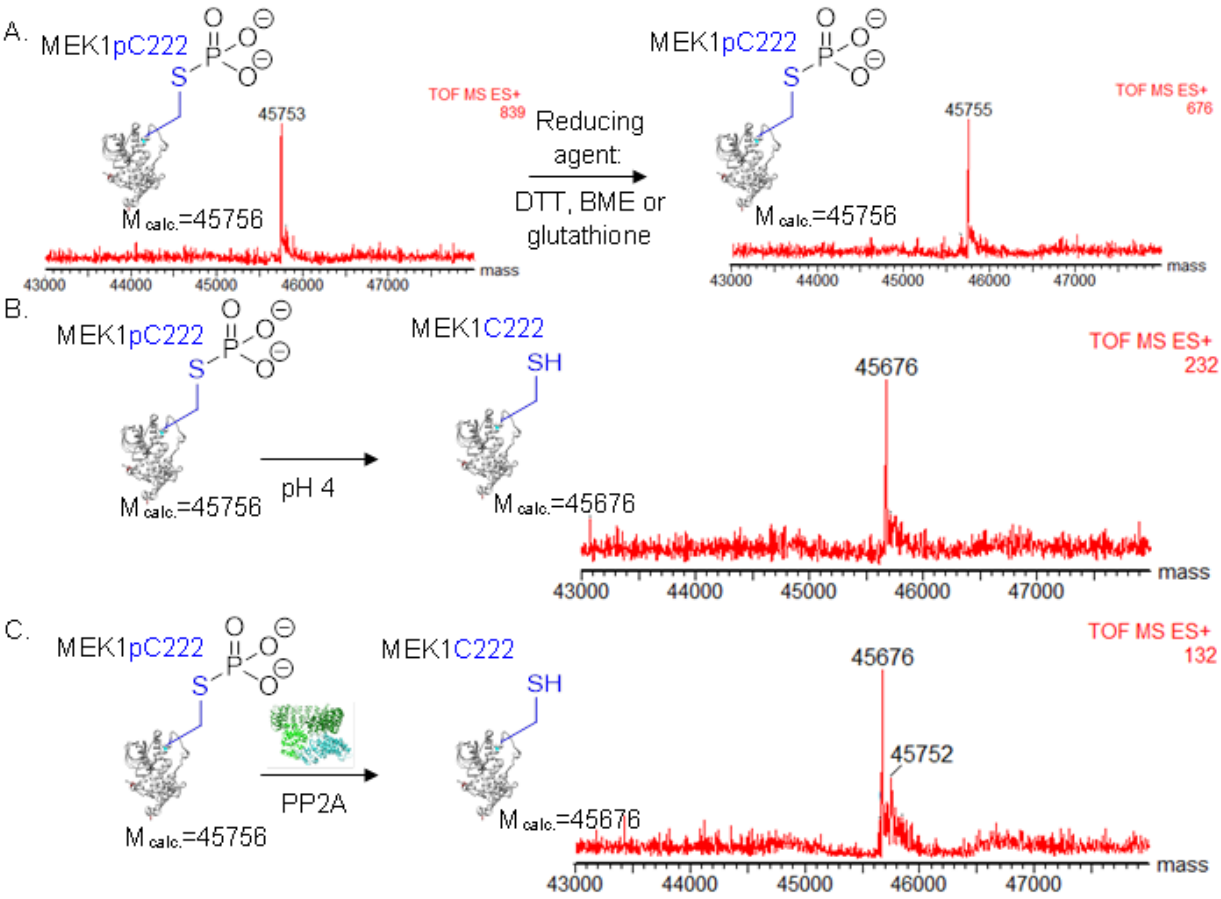
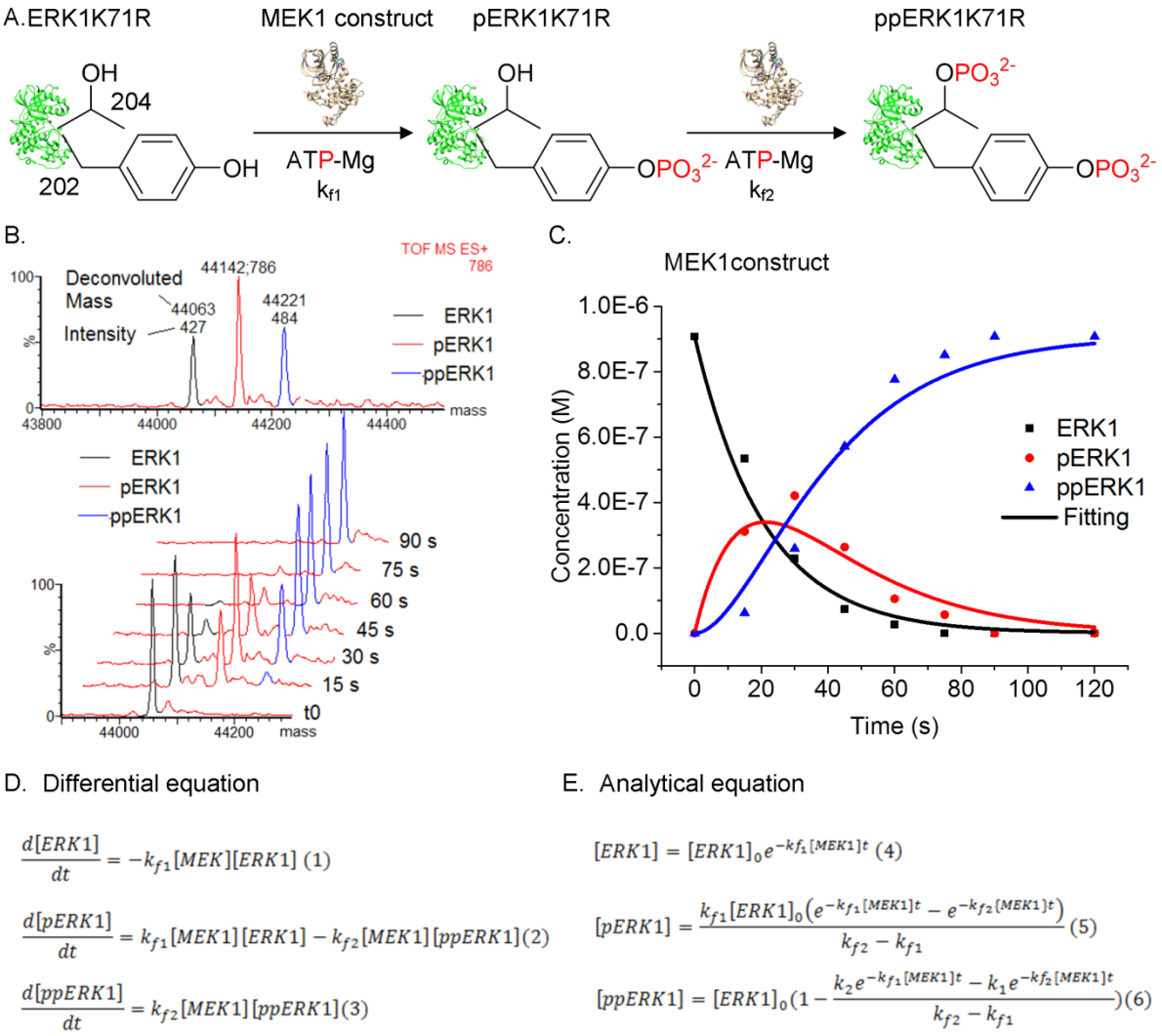


Figure 5.3. A. Reducing agents (DTT, BME or glutathione) did not hydrolysed MEK1pC222 to MEK1C222. B. MEK1pC222 was hydrolysed under acidic conditions. C MEK1pC222 was hydrolysed by the protein phosphatase PP2A.

5.1.3. Assessment of the kinase activity of MEK1X222 constructs

To study the influence of the site-selective chemical phosphorylation at position 222 on MEK1 activity, the progress of the phosphorylation of the MAP kinase, ERK1, the only known natural substrate of MEK1,[19, 215-217] was monitored by intact protein LCMS. ERK1K71R was chosen for our study as this construct did not autophosphorylate and display any kinase activity.[218] LCMS monitoring of the reaction of phosphorylation of ERK1K71R catalysed by MEK1 constructs allowed the visualisation of both mono and diphosphorylated ERK1 (Figure 6.4.A). The phosphorylation of ERK1 was reported to be sequential and distributive.[219] The phosphorylation of Y204 triggered the second phosphorylation of T202 (Figure 5.4.B). Therefore, the data was fitted to a sequential two-step first order reaction (Figure 5.4.C and D).[35, 36] The analytical equation provided both phosphorylation rate constants k_{f1} and k_{f2}, corresponding to the formation of the monophosphorylated ERK1 (pERK1) and diphosphorylated ERK1 (ppERK1), respectively (Figure 5.4.E). These constants were used to compare the activity of the different MEK1 constructs.



D. Differential equation

$$\frac{d[ERK1]}{dt} = -k_{f1}[MEK][ERK1] \quad (1)$$

$$\frac{d[pERK1]}{dt} = k_{f1}[MEK1][ERK1] - k_{f2}[MEK1][ppERK1] \quad (2)$$

$$\frac{d[ppERK1]}{dt} = k_{f2}[MEK1][ppERK1] \quad (3)$$

E. Analytical equation

$$[ERK1] = [ERK1]_0 e^{-kf_1[MEK1]t} \quad (4)$$

$$[pERK1] = \frac{k_{f1}[ERK1]_0\left(e^{-k_{f1}[MEK1]t} - e^{-k_{f2}[MEK1]t}\right)}{k_{f2} - k_{f1}} \quad (5)$$

$$[ppERK1] = [ERK1]_0\left(1 - \frac{k_2 e^{-k_{f1}[MEK1]t} - k_1 e^{-kf_2[MEK1]t}}{k_{f2} - k_{f1}}\right) \quad (6)$$

Figure 5.4. A. Sequential and distributive proposed mechanism of ERK1 phosphorylation by MEK1. B. The phosphorylation of ERK1K71R by ERK1 construct was monitored by LCMS. The intensities of the deconvoluted peaks corresponding to ERK1, $M_{calc.}$=44058, pERK1, $M_{calc.}$=44138, and ppERK1 , $M_{calc.}$=44218 were recorded over time. C. The conversions were calculated from these intensities and plotted. D. Differential equation defined the phosphorylation constants k_{f1} and k_{f2}. E. These constants were calculated by fitting the analytical equations.

As described in early work,[23] MEK1 wild-type, MEK1S218S222* (construct with S277 and S376 not mutated), displayed some basal activity (Table 5.1, entry 1). Significantly, the mutation at the C-terminal of C277S and C376S (MEK1S218S222), essential for the site-selective chemical phosphorylation of MEK1, did not affect the basal-activity (Table 5.1, entry 2). However, 2-fold to 3-fold increase of basal activity of MEK1 was observed with S222C and S222Dha mutations (Table 5.1 entry 3 and 4). Mutations of various residues in MEK1 have been shown to influence the level of its basal activity both *in vitro* and *in*

vivo.[23,220] In these reports, the increased basal activity was imputed to the potential autophosphorylation of MEK1 constructs.

Table 5.1. Activity of MEK1X222 constructs

Entry	MEK1 construct	$k_{f1}x10^3$ $M^{-1}.s^{-1}$	$k_{f2}x10^3$ $M^{-1}s^{-1}$
1	MEK1S218S222*	0.46±0.03	0.50±0.03
2	MEK1S218S222	0.41±0.03	0.47±0.03
3	MEK1C222	0.84±0.03	0.82±0.03
4	MEK1Dha222	1.37±0.03	1.41±0.03
5	MEK1pC222	10.90±0.05	9.20±0.05
6	MEK1E222	3.6±0.2	3.0±0.1

* MEK1wild-type with S277 and S376 not mutated.

As the autophosphorylation of MEK1C222 was observed during the optimisation of the site-selective formation of MEK1Dha222 (Chapter 2), we verified whether autophosphorylation was the cause of the basal activity of our MEK1 constructs. Indeed, the phosphorylation of both ERK1K71 and MEK1 could be monitored by intact protein LCMS (Figure 5.5.A). LCMS confirmed the multiple phosphorylation of MEK1X222 in the presence of ATP-Mg (Figure 5.5.B). However, the rate of autophosphorylation of MEK1X2 was significantly slower than the rate of phosphorylation of ERK1 (Figure 5.5.C and D). The autophosphorylation of MEK1X222 was detectable as soon as half of phosphorylatable ERK1 was consumed (Figure 5.5.C and D). Therefore, MEK1 autophosphorylation was unlikely the cause of the observed basal activity. Furthermore, ERK1 seemed a better substrate than MEK1 itself.

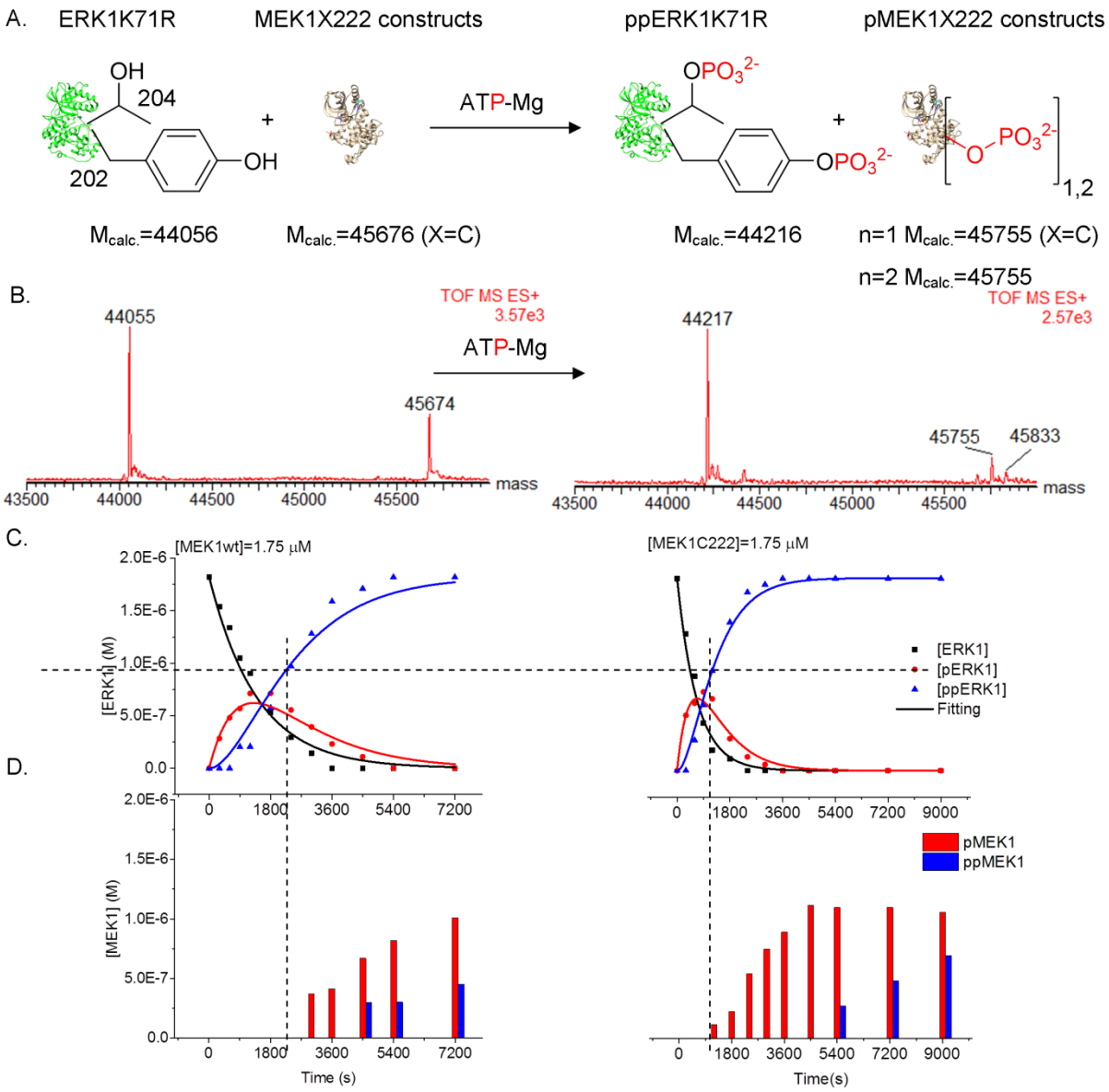


Figure 5.5. A. ERK1K71R and MEK1X222 constructs phosphorylated in the presence of ATP-Mg. B. LCMS monitoring of the reaction of ERK1K71 showed the double phosphorylation of ERK1, $M_{calc.}$=44218, $M_{obs.}$=44216 and the monophosphorylation of MEK1, $M_{calc.}$=45756, $M_{obs.}$=45755. C. Progress curve of the phosphorylation of ERK1 in function of time (s). D. Progress curve of the phosphorylation of MEK1 in function of time.

The chemical phosphorylation at position 222 with MEK1pC222 increased significantly MEK1 basal activity of 20-fold (Entry 5). While it has been reported that the phosphorylation of both S218 and S222 were necessary to activate MEK1,[23] the kinase activity of MEK1pC222 would infer otherwise. The recently reported site-selective formation of MEK1pS222 by amber-stop codon technology and its demonstrated kinase activity confirmed our result.[63] The phosphorylation of only one serine residue of MEK1 was enough to trigger an increase in MEK1 kinase activity.

Finally, MEK1pC222 was more active than the constitutively active MEK1E222. MEK1pC222 appeared to be a better mimic of MEK1pS222. The geometry of the phosphate moiety as well as the two negative charges is necessary to dramatically enhance the activity. This result validated our chemical phosphorylation methodology.

5.1.4. Trial crystal structure MEK1pC222-ERK1-nucleotide Mg complex

With an active phosphorylated MEK1 in hand, the crystallisation of the MEK-ERK-nucleotide complex was attempted (Figure 6.6.A). MEK1 complexes, with a protein partner, either with its kinase, BRAF[221] or the scaffold protein known as kinase suppressor of Raf (KSR) that facilitated MEK1 phosphorylation by Raf,[222] have been previously crystallised. However, MEK-ERK complex structure has not been solved. The reported K_D values of constitutively active MEK1E218D222-ERK2 and pMEK1-ERK2 complexes are respectively 58 nM *in vitro*,[223] and 100 nM *in vivo*.[34] These values are <1 µM, which indicates that the binding of the complex is not weak. Therefore, in collaboration with the Structural Genomics Consortium (SGC)[**], various conditions for the crystallisation of MEK1pC222-ERK2-AMP-PNP[††] complex were screened (Figure 5.6.A).

Firstly, the synthesis of MEK1pC222 from MEK1C222 was scalable. 7 mg of pure protein were obtained without further purification. Initial size exclusion chromatography coupled to multi-angle light scattering analysis (SEC-MALS) of a 1:1 mixture of MEK1pC222-ERK1 in a running buffer containing ADP-Mg supported the formation of a complex (Figure 5.6.B). The actual MEK1pC222-ERK2 complex was formed and isolated by size exclusion

[**] The crystallisation screen was designed and performed in collaboration with Dr. Apirat Chaikuad from the Structural Genomics Consortium, Oxford. The on-going crystallisation studies have been transferred to AstraZeneca.

[††] AMP-PNP is an analogue of ATP. It contains one nitrogen in place of one oxygen atom on the -phosphate, which makes it more stable to hydrolysis than ATP. It is considered structurally equivalent to ATP.

chromatography (Figure 5.6.B and D). The isoform ERK2 was chosen as the crystallisation of ERK2 have been extensively performed compared to ERK1. Indeed, the RCSB protein data bank (www.rcsb.org/) reported more than 100 crystal structures of ERK2 for a dozen ERK1 structures. Although various conditions were tested for the crystallisation of MEK1pC222-ERK2-AMP-PNP (Figure 5.6.C), no hits were observed.

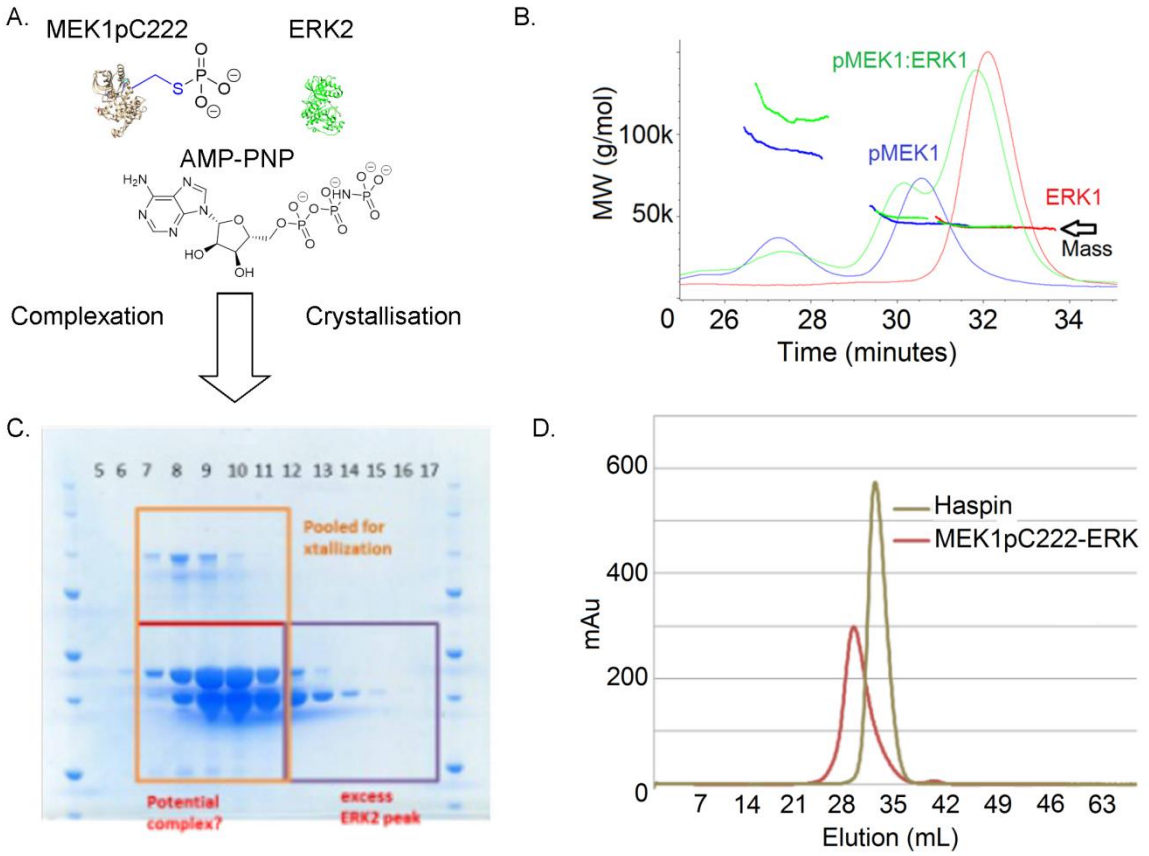


Figure 5.6.A. Crystallisation of MEK1pC222-ERK-AMPNP. B. SEC-MALS using ADP-Mg buffer of ERK1 (red) showed one peak at 32.5 minutes with a mass of 45kDa (Mcalc.=44068). SEC-MALS using ADP-Mg buffer of MEK1pC222 (blue) showed two peaks at 30.5 minutes with a mass of 45kDa corresponding to MEK1pC222 monomer (Mcalc.=45756) and at 27.25 minutes with a mass of 90 kDa corresponding to MEK1pC222 dimer. SEC-MALS using ADP-Mg buffer of a 1:1 MEK1pC222: ERK1 mixture (green) showed three peaks at 32 minutes corresponding to ERK1, at 30 minutes corresponding to MEK1pC222 and at 28 minutes with a mass of 110 kDa corresponding to an heteromeric mixture of dimer and tetramer. C. Electrophoretic gel of a first SEC to isolate the MEK1pC222-ERK2 complex. D. SEC of MEK1pC222-ERK2 complex compared to SEC of harpin, which mass was equivalent to ERK2.

Due to the failure of the crystallisation of full-length MEK1pC222 in complex with ERK2, the crystallisation of a truncated construct of MEK1 was envisaged. Indeed, the first MEK1

crystal structures described were from truncated MEK1.[224] Based on homology modelling studies predictions with active PKA (PDB 1CDK),[220] the first example of MEK1 crystal structure construct was lacking its first 60-*N*-terminal amino acids loop.[225] This Δ60-MEK1 construct was used extensively in crystallisation studies.[33, 221, 226-231] Other deletions were also made: the loop 271-302[31, 41, 232], 262-304[31]. The deletion of 1-36 *N*-terminal deletion and the triple mutation S298NS299FS300K afforded better constructs.[156] Variants of this construct were therefore used extensively.[155, 233, 234] We decided to use the truncated Δ60MEK1, further referred as ΔMEK1, for our study. The expression of ΔMEK1 was more efficient as it yielded 5 mg/L of culture compared to the 2 mg/L for the full-length construct. The chemistry developed for the site-selective chemical phosphorylation was also transferable from the full-length MEK1 to ΔMEK1. However, the synthesis of ΔMEK1Dha222 was performed using MDBP at room temperature as the truncated construct precipitated at 37 °C. Although the formation of ΔMEK1pC222 was detected in preliminary trials, the thiamichael reaction needed to be optimised.

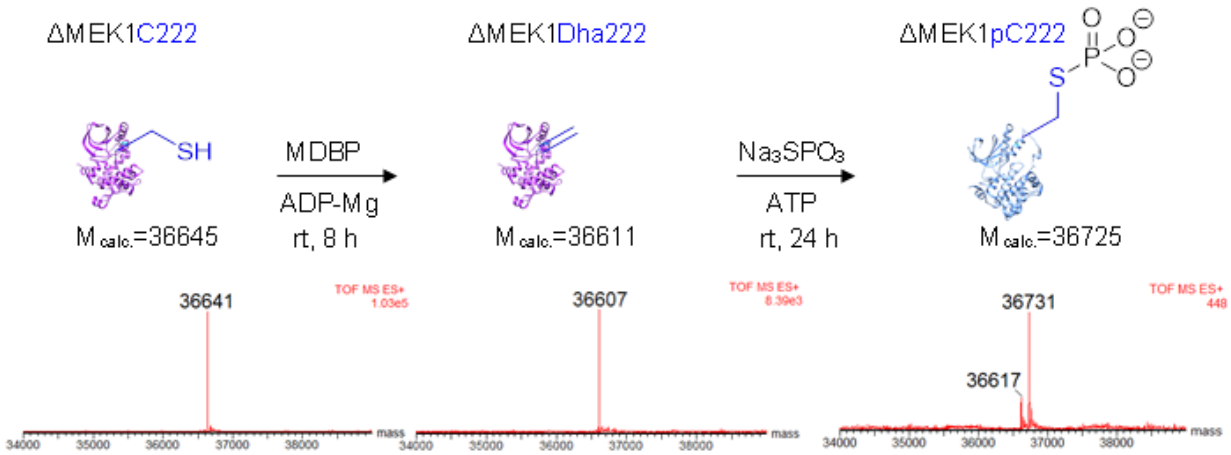


Figure 5.7. The treatment of ΔMEK1C222 with MDBP in the presence of ADP-Mg afforded ΔMEK1Dha222, $M_{calc.}$=36611, $M_{obs.}$=36607 after 8 hours. The treatment of ΔMEK1Dha222 with sodium thiophosphate in the presence of ATP was partially converted to ΔMEK1pC222 (85% conversion), $M_{calc.}$=36725, $M_{obs.}$=36731 after 24 hours.

5.2. Chemical activation of MEK1 at position 218, mimics of MEK1pS218 and MEKpS218pS222

5.2.1. Chemical phosphorylation of MEK1 at site 218

Similarly to the formation of MEK1Dha218, the reaction of MEK1pC218 was more challenging as the position 218 was less accessible. The one-pot reaction with ADP-Mg as additive was not successful. LCMS signal was not interpretable after 6 hours at 37 °C. A larger excess of sodium thiophosphate, 40000 equivalents instead of the 20000 equivalents used in the formation of MEK1pC222 in a buffer containing ATP, afforded MEK1pC218 after 24 hours at room temperature (Figure 5.8).

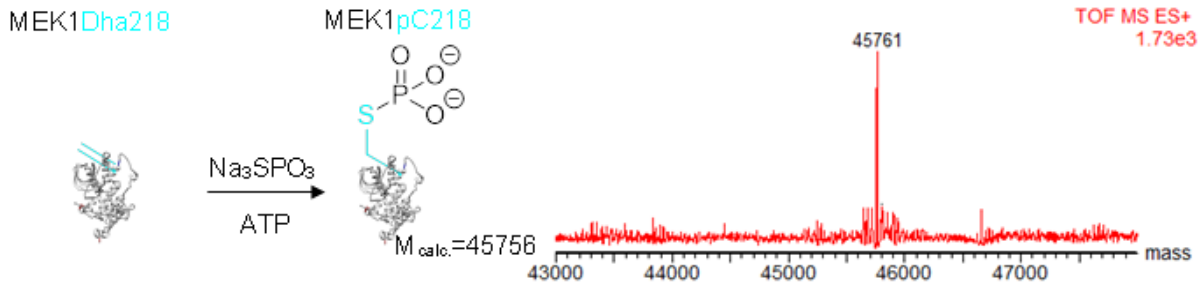


Figure 5.8. Treatment of MEK1Dha in buffer containing ATP with sodium thiophosphate afforded MEK1pC218, $M_{calc.}$=45756, $M_{obs.}$=45761.

However, the reaction was capricious, and not reproducible. The quality of the commercially available sodium thiophosphate hydrate batches was variable. Not only their water content fluctuated, but their purity was inconsistent from batches to batches of the same supplier (Figure 5.9.A and B). Visual and olfactory inspection allowed the detection of sulfury impurities (Figure 5.9.A), likely sodium sulphide, Na_2S, which reacted with MEK1Dha218 to afford MEK1C218 (Figure 5.9.C). Analysis of the sodium thiophosphate by ^{31}P NMR permitted the estimation of hydration level as well as the detection of an uncharacterised impurity at -5.82 ppm (Figure 5.9.A and B).

To control the concentration of sodium thiophosphate and the pH of the reaction, anhydrous sodium thiophosphate was prepared by triturating the commercial batches in dry methanol (Chapter 3). This procedure, even repeated several time, did not eliminate the sulfury impurity. The use of a 0.4 M sodium thiophosphate pH 8.7 solution, made from a pure anhydrous batch of the salt, was optimal for the transformation of MEK1Dha218 to MEK1pC218 (Figure 5.9.D).

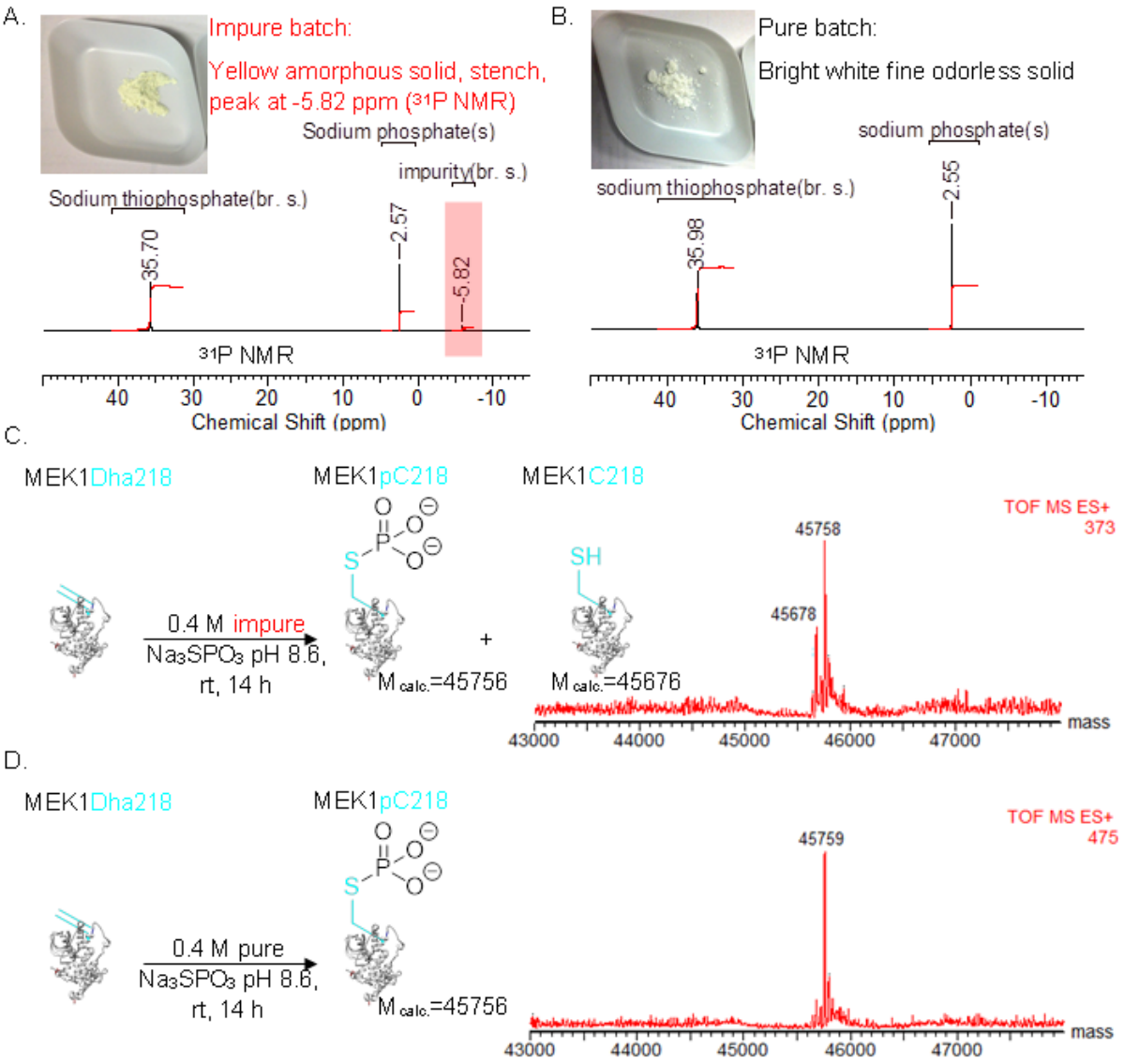


Figure 5.9 A. Impure batch of anhydrous sodium thiophosphate was yellow and 31P NMR showed an unknown peak at -5.82 ppm. B. Pure batch of anhydrous sodium thiophosphate was a bright white fine “icing sugar” type solid.C. Reaction of MEK1Dha218 with 0.4 M impure sodium thiophosphate at pH 8.6 afforded a mixture of desired MEK1pC218 $M_{calc.}$=45756, $M_{obs.}$=45758 and MEK1C218, $M_{calc.}$=45676, $M_{obs.}$=45678. D. Reaction of MEK1Dha218 with 0.4 M pure sodium thiophosphate at pH 8.6 afforded MEK1pC218 $M_{calc.}$=45756, $M_{obs.}$=45759.

Finally, using the same optimised conditions, MEK1Dha218Dha222 was treated with sodium thiophosphate. As expected, the first phosphorylation was relatively fast, and only MEK1Dha222pC218 was detectable after 6 hours. In comparison, the second phosphorylation was particularly sluggish. But full conversion to MEK1pC218pC222 was observed after 24 hours (Figure 5.9). Unfortunately, the construct could not be isolated as the apo-form for activity studies.

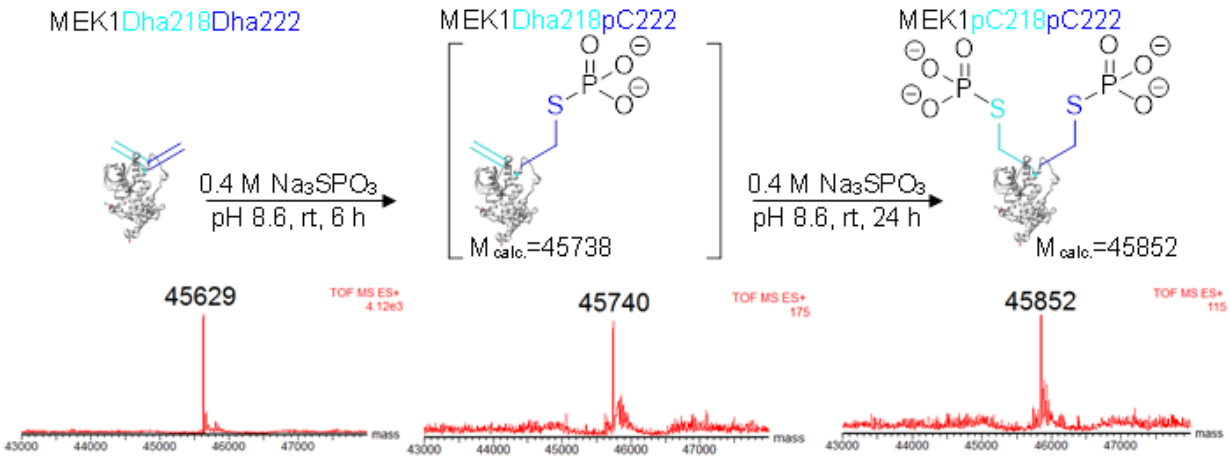


Figure 5.10. Reaction of MEK1Dha218Dha222 with 0.4 M pure sodium thiophosphate at pH 8.6 afforded successively MEK1Dha222pC218, $M_{calc.}$=45739, $M_{obs.}$=45740, and MEK1Dha222Dha218, $M_{calc.}$=45852, $M_{obs.}$=45852.

Like MEK1pC222, both MEK1pC218 and MEK1pC218pC222 were subjected to tryptic/(AspN) digest conditions. Unfortunately, the desired phosphorylated peptides were not detected by LCMS-MS with both CID and ETD fragmentation. High-energy collision dissociation (HCD) has been employed for the detection of phosphopeptides.[235] However, the desired peptides were not observed using this fragmentation method. The same challenges were encountered for the LCMS-MS study of MEK1pSs,[63] which would indicate that the nature of the digested peptide is the cause.

5.2.2. Activation of MEK1 at position 218

Both MEK1C218 and MEK1C218C222 ((Table 5.2, Entry 2 and 3) showed a reduced basal activity compared to MEK1S218S222 (Table 5.2, Entry 1). MEK1pC218 (Table 5.2,

Entry 4) displayed a similar activity as MEK1pC222 (Table 5.1, Entry 5), which was also higher than the corresponding constitutively active MEK1E218 (Table 5.2, Entry 5).

Table 5.2. Activity of MEK1X218 constructs

Entry	MEK1 construct	$K_{f1}x10^3$ $M^{-1}.s^{-1}$	$K_{f2}x10^3$ $M^{-1}s^{-1}$
1	MEK1S218S222	0.41±0.03	0.47±0.03
2	MEK1C218	0.17±0.01	0.31±0.03
3	MEK1C218C222	0.25±0.01	0.30±0.01
4	MEK1pC218	16.5±0.7	9.2±0.5
5	MEK1E218	1.40±0.07	1.30±0.09

Interestingly, MEK1Dha218 and MEK1Dha218Dha222 displayed high kinase activity (Table 5.3, Entry 4 and 5). As MEK1S218A and MEK1S222A were reported inactive,[23] the kinase activity of MEK1 containing Dha at position 218 could only be explained by the introduction of a sp^2 centre in place of a sp^3 hybridized α-C in the protein chain.[167] The difference of MEK1Dha218 and Dha222 (Table 5.3, Entry 3 and 4) was explained by the fact that 218 belonged to an α-helix and 222 to a more flexible loop. The influence of a geometrical change was more marked for position 218. However, variations in folding were not detected by CD (Chapter 2, Figure 2.13), which indicated that those changes were particularly localised.

Table 5.3. Activity of MEK1Dha constructs

Entry	MEK1 construct	$K_{f1}x10^3$ $M^{-1}.s^{-1}$	$K_{f2}x10^3$ $M^{-1}s^{-1}$
1	MEK1S218S222	0.41±0.03	0.47±0.03
2	MEK1C218C222	0.25±0.01	0.30±0.01
3	MEK1Dha222	1.37±0.03	1.41±0.03
4	MEK1Dha218	11.99±0.54	10.00±0.61
5	MEK1Dha222Dha218	27.32±2.19	26.37±2.38

The physiological relevance of the activity of MEK1Dhas is unknown. Nevertheless, activity of kinases containing Dha is intriguing as Dha is naturally a product of the degradation of phosphorylated residues in proteins.[104, 105] These Dhas residues further react with cellular nucleophiles like glutathione to irreversibly inactivate the enzyme. In addition, some microorganism possesses enzymes, lyases, able to cleave MAPK phosphothreonine to Dha in the host cell and therefore inhibiting their activity.[106, 236] Moreover, Dha is also present and necessary for the activity of diverse enzyme like histidine ammonia lyases, which catalyses the deamination of histidine to transurocanic acid,[237] and phenylalanine ammonia lyases, which cleave the amino acid to form trans-cinnamate. It would be particularly interesting to examine whether kinases bear a Dha residue *in vivo.*

5.2.3. Crystallisation of MEK1 activated at position 218

As explained for the crystallisation of MEK1pC222, the truncated ΔMEK1pC218 and ΔMEK1pC218pC222 were chosen to be crystallised. In addition, in view of the activity of both MEK1Dha218 and MEK1Dha218Dha222, the crystallisation of their corresponding truncated constructs was also envisaged. 10 mg of ΔMEK1Dha218 were readily produced from ΔMEK1C218 (Figure 5.11). The synthesis of ΔMEK1pC218 is on-going. Preliminary experiments with an impure batch of sodium thiophosphate were encouraging as the corresponding mass was detected (Figure 5.11).

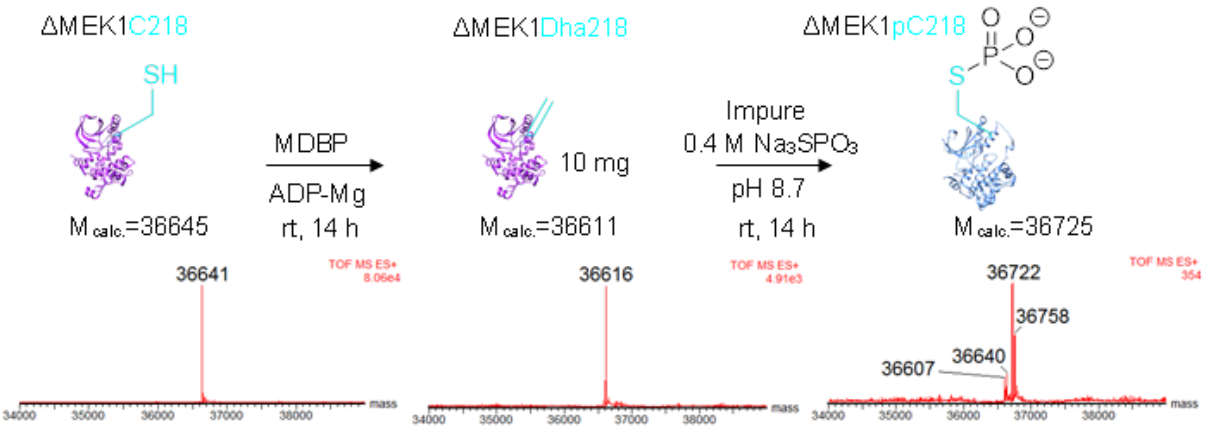


Figure 5.11. The treatment of ΔMEK1C218 with MDBP in the presence of ADP-Mg afforded ΔMEK1Dha218, $M_{calc.}$=36611, $M_{obs.}$=36616 after 14 hours. The reaction was scalable up to 10 mg. The treatment of ΔMEK1Dha218 with a 0.4 M impure sodium thiophosphate pH 8.7 was partially converted to ΔMEK1pC218 (85% conversion), $M_{calc.}$=36725, $M_{obs.}$=36722 after 24 hours. Oxidised protein (+36) was also detected.

The activity of the truncated constructs ΔMEK1C218 and ΔMEK1Dha218 was assessed. ΔMEK1C218 displayed particularly low basal activity (Table 5.4, Entry 2). As expected, ΔMEK1Dha218 was still more active than ΔMEK1C218 (Table 5.4, Entry 4). However, its activity was negligible compared to the one found for the full length MEK1Dha218 (Table 5.4, Entry 3). Furthermore, in contrast with full length MEK1 constructs, the rate of the second phosphorylation, k_{f2}, appeared significantly lower than the rate of the first phosphorylation, k_{f1}. As the residues 1-11 of MEK1 are associated with ERK1 docking,[37, 238, 239] these results appeared consistent with the effect of the truncation of the first 60 amino acids. They also showed that the loss of these residues affected the recognition of monophosphorylated ERK1.

Table 5.4. Comparative activity of full-length MEK1 and truncated ΔMEK1Dha constructs

Entry	MEK1 construct	$K_{f1}x10^3$ $M^{-1}.s^{-1}$	$K_{f2}x10^3$ $M^{-1}.s^{-1}$
1	MEK1C218	0.25±0.01	0.30±0.01
2	ΔMEK1C218	0.100±0.003	0.010±0.003
3	MEK1Dha218	11.99±0.54	10.00±0.61
4	ΔMEK1Dha218	0.70±0.05	0.20±0.01

A site-selectively chemical phosphorylation of MEK1 has been described. It relied on the regioselective incorporation of an electrophilic Dha “tag”, and the subsequent creation of phosphocysteine, a phosphoserine mimics by thiamichael addition with sodium thiophosphate. However, the limitations of the thiamichael addition were evident. Not only was the reactivity of the thiophosphate poor, but, the reaction also afforded an epimeric mixture as eluded chapter 1. Finally, the stability of the S-P against acid and phosphatases limits the potential application of the methodology.

Alternative reactions were therefore considered to solve these issues.

5.3. Alternative approaches towards chemical phosphorylation of MEK1

5.3.1. Alternative thiamichael reagent

Thiamichael addition to protein with negatively charged thiol is notoriously not efficient.[95] Therefore, the strategy was to mask some of the charges with chemical group that could be removed after the thiamichael reaction (Figure 5.12.A). Potassium diethylthiophosphate was chosen as our model as it was readily available, and has been described to add to small molecule Michael acceptor.[240] In our hands, potassium diethylthiophosphate did not react with MEK1Dha222 (Figure 5.12.B).

Not only is the thiamichael reaction with thiophosphate impaired by the presence of the two charges but its ambident nucleophilicity is also an issue.[241] A way to shift the equilibrium to S nucleophile would be particularly useful as O nucleophile does not react in conjugated addition. The counter ion should also been investigated as the ammonium diethylthiophosphate reacts with an alkyl halide.

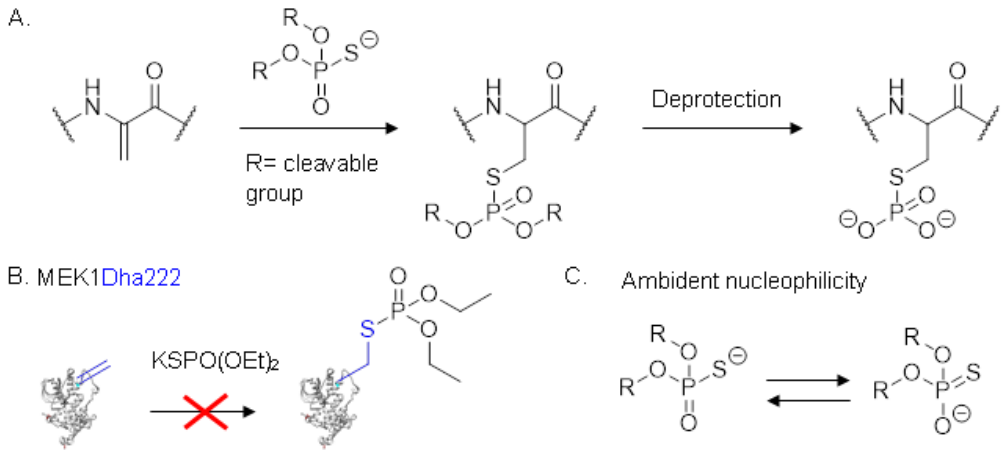


Figure 5.12. A. Strategy of phosphocysteine formation in two steps: thiamichael with a O-protected thiophosphate ester followed by deprotection. B. The formation of the chemically phosphorylated MEK1 was not detected during the reactions of MEK1Dha222 with potassium diethylthiophosphate, $KSPO(OEt)_2$. C. Ambident nucleophilicity of thiophosphate derivatives.

5.3.2. Stereo-and-site-selective phosphocysteine formation approach

We have demonstrated that phosphocysteine residues were mimics of phosphoserine and phosphothreonine residues. However, our methodology afforded epimeric mixtures. On an amino acid level, chiral Dha precursors were built to resolve the stereochemistry in a thiamichael addition.[242] However, the complexity of a protein makes it more challenging. Therefore, a non-thiamichael strategy towards the stereoselective formation of a phosphocysteine residue was envisaged. It combined two reactions reported on peptides and proteins (Figure 5.13.A): the Staudinger-phosphite reaction, the reaction between an azide and a phosphite ester, which leads to the formation of a phosphorylated amine after the hydrolysis of the phosphoester bonds,[85, 86] and the reduction of disulphide bonds by electron-rich phosphines.[117] Ellman's reagent reacts readily and selectively with cysteine residues to form assymetrical disulphides. We anticipated that these disulphides could then be reduced by phosphite esters to yield the desired phosphocysteine residues stereoselectively after the hydrolysis of the phosphoester bonds. The successful use of this

strategy was published by Hackenberger group during the preparation of this manuscript. The reaction was successful in organic solvents, and a wide range of phosphite esters was tested (Figure 5.13.C). These esters were protected by photocleavable or acidic labile groups. The reaction in aqueous media was poor although the yield could be improved by adding an organic solvent.

The strategy was tested on MEK1C218C222. As exemplified in previous experiments (Chapter 3), Ellman's reagent reacted with MEK1 cysteines. However, a selective and controlled addition proved challenging. An optimal addition of three adducts to MEK1C218C222 was achieved (Figure 5.13.B). Triethylphosphite ester was used as our control phosphite as it was readily available. Nonetheless, the treatment of this construct with triethylphosphite ester only provided MEK1C218C222. Indeed, in aqueous buffer, the hydrolysis of the phosphonium ion (1) was faster than the phosphate ester formation (2) (Figure 5.13.B). Hackenberger drew similar conclusions. They also observed the formation of monothiophosphate esters, which indicates a possible nucleophilic attack on the benzylic position. In my case, the mass of the monoethylthiophosphate ester was not detected. The strategy was not pursued further. Indeed, a selective addition of Ellman's reagent on protein containing multiple cysteine residues is challenging. Moreover, finding the ideal phosphite ester with the right balance between reactivity and lability is also tricky.

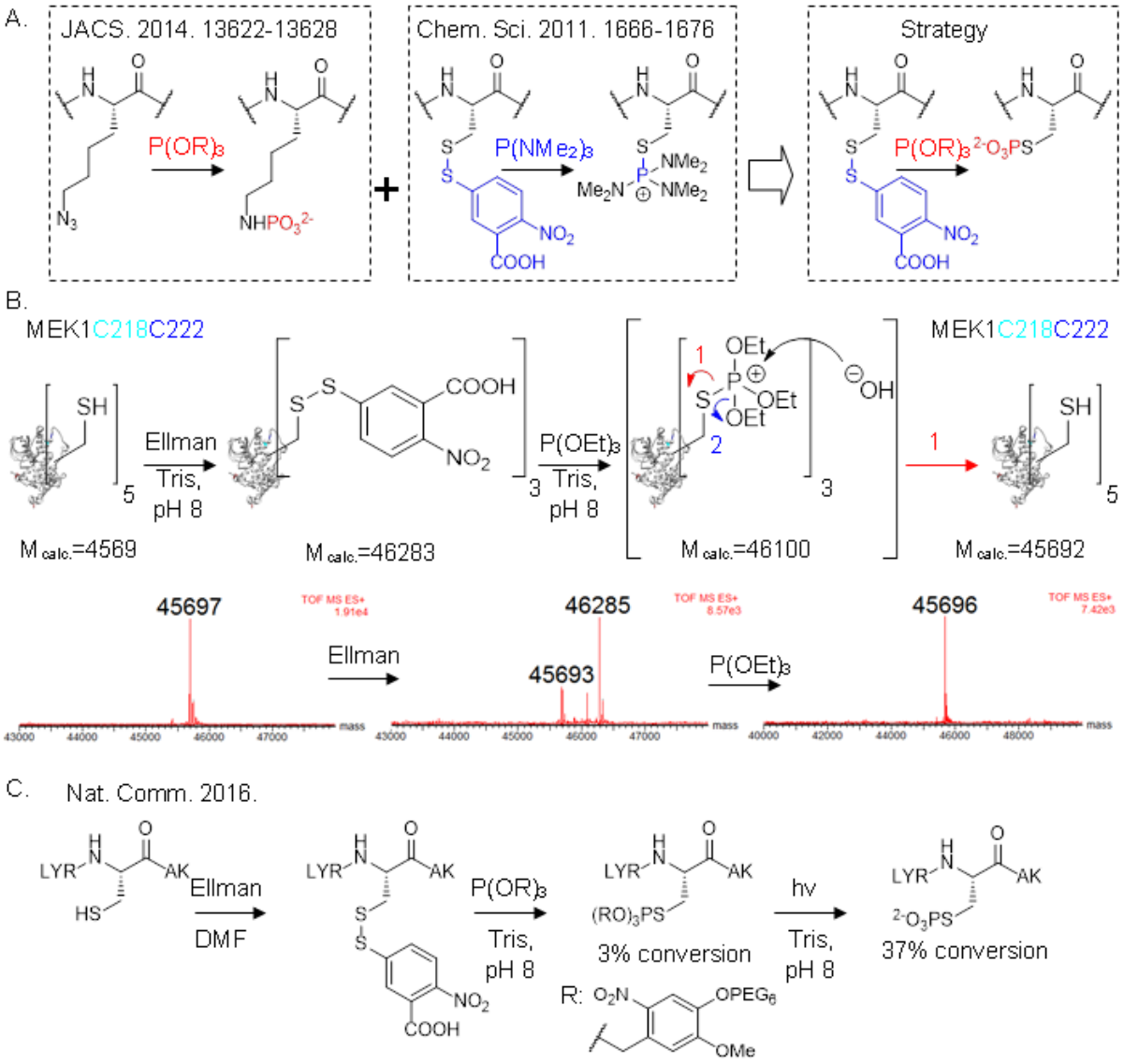


Figure 5.13. A. The strategy of the site-and-enantioselective pC formation involved the formation of a disulphide with Ellman's reagent followed by hydrolysis and ester deprotection, which was inspired by both the Staudinger-phosphite reaction and the reaction of disulphide with electron rich phosphine. B. MEK1C218C222 reacted with Ellman's reagent to optimally form MEK1C$_{NTB}$x3, $M_{calc.}$=46283, $M_{obs.}$=46285 . The subsequent reaction with phosphite ester afforded MEK1C218C222, $M_{calc.}$=45692, $M_{obs.}$=45696 due to hydrolysis of the S-P bond (1). C. This methodology was used successfully on a peptide.[92]

5.3.3 Site-and-stereoselective and stable phosphocysteine mimics formation approach

A simple and selective alkylation of cysteine residues with phosphonate bearing halide would provide stable phosphate mimics[43] without scrambling the stereochemistry (Figure 5.14). The alkylation of thiols, and even cysteine amino acid, with (iodomethyl) phosphonate ester has been described on small molecule.[243-245]

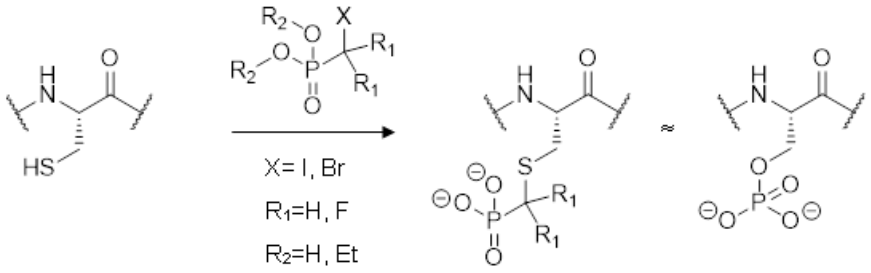


Figure 5.14. Strategy of phosphocysteine formation in a one-step alkylation with (iodomethyl) phosphonic acid (X=I, R_1=R_2=H) and (bromodifluoromethyl) phosphonic acid (X=Br, R_1=F, R_2=H) or in two-step alkylation/hydrolysis with the corresponding O-diethyl esters.

However, the reaction of MEK1C222 with (iodomethyl) phosphonic acid (IMPA) or with (bromodifluoromethyl) phosphonic acid (BFPA) and their respective ester precursors did not lead to the formation of phosphorylated MEK1 (Table 5.5, Entry 4, 5, 6 and 7). Indeed, only starting MEK1C222 was recovered. Iodoacetamide and the negatively charged iodoacetic acid were used as control. LCMS showed that both of them reacted with MEK1C222 (Table 5.5, Entry 2 and 3).

Table 5.5. Reaction of MEK1C222 ($M_{calc.}$=45676) with diverse alkylating agents[a]

Entry	Alkylating agent	Time	$M_{obs.}$ in Da	Product ($M_{calc.}$ in Da)	Conversion
1	DBHDA	5 h	45645	oaMEK1Dha222 (45642)	100%
2	IAA	1 h	45902	$MEK1C_{CAM}$x 4 (45904)	45%
			45962	$MEK1C_{CAM}$x5 (45961)	55%
3	IAAc	1 h	45795	$MEK1C_{COM}$x2 (45792)	50%
			45851	$MEK1C_{COM}$x3 (45850)	50%
4	IMPA ester	5 h	45676	No reaction	0%
5	BFPA ester	5 h	45676	No reaction	0%
6	IMPA	5 h	45676	No reaction	0%
7	BFPA	5 h	45645	No reaction	0%

[a] 500 equivalents was added for each reaction.

This method was also abandoned, however, IMPA and BFPA were used in the methodology described below.

5.3.4. Site-selective formation of stable phosphylated residues by radical reaction

Radical chemistry was being developed in our group to form C-C bonds between Dha and alkyl halide (Figure 5.15.A).[98, 109] The reaction was fast and mild for the protein examples. IMPA and BFPA were well suited for radical chemistry. Interestingly, their reaction with Dha would form two mimics of phosphoserine residues, respectively phosphohomoalanine (Pha) and phosphodifluorohomoalanine (Pfa) (Figure 5.15.B). Pfa was a better mimic of pS, as the CF_2 was a bioisostere of oxygen, and its pK_a was near identical to a phosphate moiety. The major advantage of this chemical phosphorylation strategy was the stability of the C-P bond against phosphatases. Therefore the incorporation of both Pha and Pfa into protein would provide ideal tools for biological studies e.g. discovering new phosphatases using proteomic techniques.

The reactions were successfully tested on a model protein histone H3Dha10.[‡‡] Both H3Pha10 and H3Pfa10 were produced with good conversion (Figure 5.16.C and D). Both were stable to the phosphatase PP1. However, the chemistry on histone protein was performed in denaturing conditions. Therefore, the incorporation of these moieties on a kinase appeared challenging. Nevertheless, initial trials on p38αDha180 were encouraging (Figure 5.16.E), but the optimisation of the conditions was still necessary to transfer the chemistry to MEK1.

[‡‡] This work was performed in collaboration with Thomas Wright and Dr. Billy Ng.

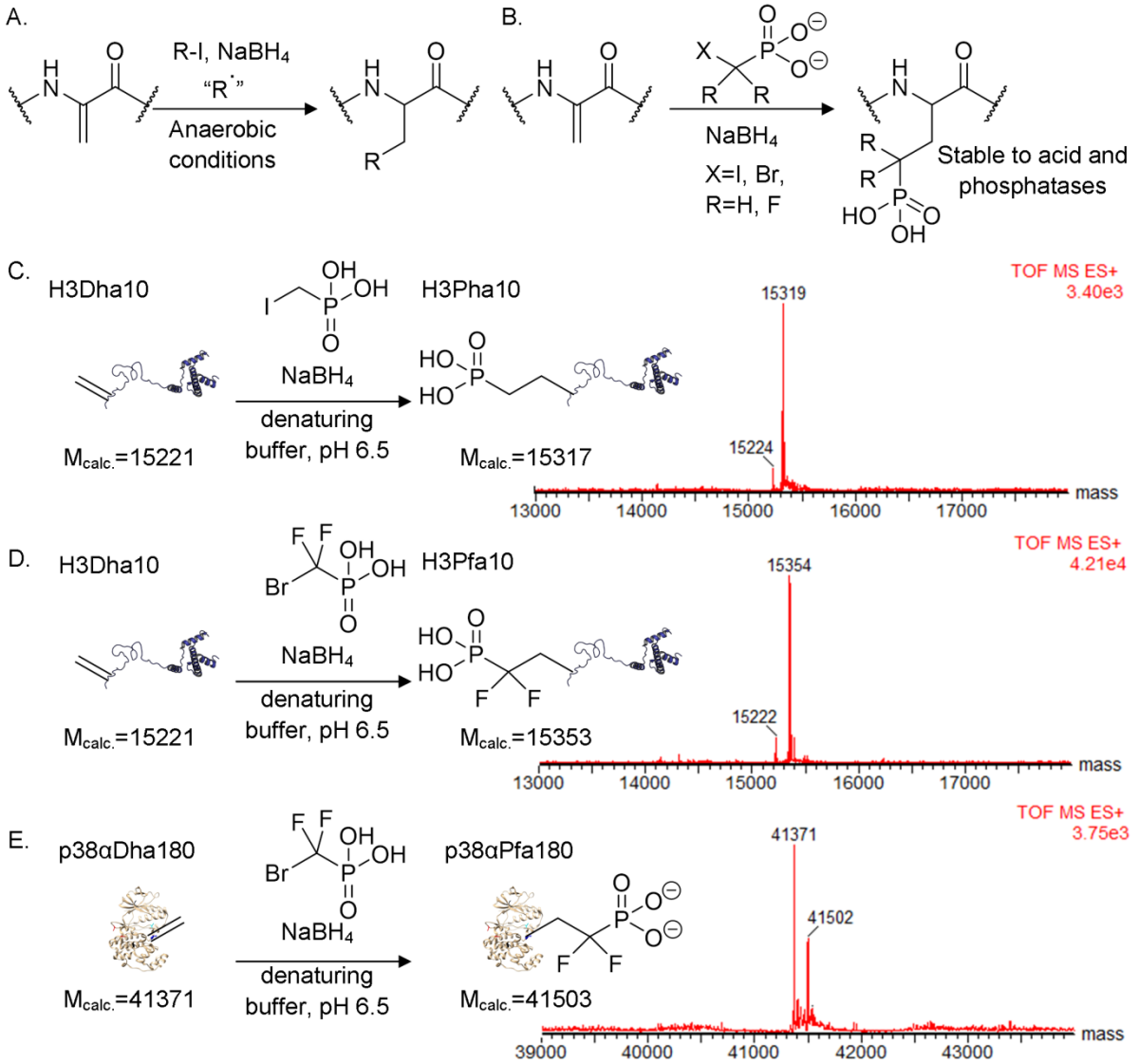


Figure 5.15. A. The reaction of Dha with iodoalkyl compounds in the presence of sodium borohydride involved a radical mechanism and allowed the incorporation of a various R groups to a protein B. Strategy to incorporate phosphorylated residue involved the use of IMPA (X=I, R=H) and BFPA (X=Br, R=F). C. The treatment of H3Dha10 with IMPA in the presence of sodium borohydride afforded >85%, H3Pha10, $M_{calc.}$=15317, $M_{obs.}$=15319. D. The treatment of H3Dha10 with BFPA in the presence of sodium borohydride afforded >85%, H3Pfa10, $M_{calc.}$=15353, $M_{obs.}$=15354.E. The treatment of p38αDha180 with BFPA in the presence of sodium borohydride afforded p38αPfa180, $M_{calc.}$=41371, $M_{obs.}$=41503.

CONCLUSION

Chapter 6: Conclusion

Post-translational phosphorylation of proteins affects dramatically their structure and function. However, the effects of phosphorylation at a molecular level are not well understood. Site-selective phosphorylation of proteins, and in particular kinases, therefore appears as a valuable tool towards the elucidation of these intimate effects.

The site-selective chemical phosphorylation of the p38α kinase was developed by our group and described in JACS.2014.1698-1701 as “Synthetic Phosphorylation of p38α Recapitulates Protein Kinase Activity” that I co-authored.* The chemical phosphorylation was achieved in three steps:

1. The site-directed mutagenesis of a cysteine residue in place of the “phosphorylable” residues of the kinase.
2. The alkylation of this cysteine residue with a bis-alkylating/elimination agents leading to the formation of dehydroalanine as an electrophilic handle or “tag”.
3. The chemical phosphorylation by addition of sodium thiophosphate affording a phosphocysteine residue, biomimetic of phosphoserine and phosphothreonine residue.

This chemical phosphorylation methodology was validated on the kinase as its enzymatic activity was recapitulated *in vitro* against its natural substrate. However, no fundamental application of this methodology has been described yet. We proposed to use and optimise this methodology to crystallise a phosphorylated and active MEK1 kinase. MEK1 is activated by the phosphorylation of two residues, S218 and S222. In this thesis, the formation of MEK1pC218, MEK1pC222 and MEK1pC218pC222, as mimics of the corresponding MEK1pSs is discussed.

Site-directed mutagenesis to introduce cysteine residues at position 218 and 222 was straightforward. In contrast with previous examples using the methodology, MEK1 naturally possesses 5 cysteine residues, some of them potentially reactive and important for the overall kinase activity so that they could not be mutated out.

The bis-alkylation/elimination step was therefore optimised towards the reaction at one cysteine residue over the others. The structural study of non-phosphorylated MEK1 was essential to engineer a dual strategy based on masking the reactivity of undesired cysteine residues, and on enhancing the accessibility and reactivity of the bis-alkylating/elimination agents. The discovery of ADP-Mg complex as a masking and stabilising additive was crucial to perform the chemistry on the protein in a mild manner. In addition, the design of more reactive and selective alkylating agents promoted the site-selective and scalable formation of MEK1Dha218 and MEK1Dha222. This strategy was successfully applied to the simultaneous formation of double Dha at position 218 and 222, to provide MEK1Dha218Dha222. This emerged as the first example of multiple Dha incorporation onto a protein substrate.

During the optimisation of the bis-alkylation-elimination step, some observations prompted us to investigate the chemical mechanism of the sulfonium ion elimination more in detail. Caddick had previously described the use of the stable sulfonium ion as an orthogonal handle to Dha. Therefore, in a joint effort, we believed that the elucidation of this mechanism would lead to the design of novel bioorthogonal handles. The reactions with DBHDA, our standard bis-alkylating/elimination agent and the commercial diiodobutane (DIB) were therefore compared. A method to site-selectively label cysteine residues with deuterium on amino acid and protein template was conceived. The deuteration of cysteine residues on the α-carbon ultimately enabled the elucidation of the mechanism. This mechanism was indeed dependent on the reagent. The use of DBHDA entailed a two-step elimination relying on the formation of a sulfonium ylid, while the use of DIB was associated with direct one-step elimination. The manuscript describing this work is currently under revision.*

The presence of electrophilic handles at the desired positions eventually enabled the chemical phosphorylation of MEK1 at position 218 and 222. The three constructs MEK1pC218, pC222 and pC218pC222 were therefore available. The activation of MEK1 necessitating the phosphorylation of either one or two residues is still debatable. The assessment of the kinase activity of our phosphorylated constructs confirmed that the phosphorylation of only one site was necessary. Interestingly, our phosphorylated constructs were significantly more active than the corresponding constitutively active constructs. This validated their potential application in biological studies since the standard constitutively active constructs had been used extensively in the study of MEK1 function. Provocatively, the formation of Dha at position 218 considerably activated the enzyme. A negative charge was therefore not necessary to activate the kinase. As proteins containing Dha are present in cells, usually *via* the degradation of a phosphorylated residue, it is not unconceivable that kinases are irreversibly activated by the formation of Dha. It would be therefore interesting to include Dha in future proteomics searches. A manuscript on MEK1Dha and its biological relevance is in preparation.

Our site-selective chemical phosphorylation methodology was mild, scalable and transferable to alternative MEK1 constructs. Structural studies of active MEK1 is therefore accessible. The crystallisation of active MEK1 is on-going.

Finally, the ability to segregate the reactivity of cysteine residues within a protein is powerful. Moreover, the site-selective Dha formation is applicable on high profile, structurally complex and functional proteins, which could be combined to new technologies employing Dha as a bioorthogonal handle on proteins. An example of this application has been published in Science.2016 as “Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity” that I co-authored.*

*During the time of submission (Oct. 2016), the viva (June 2017), the granting leave to supplicate (Sept. 2017) and the printing of this thesis (Nov.2018), the work on the site-selective deuteration and mechanism determination was published as:

Galan, S. R. G.; Wickens J. R.; Dadová J.; Ng W.; Zhang X.; Simion R. A.; Quinlan R.; Pires E.; Paton R.S.; Caddick S.; Chudasama V.; Davis B. G. Post-translational Site-elective Protein Backbone α-deuteration. *Nat. Chem. Biol.* **2018**, 14, 955-963.

A review on Dha formation and application on proteins was also written as:

Dadová J.; Galan, S. R. G.; Davis B. G. Synthesis of Modified Proteins via Functionalization of Dehydroalanine. *Current Opinion in Chemical Biology.* **2018**, 46, 71–81.

They are in addition of the two publications cited above:

Chooi, K. P.; Galan, S. R. G.; Raj, R.; McCullagh, J.; Mohammed, S.; Jones, L. H.; Davis, B. G. Synthetic Phosphorylation of p38α Recapitulates Protein Kinase Activity. *J. Am. Chem. Soc.* **2014***, 136*, 1698-1701.

Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J. L.; Wiewiora, R.; Ng, W.; Raj, R.; Faulkner, S.; Vallée, M. R. J.; Phanumartwiwath, A.; Coleman, O. D.; Thézénas, M.; Khan, M.; Galan, S. R. G.; Lercher, L.; Schombs, M. W.; Gerstberger, S.; Palm-Espling, M. E.; Baldwin, A. J.; Kessler, B. M.; Claridge, T. D. W.; Mohammed, S.; Davis, B. G. Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science.* **2016**, *354*, 6312, aag1465.

EXPERIMENTAL

Chapter 7: Procedures for the Site-Selective Chemical Activation of MEK1

7.1. Protein visualisation tools and general computational calculations

7.1.1. Molecular graphics

Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).[246]

The crystal structures were retrieved from the RSCB protein database bank.

PDBs used were: 1R3C[247] (p38α), 3W8Q (MEK1-ATP), 3EQD[248] (MEK1-ATP-Mg), 3EQI[248] (MEK1-ADP-Mg), 1S9J[225] (ΔMEK1), 1ST3[249] (SBL), 2IAE[214] (PP2A) and 4QTB[250] (ERK1).

Homology model of ATF2 and p38αC180C182 were calculated using Phyre2.[251] ATF2 sequence was based on the sequence of amino acids 1–287 from Bos Taurus (sequence ID: AA133291.1) obtained by BLAST search[252] from PDB 1BHI. p38α sequence was based on PDB1R3C.

7.1.2. Protein properties calculation

Distances between amino-acid residues were calculated using Chimera.

The relative side-chain accessibility was calculated using Naccess ((Hubbard,S.J.& Thornton, J.M. (1993), 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London) as the % accessiblity compared to the accessibility of that residue type in an extended ALA-x-ALA tripeptide (for amino acids).[253] By default, Naccess calculation does not takes into account the ligands. The ligand can be added with the command –h.

Appendix A shows the residue accessibility file (.rsa) output of a given calculation, which contains summed atomic accessible surface areas over each protein, the relative accessibility of each residue. The data is summed over residues, and split into 5 classes. Total (all atoms), Non Polar Sidechain (all non-oxygens and non-nitrogens in the sidechain), Polar Sidechain (all oxygens and nitrogens in the sidechain), total sidechain, and mainchain. Absolute sums over the whole chain are also given.

7.2. Protein reaction monitoring and characterisation

7.2.1. Intact protein mass spectrometry

Two different LC- MS-TOF systems were used in this study.

LCT classic system

Equipment

Mass spectrometer Waters Micromass LCT (ESI-TOF-MS) was coupled to a Shimadzu HPLC system composed of a Shimadzu LC-20AD SP, a DGU-20A5 degasser, a SIL-20Ac autosampler-cooler, SCL-10A VP system controller.

LC methods

Thermo Scientific, Proswift™ RP-2H (4.6x50 mm SS) column was run with eluents A, water/acetonitrile/formic acid (95:5:0.1), B: acetonitrile/water/formic acid (95:5:0.1). The method was programmed as follows: 100%A (1 minutes, flow 0 to 0.4 mL/min), 100% A to 100% B (6 minutes), 100% B (4 minutes), 100% B to 100% A (1 minute) and 100% A (5 minutes).

Thermo Scientific, Proswift™ RP-4H Analytical (1x50 mm) column was with eluents A, water/acetonitrile/formic acid (95:5:0.1), B: acetonitrile/water/formic acid (95:5:0.1). The method was programmed as follows: 100% A to 50% A (4 minutes flow 0 to 0.4 mL/min),

50% A to 100% B (2 minutes), 100% B (2 minutes), 100% B to 100% A (1 minute) and 100% A (4.2 minutes).

Merk Millipore Chromolith® FastGradient RP-18e (50x2 mm) column coupled with a RP18e (5x2 mm) guard cartridge was run with eluents A, water/formic acid (100:0.1), B: acetonitrile/ /formic acid (100:0.1). The method was programmed as follows: 95% A (2 minutes flow 0.4 mL/min), 95% A to 95% B (6 minutes), 95% B (3 minutes), 95% B to 95% A (0.5 minute) and 95% A (2.5 minutes).

MS method

The electrospray source of LCT was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L hr-1.

Sample preparation

The protein mixture was diluted to the desired concentration with MilliQ water, typically 0.03-0.05 mg/mL. 8-10 µL were injected.

LCT premier system

Waters LCT Premier XE was coupled to a Waters 1525 Micro HPLC.

LC method

RP18e (5x2 mm) guard cartridge was run with eluents A, water/formic acid (100/0.1) and B, acetonitrile). The method was programmed as follows: initial 95% A (flow 0.3 mL/min), 95% A (0.5 min), 95% A to 40% A (1 min), 40% A (1.5 min, flow 0.3 mL/min and 1.5 min, flow 0.75 mL/min), 2% A (0.5 min, flow 0.75 mL/min).

MS method

The electrospray source of LCT was operated with a capillary voltage of 3 kV and a cone voltage of 100 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 300 L hr-1.

Sample preparation

The protein mixture was diluted to the desired concentration with MilliQ water, typically 0.01-0.02 mg/mL. 10 µL were injected.

Processing of intact protein LCMS spectra

MassLynx4.1 (Waters) software was used to process the LCMS data (Figure 7.1). The peak of interest in the chromatogram was first integrated to give an ion series. For LCT classic, the ion series were calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin, obtained with the same LCMS settings. For LCT-premier, the ion series was autocalibrated. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.1 from Waters). Typically, for MEK1 construct, the deconvolution was set up as follow: output mass, ranges 20000:60000, resolution 1 Da/channel; damage model, uniform Gaussian width at half height 0.75 Da, minimum intensity ratios left 33% and right 33% and completion options, iterate to convergence. The deconvoluted spectra were then zoomed accordingly.

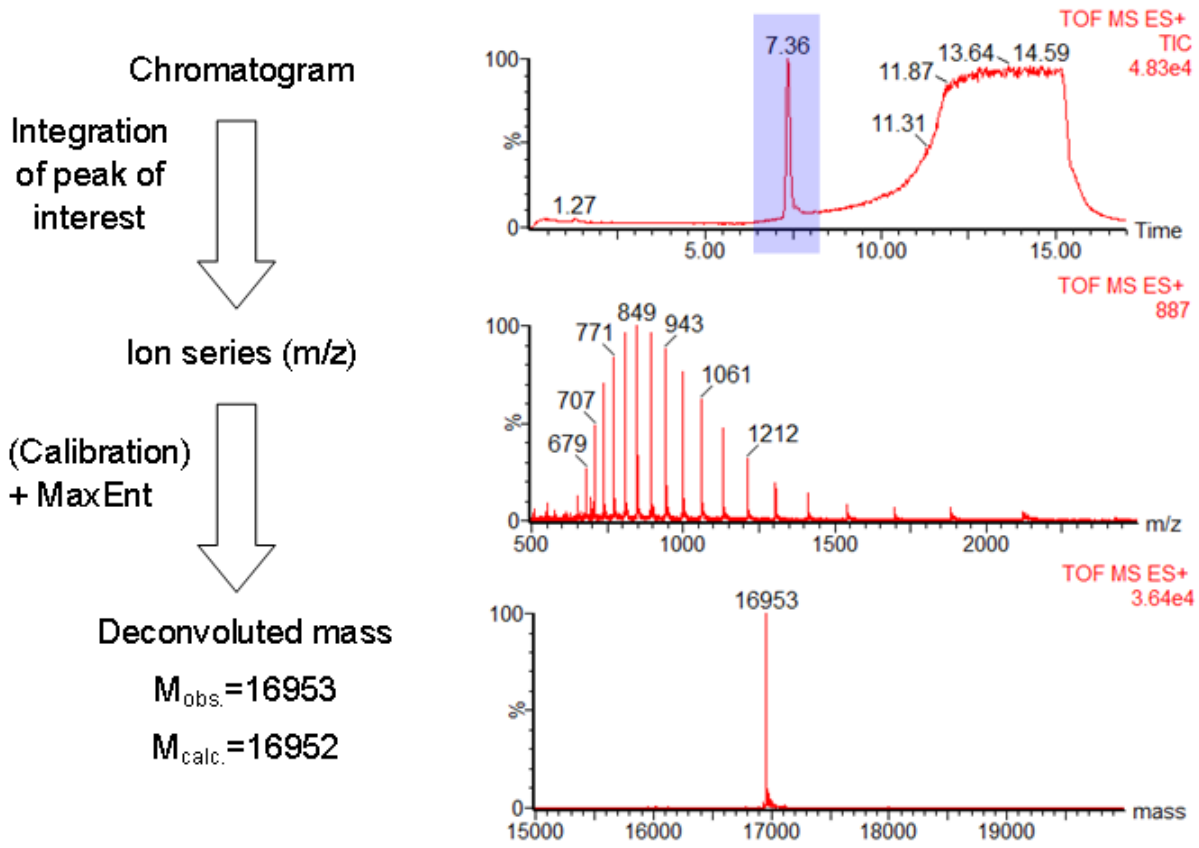
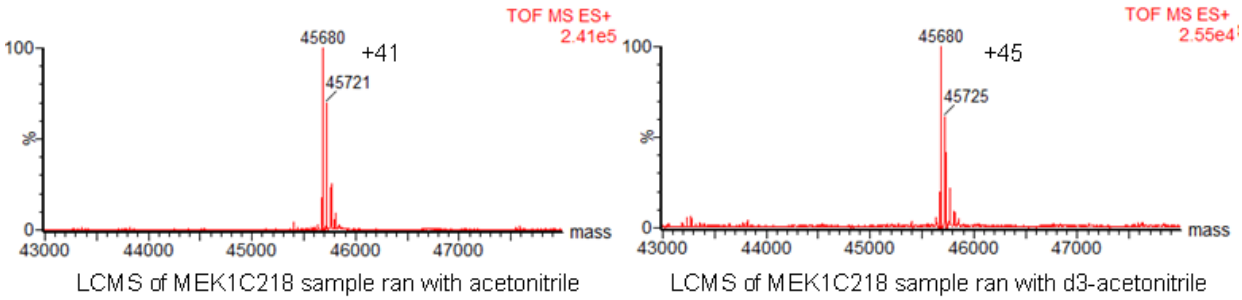


Figure 7.1. Work flow of LCMS processing using Masslynx with myoglobin $M_{calc.}$=16952, $M_{obs.}$=16953, as an example.

LCT-premier +41 adduct

A +41 adduct was visible by LCMS when LCT-premier was used. The intensity of this adduct was dependent on the protein. The adduct corresponded to the addition of acetonitrile during the LCMS process. Indeed, when the LCMS was run with acetonitrile-d3 in place of non-deuterated acetonitrile, the mass of the protein was accompanied by its mass +45.



7.2.2. Analysis of digested proteins by LCMS-MS

Preparation of digested protein samples for LCMS-MS

Two methods of digestions were used.

In-solution digest[254]

The sample (25 µL, around 1 mg/mL) was diluted with 8 M urea solution (75 µL). 200 mM aqueous DTT (2 µL) was then added and the mixture was heated at 56 °C for 15 minutes. The mixtures were cooled down to room temperature and treated with a 36 mg/mL iodoacetamide in 50 mM ammonium carbonate solution (2 µL). The mixture was left standing in the dark for 30 minutes. 200 mM aqueous DTT (2 µL) was then added to neutralize the iodoacetamide. The mixture was diluted 4 times with 50 mM ammonium bicarbonate solution and treated with trypsin (1:50 w/w ratio enzyme:protein). The mixture was incubated at 37 °C overnight. Some samples were also treated with AspN (1:50) and incubated at 37 °C for 4 hours.

The sample was diluted down to 200 fmol/µL with, either milliQ water or 1% formic acid before analysis.

In-gel digest[134]

Excision of protein bands from polyacrylamide gels[§§]

The gel slab was washed with water (2 changes, 10 min each). The bands of interest were excised with a scalpel. As a control, a gel piece was also excised. The gel pieces were cut into roughly 1 mm^3 cubes, and transferred to 0.5 or 1.5 mL eppendorfs.

De-staining gel pieces from Coomassie stained bands/spot

[§§] Gel were run in reducing conditions (β-mercaptomethanol, BME was present).

The gel pieces were washed with water (15 min shaking). 100 µL of 100 mM ammonium bicarbonate (10 min shaking) was added. The supernatant was pipetted out, and 50 µl of 1:1 50 mM ammonium bicarbonate/acetonitrile was added to the gel pieces (10 min shaking). The last two steps were repeated until the colour disappeared. The supernatant was then removed, and 100% acetonitrile was added to cover the gel pieces (30 min at 37°C shaking). After the gel pieces have shrunk (they become white and stick together), the acetonitrile was removed.

Reduction and alkylation

The gel pieces were swollen by addition of 30 µl of 10 mM DTT in 100mM NH_4HCO_3. The mixture was incubated for 45 min at 56°C with agitation. The supernatant was pipetted out. 30 µL of 55 mM of iodoacetamide in 100 mM NH_4HCO_3 was then added. The mixture was left standing in the dark for 30 min at room temperature with agitation. The supernatant was pipetted out. The gel pieces were washed with 100 µL of 100 mM NH_4HCO_3 for 5 min at room temperature and agitation. The supernatant was pipetted out. The gel pieces were then washed with 100 µL of 100% acetonitrile for 10 min at room temperature and agitation. The supernatant was pipetted out, and dry completely in a vacuum centrifuge.

Trypsin digestion

The gel pieces were rehydrated with 30 µL (add enough liquid to cover the gel pieces) of digestion buffer (50 mM NH_4HCO_3 and 20 ng/µl of trypsin), for 30 min and at 4°C. After 15 min, more digestion buffer was added if the initial volume has been absorbed by the gel pieces. Excess enzyme solution was removed. 50mM NH_4HCO_3 buffer was added to cover the gel pieces. The mixture was incubated overnight at 37 °C. After overnight incubation, the supernatant containing digested peptides were recovered to a new tube and store at -20°C.

To extract larger peptides stuck to the gel, 100 µL of 1% formic acid were added and incubated for 15 min at RT with shaking, then transfer to the vial containing supernatant. 100% acetonitrile was then added to the gel pieces. The mixture was incubated for 15 min at

room temperature with shaking. The supernatant was pipetted out, and combined with the solution containing the peptides. This step was repeated with 100 µL of 1:1 acetonitrile/water and 100 µL of 1% formic acid in acetonitrile combining the supernatants to the solution containing peptides. For MALDI MSMS analysis, this solution was then dried, re-dissolved in 0.1% formic acid and purified using Zip Tip.

Analysis of digested protein by LCMS-MS

Two different LCMS-MS systems were used in this study.

Synapt[***]

Samples were analysed using a nanoACQUITY UPLC coupled to SYNAPT HDMS interfaced with a nanoelectrospray source (Waters Corporation, Milford, MA, USA). Protein digests were injected on a 5mm symmetry C18 trap column (180 mmx20 mm) and washed for 1 min at 15 mL/min with 0.1% (v/v) formic acid. Peptides were then separated and eluted for MS analysis using a gradient of acetonitrile containing 0.1% (v/v) formic acid at 300 nL/min over 23 min on a nanoACQUITY UPLC column (BEH130 C18 1.7 µm particle size, 75mmx250mm). The column temperature was set at 35 °C. The reference for the nanolockspray was set to the doubly charged peak of Glu-Fiprinopeptide B at a concentration of 500 fmol/µL, flowing at 400 nL/min. The reference was constantly infused and sampled at 30 s intervals. The eluted peptides were analyzed in the positive ionization mode over a mass range of 50–1990 m/z with a scan time of 0.6 s. The online-eluted peptides were analysed using an MSE method collecting MS/MS data using collision energy ramping from 15–35 eV.

Orbitrap[†††]

The samples were analysed on an Orbitrap Elite (Thermo Fisher Scientific, DE) connected to a UHPLC Proxeon EASY-nLC 1000 and an EASY-Spray nano-electrospray ion

[***] In collaboration with Lingzhi Gong, Mass spectrometry group, Chemistry Department, Oxford.
[†††] In collaboration with Ritu Raj and Prof. Shabaz Mohammed, Biochemistry Department, Oxford.

source with EASY-Spray column (Thermo Fischer Scientific, DK). Peptides were trapped on an Acclaim PepMap® trapping column (100 µmx20 mm, 5 µm C18) and separated on an EASY-spray Acclaim PepMap® analytical column (75 µmx500 mm, RSLC C18, 2 µm, 100 A°). Samples were loaded at a pressure of 500 bar with 100% solvent A (0.1% formic acid in water) and the peptides separated by a linear gradient (length: 15 min, 7% to 30% solvent B (0.1% formic acid in acetonitrile), flow rate: 200 nL/min). Full 99 scan MS were acquired in the Orbitrap (350-1500 m/z, resolution 120 000, AGC target 1_106, maximum injection time 250 ms). CID and ETD spectra were acquired in Ion Trap (resolution 7500, AGC cation target 3 _ 104, AGC Anion target 2 _ 105, maximum injection time 100 ms). After the MS scans, the 5 most intense peaks were selected for fragmentation based on data-dependent decision tree (DDDT). The signal threshold for fragmentation of parent ion was set to 500 ion counts. For CID fragmentation the normalized collision energy and default charge state was set to 35% and 2 respectively. For ETD fragmentation, reaction time was set to 50 ms and supplemental activation was enabled. Charge state screening was enabled; parent ion with unassigned charge state and charge state 1 was rejected. Dynamic exclusion was enabled (exclusion size list 75, exclusion duration 5 s). The setting for DDDT was as such: ETD fragmentation was followed for charge states 3, 4 and 5 with m/z less than 750. For charge states greater than 5, ETD fragmentation was always performed. In all other cases, CID activation was performed. In a separate experiment, HCD was used. All the settings for the instrument were kept the same, except HCD spectra were acquired in Orbitrap and the 3 most intense peaks were selected for fragmentation and normalized collision energy was set to 32%.

Processing

Raw MS data were used for MS data analysis and the characterisation of chemical modifications on proteins. Searches were performed using PEAKS software v7.0 including the following parameters: Data refinement was applied by correcting the precursor's mass. A precursor mass tolerance of 10 ppm and a fragment ion tolerance of 0.5 Da were applied for

the De Novo and the PEAKS searches. The enzyme used was trypsin or AspN. No unspecific cleavages were allowed in the case of trypsin. A maximum of 2 or 3 missed cleavages were allowed when trypsin was used. The PTM search settings were as follows: fixed modification for cysteine carbamidomethyl and variable modification for cysteine desulfurisation, dehydroalanine (-34), dehydroalanine reacted with dithiothreithol (DTT) and β-mercaptoethanol (BME), phosphorylation (+80). Expected mass shifts for novel chemical modifications were entered manually. Data was searched against the specific database containing the protein of interest (MEK1 constructs). The false discovery rate (FDR) was calculated using the decoy fusion method which uses a concatenation of the decoy and target sequences of the same protein together as a "fused" sequence enabling a better assignment of the decoy. In peaks, a score is given by the calculated value -10logP (P representing the P-value). 20 is a good threshold for -10lgP.

The raw data files generated were processed using MaxQuant software (Version 1.4.1.2), integrated with Andromeda search engine as described elsewhere.[255, 256]

Further analyses of the LCMS-MS chromatogram were performed using Thermo Scientific™ Xcalibur™ software.

7.2.3. Circular dichroism (CD) spectroscopy

Circular dichroism (CD) data were collected on a Chiralscan. Experiments were performed in a 1-mm cuvette at a protein concentration of 5 μM in 20 mM Tris-HCl pH 8.0 and 100 mM NaCl. Far-ultraviolet scans were collected from 200 nm to 250 nm. All data are reported in units of mean residue ellipticity, which was calculated from $[\theta]_{MRE}=(\theta_{raw}xMRW)/10cl$, where θ_{raw} is the ellipticity in degrees, MRW is molecular mass in dalton divided by (number of residues − 1), c is the protein concentration ($g.ml^{-1}$), and l is the path length of the cuvette (cm). The concentrations were measured using Nanodrop.

7.3. Small molecule <500 Da reaction monitoring and characterisation

7.3.1. Nuclear magnetic resonance (NMR) spectroscopy

NMR Spectroscopy General

Deuterated solvents were used as the lock, and the residual protonated solvent as the internal reference peak. Spectra were analysed using ACDLabs. Correlation spectroscopy (COSY), heteronuclear multiple bond correlation (HMBC) and heteronuclear single quantum correlation (HSQC) were employed when required for signal assignment. Samples were prepared by dissolving ~5 mg of sample (^{1}H) or ~20 mg (^{13}C) in 500 µL of solvent, filtered if required.

^{1}H NMR Spectroscopy

Chemical shift information (δH) is given in parts per million (ppm) relative to the centre of the solvent signal ($CDCl_3$ 7.26 ppm, DMSO-d6 2.50 ppm, MeOH-d4 3.31 ppm, Acetone-d6 2.05 ppm or D_2O 4.79 ppm). Spectra were recorded on a Bruker AVIIIHD 400 nanobays (400.1 MHz) or AVII 500 (500.1 MHz). The number of protons for a resonance signal is given by nH based on integral values. Multiplicity is denoted by the bracketed symbol. Coupling constants (J) are reported to the nearest 0.1 Hz and are determined using ACDlabs.

^{13}C NMR Spectroscopy

Chemical shift information (δC) is given in parts per million (ppm) relative to the centre of the solvent signal ($CDCl_3$ 77.16 ppm, DMSO-d6 39.52 ppm, MeOH-d4 49.00 or Acetone-d6 29.84 ppm). Spectra were recorded on a Bruker AVIIIHD 400 nanobays (100.6 MHz) or AVII 500 (125.7 MHz) Signals are assumed to be singlets unless otherwise stated. Distortionless enhancements by polarisation transfer with retention of quaternaries (DEPTQ) pulse sequence with broadband proton decoupling were employed for spectrum recording.

^{19}F NMR Spectroscopy

Chemical shift information (δF) is given in parts per million (ppm) relative to the centre of the external standard CFCl3 Spectra were recorded on a Bruker AVIIIHD 400 nanobays (376.4 MHz). Signals are assumed to be singlets unless otherwise stated. Pulse sequences with broadband proton decoupling were employed for spectrum recording, unless stated.

^{31}P NMR Spectroscopy

Chemical shift information (δP) is given in parts per million (ppm) relative to the centre of the external standard 85 % phosphoric acid. Spectra were recorded on a Bruker AVIIIHD 400 nanobays (161.9 MHz). Signals are assumed to be singlets unless otherwise stated. Pulse sequences with broadband proton decoupling were employed for spectrum recording, unless stated.

7.3.2. IR Spectroscopy

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform (FT)-spectrophotometer using a KBr disk or attenuated total reflectance (ATR) attachment – oil or crystals of the compound were placed directly against the stage. A background measurement was subtracted from each spectrum and a new background measurement was recorded before each spectrum was recorded.

7.3.3. Mass spectrometry

Low-resolution mass spectra were recorded on an Agilent 6120 Quadrupole spectrometer using chemical ionisation (CI), positive and negative mode are described where appropriate. High-resolution mass spectra were recorded on a Bruker MicroTOF-ESI mass spectrometer.

7.4. Plasmids and protein expression and purification

7.4.1. General consideration for protein expression and purification

For protein purification and characterisation spectra as well as protein sequences see Appendix B.

Expression and cloning

All bacterial handling was done in a sterile environment inside a biological safety cabinet. LB was bought as pre-formulated dry granules (Melford) and diluted according to the manufacturer's specifications. SOC media were prepared from biological laboratory reagents. All other biological laboratory buffer and media reagents were bought from common suppliers and used as purchased. BL21(DE3) and XL10 E.coli chemically competent cells (Novagen) were handled according to the manufacturer's instructions, thawing on ice before incubation with DNA. XL10 was used for cloning, while BL21(DE3) was used for protein expression.

Site-directed mutagenesis was done using PfuUltra DNA polymerase AD (Agilent), dNTPs (Sigma Aldrich) and DpnI (NEB). Primers for mutagenesis were designed using the Agilent QuickChange primer design, which follows the guidelines given by Stratagene. The sequences were then sent to Life Technologies (Invitrogen) for custom oligonucleotide synthesis. Plasmid purification was done by using the QIAprep® Miniprep Kit (Qiagen) according to the manufacturer's recommendations. LyseBlue reagent had been added to buffer P1 from the Miniprep kit to aid identification of cell lysis. DNA and protein concentration determination was done using a NanoDrop 1000 spectrophotometer (Thermo Scientific), which allowed the calculation of yield. The wavelengths used were a ratio of 260/230 nm and 280/260 nm respectively. The extinction coefficients for proteins were calculated using the ProtParam tool (ExPASy). Extinction coefficients are in units of $M^{-1}.cm^{-1}$ at 280 nm measured in water: MEK1 was 24410 $M^{-1}.cm^{-1}$, 43320 $M^{-1}.cm^{-1}$ for ERK1 and

49850 $M^{-1}.cm^{-1}$ for p38α assuming all cysteine residues were reduced. Other bioinformatics tools from the ExPASy website for various applications were also used.

Purification

Protein pellets were purified using Fast protein liquid chromatography (FPLC) using ÄKTA Purifier™ system (Amersham Biosciences/GE Healthcare). His-trap HP 1 and 5 mL were purchased from GE Healthcare Life Sciences. Dialysis was performed using Slide-A-lyser® MWCO 10 000 dialysis cassettes (Thermo Scientific). Concentration of protein was performed using Vivaspin 500, 15 or 6, MWCO 10 000 spin columns (Sartorius Stedim Biotech) as appropriate. Autoclaving was performed at 121 °C for 10–15 min. Solutions which were used as media or buffers and which could not be autoclaved were filtered through either 0.2 µm cellulose acetate filters, 0.2 µm polyamide filters, or 0.2 µm Minisart® filter cartridges (Sartorius Stedim Biotech) as appropriate. Shaking incubation was performed using either a New Brunswick Scientific Innova® incubator shaker or a New Brunswick Scientific Model G25 incubator shaker. Plate incubation was performed using a Heraeus B6030 incubator. Microcentrifugation was performed using an Eppendorf Centrifuge 5415R. Low speed centrifugation was performed using a Beckman Coulter Allegra™ X-12R centrifuge. High speed centrifugation of large volumes (>2 mL) was performed using a Beckman Coulter Avanti™ J-25 centrifuge. Sonication was performed using a Sanyo Soniprep 150 sonicator, equipped with a microtip. SDS-PAGE protein analysis was done using NuPAGE® gels (Invitrogen). The samples (5 µL) were prepared by mixing with equal volumes of 2xSDS gel-loading buffer (100 mMTris-HCl pH 6.8, 4% SDS, 0.2% bromophenol blue, 20% glycerol), with β-mercaptoethanol (200 mM) being added to the buffer just before use. The samples were mixed by vortexing and heated at 95 °C for 10 min. The samples were run on 10% bis-tris gels as appropriate. The gels were run in MOPS running buffer (50 mM MOPS pH 7.7, 50 mM Tris base, 0.1% SDS, 1 mM EDTA) at 200 V.

Samples (5–9 µL) were loaded as appropriate. Staining of the gels was done with Instant Blue™ Coomassie® stain for 2 hours and destained with MilliQ water for a minimum of 18 h.

7.4.2. Plasmids and cloning

pET28a-MEK1-S222C/C277S/C376S plasmid was prepared by Genscript (1 µL, 200 ng/µL). This plasmid encoded for an *N*-terminal thrombin cleavable His6 tag. In order to overcome gluconoylation, the plasmid was mutated at G(-19)F and cloned by Dr. Phin Chooi to afford MEK1G(-19)F/S222C/C277S/C376S referred as MEK1C222. All desired mutations and deletions were performed from this original plasmid.

The 3 desired *N*-terminal truncated (residues 1-61) MEK1C277SC377S referred as ΔMEK1 mutants with a C-terminal His6 tag (residues 323-329) were constructed and cloned by SGC (Dr. Apirat Chaikuad) from the MEK1 plasmids described above.

7.4.3. Protein expression and purification

Commercial and BGD group proteins

Phosphatases PP1 were bought from New England Biolabs catalogue # P0754S, and PP2A from Millipore #2246078.

Histone H3S10 was expressed and purified by Dr. Billy Ng following a known protocol.[94]

ERK1K71R and p38αC180C182 were expressed and purified by Dr Phin Chooi.

Thrombin was purchase as Thrombin CleanCleave kit (Sigma Aldrich)

Expression of MEK1 variants

Transformation

BL21(DE3) E. coli competent cells (10 µL) were thawed on ice. The plasmid encoding the desired variant (2 µL) was added. The mixture was incubated on ice for 5 min. The transformation was performed by heat shock: the mixture was incubated at 42 °C for 30 s before immediately returning on ice. After 2 minutes, SOC media (30 µL) was added to the cells. The mixture was incubated at 37 °C at 200 rpm for 45 minutes.

Bacterial Culturing

The transformed cells were streaked onto a LB Miller/agar plate supplemented with kanamycin (50 µg/mL final concentration). The plate was incubated at 37 °C for 14 hours. A colony was picked and incubated for cultured 14 hours incubation at 37 °C in LB Miller medium supplemented with kanamycin (10 mL) to provide a starter culture.

Protein Expression

The starter culture was inoculated into LB Miller medium supplemented with kanamycin (1 L). The culture was incubated at 37 °C at 200 rpm until OD600 ~ 0.4. The temperature of the incubator was decreased to 20 °C. The culture was grown further to OD600~1. Protein expression was then induced by the addition of an aqueous solution of IPTG (1 mM final concentration). The culture was incubated at 20 °C at 200 rpm for 20 h. The cells were harvested by centrifugation (9000 rpm, 10 min, 4 °C). The supernatant was discarded and the pellets placed into bags, flash frozen in liquid nitrogen and stored at −80 °C.

Purification of MEK1 variants

Cell Lysis

The cell pellets (10 g) were re-suspended in an ice cold Nickel affinity lysis buffer (50 mM HEPES, 500 mM NaCl, 10 mM imidazole, 0.5 mM TCEP, 5% glycerol, pH 7.5 buffer, 50 mL)

containing 1 pellet of protease inhibitor, DNase I (1 mg) and lysozyme (10 mg). The mixture was shaken using a rotating mixer at 4 °C for 30 minutes. The homogeneous mixture was then sonicated (25x2 s blasts, 60% amplitude, liberating 2500-3000 J total energy). The lysate was clarified by centrifugation (20 000 rpm, 40 min, 4 °C). The supernatant was filtered using 0.2 µm syringe filters.

Ni2+ Affinity Chromatography

Using an AKTA Purifier FPLC system, the clarified lysate was loaded (150 mL) onto a pre-packed 1 mL HisTrap™ HP Ni^{2+} affinity column equilibrated with Nickel affinity lysis buffer. The desired protein was eluted using a linear gradient from lysis buffer to Ni2+ affinity elution buffer (same as lysis buffer but with 500mM imidazole). The fractions containing the desired protein were determined from the UV trace of the FPLC report file and from SDS-PAGE analysis, and collected.

Size Exclusion Chromatography and Storage

The collected fractions were concentrated by Vivaspin 20 (MWCO 10 000) to a final volume of ~1.5 mL. The concentrated protein solution was loaded onto a HiLoadTM 16/60 Superdex 75 gel filtration column and filtered (180 mL, 1.5 CV, 1.2mL fractions) into MEK1 storage buffer (50 mM HEPES, 150mM NaCl, 5% glycerol, 0.5mM TCEP, pH 7.5 buffer). The fractions containing the desired protein were determined from the UV trace of the FPLC report file and their purity was determined from SDS-PAGE analysis. The pure and impure fractions of the desired protein were collected separately, re-concentrated by Vivaspin (MWCO 10 000, 3.2 mg/mL final concentration), divided into aliquots, flash frozen in liquid nitrogen and stored at -80 °C. Yields were 2-5 mg/L.

Circular dichroism spectra of full length MEK1 variants can be found in appendix C.

Table 7.1. List of expressed and purified MEK1 variants

Entry	Protein	Referred as	$M_{calc.}$/Da	$M_{obs.}$=/Da
1	MEK1G(-19)F/S222C/C277S/C376S	MEK1C222	45676	45678
2	MEK1G(-19)F/S218C/C277S/C376S	MEK1C218	45676	45671
3	MEK1G(-19)F/S218C/S222C/C277S/C376S	MEK1C218C222	45692	45687
4	MEK1G(-19)F	MEK1wt*	45692	45690
5	MEK1G(-19)F/C277S/C376S	MEK1wt	45660	45657
6	MEK1G(-19)F/E218	MEK1E218	45735	45735
7	MEK1G(-19)F/E222	MEK1E222	45735	45736
8	Δ60MEK1 S222C/C277S/C376S	Δ60MEK1C222	36645	36641
9	Δ60MEK1 S218C/C277S/C376S	Δ60MEK1C222	36645	36641
10	Δ60MEK1 S218C/S222C/C277S/C376S	Δ60MEK1C222	36661	36656

Determination of melting temperatures and thermal shift of MEK1C222 with or without nucleotides and magnesium[257, 258]

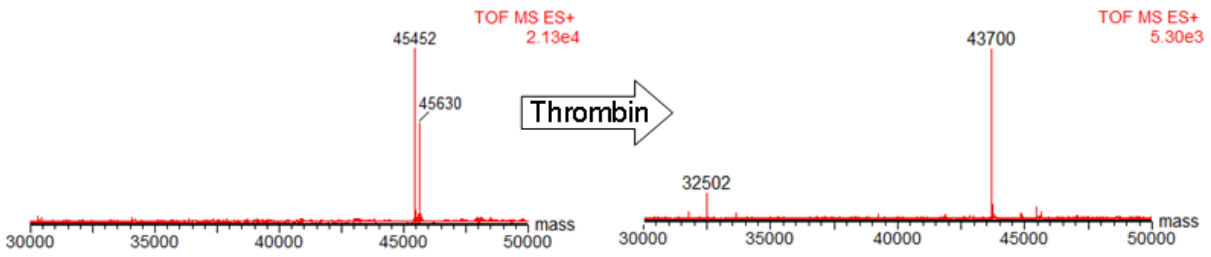
Differential Scanning Fluorimetry assays (DSF) were performed in white 8 x 6 PCR plates using a MiniOpticon™ Real-Time PCR Detection System (*BioRad*). Reactions were performed in 50 µl. Protein unfolding was monitored by measuring the fluorescence of the SYPRO® Orange dye (*Invitrogen*). Temperature was increased from 25 to 85 °C, 1 °C increment every 30 seconds. The dye stock (5000x concentrate) was first diluted by adding 1 µl to 2.5 ml of 20 mM Tris, 100 mM NaCl pH 8 metal free buffer to produce a 2x concentrated solution. Nucleotides were tested at 1 mM final concentration, and magnesium chloride at 2.5 mM final concentration. The protein was diluted in SYPRO Orange buffer to obtain a final assay concentration of 2 µM. Measurements were performed in triplicate. For determination of T_m values, melting curves for each triplicate data set were analyzed using GraphPad Prism.

Table 7.2. Melting temperatures and thermal shifts of MEK1C222 with or without nucleotides and magnesium estimated by DSF assay

Protein	Entry	Nucleotide	Additive	Average Tm±S.D. in °C	Average ΔTm±S.D. (Entry-Entry 0) in °C
MEK1C222	0	None	None	47.84±0.19	0
	1	None	$MgCl_2$	47.12±0.14	-0.72±0.18
	2	ATP	None	48.53±0.20	0.69±0.25
	3	ATP	$MgCl_2$	52.94±0.39	5.10±0.48
	4	ATPγS	None	48.92±0.34	1.08±0.42
	5	ATPγS	$MgCl_2$	53.38±0.14	5.54±0.17
	6	ADP	None	47.84±0.12	-0.01±0.15
	7	ADP	$MgCl_2$	52.79±0.10	4.95±0.13

Gluconoylation and thrombin miscleavage

MEK1S222C/C277S/C376S ($M_{calc.}$=45455, $M_{obs.}$=45452) expressed from the original genscript plasmid gluconoylated ($M_{calc.}$=45633, $M_{obs.}$=45630).[259] This modification could interfere with the crystallisation of MEK1.[260] The variant was designed with a thrombin cleavable site LVPRGS at R (Appendix B). However, the cleavage of the gluconoylated protein was not selective. The desired cleavage ($M_{calc.}$=43704, $M_{obs.}$=43700) was slow and accompanied by further cleavage at position R291 ($M_{calc.}$=32506, $M_{obs.}$=32502). Different amounts of thombin were tested but did not reduce the miscleavage.



The gluconoylation was prevented by G(-19)F mutations.

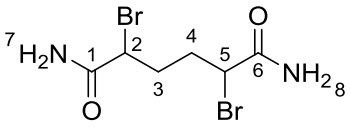
7.5. Synthesis of small molecules as tools in the site-selective chemical phosphorylation of protein

7.5.1. General chemical considerations

Chemicals and solvents were purchased from Sigma-Aldrich UK, Acros UK, Alfa Aesar UK, or Fischer UK and were used as delivered unless stated otherwise. Thinlayer chromatography (TLC) was carried out using plastic 0.20 mm Polygram® Silg/UV254 plates that were dried using a heat gun and visualised under UV (λmax 254 nm or 366 nm) or by use of potassium permanganate, iodine dip. Flash column chromatography was performed using Geduran® Si 60.8.2 or a Biotage Flash Purification System with a Kinesis Telos or Biotage Snap columns.

7.5.2. Synthesis of bisalkylating/elimination agents

Synthesis of 2,5-dibromohexanediamide, DBHDA[117]



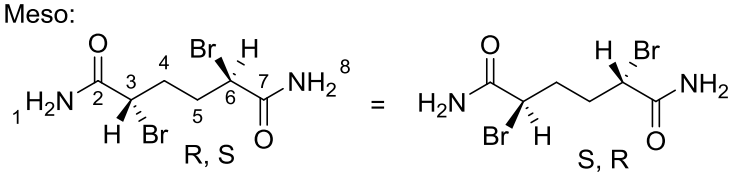
Adipic acid (25 g, 171.1 mmol) was suspended in thionylchloride (75 mL, 1034 mmol). The reaction was heated to reflux (open to air, bath temp 80 ºC). After 30 minutes at reflux, the solid had dissolved. The reaction was stirred for an additional 60 minutes at reflux and then cooled to room temperature. Carbontetrachloride (100 mL) was added to the reaction followed by N-bromosuccinimide (73 g, 411 mmol). The reaction was stirred vigorously and 10 drops of HBr (48% aq.) was added by pipette. The reaction was heated to reflux, again open to air. The reaction gradually turns from red to black over the course of an hour. After

2 hours at reflux, the reaction was cooled to room temperature and then to 0 ºC. The mixture was stirred at 0 ºC to ensure all succinimide had precipitated. The solid was removed by filtration. Diethylether (50 mL) was used to rinse and complete the filtration. The filtrate was concentrated under reduced pressure to give a thick, dark red liquid. In a 500 mL round bottom flask, 200 mL of NH4OH (25% aq.) was cooled to 0 ºC. The crude acid chloride was added dropwise over 20 minutes to the ammonia solution with rapid stirring. After the addition was complete, the reaction was stirred vigorously at 0 ºC for 1 hour. The bis-amide product precipitated from the reaction mixture. The dark solid was isolated by filtration and partially dried. The product was purified by triturating in MeOH/H2O: The solid was suspended in H2O (100 mL) and MeOH (100 mL) and heated to 60 ºC. The mixture was stirred rapidly at 60 ºC for 30 minutes. After this time, the mixture was cooled to room temperature. The resulting white solid (a mixture of meso and d/l diastereomers) was isolated by filtration and washed with MeOH (200 mL). The title compound was dried under high vacuum to afford 26 g (50% yields) as an off-white solid.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 1.75-2.08 (4H, m, H3, H4), 4.28-4.36 (2H, m, H2, H5), 7.30 (2H, s, H7, H8), 7.69 (2H, s, H7, H8). ^{13}C NMR (100 MHz, DMSO): (both diastereomers reported), δ ppm 32.5 (s, C3, C4), 32.6 (s, C3, C4), 48.2 (s, C2, C5), 48.5 (s, C2, C5), 169.87 (s, C1), 169.92 (s, C5). IR (umax, KBr): 3302, 2946, 2801, 1684, 1418, 1319, 1277, 1250, 1220, 1194, 978, 876. HRMS m/z (ESI+): Found 300.9178 [M+H]+; $C_8H_{11}O_2N_2{}^{79}Br_2$ requires 300.9182.

Isolation of meso-2,5-dibromohexanediamide, meso-DBHDA

Meso:



Method A: X-Ray analysis

2,5-dibromohexanediamide (DBHDA, 50 mg, 0.16 mmol) were dissolved in boiling methanol (10 mL). The solution was filtered through a cottonwool plug. The filtrate was left standing at room temperature for one week to afford very fine crystals. One was analysed by X-Ray diffraction. Analysis showed that the crystal was the meso-R, S stereoisomer.[‡‡‡]

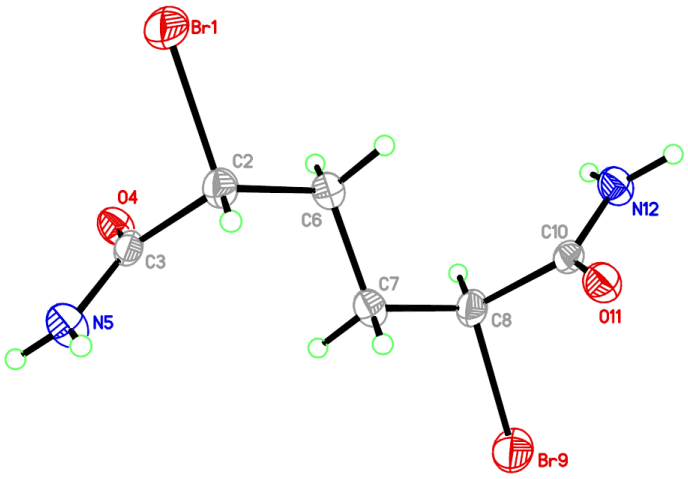


Figure 7.2. Ellipsoid representation of a crystal of meso-2,5-dibromohexanediamide

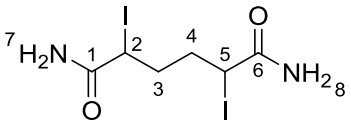
Method B: ^{1}H, ^{13}C NMR analysis

2,5-dibromohexanediamide (100 mg, 0.33 mmol) were dissolved in a minimum of boiling methanol. The solution was filtered through a cottonwool plug. The filtrate was left standing at room temperature for 24 hours to afford fine crystals. The mixture was filtered to afford 35 mg of title compound as beige crystals.

^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 1.72 - 1.91 (m, 2 H, H4, H5), 1.95 - 2.11 (m, 2 H, H4, H5), 4.20 - 4.40 (m, 2 H, H3, H6), 7.32 (br. s., 2 H, H1, H8), 7.70 (br. s., 2 H, H1, H8). ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 32.58 (s, C4, C5), 48.23 (s, C3, C6), 169.80 (s, C2, C7).

‡‡‡ Crystal structure analysis was performed by Dr. Amber L Thompson, Chemical Crystallography Service, Chemistry Department, Oxford.

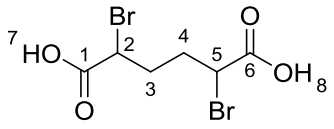
Synthesis of 2,5-diiodohexanediamide, DIHDA



A suspension of 2,5-dibromohexanediamide (DBHDA, 115 mg, 0.38 mmol) in dry acetone (15 mL) was heated under reflux. Dry DMF was added (1 mL) to give a solution before a solution of sodium iodide (1.2 g, 8.0 mmol) in acetone (5 mL) was added. The mixture was heated under reflux with stirring for 4 h, after which a white precipitate had formed and the reaction mixture had turned yellow. The mixture was filtered while hot and the solvent removed under vacuum to give a yellow, solid residue which was triturated with water (20 mL). The remaining precipitate was filtered, washed with water (10 mL) and dried under high vacuum to afford the title compound (55 mg, 36%) as a yellow powder.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 1.64–1.96 (4H, m, H3, H4), 4.25–4.35 (2 H, m, H2, H5), 7.11 (2 H, br s, H7, H8), 7.63 (2 H, br s,H7, H8),; ^{13}C NMR (101 MHz, DMSO-d6) δ ppm (both diastereomers reported) 24.6 (C3,C4), 25.4 (C3,C4), 36.6 (C2, C5), 172.5 (C1,C6); HRMS m/z (ESI+): Found 418.87286 [M+Na]+; $C_6H_{10}{}^{127}I_2N_2{}^{23}O_2$ requires 418.8724.

Synthesis of 2,5-dibromohexanedioic acid, DBHDAc[261]



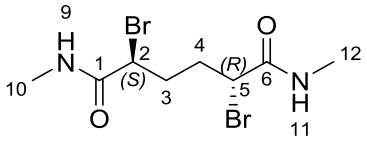
Adipic acid (2.0 g, 13.7 mmol) and thionyl chloride (2.98 mL, 41.1 mmol) were stirred at reflux for 1 hour. The mixture was cooled down to room temperature and treated with *N*-bromosuccinimide (7.3 g, 41.1 mmol). The mixture was stirred at 75 °C for 3 hours. The mixture was cooled down to 0 °C before addition of a 1:1 mixture of water/tetrahydrofuran (20 mL). The mixture was stirred vigorously for 2 hours. The mixture was filtered. The

residue was stirred in dichloromethane for 1 hour. Filtration afforded 1.3 g (32% yield) of the title compound (mixture of racemate and meso diastereoisomer) as an off-white solid.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 1.83 - 2.22 (m, 4 H, H3, H4), 4.45 - 4.60 (m, 2 H, H2, H5), 13.32 (br. s., 2 H, H7, H8). ^{13}C NMR (101 MHz, DMSO-d6) δ ppm 32.05 (s, C3, C4), 32.17 (s, C3, C4), 46.74 (s, C2, C5), 46.88 (s, C2, C5), 170.44 (s, C1, C6). HRMS m/z (ESI+): Found 300.8718 [M-H]-; $C_6H_7O_4{}^{79}Br_2$ requires 300.8717.

The compound (50 mg, 0.016 mmol) was stirred in water (1 mL) for 16 hours. The solid was filtered to afford meso-2,5-dibromohexanedioic acid.

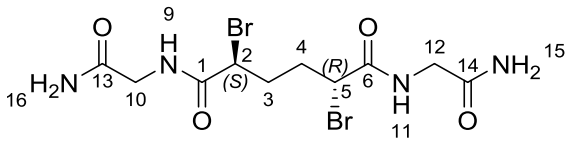
Synthesis of 2,5-dibromo-*N,N'*-dimethylhexanediamide, DBDMet



2,5-dibromohexanedioic acid (100 mg, 0.33 mmol) and thionyl chloride (239 µL, 3.3 mmol) was stirred at reflux for 2 h. The solution was cooled down to room temperature before evaporation under reduced pressure. The residue was then added to a stirred 40% wt aqueous solution of methyl amine (1 mL, 13.6 mmol) at 0 °C. The mixture was stirred vigorously for 2 h. The precipitate was filtered and stirred in a minimum of dichloromethane. Filtration afforded 35 mg (32% yield) of title compound as a white solid.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 1.67 - 2.13 (m, 4 H, H3, H4), 2.61 (d, J=4.58 Hz, 6 H, H10, H12), 4.23 - 4.45 (m, 2 H, H2, H5), 8.13 - 8.36 (m, 2 H, H9, H11). ^{13}C NMR (101 MHz, DMSO-d6) δ ppm 25.87 (s, C10, C12), 32.67 (s, C3, C4), 48.24 (s, C2, C5), 168.31 (s, C1, C6). LRMS m/z (ESI+): 354 [M+Na]+

Synthesis of meso-*N,N'*-bis(2-amino-2-oxoethyl)-2,5-dibromohexanediamide, DBDGla

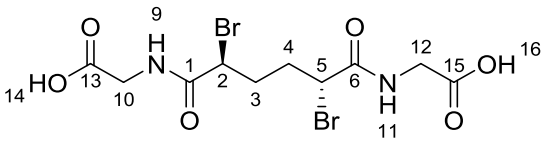


2,5-dibromohexanedioic acid (100 mg, 0.33 mmol) and thionyl chloride (100 µL, 1.38 mmol) was stirred at reflux for 1.5 h. The solution was cooled down to room temperature before evaporation under reduced pressure. The residue was then diluted in diethyl ether (1 mL). The solution was added dropwise to a stirred solution of glycinamide hydrochloride (72 mg, 0.66 mmol) in a mixture of diethyl ether (500 µL) and 2 M aqueous sodium hydroxide solution (660 µL) at 0 °C. The mixture was stirred for 1 hour at room temperature. The organic layer was separated. The aqueous was acidified to pH~2 with addition of a 1 N hydrochloric acid solution. The mixture was washed with ethyl acetate (2 mL). The aqueous layer (suspension) was separated, evaporated under reduced pressure and azeotroped twice with toluene. The residue was stirred in acetonitrile for 1 hour at room temperature. The precipitate was filtered to afford 130 mg of crude material. 40 mg of this material were stirred in water for 16 hours. Filtration afforded 7 mg of the title compound as a white solid.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 1.73 - 1.92 (m, 2 H, H3, H4), 2.02 - 2.19 (m, 2 H, H3, H4), 3.55 - 3.83 (m, 4 H, H10, H12), 4.43 - 4.63 (m, 2 H, H2, H5), 7.09 (br. s., 2 H, H14, H16), 7.39 (br. s., 2 H, H14, H16), 8.44 (t, J=5.62 Hz, 2 H, H9, H11). ^{13}C NMR (101 MHz, DMSO-d6) δ ppm 32.65 (s, C3, C4), 42.04 (s, C10, C12), 48.01 (s, C2, C3), 168.27 (s, C13, C14), 170.24 (s, C1, C6). HRMS m/z (ESI+): Found 412.9467 [M-H]-; $C_{10}H_{15}O_4N_4{}^{79}Br_2$ requires 412.9466.

Synthesis of

({(meso)-2,5-dibromo-6-[(carboxymethyl)amino]-6-oxohexanoyl}amino)acetic acid, DBDGly

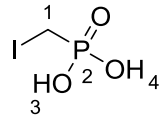


2,5-dibromohexanedioic acid (100 mg, 0.33 mmol) and thionyl chloride (100 µL, 1.38 mmol) was stirred at reflux for 1.5 h. The solution was cooled down to room temperature before evaporation under reduced pressure. The residue was then diluted in diethyl ether (1 mL). The solution was added dropwise to a stirred solution of glycine (50 mg, 0.66 mmol) in a mixture of diethyl ether (500 µL) and 2 M aqueous sodium hydroxide solution (660 µL) at 0 °C. The mixture was stirred for 1 hour at room temperature. The organic layer was separated. The aqueous was acidified to pH~2 with addition of a 1 N hydrochloric acid solution. The mixture was washed with ethyl acetate (2 mL). The aqueous layer (suspension) was separated, evaporated under reduced pressure and azeotroped twice with toluene. The residue was stirred in acetonitrile for 1 hour at room temperature. The precipitate was filtered to afford 100 mg of crude material. 40 mg of this material were stirred in water for 16 hours. Filtration afforded 4 mg of the title compound as a white solid. The solid was highly soluble in water.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 1.77 - 1.93 (m, 2 H, H3, H4), 2.01 - 2.15 (m, 2 H, H3, H4), 3.80 (dd, J=5.87, 3.42 Hz, 4 H, H10, H12), 4.38 - 4.59 (m, 2 H, H2, H5), 8.60 (t, J=5.75 Hz, 2 H, H9, H11), 12.66 (br. s, 2 H, H14, H16). ^{13}C NMR (101 MHz, DMSO-d6) δ ppm 32.69 (s, C3, C4), 40.95 (s, C10, C12), 47.56 (s, C2, C5), 168.44 (s, C13, C15), 170.78 (s, C1, C6).

7.5.3. Synthesis of alkylating phosphorylated agents

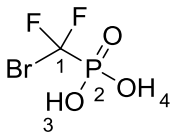
Synthesis of (iodomethyl)phosphonic acid,[262] IMPA



Diethyl (iodomethyl)phosphonate (1 g, 3.60 mmol) was stirred at reflux in a mixture of water (2.12 mL) and concentrated hydrochloric acid solution (4.25 mL) for 14 hours. The reaction mixture was cooled down to room temperature. The solution was evaporated under reduced pressure. The residue was azeotroped 3 times with toluene (10 mL). The resulting oil was flash-frozen with liquid nitrogen and freeze-dried for 12 hours to afford 750 mg (95% yield) of the title compound as a white solid.

^{1}H NMR (400 MHz, DMSO-d6) δ ppm 2.97 (d, J=10.03 Hz, 2 H, H1), 7.30 (br. s, 2 H, H3, H4); ^{13}C NMR (101 MHz, DMSO-d6) δ ppm -8.11 (d, J=145.44 Hz, C1); ^{31}P NMR (162 MHz, DMSO-d6) δ ppm 15.75 (s, 1 P, P2); HRMS m/z (ESI-): Found 220.88676 [M-H]-; $CH_3O_3I^{127}$ IP requires 220.88700 [M-H]-.

Synthesis of (bromodifluoromethyl)phosphonic acid, BFPA[263]



Diethyl (bromodifluoromethyl)phosphonate (1 g, 3.75 mmol) was treated with concentrated hydrochloric acid solution (10 mL) then heated to reflux for 18 hours. The reaction mixture was cooled to room temperature. The solution was evaporated under reduced pressure. The residue was dried azeotropically with toluene (3 x 10 mL) and dried extensively under high vacuum, to afford 720 mg (91% yield) of the title product as a colourless oil.

^{1}H NMR (400 MHz, acetone-d6) δ ppm 4.4 (s, 2 H, H1); ^{19}F NMR (376.5 MHz, acetone-d6) δ ppm -62.7 (d, J=88 Hz, F1); ^{31}P NMR (162 MHz, acetone-d6) δ ppm -2.3 (t, J=88 Hz, 1 P, P2); HRMS m/z (ESI+): Found 210.8967 [M+H]+; $CH_2{}^{19}F_2O_3{}^{79}BrP$ requires 210.8966 [M-H]+.

7.6. Chemical protein modifications towards site-selective chemical phosphorylation of MEK1

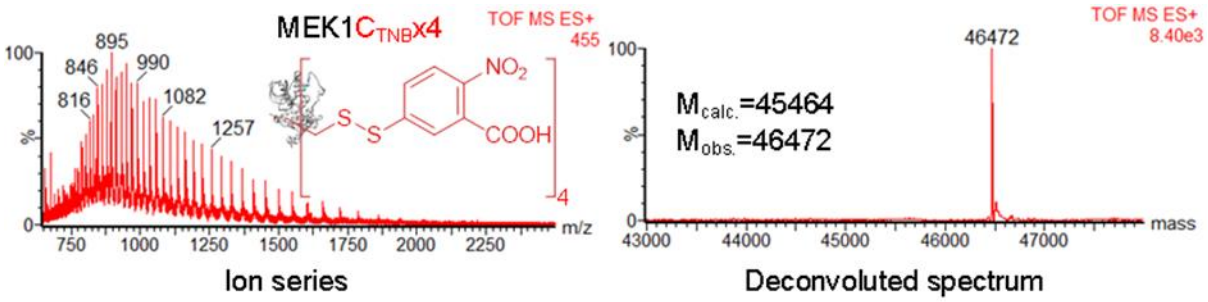
7.6.1. General protein chemistry and reagents considerations

Desalting and buffer exchange was performed using PD-SpinTrap G25, PD-MiniTrap™ G25 provided by GE Healthcare, Micro Bio-Spin™ 6 Columns (Bio-Rad), Vivaspin 500, 6, 50 Protein Concentrators MWCO 10000, 5000 provided by GE Heathcare and Slide-A-lyser® MWCO 10 000 dialysis cassettes (Thermo Scientific).

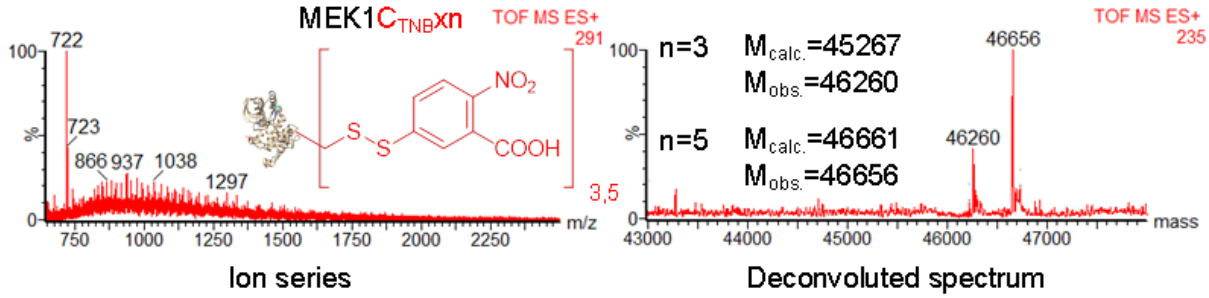
Reagents and buffer components were provided by Sigma Aldrich, Fluka and Apollo.

7.6.2. Treatment of MEKC222 with Ellman's reagent

MEK1C222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 0.14 M stock solution of 5,5'-dithiobis-(2-nitrobenzoic acid), DTNB (Ellman's reagent) in DMSO to afford a final concentration of 1.1 mM of DTNB. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF (LCT-classic-Chromolyth). Full conversion to MEK1 containing 4 TNBD adducts (MEK1[CTNB]x4) was observed after 30 minutes.



Urea was added to the mixture to afford a 20 mM Tris, 100 mM NaCl, 2 M Urea, pH 8 buffer. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF (LCT-classic-Chromolyth). A mixture of MEK1 containing 3 and 5 TNBD adducts (MEK1[CTNB]x3) and MEK1 (MEK1[CTNB]x5) was observed after 2 hours.

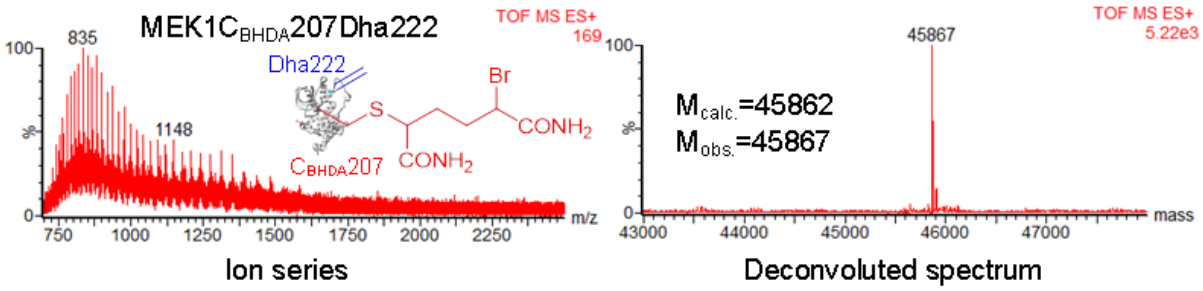


7.6.3. Experiments towards the site-selective formation of MEK1Dha222

Formation of MEK1C$_{BHDA}$207Dha222 with DBHDA

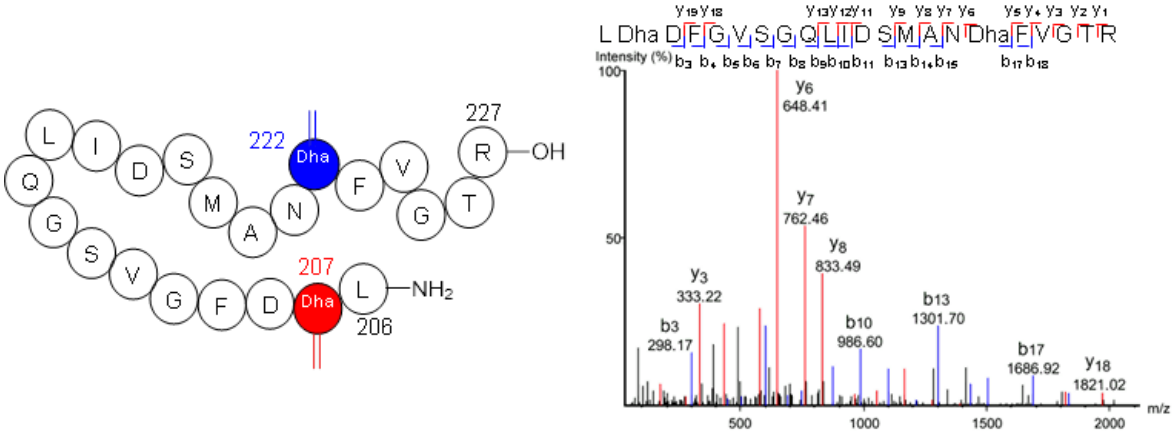
MEK1C222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 0.26 M stock solution of 1,5-dibromohexanediamide, DBHDA in DMSO to afford a final concentration of 4.4 mM of DBHDA.[§§§] The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to overalkylated MEK1Dha222 (MEK1C$_{BHDA}$207Dha222) was observed after 4 hours.

[§§§] The relative solubility of DBHDA in water is around 3.5 mM. Stock solution of DBHDA in DMSO were made to improve the solubility. However, the final amount of DMSO could not exceed 5% before precipitation of MEK1C222 was observed.



The sample was digested with trypsin[****] using the in-solution method (Paragraph 8.2.2.). The digested peptides were analysed using the Synapt method and processed by Peaks 7.0. LCMSS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at the two positions.

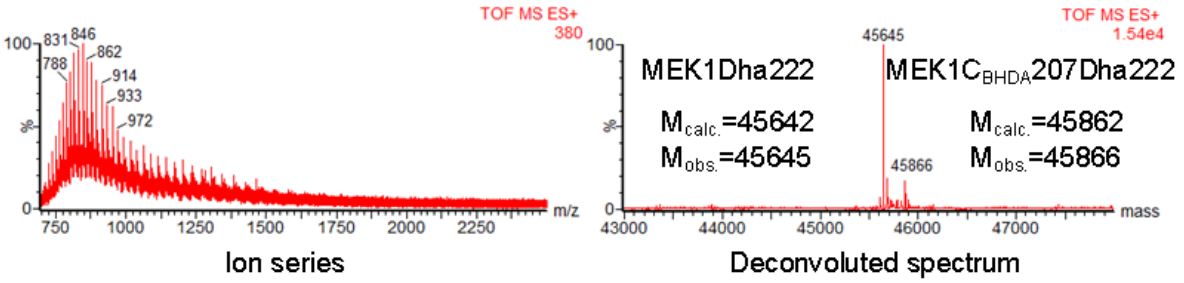
Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	44.77	755.7722 (3)	90.6	Dha222, Dha207	90.6	5E3	4E1



Variation of the concentration of MEK1C222 (11 and 44 µM), DBHDA (1.1, 2.2, 11.0 mM), temperature (4 °C, room temperature and 30 °C) and buffer (20 mM Tris, 100 mM NaCl, 25 mM $MgCl_2$, pH 8) did not change the overall overalkylation pattern.

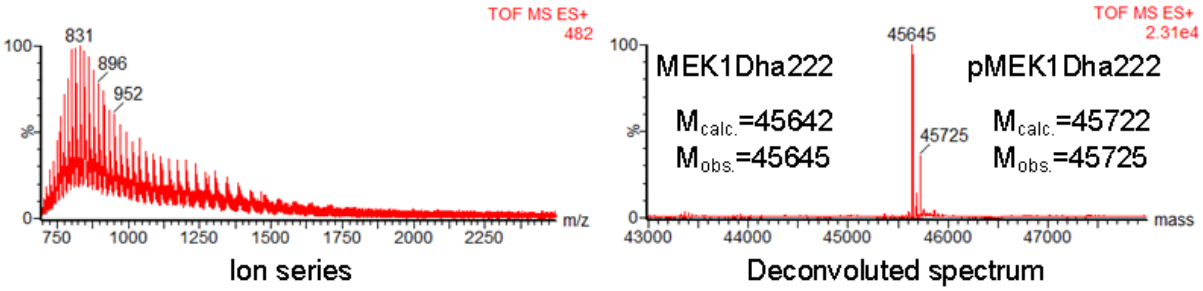
[****] Coverage of the digest with trypsin was typically 70% for MEK1 constructs. AspN was needed for 100% coverage.

However, the same reaction was performed using 20 mM Tris, 100 mM NaCl, 10 mM ATP, pH 8 buffer, a mixture of MEK1Dha222 and overalkylated MEK1Dha222 (MEK1C_{BHDA}207Dha222) was observed after 4 hours by LCMS-ESI-TOF.

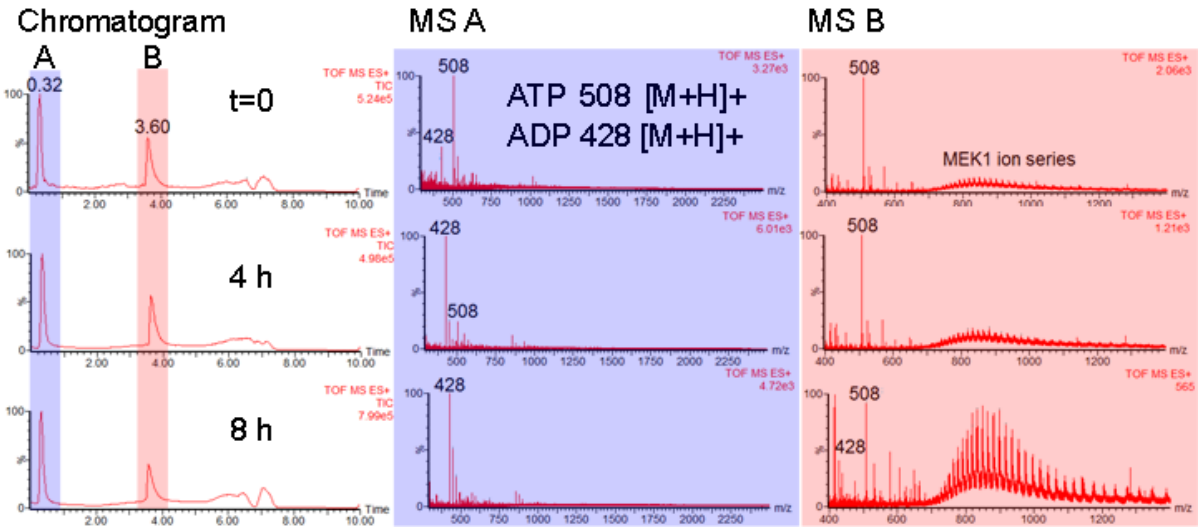


Formation of MEK1Dha222 with DBHDA in the presence of ATP-Mg

MEK1C222 (33 µM, 1.5 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ATP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.49 M stock solution of 1,5-dibromohexanediamide, DBHDA in DMSO to afford a final concentration of 4.9 mM of DBHDA. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Dha222 as a mixture 75% MEK1Dha222 ($M_{obs.}$=45645, $M_{calc.}$=45642) and 25% phosphorylated pMEK1Dha222 ($M_{obs.}$=45862, $M_{calc.}$=45866) was observed after 7 hours.



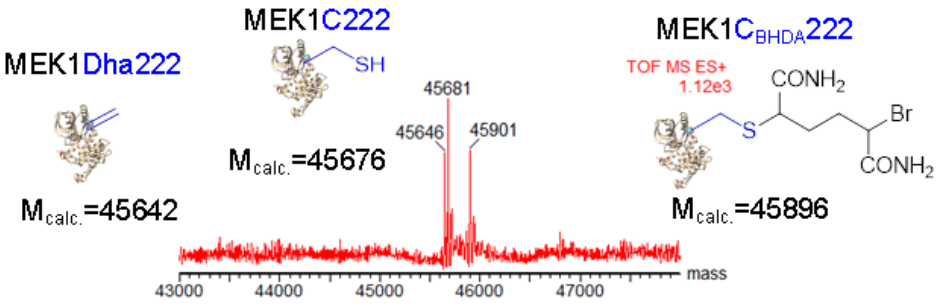
ATP hydrolysis was also monitored by LCMS-ESI-TOF (LCT-premier).



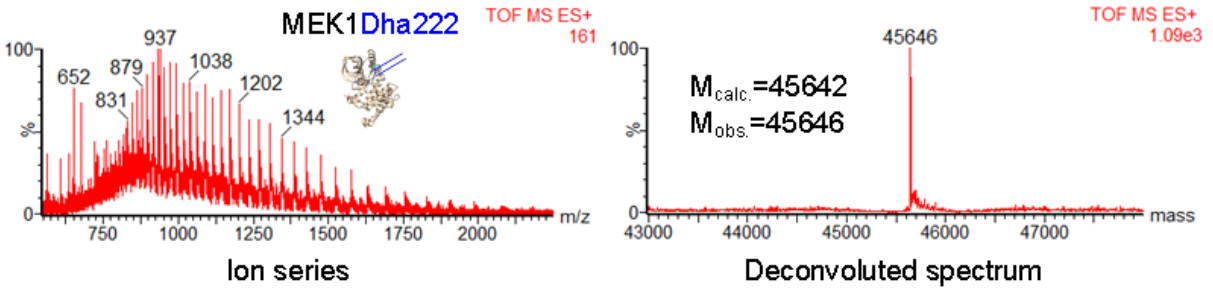
Control experiment using shaking 20 mM Tris, 100 mM NaCl, 10 mM ATP, 25 mM $MgCl_2$, pH 8 at 600 rpm at 37 °C for 4 h showed no disappearance of ATP and formation of ADP by LCMS.

Formation of MEK1Dha222 with DBHDA in the presence of ADP-Mg

MEK1C222 (33 µM, 1.5 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.66 M stock solution of 1,5-dibromohexanediamide, DBHDA in DMSO to afford a final concentration of 6.6 mM of DBHDA. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. LCMS after 1 hour showed a mixture MEK1Dha222 ($M_{obs.}$=45646, $M_{calc.}$=45642), MEK1C222 ($M_{obs.}$=45681, $M_{calc.}$=45676) and MEK1C_{BHDA} ($M_{obs.}$=45901), $M_{calc.}$=45896).

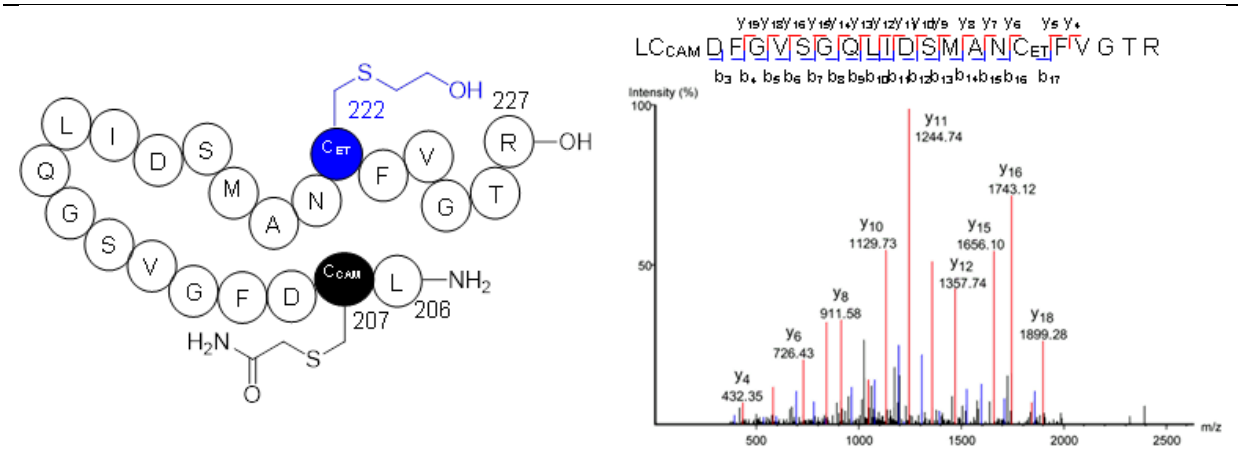


The formation of sulfonium intermediate MEK1C_{THTCD}222 ($M_{calc.}$=45707) was not observed. Full conversion to MEK1Dha222 ($M_{obs.}$=45646, $M_{calc.}$=45642) was observed after 5 hours.

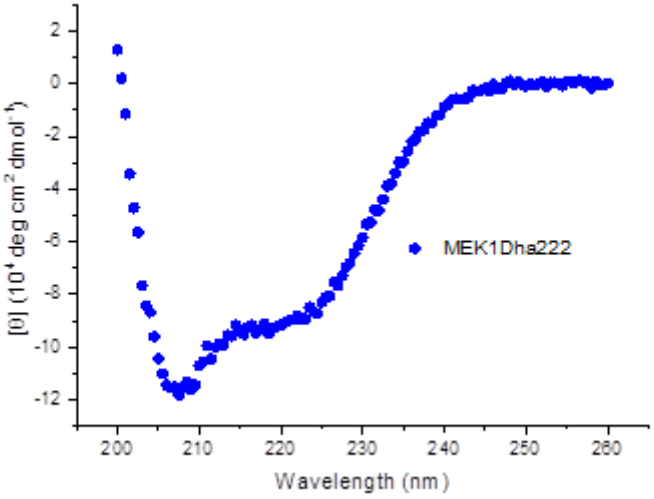


One aliquot of the reaction mixture was digested with trypsin using in-gel protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation.

Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	68.872	1217.5563 (2)	-6.0	C_{ET}222, C_{CAM}207	109.76	4.22 E-4	2.15 E3

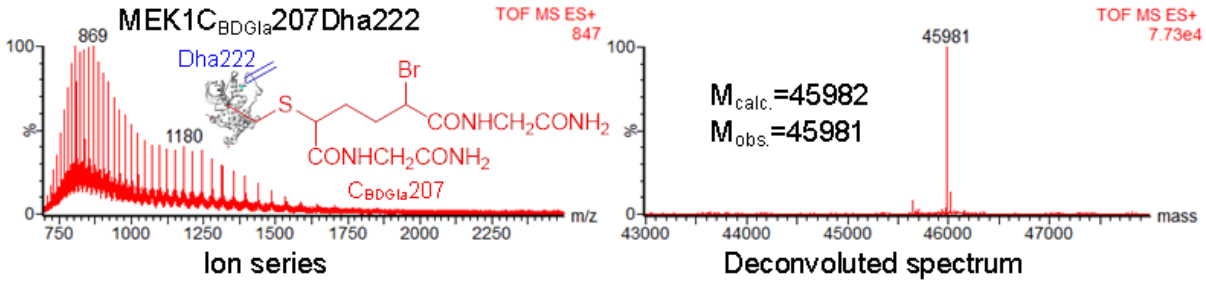


On aliquot was desalted twice with PD-G25 columns. Circular dichroism analysis was done and processed according to the procedure (Paragraph 7.2.3).



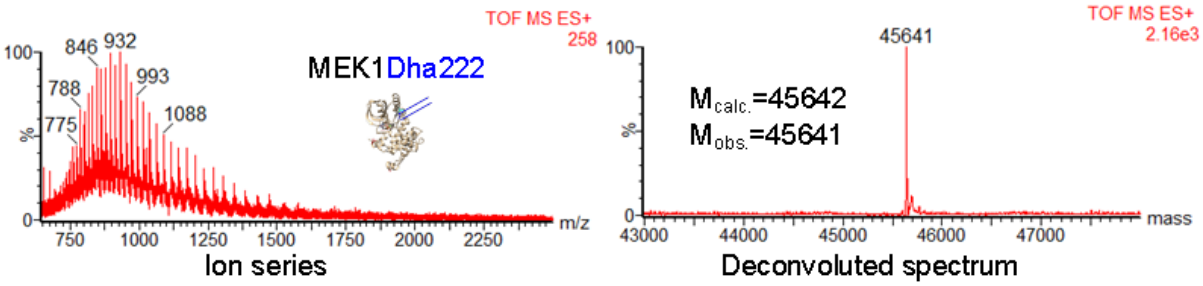
Formation of MEK1Dha222 with DBDGla

MEK1C222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 0.26 M stock solution of meso-*N,N'*-bis(2-amino-2-oxoethyl)-2,5-dibromohexanediamide, DBDGla in DMSO to afford a final concentration of 4.4 mM of DBDGla. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to overalkylated MEK1C_{BDGla}207Dha222 ($M_{obs.}$=45981, $M_{calc.}$=45982) was observed after 3 hours.



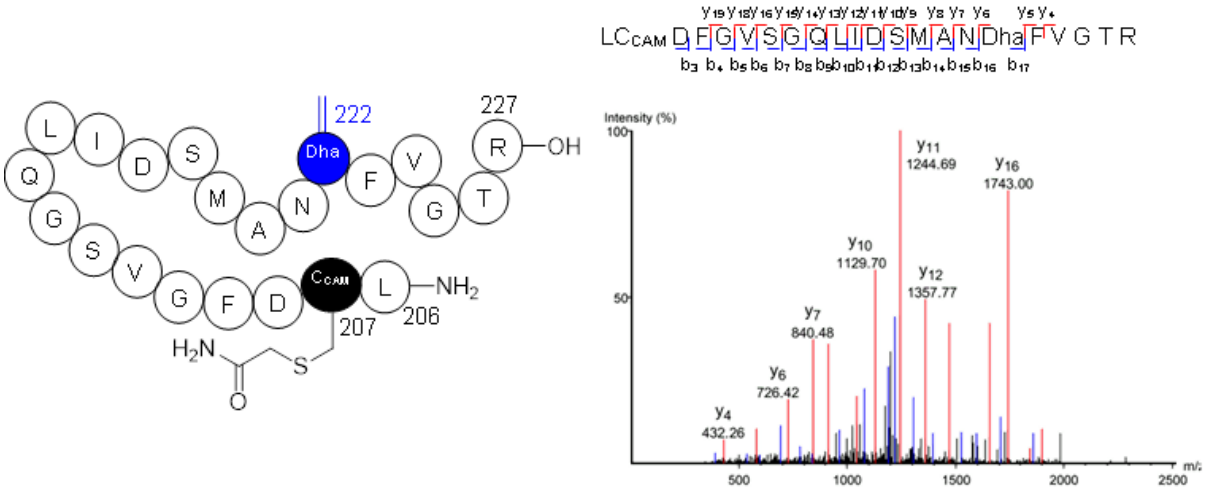
MEK1C222 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.44 M stock solution of meso-*N,N'*-bis(2-amino-2-oxoethyl)-2,5-dibromohexanediamide, DBDGla in DMSO to afford a final concentration of 8.8 mM of DBDGla. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF.

Full conversion to MEK1Dha222 ($M_{obs.}$=45981, $M_{calc.}$=45982) was observed after 4 hours.



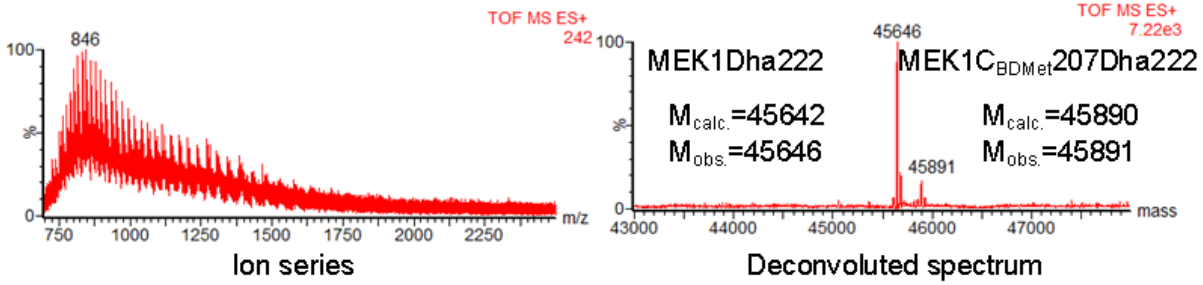
One aliquot of the reaction mixture was digested with trypsin using in-gel protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at position 222.

Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	43.80	1217.5568 (2)	-5.6	C_{ET}222, C_{CAM}207	155.94	3E4	1.3E3

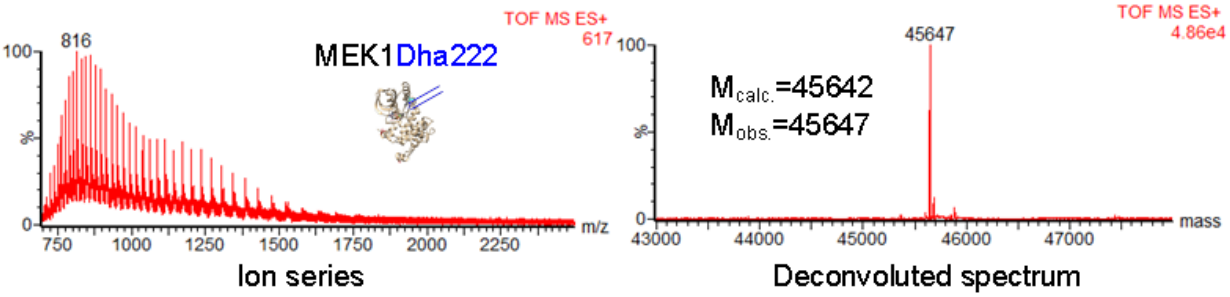


Formation of MEK1Dha222 with DBDMet

MEK1C222 (33 µM, 1.5 mg/mL) in 20 mM Tris, 100 mM NaCl pH 8 buffer (70 µL) was treated with a 0.5 M stock solution of 2,5-dibromo-*N*,*N'*-dimethylhexanediamide, DBDMet to afford a final concentration of 6.6 mM of DBDMet. The mixture was shaken at 600 rpm at 37 °C for 4 hours. The mixture was allowed to stand for 20 hours at room temperature. The mixture was then shaken at 37 °C for 3 hours. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Dha222 as a mixture 90% MEK1Dha222 ($M_{obs.}$=45646, $M_{calc.}$=45642) and 10% MEK1C_{BDMet}207Dha222 (($M_{obs.}$=45891, $M_{calc.}$=45889) was observed.



The same conditions in 20 mM Tris, 100 mM NaCl,10mM ADP, 35 mM $MgCl_2$ pH 8 buffer afforded excusively MEK1Dha222 ($M_{obs.}$=45647, $M_{calc.}$=45642).

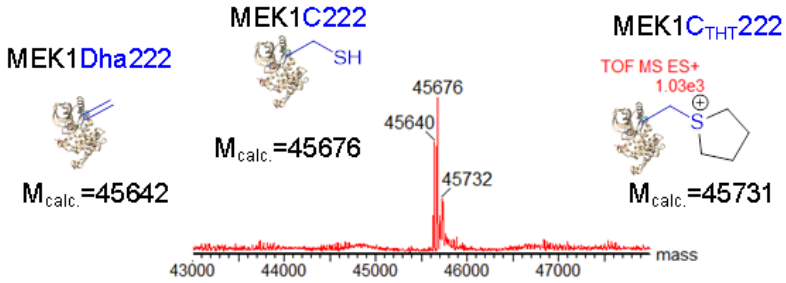


Formation of MEK1Dha222 with DIB in the presence of ADP-Mg

MEK1C222 (44 µM, 2.2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with 1,4-diiodobutane, DIB†††† (2000 Eq. added as such). The mixture was shaken at 1000 rpm at 37 °C. The reaction was monitored using LCMS-ESI-

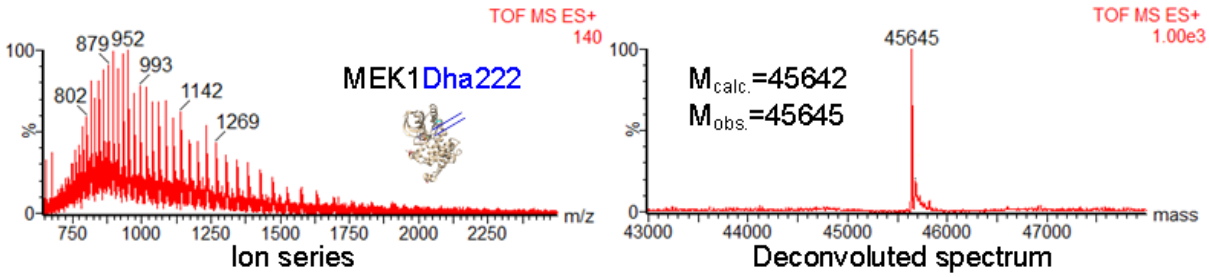
†††† DIB is a liquid not miscible in aqueous buffer. It will form a drop in the bottom of the reaction mixture. Efficient shaking is essential for the reaction to occur.

TOF. LCMS after 6 hour showed a mixture MEK1Dha222 ($M_{obs.}$=45640, $M_{calc.}$=45642), MEK1C222 ($M_{obs.}$=45676, $M_{calc.}$=45676) and the sulfonium intermediate MEK1C_{THT} ($M_{obs.}$=45732, $M_{calc.}$=45731).



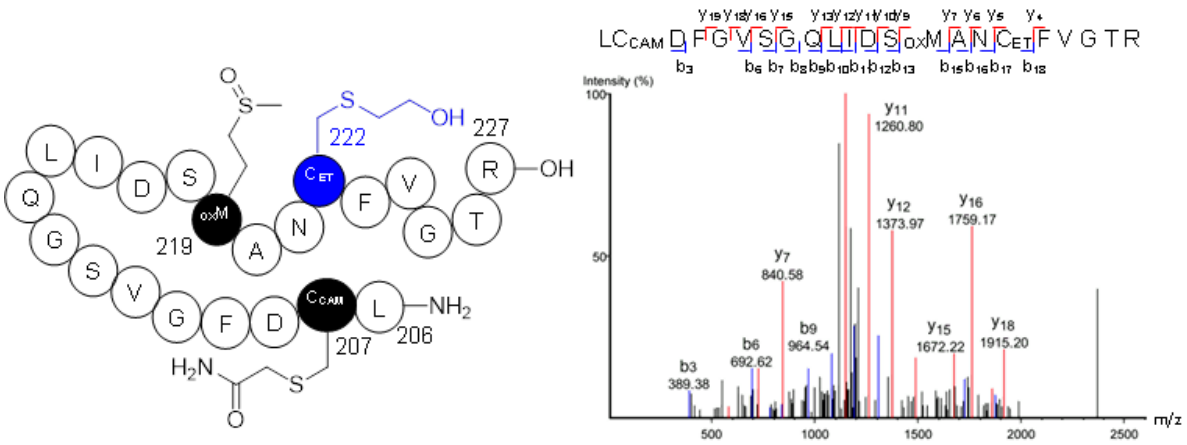
The formation of the first alkylation intermediate MEK1C_{IB}222 ($M_{calc.}$=45858) was not observed.

Full conversion to MEK1Dha222 was observed after 20 hours.



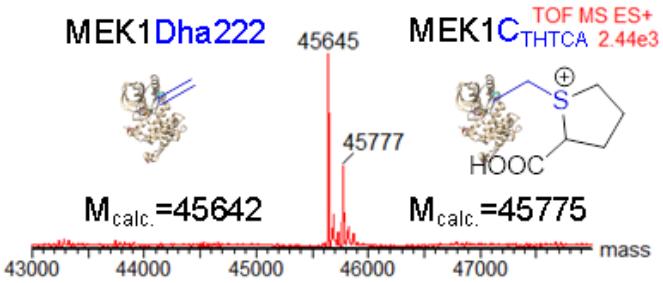
One aliquot of the reaction mixture was digested with trypsin using in-gel protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at position 222.

Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	37.74	1225.5529 (2)	-6.7	C_{ET}222, oxM219, C_{CAM}207	98.92	1.21E4	7.74E2



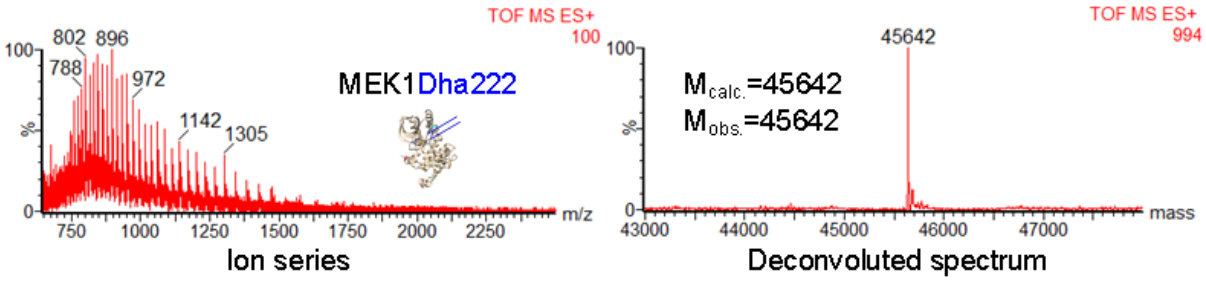
Formation of MEK1Dha222 with MDBP

MEK1C222 (44 µM, 2 mg/mL) in 20 Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.64 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.4 mM of MDBP. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. LCMS after 1 hour showed a mixture MEK1Dha222 ($M_{obs.}$=45645, $M_{calc.}$=45642), and the hydrolysed sulfonium intermediate MEK1C_{THTCA}222 ($M_{obs.}$=45777, $M_{calc.}$=45775).



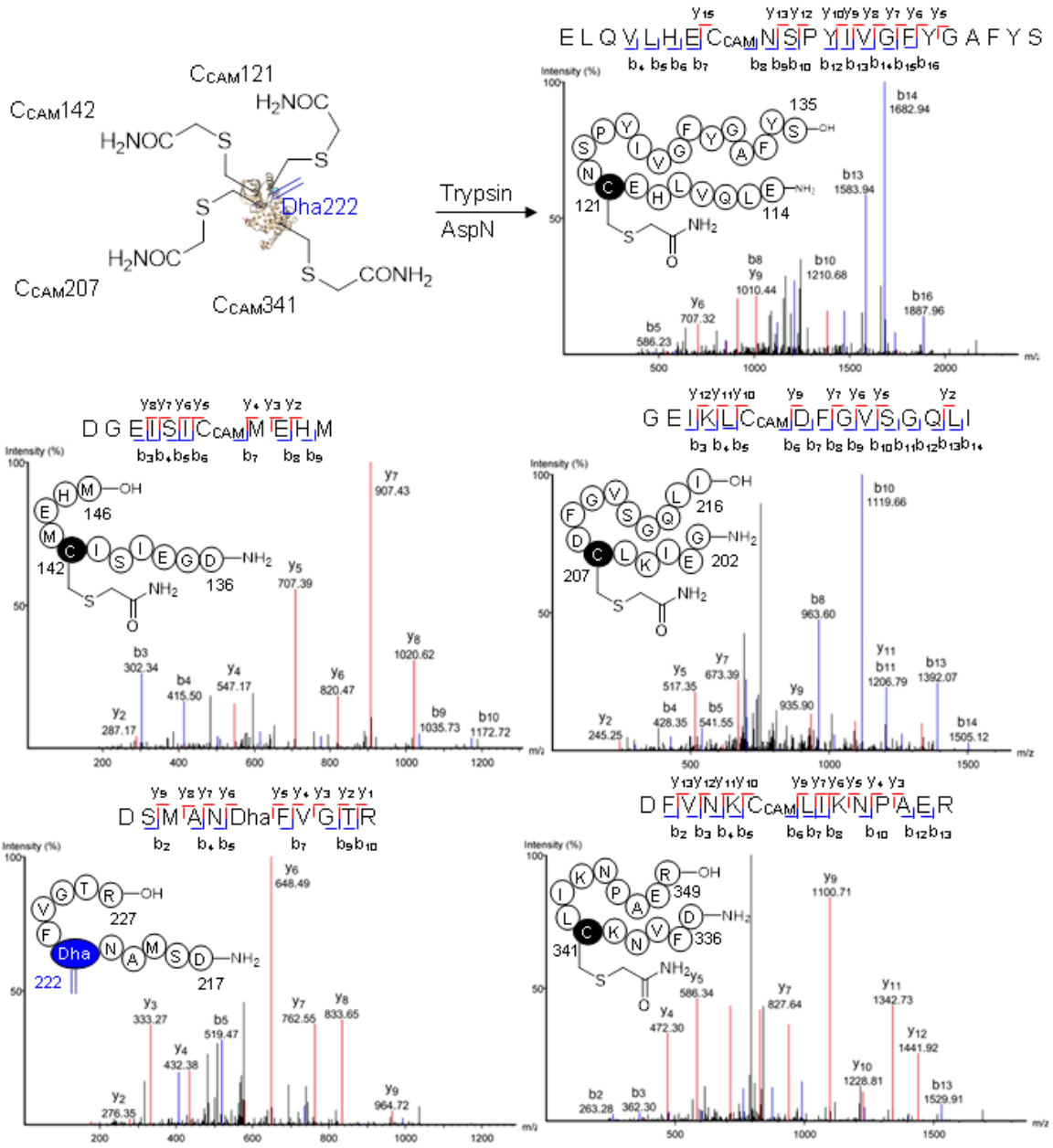
The formation of the first alkylation intermediate MEK1C_{MBP}222 ($M_{calc.}$=45868) was only observed at room temperature. The non-hydrolysed sulfonium intermediate MEK1C_{THTCE}222 ($M_{calc.}$=45789) was not detected.

Full conversion to MEK1Dha222 ($M_{obs.}$=45642, $M_{calc.}$=45642) was observed after 5 hours.



One aliquot of the reaction mixture was digested with trypsin and AspN using in-solution protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at position 222. LCMS-MS analysis of peptides containing cysteine residues C121, C142, C207 and C341 confirmed they did not react under our conditions.

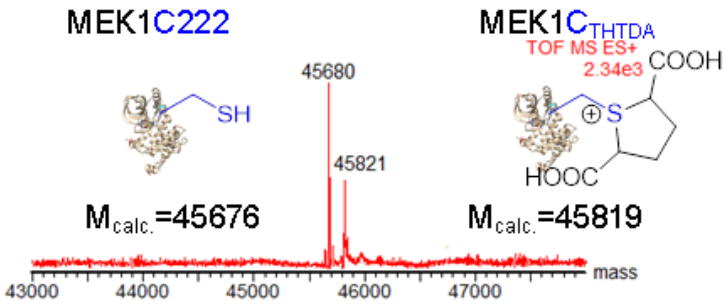
Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
114-135	68.05	1297.0984 (2)	-6.6	C_{CAM}121	70.25	1.21E4	7.74E2
136-146	47.74	661.2607 (2)	-6.4	C_{CAM}142	87.77	2.22E5	4.23E4
202-216	60.15	818.4196 (2)	-5.6	C_{CAM}207	71.97	8.72E5	7.31E4
217-227	37.24	583.7626	-6.9	Dha222	66.50	1.29E6	1.63E5
336-349	29.88	852.4431	-6.4	C_{CAM}341	87.77	8.54E4	8.63E4



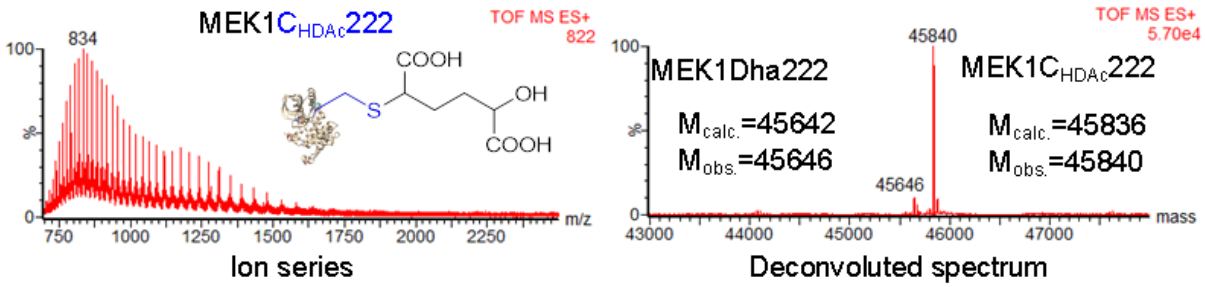
Reaction of MEK1C222 with DBHDAc

MEK1^{S222C} (1.5 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM pH 8 buffer (20 µL) was treated with a 0.5 M stock solution of 2,5-dibromohexanedioic acid, DBHDAc in water basified with a 1 M aqueous sodium hydroxide solution to pH 7. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF.

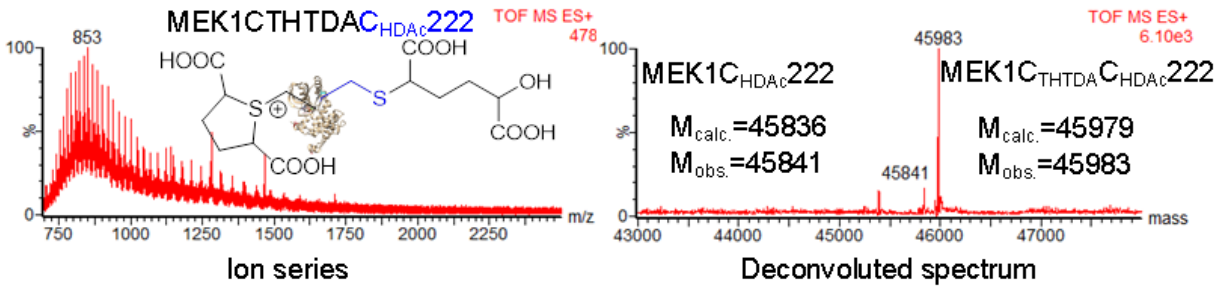
LCMS after 4 hours showed a 30% conversion of MEK1C222 ($M_{obs.}$=45680, $M_{calc.}$=45676), to the sulfonium intermediate MEK1C$_{THTDA}$222 ($M_{obs.}$=45821, $M_{calc.}$=45819).



Full conversion to 90% MEK1C$_{HDAc}$222 ($M_{obs.}$=45840, $M_{calc.}$=45836) and 10% MEK1Dha222 ($M_{obs.}$=45642, $M_{calc.}$=45642) was observed after 24 hours.



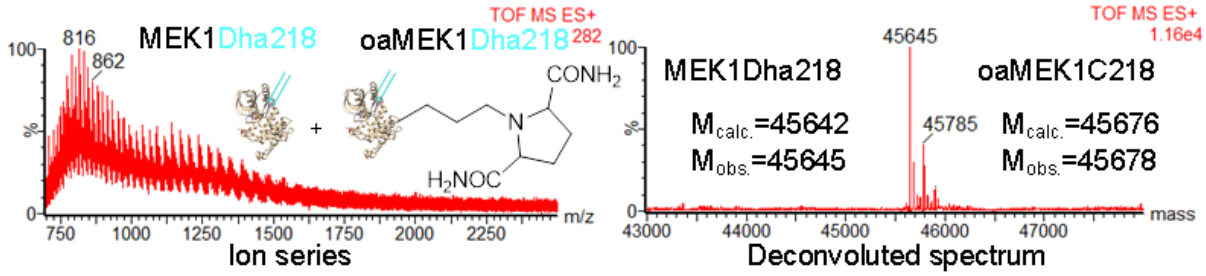
The reaction performed using 20 mM Tris, 100 mM NaCl, pH 8 buffer lead to the formation of MEK1C$_{HDAc222}$, which reacted further to afforded 90% of a compound, $M_{obs.}$=45982 corresponding to the formation of a THTDA adduct, $M_{calc.}$=45979.



7.6.4. Experiments towards the site-selective formation of MEK1Dha218

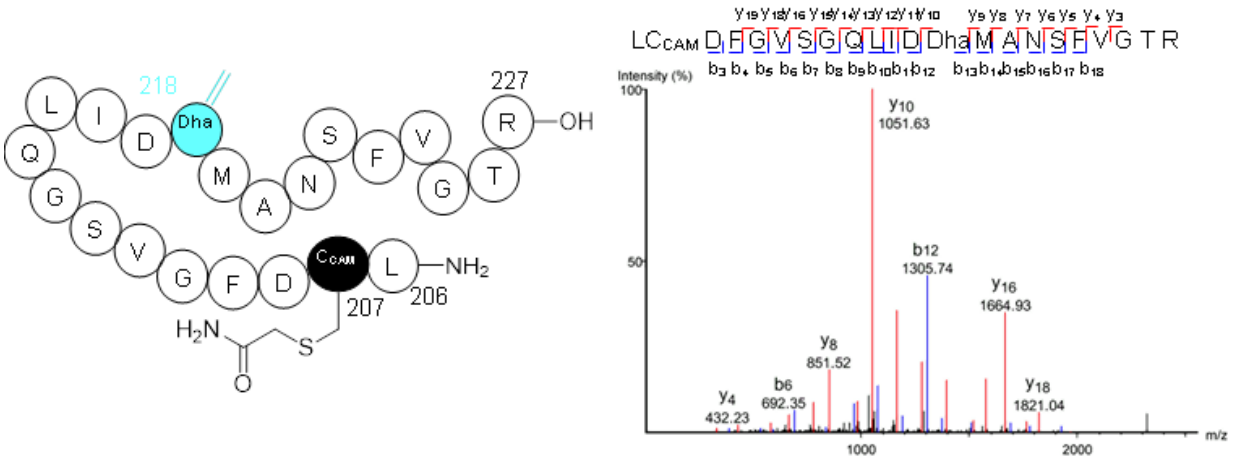
Formation of overalkylated MEK1Dha218 with DBHDA in the presence of ADP-Mg

MEK1C218 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.17 M stock solution of 1,5-dibromohexanediamide, DBHDA in DMSO to afford a final concentration of 4.4 mM of DBHDA. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Dha218 as a mixture MEK1Dha218 (($M_{obs.}$=45645, $M_{calc.}$=45642) and overalkylated oaMEK1Dha218 ($M_{obs.}$=45785, $M_{calc.}$=45782) was observed after 13 hours.

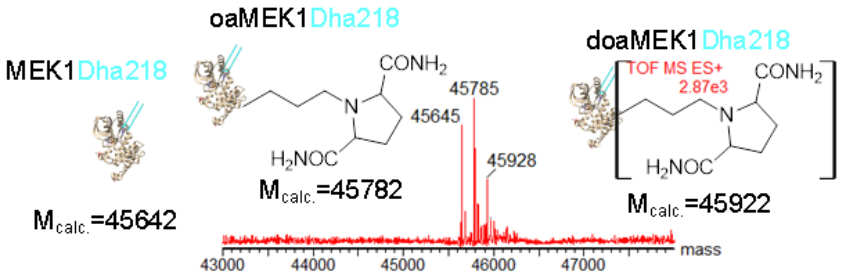


This overalkylation was not the C207 alkylation by DBHDA. One aliquot of the reaction mixture was digested with trypsin using in-gel protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 218 confirmed the position of Dha formation at position 218 only.

Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	42.78	1178.5503 (2)	-5.3	Dha218, C_{CAM}207	93.51	8.16E5	1.06E3

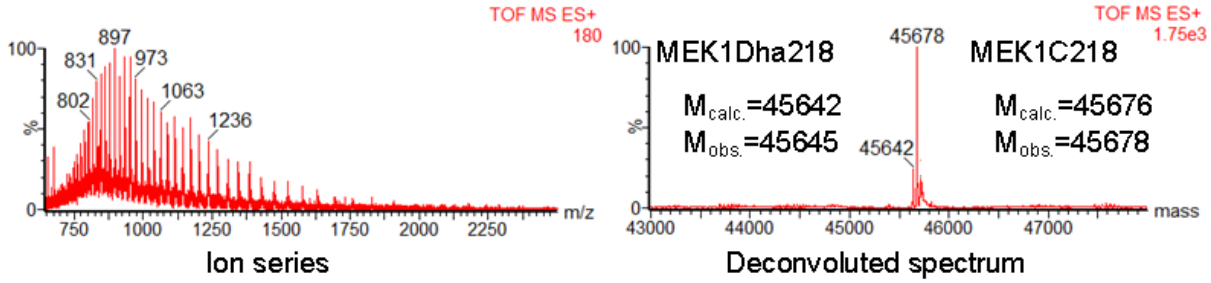


Variation of the concentration of MEK1C218 (11 and 44 µM), DBHDA (1.1, 2.2, 11.0 mM), and temperature (4 °C, room temperature and 30 °C) did not prevent the overalkylation occurring before full conversion to MEK1Dha218. Longer reaction time lead to more adducts.



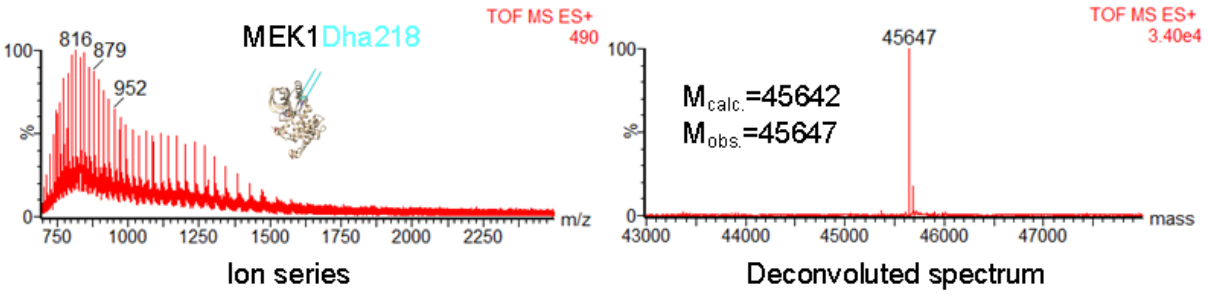
Reaction of MEKC218 with DIB in the presence of ADP-Mg

MEK1C218 (44 µM, 22 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with 1,4-diiodobutane, DIB (2000 Eq. added as such). The mixture was shaken at 1000 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. 25% conversion of MEK1C218 ($M_{obs.}$=45678, $M_{calc.}$=45676) to MEK1Dha218 ($M_{obs.}$=45642, $M_{calc.}$=45642) was observed after 26 hours.



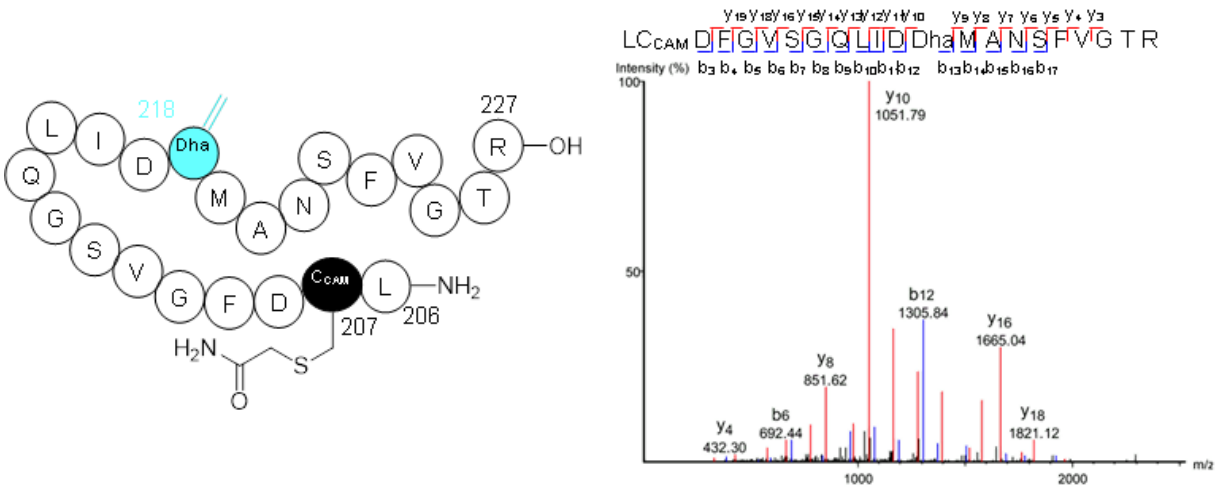
Formation of MEK1Dha218 with DBDGla in the presence of ADP-Mg

MEK1C218 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.44 M stock solution of meso-*N,N'*-bis-(2-amino-2-oxoethyl)-2,5-dibromohexanediamide, DBDGla in DMSO to afford a final concentration of 8.8 mM of DBDGla. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Dha218 ($M_{obs.}$=45647, $M_{calc.}$=45642) was observed after 20 hours.

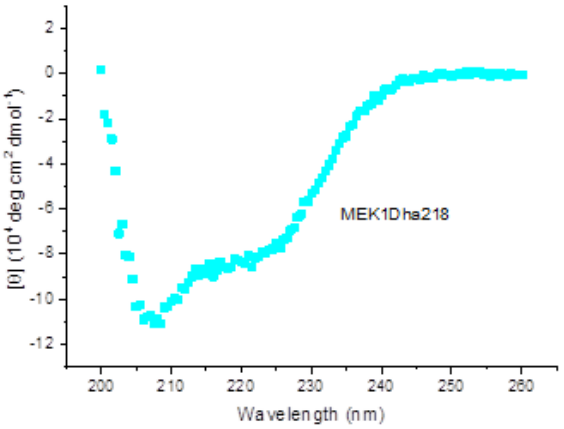


One aliquot of the reaction mixture was digested with trypsin using in-gel protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at position 218.

Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	42.61	1178.5476 (2)	-7.6	Dha218, C_{CAM}207	120.24	4.72E5	4.93E3



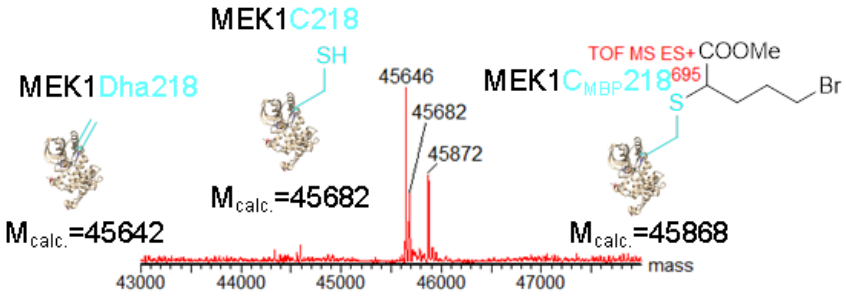
One aliquot was desalted twice with PD-G25 columns. Circular dichroism analysis was done and processed according to the procedure (Paragraph 7.2.3).



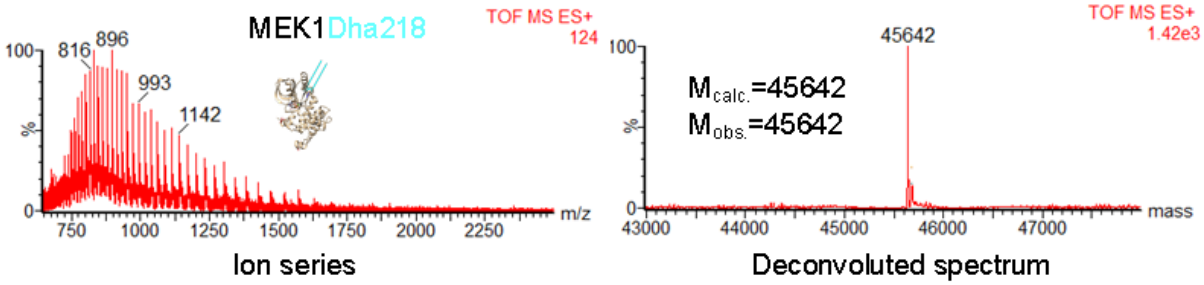
Formation of MEK1Dha218 with MDBP in the presence of ADP-Mg

MEK1C218 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.64 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.4 mM of MDBP. The mixture was shaken at 600 rpm at 37 °C.‡‡‡‡ The reaction was monitored using LCMS-ESI-TOF. LCMS after 1 hour showed a mixture MEK1Dha218 ($M_{obs.}$=45646, $M_{calc.}$=45642), MEK1C218 ($M_{obs.}$=45682, $M_{calc.}$=45676) and the first alkylation intermediate $MEK1C_{MBP}218$ ($M_{obs.}$=45872,$M_{calc.}$=45868).

‡‡‡‡ Precipitation tends to occur at 37 °C. The reaction can be performed at room temperature for around 8-10 h.



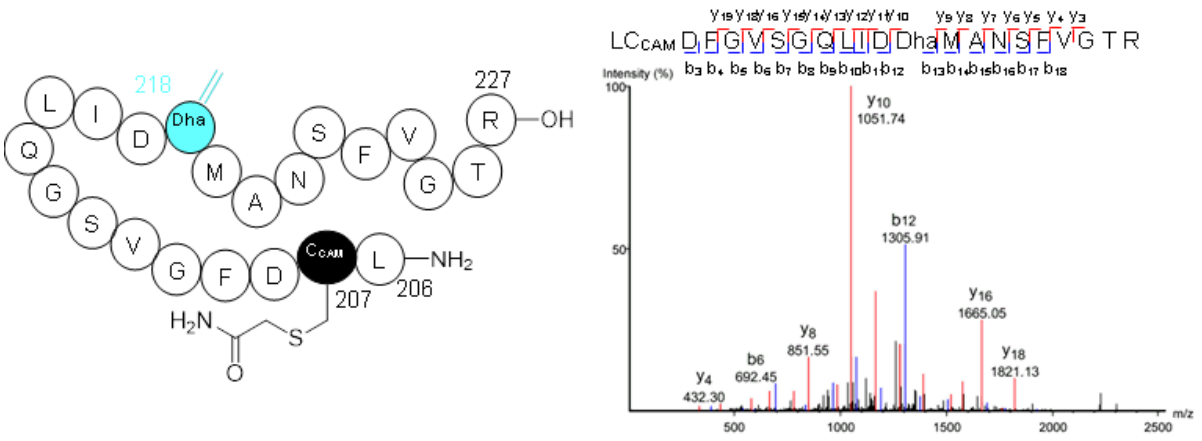
Full conversion to MEK1Dha218 ($M_{obs.}$=45642, $M_{calc.}$=45642) was observed after 4 hours.



Similarly to the formation of MEK1Dha222 with MDBP, the formation of the sulfonium intermediate MEK1C$_{THTCE}$218 ($M_{calc.}$=45789) was not detected.

One aliquot of the reaction mixture was digested with trypsin using in-gel protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at position 218.

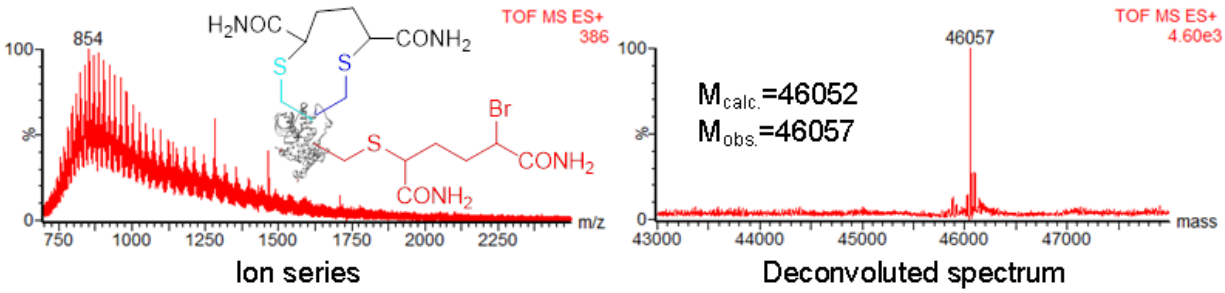
Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	67.12	1178.5476 (2)	-7.6	Dha218, C$_{CAM}$207	88.25	7.49E6	7.52E5



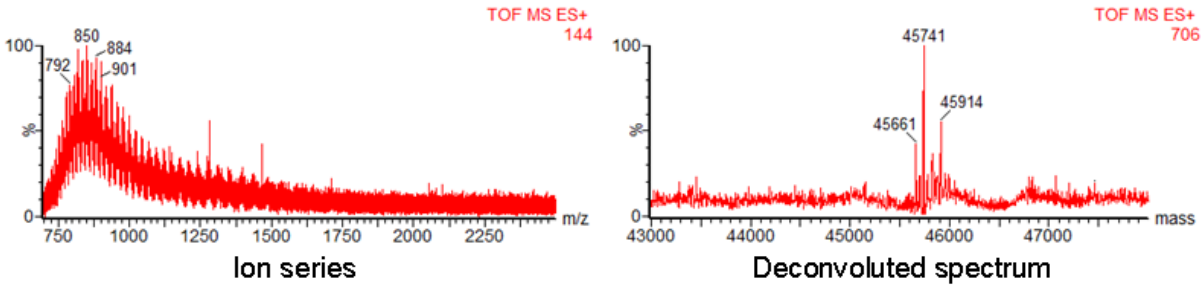
7.6.5. Experiments towards the site-selective formation of MEK1Dha218Dha222

Reaction of MEK1C218C222 with DBHDA

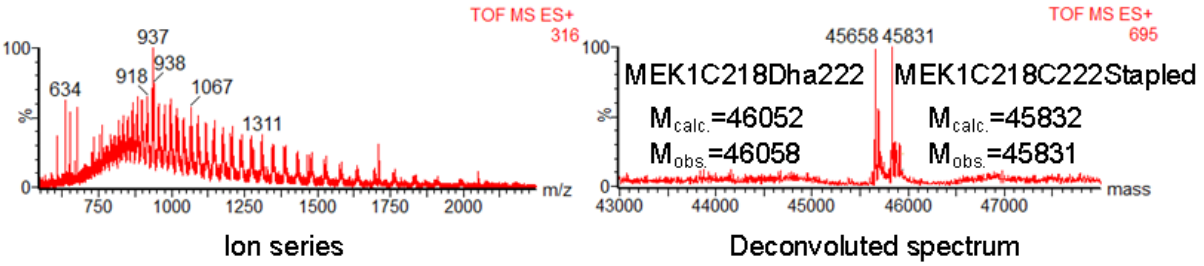
MEK1C218C222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 was treated with a 0.46 M stock solution of 2,5-dibromohexanediamide, DBHDA in DMSO to afford a final concentration of 4.4 mM of DBHDA. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF. Full conversion to an overalkylated MEK1 with C218 and C222 stapled ($M_{obs.}$=46057, $M_{calc.}$=46052) was observed after 6 hours.



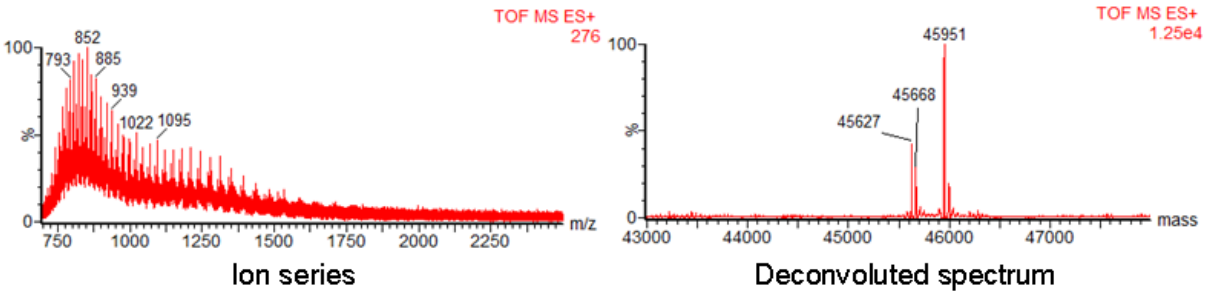
The same reaction conditions, but using 20 mM Tris, 100 mM NaCl, 10 mM ATP, 25 mM $MgCl_2$, pH 8 buffer, afforded a mixture of 60% phosphorylated pMEK1C218Dha222 ($M_{obs.}$=45741, $M_{calc.}$=45738), 20% MEK1C218Dha222 ($M_{obs.}$=45661, $M_{calc.}$=45914) and 20% phosphorylated MEK1 with C218 and C222 stapled ($M_{obs.}$=45914, $M_{calc.}$=450912).



When 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was used, a mixture 50% MEK1C218Dha222 ($M_{obs.}$=46058, $M_{calc.}$=46052) and 50% MEK1 with C218 and C222 stapled ($M_{obs.}$=45831, $M_{calc.}$=45832).

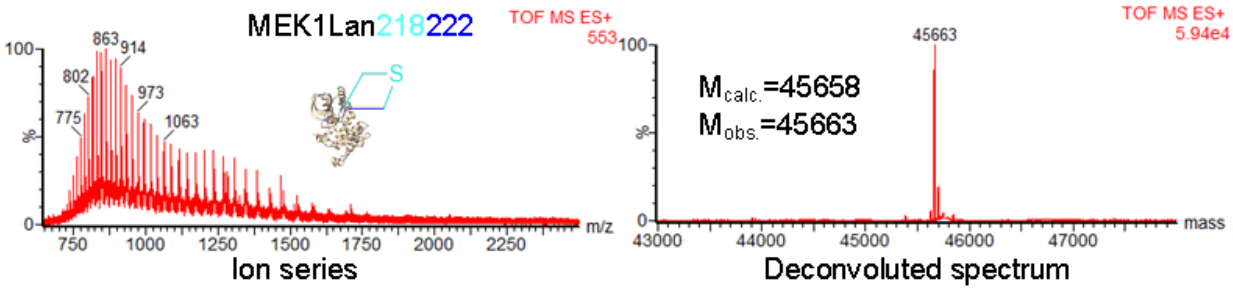


The variation of protein and reagent concentration, temperature and pH did not afford pure MEK1Dha218Dha222 and prevent the stapling between C218 and C222. The optimal conditions (MEK1C218C222 (44 µM, 2 mg/mL) and DBDGla (8.8 mM) in ADP-Mg buffer at 37 °C) produced a maximum 23% MEK1Dha218Dha222 ($M_{obs.}$=45627, $M_{calc.}$=45624) accompanied by 18% MEK1C218Dha222 ($M_{obs.}$=45668, $M_{calc.}$=45658) and 59% MEK1 with stapled C218 and C222 ($M_{obs.}$=45951, $M_{calc.}$=45946).



Reaction of MEK1C218C222 with DIB leading to thioether formation

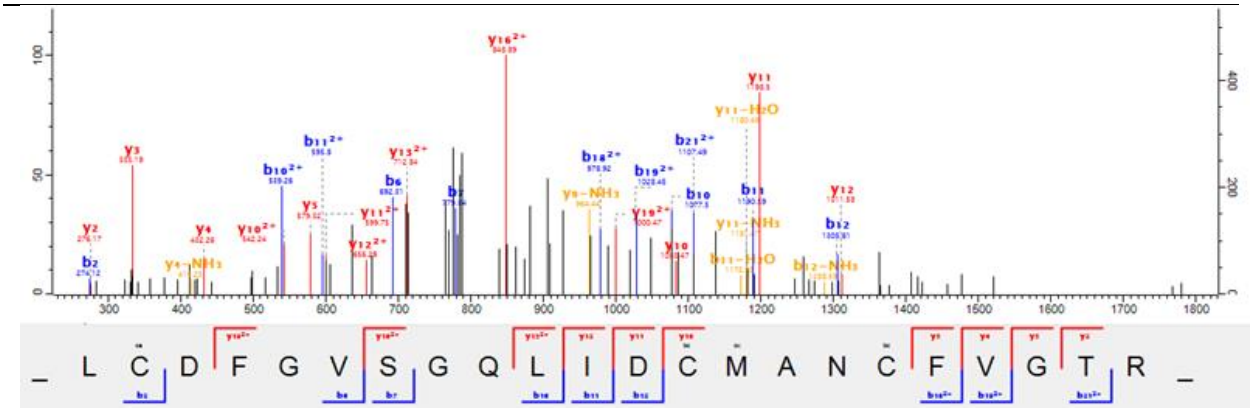
MEK1C218C222 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$ pH 8 was treated with 1,4-diiodobutane, DIB (2000 Eq. added as such). The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Lan218222 ($M_{obs.}$=46663, $M_{calc.}$=46658) after 15 hours.



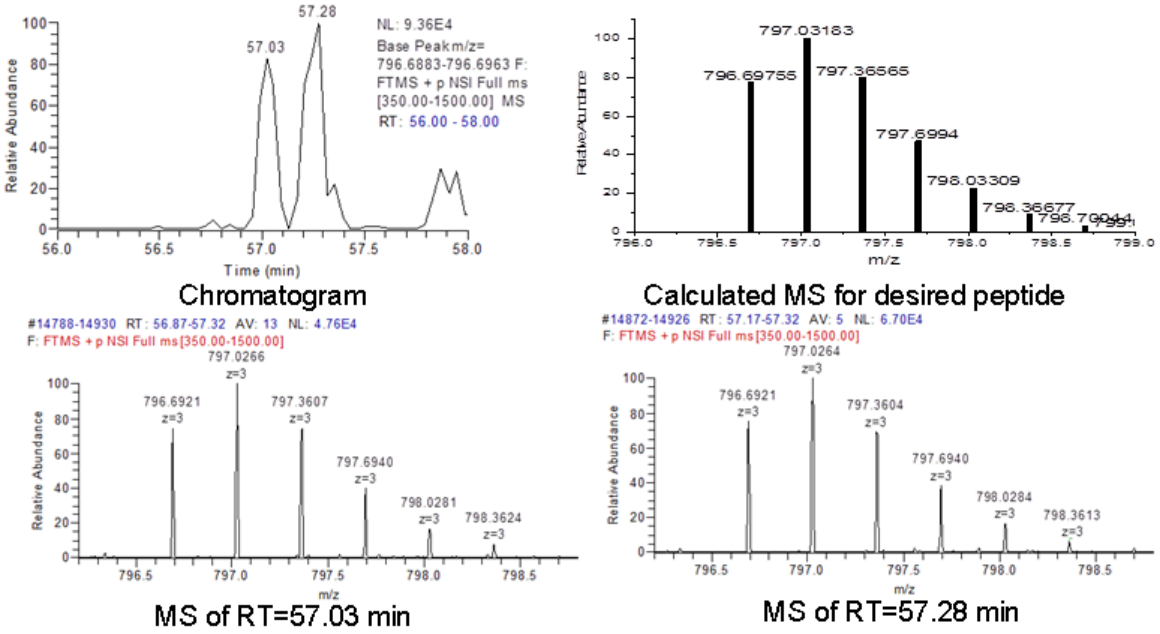
The mass did not change upon the addition of an excess of either β-mercaptoethanol or DBDGla (12 mM final concentration).

One aliquot of the reaction mixture was digested with trypsin using in-solution protocol. The digested peptides were analysed using the orbitrap method and processed by Maxquant. LCMS-MS analysis of peptide 206-227 containing the residue 207, 218 and 222 confirmed the formation of a thioether bond between both positions 218 and 222.

Peptide	Mass (Da)	Score	Modifications	Method
206-227	2348.04	101	Lan218222, oxM219	ITMS CID

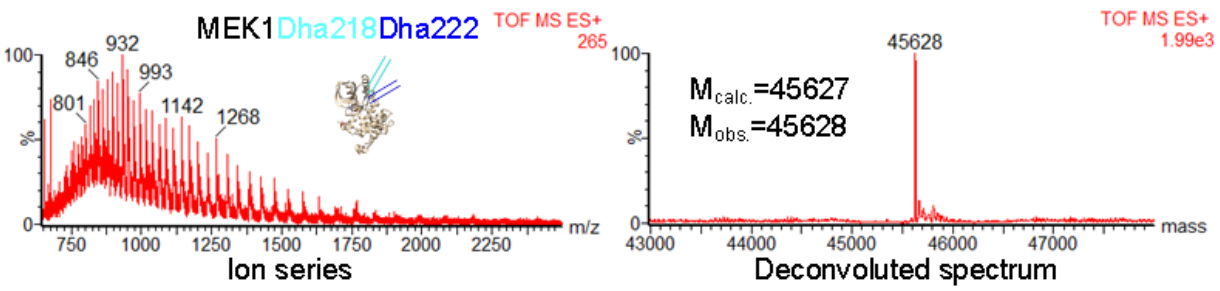


Analysis of the LC trace of the LCMS-MS analysis showed two peaks for the desired mass corresponding to a mixture of epimers.



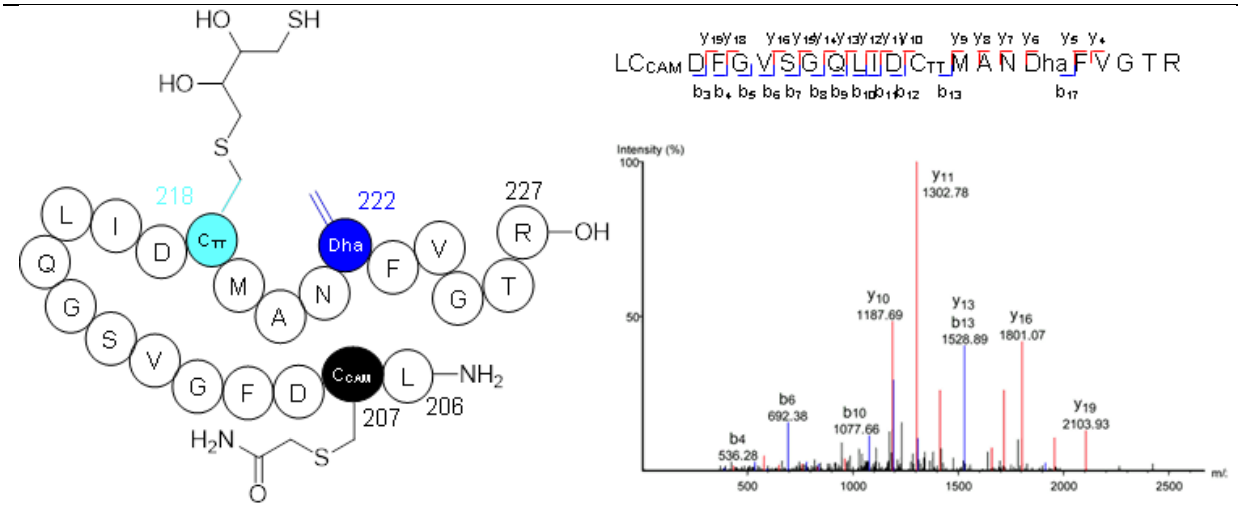
Formation of MEK1Dha218Dha222 with MDBP

MEK1C218C222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.14 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.4 mM of MDBP. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Dha218Dha222 ($M_{obs.}$=45628, $M_{calc.}$=45627) after 15 hours.

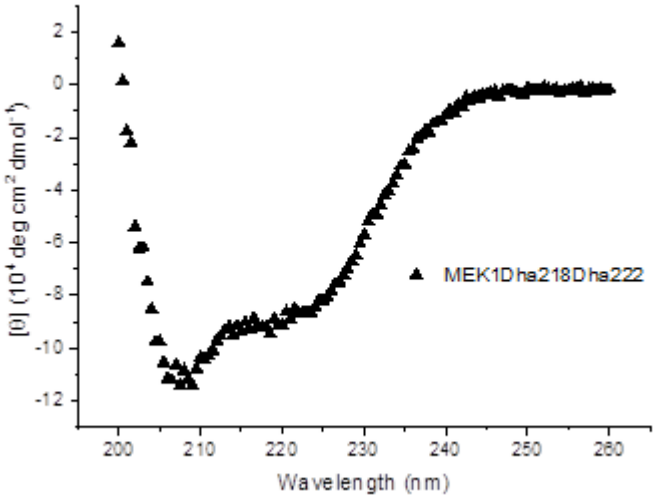


One aliquot of the reaction mixture was digested with trypsin using in-solution protocol. The digested peptides were analysed using the orbitrap method and processed by Peaks 7.0. LCMS-MS analysis of peptide 206-227 containing the residue 207 and 222 confirmed the position of Dha formation at both positions 218 and 222.

Peptide	RT/min	m/z (charge)	ppm	Modifications	-10lgP	TIC	Max Intensity
206-227	67.31	1246.5486 (2)	-7.1	Dha222, C_{DTT}218, C_{CAM}207	83.17	3.12E5	2.49E4



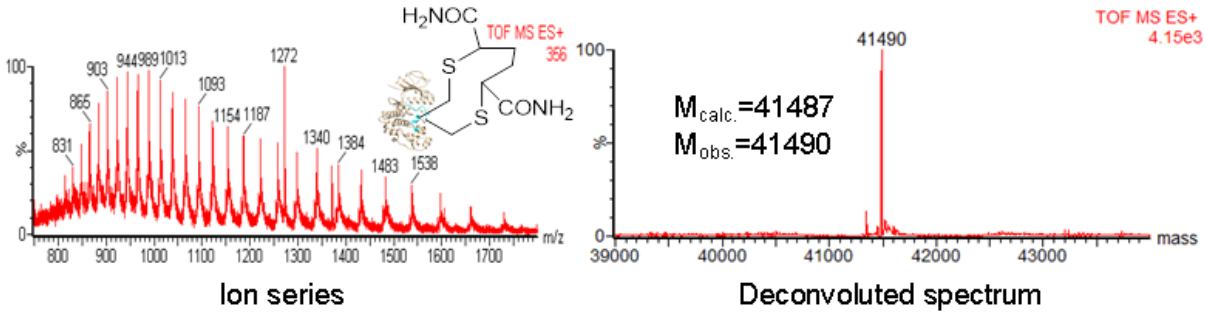
One aliquot was desalted twice with PD-G25 columns. Circular dichroism analysis was done and processed according to the procedure (Paragraph 7.2.3).



7.6.6. Formation of p38αDha180Dha182

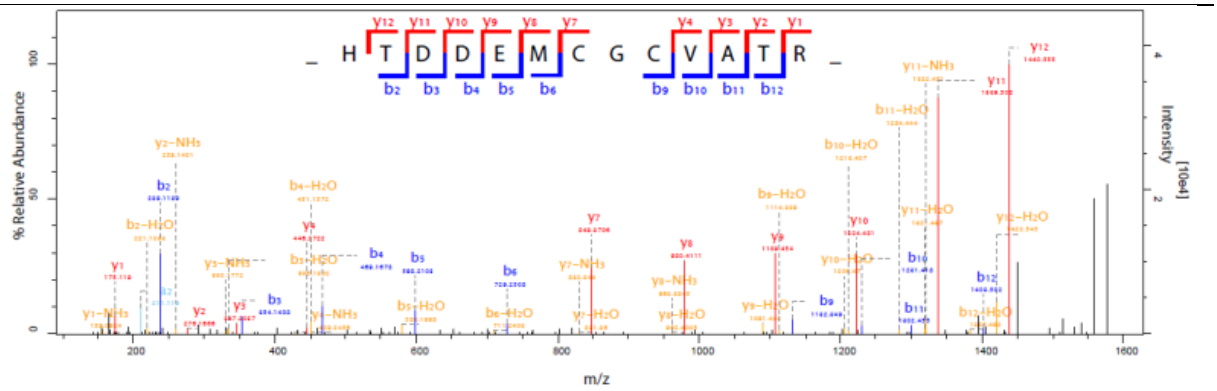
Reaction of p38αC180C182 with DBHDA

p38αC180C182 (24 μM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 2.9 M stock solution of 2,5-dibromohexane diamide, DBHDA in DMSO to afford a final concentration of 12 mM of DBHDA. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to p38α with stapled C180 and C182 ($M_{obs.}$=41490, $M_{calc.}$=41487) was observed after 2 hours.



One aliquot of the reaction mixture was digested with trypsin using in-solution protocol. The digested peptides were analysed using the orbitrap method and processed by Maxquant. LCMS-MS analysis of peptide 174-186 containing the residue 180 and 182 confirmed the stapling of both cysteine residues.

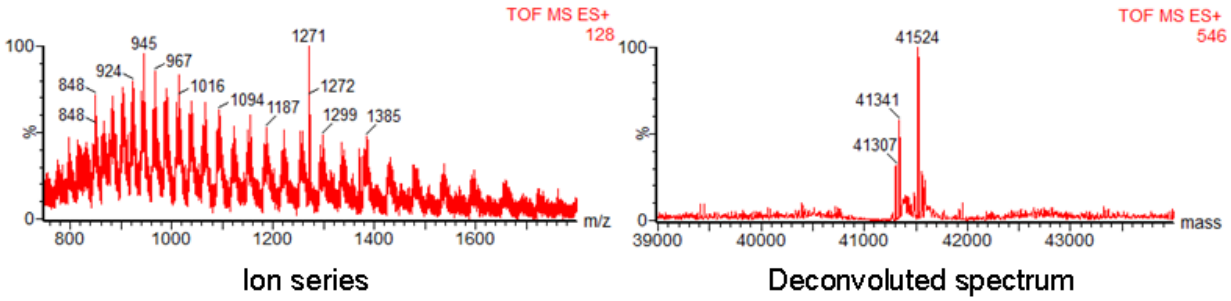
Peptide	Mass (Da)	Score	Modifications	Method
174-186	1576.60	262.11	Stapled C218 & C222	ITMS HCD



The variation of the protein and reagent concentration, temperature, and pH could not prevent the stapling.

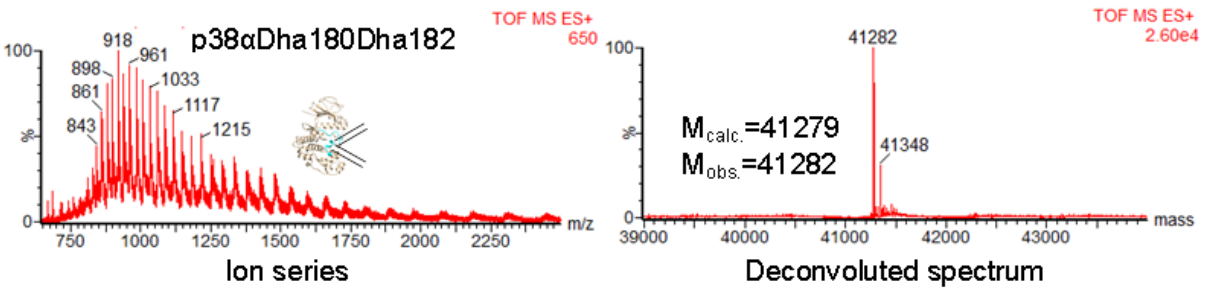
Reaction of p38αC180C182 with DIB

p38αC180C182 (48 μM, 2 mg/mL) in 50 mM Hepes,50 mM NaCl, 5% glycerol pH 8 buffer was treated with 1,4-diiodobutane, DIB (2000 Eq. added as such). The mixture was shaken at 1000 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. 25% conversion of MEK1C218 ($M_{obs.}$=45678, $M_{calc.}$=45676) to MEK1Dha218 ($M_{obs.}$=45642, $M_{calc.}$=45642) was observed after 26 hours. The reaction was monitored using LCMS-ESI-TOF. After 18 hours, 70% conversion was observed to a mixture of p38αDha ($M_{obs.}$=41307, $M_{calc.}$=41313) and monoalkylated p38αC_{IB} (($M_{obs.}$=41529, $M_{calc.}$=41524).



Formation of p38αDha180Dha182 with MDBP

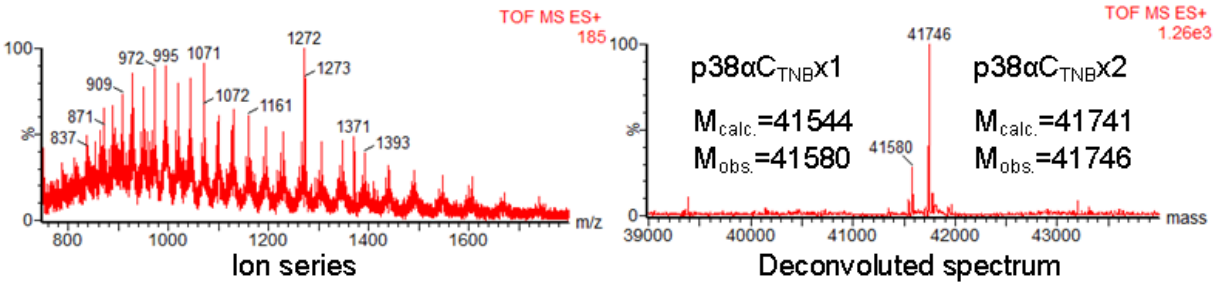
p38αC180C182 (48 μM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 0.10 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.8 mM of MDBP. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF. Full conversion to a mixture of 75% p38αDha180Dha182 ($M_{obs.}$=41282, $M_{calc.}$=41279) and 25% oxidised p38αC180C182 (($M_{obs.}$=41351, $M_{calc.}$=41345) was observed after 2 hours.



7.6.7. Alternative methods towards the formation of p38αDha180Dha182

Reaction of p38αC180C182 with Ellman's reagent followed by elimination

p38αC180C182 (48 µM, 2 mg/mL) in 50 mM Hepes,50 mM NaCl, 5% glycerol pH 8 buffer was treated with a 0.38 M stock solution of 5,5'-dithiobis-(2-nitrobenzoic acid), DTNB (Ellman's reagent) in DMSO to afford a final concentration of 4.8 mM of DTNB. The mixture was shaken at 600 rpm at 37 °C for 1 hour. LCMS-ESI-TOF showed full conversion to a mixture of 20% p38α with one Ellman's adduct p38αC_{TNB}x1 ($M_{obs.}$=41580, $M_{calc.}$=41544) and 80% p38α with two Ellman's adducts p38αC_{TNB}x2 ($M_{obs.}$=41746, $M_{calc.}$=41741).



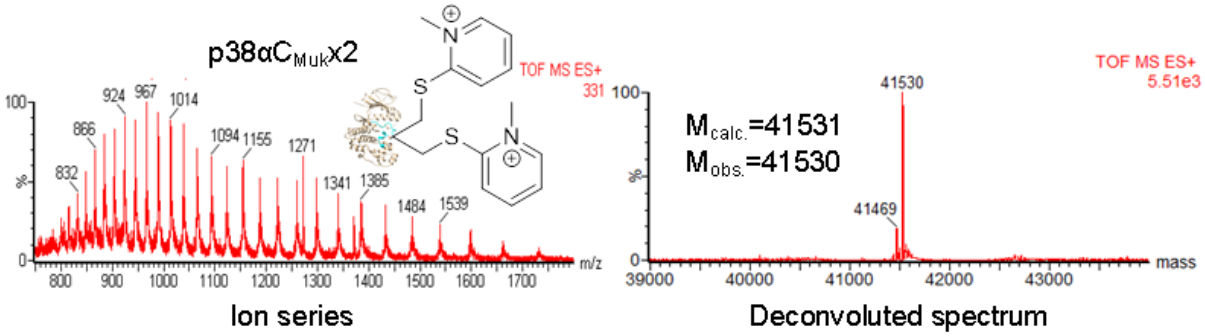
The addition of Ellman's reagent could not be controlled despite the variation of the protein and reagent concentrations, temperature, and pH.

The formation of p38αDha180Dha182 was not observed by LCMS after the treatment of the mixture with either sodium hydroxide or hexamethylphosphorus triamide.

Reaction of p38αC180C182 with Mukaiyama's reagent followed by elimination

p38αC180C182 (24 µM, 1 mg/mL) in 50 mM Hepes,50 mM NaCl, 5% glycerol pH 8 buffer was treated with a 0.14 M stock solution of 2-chloromethylpyridinium iodide (Mukaiyama

reagent) in 50 mM Hepes,50 mM NaCl, 5% glycerol pH 8 to afford a final concentration of 1.2 mM of Mukaiyama reagent. The mixture was left standing on ice for 1 hour. LCMS-ESI-TOF showed full conversion to mainly p38αC$_{Muk}$x2 ($M_{obs.}$=41530, $M_{calc.}$=41531).



The formation of p38αDha180Dha182 was not observed by LCMS after subsequent basification of the buffer with sodium hydroxide.

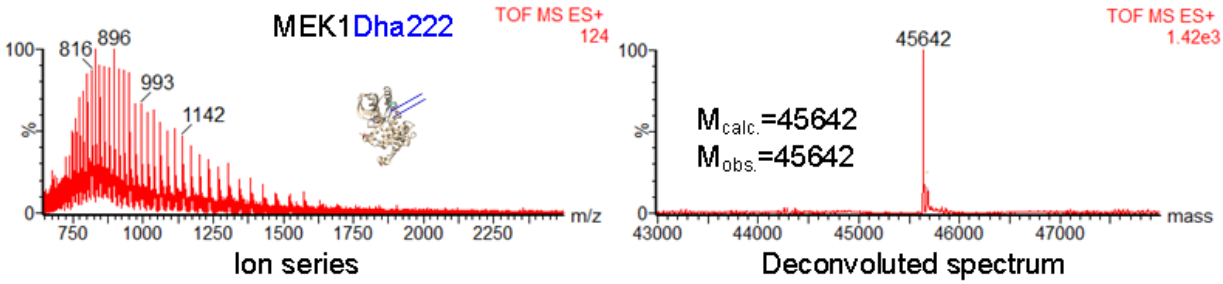
Reaction of p38αC180C182 with MSH

p38αC180C182 (24 µM, 1 mg/mL) in 50 mM NaPi pH 8 buffer was treated with a 0.24 M stock solution of *O*-mesitylenesulfonylhydroxylamine (MSH) in DMF to afford a final concentration of 2.4 mM of MSH. The mixture was left standing on ice. The reaction was monitored using LCMS-ESI-TOF. The formation of p38αDha180Dha182 was not observed by LCMS.

7.6.8. Formation of MEK1pC222

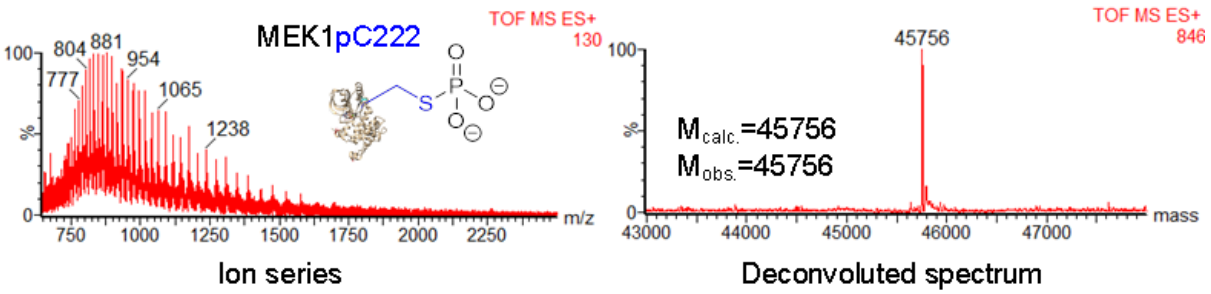
One-pot formation of MEK1pC222 from MEK1C222

MEK1C222 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.42 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.4 mM of MDBP. The mixture was shaken at 600 rpm at 37 °C. The reaction was monitored using LCMS-ESI-TOF. Full conversion to MEK1Dha222 ($M_{obs.}$=45642, $M_{calc.}$=45642) was observed after 3 hours.



A 3.8 M solution of sodium thiophosphate was prepared with commercial sodium thiophosphate hydrate. Typically, 13.7 mg of sodium thiophosphate was suspended in 5.5 µL of MilliQ water. The suspension was acidified with 6 µL of a 5 M hydrochloric acid solution. The solid solubilised. The pH 8-8.5 was checked with pH paper. The concentration was calculated taking into account the final volume of the solution.

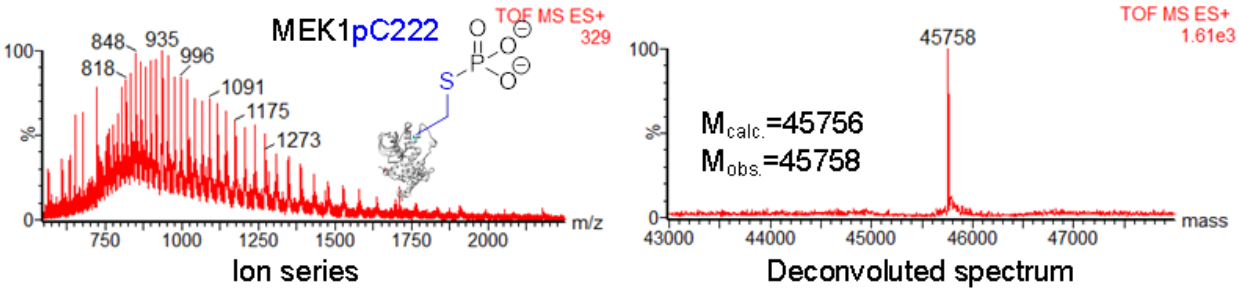
To the mixture containing MEK1Dha222, the 3.8 M stock solution of sodium thiophosphate was added portion-wise, in four aliquots every 2 minutes to afford a final concentration of 0.6 M of sodium thiophosphate and 38 µM of MEK1Dha222. The mixture was then shaken at 600 rpm at 37 °C. The reaction was monitored by LCMS-ESI-TOF. Full conversion to MEK1pC222 ($M_{obs.}$=45756, $M_{calc.}$=45756) was observed after 4 hours.



Large scale formation of MEK1pC222

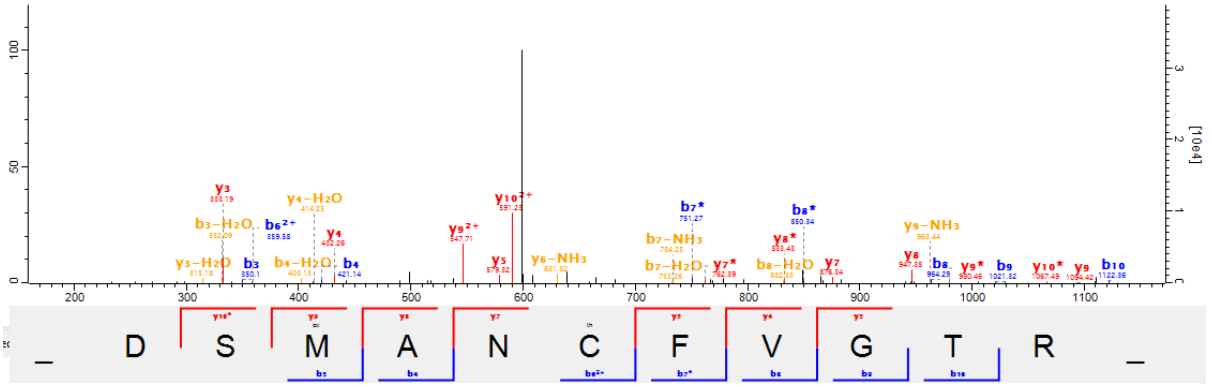
MEK1Dha222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ATP, pH 8 buffer was treated with a 3.8 M stock solution of sodium thiophosphate (see above) to afford a final concentration of 0.5 M sodium thiophosphate solution and 19 µM of MEK1Dha222. The mixture was left standing at room temperature. The reaction was monitored by LCMS-ESI-TOF. Full conversion to MEK1pC222 ($M_{obs.}$=45758, $M_{calc.}$=45756) was observed

after 18 hours. The mixture was dialysed twice (two cycles of 2 hours+2 hours+overnight) at 4 °C against 20 mM Tris, 100 mM NaCl, pH 8 buffer to fully remove the sodium thiophosphate and the ATP.[§§§§] The concentration was estimated with nanodrop. The reaction was 55% yield.



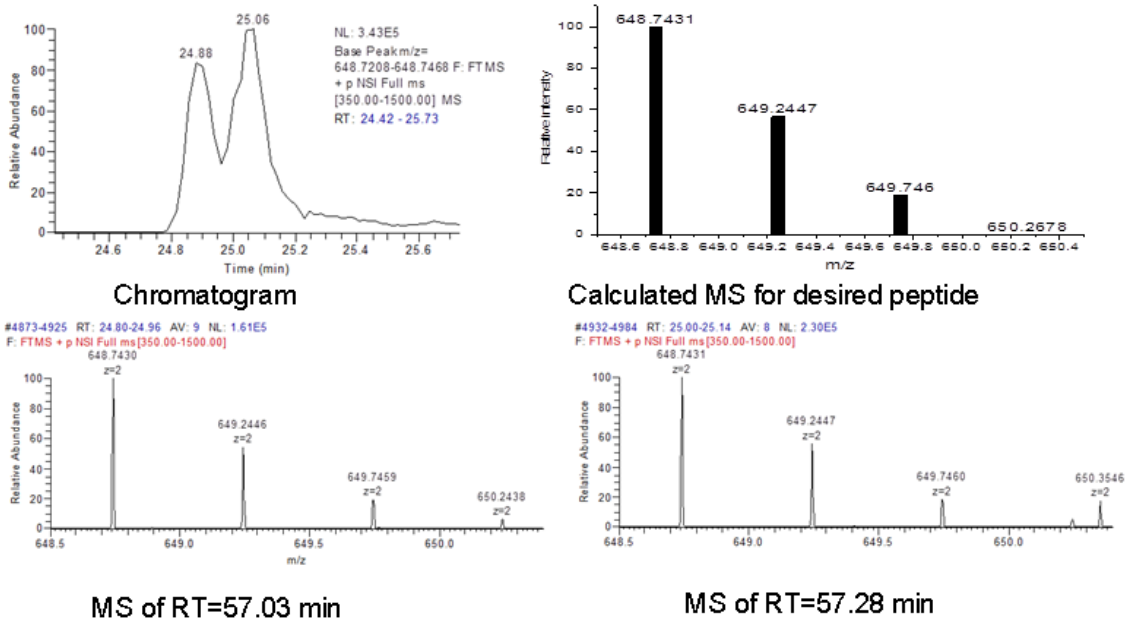
One aliquot of the reaction mixture was digested with trypsin and AspN using in-solution protocol. The digested peptides were analysed using the orbitrap method and processed by Maxquant. LCMS-MS analysis of peptide 216-227 containing the residue 207 and 222 confirmed the position of the pC formation at position 222.

Peptide	Mass (Da)	Score	Modifications	Method
216-227	1199.51	99.8	pC222 (neutral loss), C_{CAM}218	ITMS CID

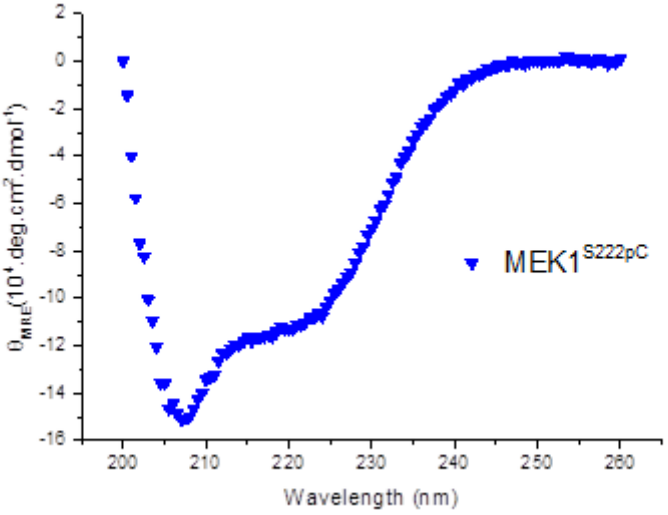


Analysis of the LC trace of the LCMS-MS analysis showed two peaks for the desired mass corresponding to a mixture of epimers.

§§§§ The presence of ATP was monitored on the nanodrop as it was visible at 260 nm.



One aliquot was analysed by circular dichroism according to the procedure (Paragraph 7.2.3).



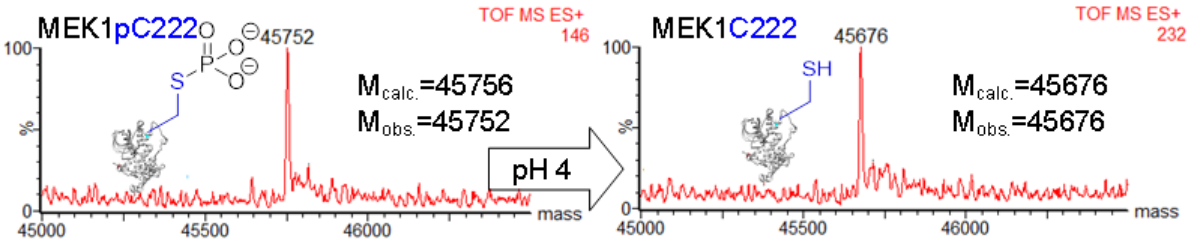
7.6.9. Stability of MEK1pC222

Stability to reducing agents

MEK1pC222 (11 µM, 0.5 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 430 mM aqueous solution of reducing agent (DTT, BME, TCEP and glutathione[*****]) to afford a final concentration of 10 mM of reducing agent. The mixtures were shaken at 600 rpm at 37 °C. The reaction was monitored by LCMS-ESI-TOF. No hydrolysis of MEK1pC222 to MEK1C222 ($M_{calc.}$=45676) was observed after 12 hours.

Stability to acidic conditions

MEK1pC222 (11 µM, 0.5 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 430 mM aqueous solution of glutathione to afford a final concentration of 10 mM of glutathione. The pH of the mixture was around 4. The mixtures were shaken at 600 rpm at 37 °C. The reaction was monitored by LCMS-ESI-TOF. Full conversion to MEK1C222 ($M_{calc.}$=45676, $M_{obs.}$=45676) was observed after 3 hours.

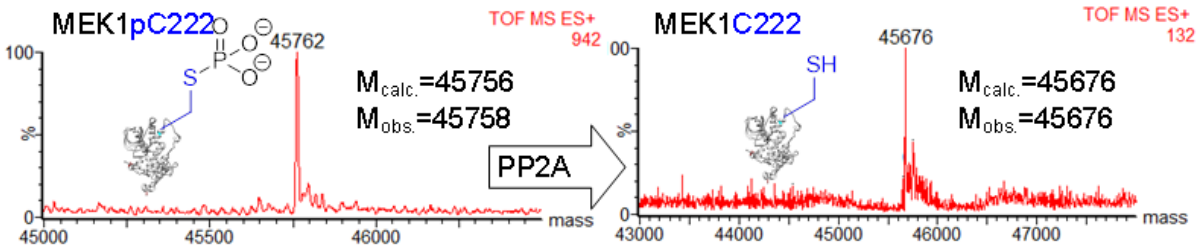


Stability to phosphatase PP2A

MEK1pC222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with 5 µL of a 1 unit/20 µL PP2A solution (according to the manufacturer instruction). The mixture was shaken at 600 rpm at 30 °C. The reaction was monitored by LCMS-ESI-TOF.[†††††] After 2 hours, near full hydrolysis of MEK1pC222 to MEK1C222 ($M_{obs.}$=45676, $M_{calc.}$=45676) was observed.

[*****] The pH of the glutathione stock solution was basified to pH 8 before use.

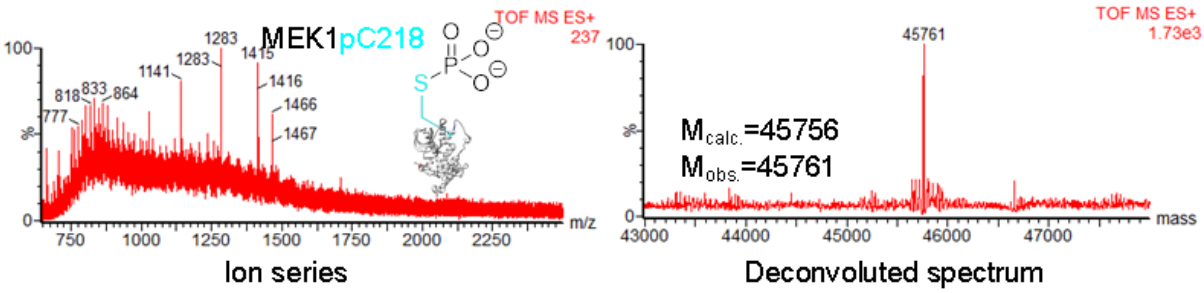
[†††††] BME, present in excess in the PP2A buffer formed adduct +78 with MEK1pC222. Monitoring the dephosphorylation -80 of MEK1 was therefore difficult. To prevent the adduct formation in the LCMS, the LCMS samples were prepared in a 10 mM DTT solution.



7.6.10. Formation of MEK1pC218 and MEK1pC218pC222

Formation of MEK1pC218

MEK1Dha218 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ATP, pH 8 buffer was treated with a 3.8 M stock solution of sodium thiophosphate (see above) to afford a final concentration of 0.9 M sodium thiophosphate solution and 19 µM of MEK1Dha222. The mixture was left standing at room temperature. The reaction was monitored by LCMS-ESI-TOF. Full conversion to MEK1pC218 ($M_{obs.}$=45761, $M_{calc.}$=45756) was observed after 24 hours.

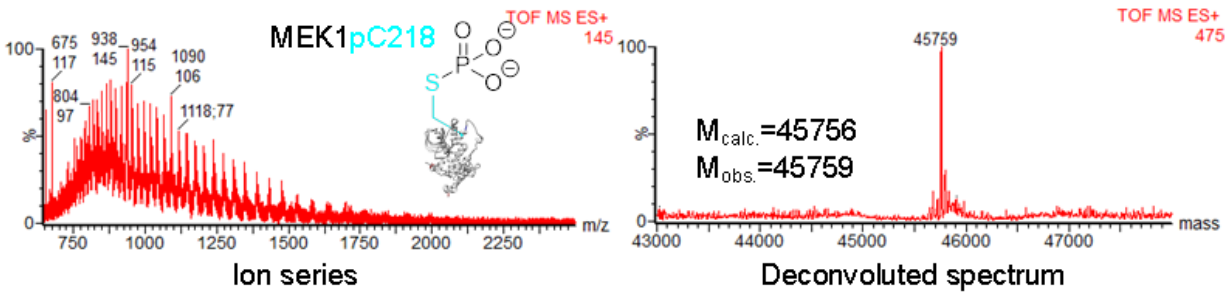


This reaction was not reproducible and depended on the quality of sodium thiophosphate (see chapter 5, 5.2.1). The formation of MEK1pC218 was therefore performed as follow.

A 0.4 M aqueous solution pH 8.6-8.7 of pure sodium thiophosphate‡‡‡‡‡ was prepared by diluting anhydrous sodium thiophosphate[203] in the desired amount of MilliQ water. The pH was adjusted with a 6 M hydrochloric acid solution down to pH 9 and 1 M hydrochloric acid solution down to pH 8.6-8.7.

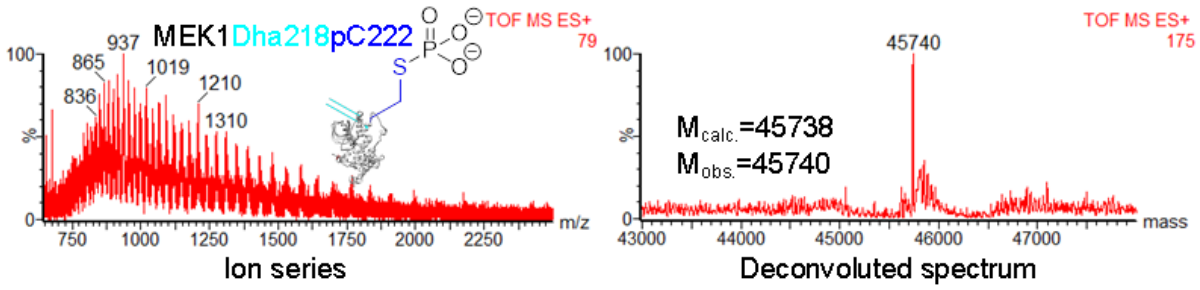
‡‡‡‡‡ Aldrich batches are not consistent. The powder has to be white and fine. ^{31}P NMR should not show a peak at -5.82 ppm. Alpha Aesar batches are crystalline but very hydrated.

MEK1Dha218 (22 µM, 1 mg/mL) in 0.4 M pure sodium thiophosphate solution pH 8.6 was left standing at room temperature. The reaction was monitored by LCMS-ESI-TOF. Full conversion to MEK1pC218 ($M_{obs.}$=45759, $M_{calc.}$=45756) was observed after 16 hours.[§§§§§]



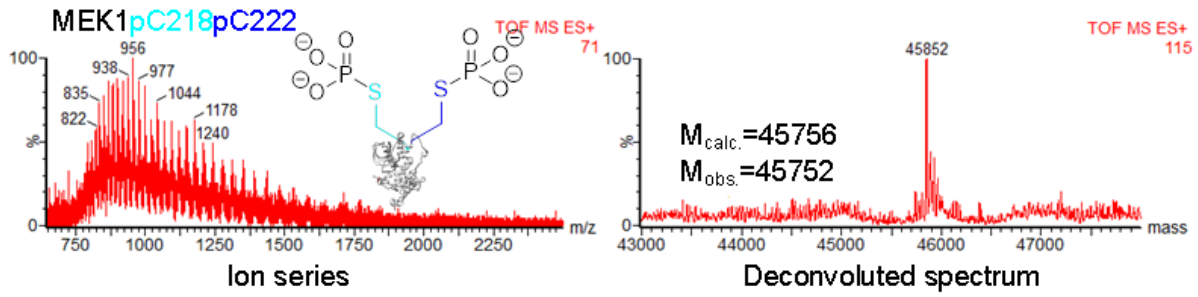
Formation of MEK1pC218pC222

MEK1Dha218Dha222 (22 µM, 1 mg/mL) in 0.4 M pure sodium thiophosphate solution pH 8.6 was left standing at room temperature. The reaction was monitored by LCMS-ESI-TOF. LCMS after 6 hours showed full conversion to MEK1Dha218pC222 ($M_{obs.}$=45646, $M_{calc.}$=45642).



Full conversion to MEK1pC218 ($M_{obs.}$=45759, $M_{calc.}$=45756) was observed after 16 hours (the LCMS sample was prepared in 10 mM DTT see footnote).

§§§§§ The LCMS sample was prepared in 10 mM DTT as the spectrum of a sample in water showed adducts of thiophosphate.



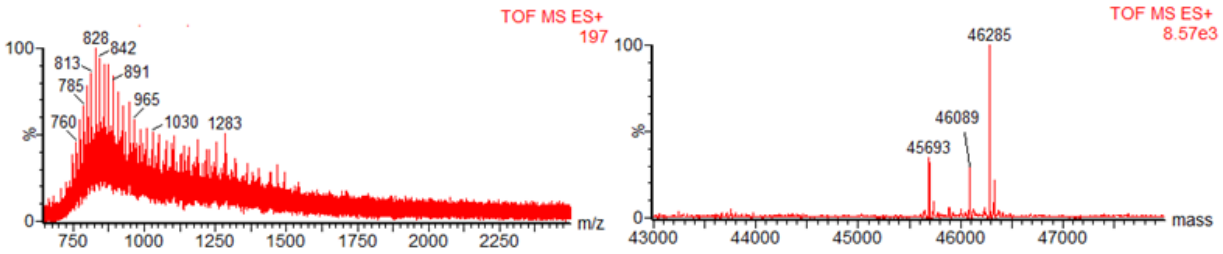
7.6.11. Alternative methods of chemical phosphorylation of MEK1

Use of potassium *O*,*O*-diethylthiophosphatesodium thiophosphate

MEK1Dha218 (44 µM, 1 mg/mL) in 0.3 M potassium *O*,*O*-diethylthiophosphatesodium thiophosphate solution pH 8.3 was shaken at 600 rpm at 37 °C. The reaction was monitored by LCMS-ESI-TOF. LCMS after 18 hours showed no reaction occurred.

Use of Ellman's reagent followed by addition of substituted phosphite

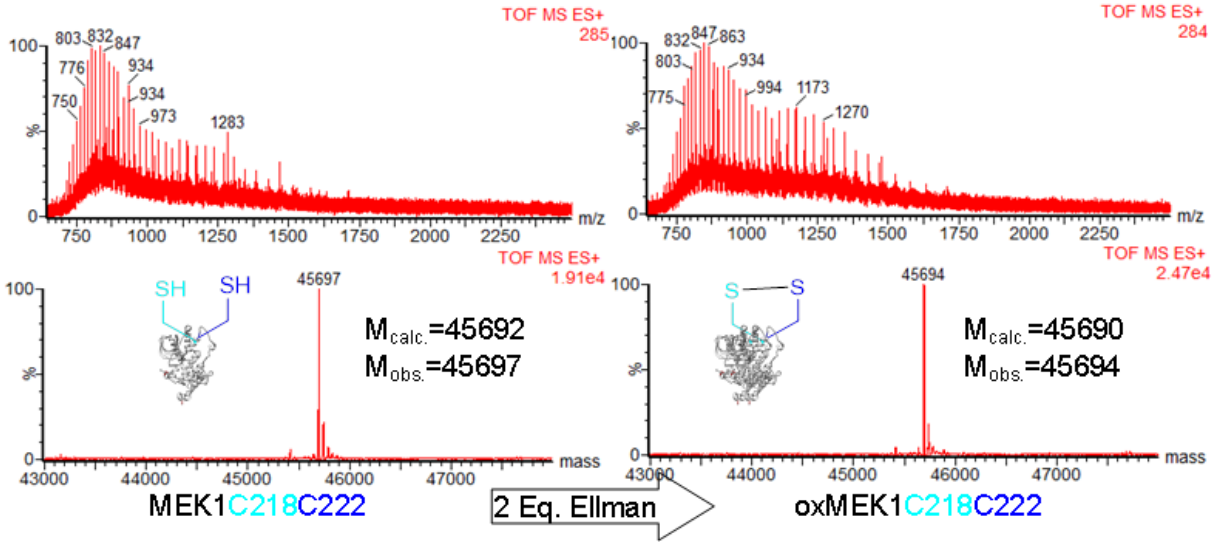
MEK1C218C222 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 buffer was treated with a 2.5 mM stock solution of 5,5'-dithiobis-(2-nitrobenzoic acid), DTNB (Ellman's reagent) in DMSO to afford a final concentration of 264 µM of DTNB. The mixture was shaken at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TO. Full conversion to MEK1 majorically containing 3 TNBD adducts MEK1[CTNB]x3 ($M_{obs.}$=46285, $M_{calc.}$=46283) was observed after 15 minutes.



The treatment with triethylphosphite ester (final concentration 4.4 mM) afforded MEK1C218C222 after 1 hour at 4 °C.

The use of lower amounts of Ellman's reagent did not provide a single compound, although a unique mass corresponding to oxidised MEK1 ($M_{obs.}$=45694, $M_{calc.}$=45690) was observed

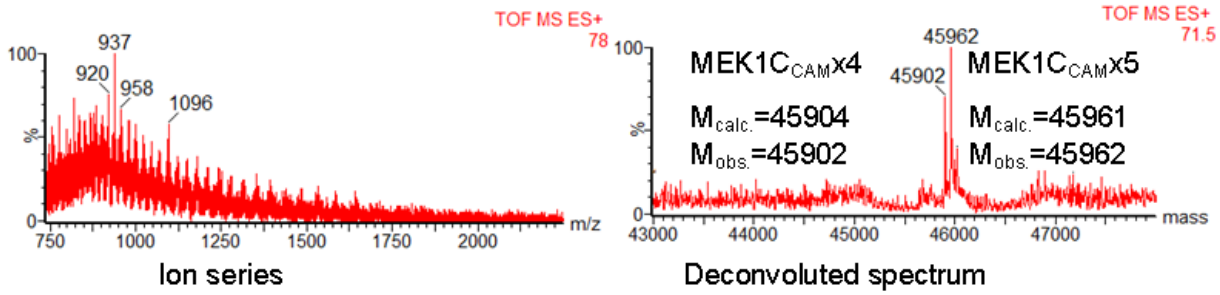
by LCMS when the final concentration of DTNB was 88 µM. The charge state distribution was also slightly different. The formation of a disulphide bond between C218 and C222 was therefore assumed.



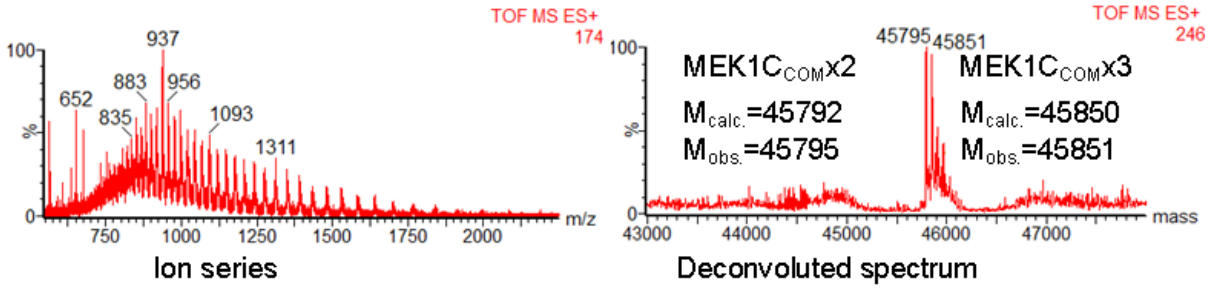
Alkylation of MEKC222

MEK1C222 (22 µM, 1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 8 was treated with 0.65 M aqueous solutions pH 8 of various alkyl iodide to afford a final concentration of 11 mM of alkyl iodide. The mixtures were shaken at 600 rpm at 37 °C. The reactions were monitored using LCMS-ESI-TOF.

With iodoacetamide, full conversion to 55% of MEK1C_{CAM}x4 ($M_{obs.}$=45902, $M_{calc.}$=45904) and 45% of MEK1C_{CAM}x5 ($M_{obs.}$=45962, $M_{calc.}$=45961) was observed after 1 hour.



With iodoacetic acid, full conversion to 50% of MEK1C$_{COM}$x2 ($M_{obs.}$=45795, $M_{calc.}$=45792) and 50% of MEK1C$_{COM}$x3 ($M_{obs.}$=45851, $M_{calc.}$=45850) was observed after 1 hour.



Only starting material MEK1C222 ($M_{calc.}$=45676) was observed by LCMS with iodomethylenephosphonic acid, bromodifluoromethylenephosphonic acid and their ethyl ester after 5 hours.

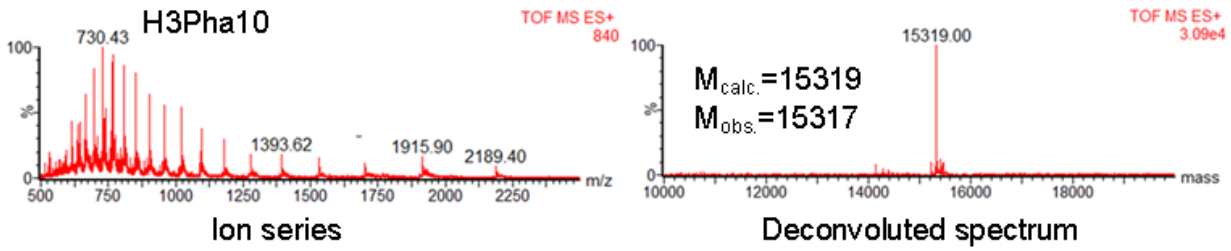
7.6.12. Radical addition of $CH_2PO_3^{2-}$ and $CF_2PO_3^{2-}$[******]

Formation of H3Pha10

H3Dha10 (58 µM, 0.88 mg/mL, 160 µL) in 500 mM NH_4OAc, 5M Gdn.HCl, pH 6 was put in a N_2 atmosphere glovebox and left overnight at room temperature. Within the glovebox, the degassed protein solution was treated with 2.1 mg of iodomethylenephosphonic acid**.** The resultant solution was gently mixed by pipette. The mixture was then added to a separate eppendorf tube containing 3.5 mg powdered $NaBH_4$ at 4°C.[††††††] The mixture was gently shaken by hand, causing vigorous effervescence. The reaction was shaken at 300 rpm at 4°C for one hour before being removed from the glovebox and quenched by buffer exchange using a PD SpinTrap™ G-25 desalting column (pre-equilibrated with fresh reaction buffer). Full conversion to H3Pha10 ($M_{obs.}$=15319, $M_{calc.}$=15317) was observed by LCMS.

[******] This work that has been published in Science.2016.has been done in collaboration with Thomas Wright and Dr.Billy Ng.

[††††††] For this reaction the equivalents of $NaBH_4$ used was 10 times greater than proved necessary for other halide reagents.

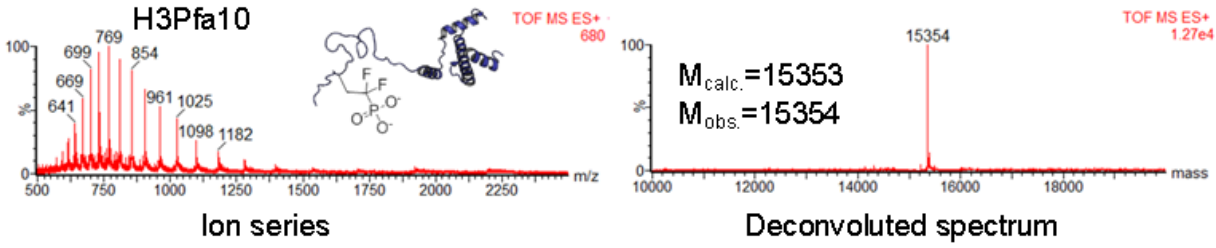


This protein was stable to protein phosphatase PP1.

Formation of H3Pfa10

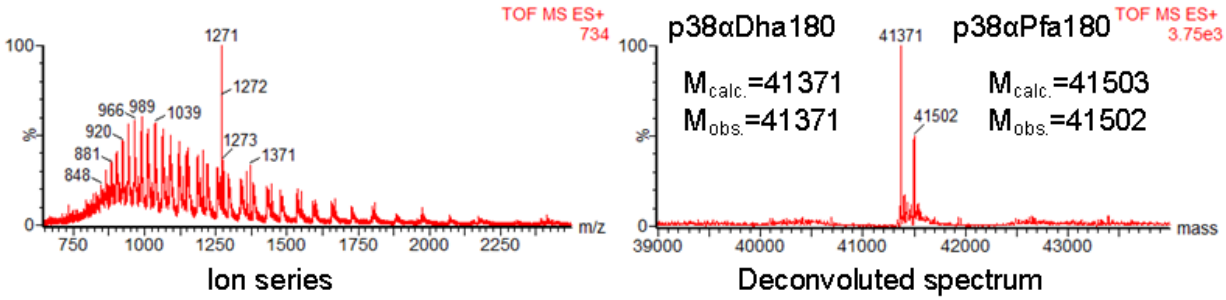
H3Dha10 (66 μM, 1 mg/mL, 200 μL) in 500 mM NH_4OAc, 5M Gdn.HCl, pH 6.0, 180 μL was put in a N_2 atmosphere glovebox and left overnight at room temperature. A stock solution of bromodifluoromethyl)phosphonic acid, BFPA was prepared by dissolving 55 mg in 275 μL water and adjusting the pH to 6.0) and also put into the glovebox and left standing overnight at room temperature. Within the glovebox, 2.5 μL of the above stock solution of BFPA was pipetted into the protein solution, and the resultant solution gently mixed by pipette. The the protein solution was then added to a separate tube containing 1 mg of solid $NaBH_4$ at 4°C. Sustained effervescence was observed. The reaction was incubated at 4°C for 30 minutes before being removed from the glovebox and quenched by elution through a SpinTrap™ G-25 desalting column (pre-equilibrated with fresh reaction buffer), affording the desired chemical mutagenesis product.

Full conversion to H3Pha10 ($M_{obs.}$=15353, $M_{calc.}$=15354) was observed by LCMS.



Formation of p38αPfa180

p38αDha180 (0.85 mg/mL, 100 µL) solution in 50mM HEPES pH 7.8 was transferred to an anaerobic atmosphere glovebox and incubated for 12 hours at 4 °C. A stock solution of (bromodifluoromethyl)phosphonic acid (prepared by dissolving 55 mg in 275 µL water and adjusting the pH to 6.0) was also added into the glovebox. The following morning, within the glovebox, 2.5 µL of the above stock solution was pipetted into the protein solution, and the resultant solution gently mixed by pipette. The entire volume (~200 µL) of the protein solution was then added to a separate 2 mL eppendorf tube, kept at 4 °C and containing 0.5 mg of solid $NaBH_4$. The reaction was gently shaken at 4 °C for 30 minutes before being analysed by LC-MS.



7.7. Kinase activity of MEK1 constructs

7.7.1. Progress curve assay

For in vitro kinase assay, ERK1K71R with MEK1 variants was incubated in *in vitro* kinase buffer (50 mM Tris [pH 7.5], 10 mM $MgCl_2$, 0.2 mM DTT). The reaction was started by adding 5 x ATP solution (5 mM ATP, 50 mM $MgCl_2$, 25 mM Tris [pH 7.2], 150 mM NaCl) and stopped at the indicated time either by addition of 50 mM EDTA quenching buffer or by the LCMS autosampler. Duplicates were made for all.

7.7.2. Processing and results

Each experiment included either 0.5 or 1 µM ERK1K71R concentration. These eluates were subjected analysed by LCMS. The amounts of phospho-isoforms of ERK were quantified using MS intensity and fitted by global least-square regression algorithm to solutions of following analytical equations with Originpro9.1:

Rate differential equations:

$$\frac{d[ERK1]}{dt} = -k_{f1}[MEK][ERK1] \ (1)$$

$$\frac{d[pERK1]}{dt} = k_{f1}[MEK1][ERK1] - k_{f2}[MEK1][ppERK1](2)$$

$$\frac{d[ppERK1]}{dt} = k_{f2}[MEK1][ppERK1](3)$$

Analytical (integrated) equations:

$$[ERK1] = [ERK1]_0 e^{-kf_1[MEK1]t} \ (4)$$

$$[pERK1] = \frac{k_{f1}[ERK1]_0\left(e^{-k_{f1}[MEK1]t} - e^{-k_{f2}\{MEK1\}t}\right)}{k_{f2} - k_{f1}}(5)$$

$$[ppERK1] = [ERK1]_0(1 - \frac{k_2 e^{-k_{f1}[MEK1]t} - k_1 e^{-kf_2[MEK1]t}}{k_{f2} - k_{f1}})(6)$$

where [MEK1], [ERK1], [pERK1], and [ppERK] are concentration of MEK, ERK1, pERK1, and ppERK1, respectively. k_{f1} and k_{f2} are phosphorylation rate constants [$M^{-1}.s^{-1}$], which are enzymatically identical to k_{cat}/K_m at substrate concentrations significantly lower than Km. 1 µM of ERK1 was assumed to be lower than reported K_M for ERK.[35]

Entry	MEK1 construct	k_{f1} x10^3 $M^{-1}.s^{-1}$	k_{f2} x10^3 $M^{-1}.s^{-1}$	R^2
1	MEK1**wt***	0.46±0.03	0.50±0.03	96.93
2	MEK1**wt**	0.41±0.02	0.47±0.03	97.58
3	MEK1**C218**	0.17±0.01	0.31±0.03	96.36
4	MEK1**C222**	0.84±0.03	0.82±0.04	99.25
5	MEK1**C218C222**	0.25±0.01	0.30±0.01	98.42
6	MEK1**Dha218**	11.99±0.54	10.00±0.61	97.08
7	MEK1**Dha222**	1.37±0.06	1.41±0.08	98.37
8	MEK1**Dha218Dha222**	27.32±2.19	26.37±2.38	97.63
9	MEK1**pC218**	16.5±0.7	12.7±0.6	96.28
10	MEK1**pC222**	10.9±0.5	9.2±0.5	93.97
11	MEK1**E218**	1.40±0.07	1.30±0.09	97.41
12	MEK1**E222**	3.64±0.16	3.00±0.14	98.92
17	**Δ**MEK1**C218**	0.100±0.004	0.010±0.003	97.95
18	**Δ**MEK1**Dha218**	0.700±0.048	0.201±0.011	96.57

The spectra are attached in Appendix F.

7.7.3.Optimisation of the assay

Temperature

The temperature of the assay was set at 30 °C as it was controllable in heated shaker.

Quenching considerations

Different quenching methods were used. Use the HPLC autosampler as a mean to quench the reaction was practical. The autosampler could be at the right temperature. However, it was limited to the optimal reaction concentration. Furthermore, quick contamination of the column occurred due to the high concentration of all reagents. The quenching was also dependent on the run time. Similarly to the assay set for p38α[96], urea and guanidine effectively quenched the reaction in a reproducible manner. However, adducts were detected during LCMS analysis, which hamper the processing/analysis of the data.

50 mM EDTA solution at pH 8 solution was therefore used as no adduct was detected.

Concentrations of MEK1 and ERK1

The concentration of ERK1 was dictated by the detection limits of the mass spectrometer. Typically, a minimum 200 ng of protein was detectable, which corresponded to 10 µL injection of a 0.5 µM, 0.02 mg/mL ERK1 solution.

7.8. Active MEK1 crystallisation

7.8.1. Attempt of crystallisation of MEK1pC222-ERK2-AMPNP[‡‡‡‡‡‡]

5 mg of MEK1pC222 were given to Dr. Apirat Chaikuad, SGC Oxford. The complex with ERK2 was purified using size exclusion chromatography. The crystallisation screen was set up at the SGC.

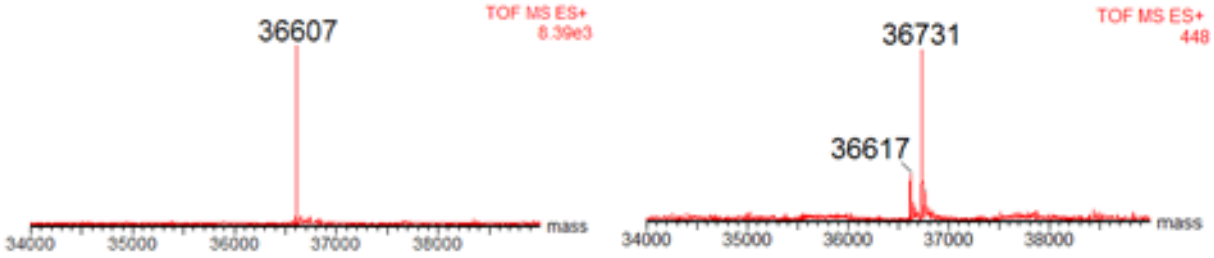
7.8.2. Crystallisation of truncated ΔMEK1 constructs

Formation of ΔMEK1pC222

ΔMEK1C222 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.64 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.4 mM of MDBP. The mixture was shaken

[‡‡‡‡‡‡] Work performed in collaboration with Dr. Apirat Chaikuad from the Structural Genomic Consortium, SGC, Oxford.

at 600 rpm at room temperature. The reaction was monitored using LCMS-ESI-TOF. Full conversion to ΔMEK1Dha222 ($M_{obs.}$=36607, $M_{calc.}$=36611) was observed after 8 hours. ΔMEK1Dha222 was buffer exchanged using a trap column in 20 mM Tris, 100 mM NaCl, 10 mM ATP, pH 8 buffer was treated with a 3.8 M stock solution of sodium thiophosphate (see above MEK1pC222 formation) to afford a final concentration of 0.5 M sodium thiophosphate solution and 19 µM of ΔMEK1Dha222. The mixture was left standing at room temperature. The reaction was monitored by LCMS-ESI-TOF. Full conversion to MEK1pC222 ($M_{obs.}$=36731, $M_{calc.}$=36725) was observed after 18 hours.



Formation of ΔMEK1pC218

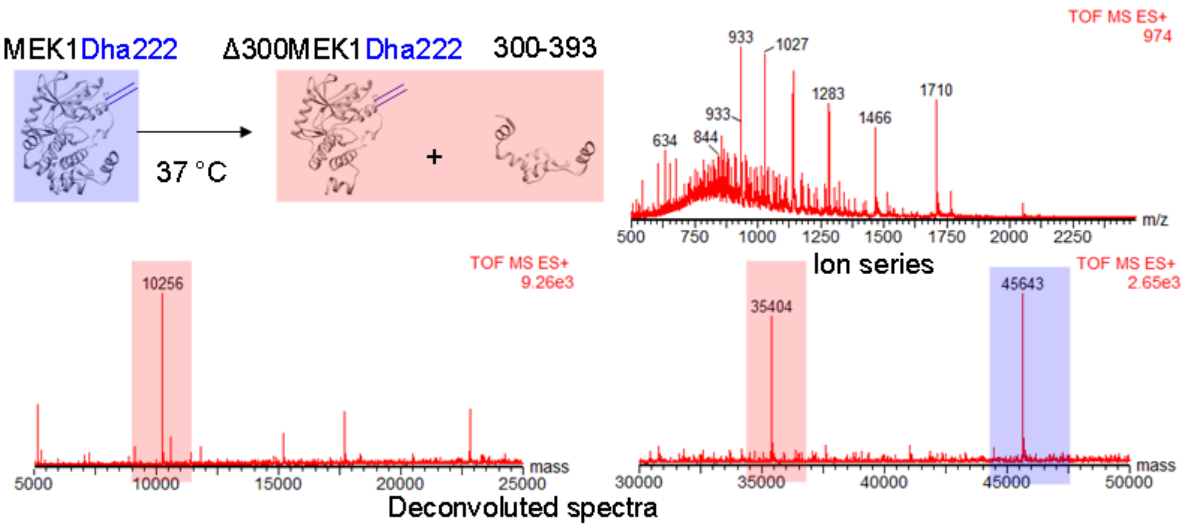
MEK1C218 (44 µM, 2 mg/mL) in 20 mM Tris, 100 mM NaCl, 10 mM ADP, 25 mM $MgCl_2$, pH 8 buffer was treated with a 0.64 M stock solution of 2,5-dimethyldibromopentanoate, MDBP in DMSO to afford a final concentration of 4.4 mM of MDBP. The mixture was shaken at 600 rpm at 37 °C.[§§§§§§] The reaction was monitored using LCMS-ESI-TOF. Full conversion to ΔMEK1Dha218 ($M_{obs.}$=36616, $M_{calc.}$=36611) was observed after 12 hours.

7.8.3. Proteolytic cleavage of MEK1 at Y300

Cleavage of full length MEK1 at Y300 was noticed while doing reactions, either during Dha formation or phosphorylation with sodium thiophosphate.[248] This cleavage is batch

§§§§§§ Precipitation tends to occur at 37 °C. The reaction can be performed at room temperature for around 8-10 h.

dependent (not every batch cleaves). The cause was not determined. Protease inhibitors have been used without success.



Chapter 8: Procedures for the Sulfonium Elimination Mechanistic Elucidation on Protein

8.1. Protein reaction monitoring and characterisation

Four different LC- MS-TOF systems were used in this study.

LCT classic system (Chapter 7. Paragraph 7.2.1)

LCT premier system (Chapter 7. Paragraph 7.2.1)

Xevo system

Instrumentation

A Xevo G2-S Q-ToF high resolution mass spectrometer (Waters, Wilmslow, UK) was coupled with an Acquity UPLC system. The column was a Chromolith RP18 2x50mm (Merck). Elution solvents were: (A) deionised water (18MΩ resistivity) with 0.1% (v/v) formic acid (analytical grade, Sigma UK), and (B) Acetonitrile (HPLC grade, Sigma UK), isocratic 95% A for 0.5 minutes followed by gradient to 5% A for 2.5 minutes, isocratic for 0.5 minutes, then gradient for 0.05 minutes to 95% A followed by isocratic for 1.45 minutes. Instrument control, data acquisition, and data processing were performed using Masslynx 4.1 software (Chapter 7, paragraph 7.2.1). The mass spectrometer was operated in positive electrospray mode, source conditions were adjusted to maximise sensitivity. Lockspray ™ was used during analysis to maintain mass accuracy.

Deuterium incorporation estimation using LCMS on intact protein

Aliquot of 0.5 mg/mL aliquots of the desired protein were prepared in water. The aliquots were analysed using Xevo G2-S Q-Tof. The data was process using MaxLynx v4.1. Deconvolution was set up: output mass, ranges 10000:20000, resolution 0.04 Da/channel;

damage model, uniform Gaussian width at half height 0.5 Da, minimum intensity ratios left 33% and right 33% and completion options, iterate to convergence. 2 independent runs were performed on each sample.

Percentage deuterium incorporation (%D) on intact protein was determined as follow:

$M_{obs.}=M_{deut.}\times\%D+M_{non\text{-}deut.}\times(1-\%D)$

with $M_{obs.}$= Mass observed, $M_{deut.}$= Mass of deuterated protein, %D= percentage of deuterated protein or percentage deuterium incorporation, $M_{non\text{-}deut.}$= mass of non-deuterated protein.

This equation was equivalent to:

$\%D=(M_{obs.}-M_{non\text{-}deut.})/(M_{deut.}-M_{non\text{-}deut.})$

As $M_{deut.}-M_{non\text{-}deut.}=1$, the percentage incorporation could be estimated as $\%D=M_{obs.}-M_{non\text{-}deut.}$.

SQD system

A Thermo Scientific uPLC was coupled to MSQ Plus Single Quad Detector (SQD). Column: Hypersil Gold C4, 1.9 µm, 2.1 x 50 mm. Wavelength: 254 nm. Mobile Phase: 99:1 Water (0.1% formic acid): MeCN (0.1% formic acid) to 1:9 Water (0.1% formic acid): MeCN (0.1% formic acid) gradient over 4 min. Flow Rate: 0.3 mL/min. MS Mode: ES+. Scan Range: m/z = 500-2000. Scan time: 1.5 s. Data obtained in continuum mode. The electrospray source of the MS was operated with a capillary voltage of 3.5 kV and a cone voltage of 50 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L/h. Ion series were generated by integration of the total ion chromatogram (TIC) over the 3.4-4.8 min range. Total mass spectra for protein samples were reconstructed from the ion series using the pre-installed ProMass software using default settings for large proteins in m/z range 500-1500.

LCMS-MS

LCMS-MS on tryptic peptides was performed accordingly to the procedures described Chapter 7, Paragraph 7.2.2). LCMS-MS allowed the confirmation of the deuteration site. Further analysis of the LC trace of the peptide containing the deuterated cysteine confirmed the level of deuterium incorporation.

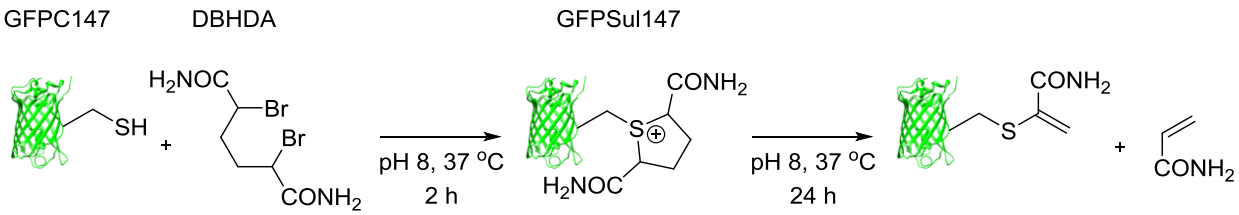
8.2. Chemical protein modification towards the elucidation of the mechanism of elimination in Dha formation

8.2.1. General protein and reagents considerations

The proteins H3, SBL and GFP were kindly provided Dr. Billy Ng (Oxford) and Dr. Vijay Chudasama (UCL). PP1 was bought (Biolabs, P0754S). Deuterated solvents and reagents were provided by Aldrich.

8.2.2. Decomposition of GFPSul147

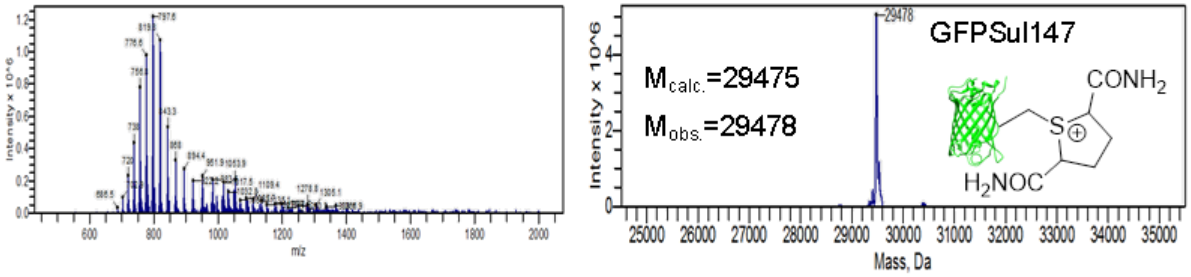
General Scheme for the decomposition of GFPSul147



Formation of GFPSul147

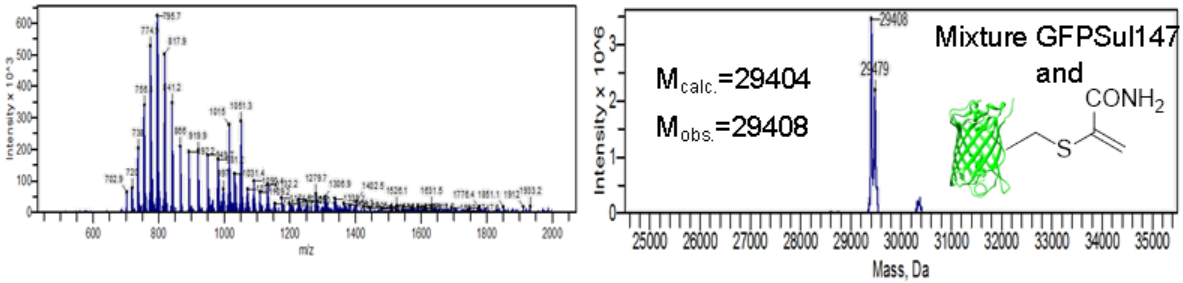
Dithiothreitol (10 µL, 340 mM as a solution in water, 1000 equivalents per 100 µL of GFP(S147C) (1 mg/mL solution)) was added, incubated at 21 °C for 1 h and the excess reducing agent removed by repeated diafiltration into fresh buffer (100 mM sodium phosphate, pH 8.0) using VivaSpin sample concentrators (10,000 MWCO). Following this, 2,5-dibromohexanediamide (10 µL, 17 mM solution in DMF, 50 equivalents) was added to a

solution of reduced GFPC147 (1.0 mg/mL) in sodium phosphate (100 mM, pH 8.0) at 21 °C. The mixture was vortexed for 1 s, maintained at 37 °C for 2 h and analysed by LCMS (SQD system) to reveal formation of GFPSul147.



Decomposition of GFPSul147

GFP(S147C, sulf) (100 µL, 17 mM solution in sodium phosphate (100 mM, pH 8.0)) was incubated at 37 °C. The mixture was monitored by LCMS. After 24 h, recordings were not clean due to partial degradation of protein.



8.2.3. Formation of perdeuterated SBLDha156

Preparation of perdeuterated[*******] SBLC156

The exchangeable protons of the protein were first deuterated: lyophilized SBLC156 was diluted in deuteroxide and lyophilized again. The procedure was repeated five times to afford perdeuterated SBLC156.

[*******] In our case, perdeuterated protein is defined by the full deuteration of the exchangeable protons of the protein.

Preparation of deuterated buffer

A 50 mM sodium phosphate pH 8 buffer was lyophilized. The solid was diluted in deuteroxide, and lyophilized. The procedure was repeated five times to afford deuterated 50 mM sodium phosphate pH 8.

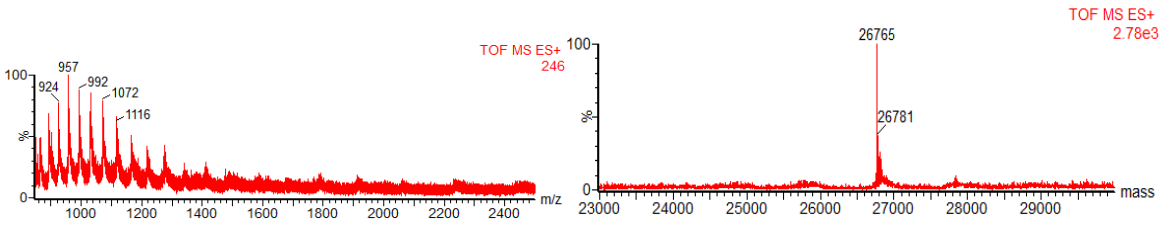
Procedure for the formation of perdeuterated SBLDha156

Perdeuterated SBLS156C (5 mg) in deuterated 50 mM sodium phosphate buffer pH 8 (reaction buffer) was treated with a 1 M solution of DTT in reaction buffer. The mixture was shaken at 600 rpm at room temperature. The mixture was desalted using a Miditrap preequilibrated with reaction buffer to afford 1.5 mL of a 2 mg/mL solution. The solution was treated with dry DBHDA in solution in dry DMF (200 Eq. in 1 µL of DMF). The mixture was shaken at 600 rpm at room temperature for 30 minutes and at 37 °C for 2 hours. The reaction was monitored by LCMS.

8.2.4. Formation of SBLdAcc156 and attempt of deprotection to SBLdC156

Formation of SBLdAcc156

To a solution of SBLDha156 in 50 mM deuterated sodium phosphate buffer (pH 8, 1.90 mg/mL) was added an aqueous solution of potassium thioacetate (0.723 mg). The solution was shaken at room temperature for 1 hour followed by 37 ^{o}C for 6 hours. The reaction was monitored by LCMS to ensure complete conversion. Potassium thioacetate was removed by filtering twice through a PD SpinTrap™ G-25 column pre-equilibrated with 50 mM sodium phosphate buffer, pH 8) to afford SBLdAcc156.

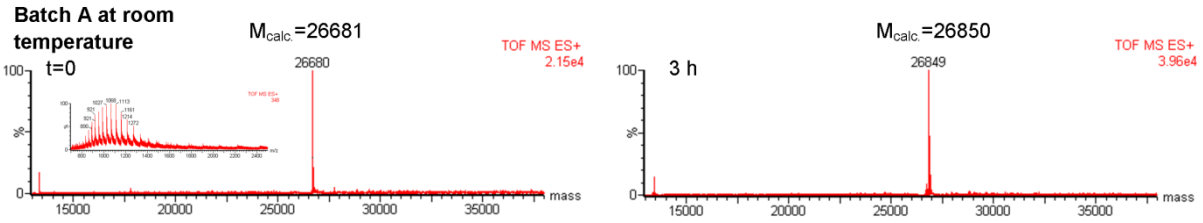


Attempt deprotection of SBLdAcc156 to SBLdC156

SBLdAcc156 in 50 mM sodium phosphate buffer (pH 8, 1 mg/mL) was treated with 1 µL of methoxyamine hydrochloride in 50 mM sodium phosphate buffer (pH 8, 80 mg/mL). The reaction was shaken at room temperature for 4 h. The reaction was monitored by LCMS. Despite disappearance of starting material, deprotection of the acetate was not observed.[193]

8.2.5. Formation of SBLNbc156

SBLDha156 was desalted using a pre-equilibrated PD SpinTrap cartridge G-25 to afford 180 µL of a 0.75 mg/mL solution of SBLS156Dha in 50 mM sodium phosphate buffer pH 8. The mixture was aliquoted in 2x80 µL batches A and B. Both batches were treated with 1.3 mg of 2-nitrobenzyl carbamimidothioate hydrobromide (1.3 mg, 0.44 mmol, 2000 Eq.). A was shaken at 700 rpm at room temperature, B at 37 °C. The reactions were monitored by LCMS.



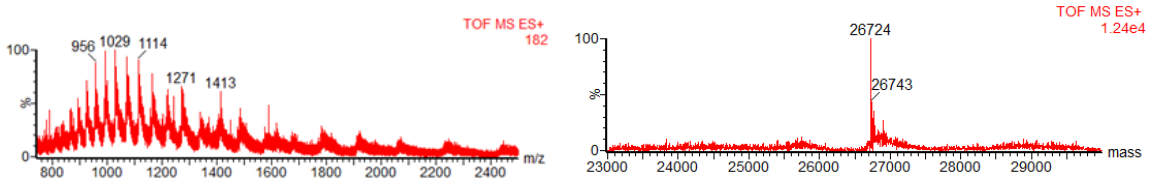
8.2.6. Formation of SBLdC156

Formation of SBLdC156

To a solution of perdeuterated SBLDha156 in 0.1 M deuterated sodium phosphate buffer (pH 7, 1.43 mgs/mL, 0.15 mL) was added an aqueous solution of Na_2S (6 µL, 50 µg/µL, dissolved in pH 7 deuterated sodium phosphate buffer). The solution was briefly vortexed and shaken at 15 ^{o}C for 10 mins. At 10, 20, 30, 40, and 50 mins, 6 µL of the Na_2S solution was added. The solution was kept shaking in between each addition. After 10 mins of the

last addition, the reaction mixture was analysed by protein MS to ensure complete conversion. Na_2S was removed by filtering through a Micro Bio-Spin™ 6 Columns (Bio-Rad) (pre-equilibrated with 50 mM sodium phosphate buffer, pH 8). The filtrate was filtered again using another pre-equilibrated Micro Bio-Spin™ 6 Columns.

For protein MS analysis, 3 uL of the reaction mixture was taken and transferred to a MS vial. 50 uL of MQ-water and 3 uL of 1M HCl solution were then added. The mixture was then analyzed by LC-MS. Full conversion to SBLdC156 was observed. Addition of HCl solution to the sample before MS analysis was essential to obtain an informative MS spectrum.



LCMS-MS of the digested protein peptides was not conclusive.[†††††††]

8.2.7. Formation of H3dC10

Formation of perdeuterated H3Dha10

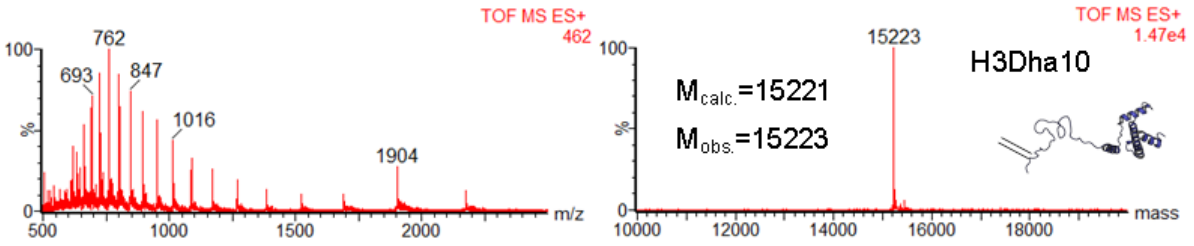
Perdeuterated H3C10 was prepared as following. The exchangeable protons of the protein were first deuterated: lyophilized H3C10 was diluted in deuteroxide and lyophilized again. The procedure was repeated five times to afford perdeuterated H3C10.

Deuterated buffer was prepared following the procedure described in paragraph 8.6.3.

The perdeuterated H3C10 (5 mg) was solubilised in deuterated 50 mM Tris, 5 M GdmCl, pH 8 buffer (500 µL). DTT solid (1 mg) was added to the solution to reduce the dimer and the β-mercaptoethanol adduct formed during purification and storage. The mixture was shaken at 600 rpm at room temperature for 30 minutes. The mixture was desalted using PD-Minitrap™ G25 preequilibrated with reaction buffer to afford 1 mL of solution. A stock solution of dry

††††††† Digest and characterisation by LCMS-MS was performed by Marie-Laetitia Thevenaz from Prof. Benedikt Kessler group, Oxford.

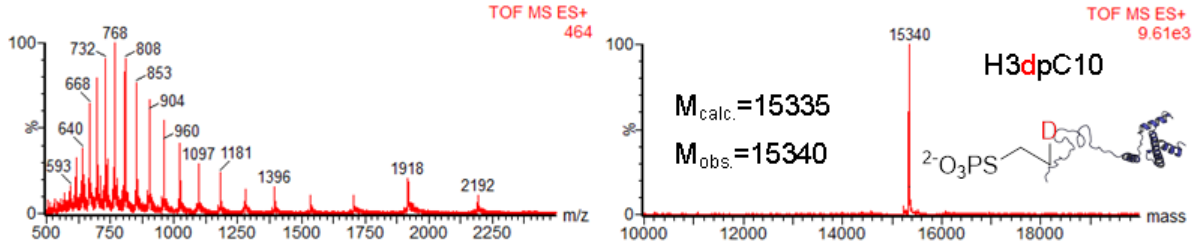
DBHDA in dry DMF was added. The mixture was shaken at 600 rpm at room temperature for 30 minutes before heating up to 37 °C for typically 3 hours to afford perdeuterated H3Dha10. The reaction was monitored using LCMS (Method A).



Formation of H3dpC10 in deuterated buffer

Deuterated thiophosphate was first prepared. Anhydrous sodium thiophosphate was dissolved in deuterated 5 M guanidinium hydrochloride solution. The pH of the solution was acidified to pH 8.7 with a deuterated DCl solution to afford a deuterated 0.4 M Na_3SPO_3, 5 M GdmCl pH 8.7 buffer.

Perdeuterated H3Dha10 prepared previously (500 µL) was desalted and buffer exchanged using PD-Minitrap G25 preequilibrated with to afford 1 mL of solution. The mixture was shaken at 600 rpm at 37 °C for typically 2 hours. One aliquot of the mixture (500 µL) was desalted and buffer exchanged using PD-Minitrap™ G25 preequilibrated with a 50 mM Tris, 5 M GdmCl pH 8 buffer to afford 1 mL of solution containing perdeuterated H3dpC10 used directly without further purification. The reaction was monitored using LCMS.

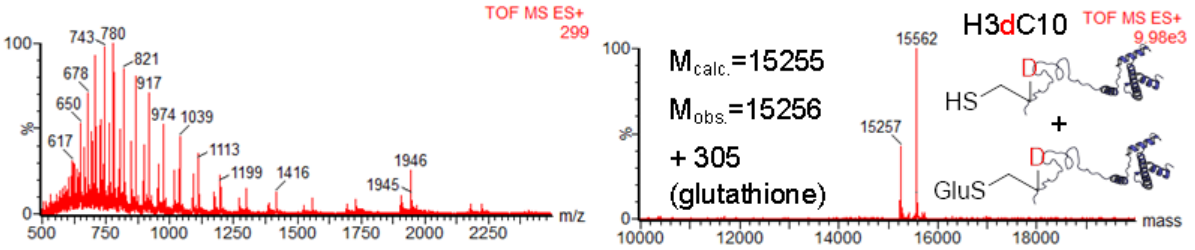


Formation of H3dC10

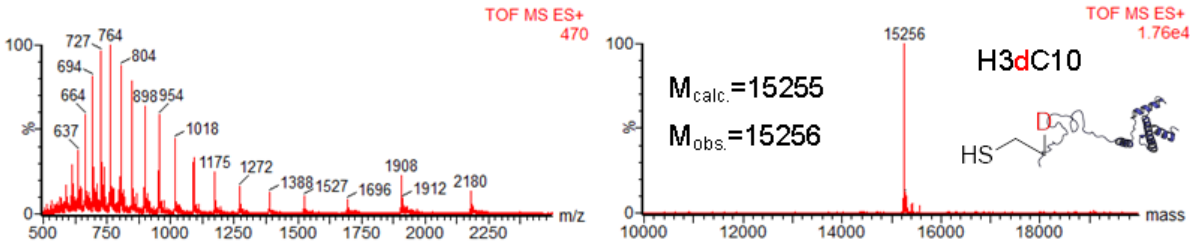
Two distinct methods were used for the phosphocysteine hydrolysis:

Method A: acid hydrolysis

H3dpC10 in 50 mM Tris, 5 M GdmCl pH 8 was acidified to pH 4 with addition of glutathione. The mixture was shaken at 600 rpm at 37 °C for typically 1 hour to afford the desired H3dC10. The reaction was monitored by LCMS. The solution was a mixture of free and glutathione (Glu) adduct (+305) H3dC10.

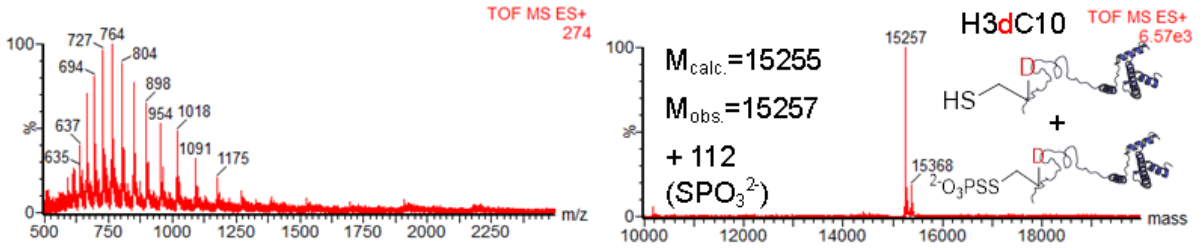


Treatment with a 10 mM DTT aqueous solution and shaking at 600 rpm at room temperature for 30 minutes at afforded the desired H3dC10.

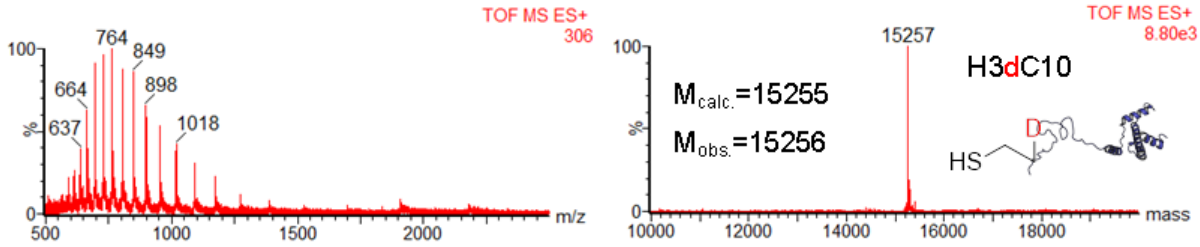


Method B: Dephosphorylation with protein phosphatase

The mixture was buffer exchanged using a PD-Spintrap G25 equilibrated with 50 mM Tris, 100 mM NaCl, pH 8 buffer. The enzymatic reaction was set up according to PP1 provider: 4 µL of H3dpC10 (1mg/mL)+ 4 µL of NE buffer (x10)+ 4 µL of MnCl2 buffer (x10)+0.4 µL PP1 (2500 U/mL). The volume was adjusted to 40 µL with Milli-Q® water and incubated at 30 °C. The reaction was monitored by LCMS. The reaction was completed after 30 minutes for H3dC10. An adduct was observed by LCMS corresponding to thiophosphate (+112)



Treatment with a 10 mM DTT aqueous solution afforded the desired H3dC10.

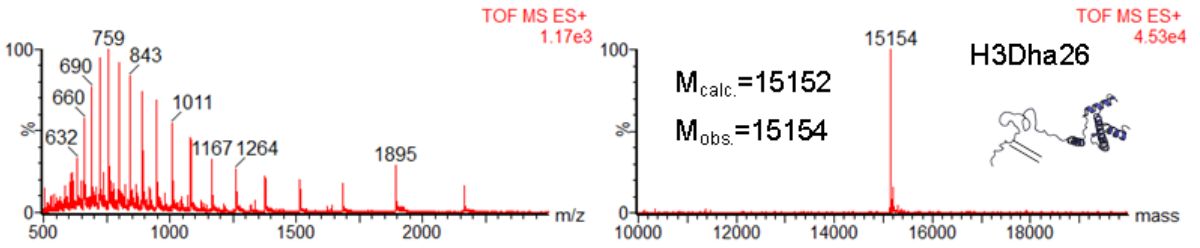


In both cases, the mixtures were then dialysed using refolding buffer 20 mM Tris, 250 mM KCl, 1 mM EDTA to afford the desired H3dCX and DTT was added to afford a 20 mM Tris, 250 mM KCl, 1 mM EDTA, 5 mM DTT solution to afford mono deuterated H3dC10, which was characterised by HR-LCMS and LCMS-MS method see appendix H.

8.2.8. Formation of H3dC26

Formation of perdeuterated H3Dha26 in deuterated buffer

Perdeuterated H3Dha26 was prepared according to the procedure described for the formation of perdeuterated H3Dha10.



Formation of H3dpC26 in deuterated buffer

Perdeuterated H3dpC26 was prepared according to the procedure described for the formation of perdeuterated H3dpC10. LCMS was not conclusive.

Formation of H3dC26

Monodeuterated H3dC26 was prepared according to the procedure described for the formation of monodeuterated H3dC10. In method B, the full dephosphorylation of H3dpC26 to H3dC26 was observed after 2 hours. For HRMS and MSMS see appendix H.

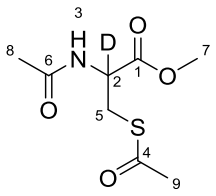
8.2.9. Optimisation of the deuterium incorporation

For optimisation see Appendix H.

8.3. Small molecule synthesis used in the site-selective chemical phosphorylation of protein

8.3.1. Synthesis of AcNHdCysOMe

Synthesis of methyl 2-(acetylamino)-3-(acetylsulfanyl)(2-^{2}H)propanoate, AcNHAccOMe-d1

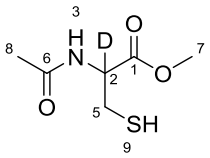


A suspension of potassium thioacetate (597 mg, 5.2 mmol) in deuteroxide (5 mL) were treated with a 35% sodium deuteroxide solution up to pH 7 at 0 °C. Methyl 2-(acetylamino)prop-2-enoate (AcNHDhaOMe, 500 mg, 3.5 mmol) was added at 0 °C. The mixture was allowed to warm to room temperature under stirring for 2 hours. The reaction was the stirred at 37 °C for 20 hours. The mixture was extracted twice with ethyl acetate (5 mL). The organic layers were combined, dried over sodium sulphate and evaporated

under reduced pressure. The residue was purified by column chromatography (eluant dichloromethane:methanol from 0 to 10% methanol) to afford 253 mg (33 % yield) of the title compound as a pale beige solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.02 (s, 3 H, H8), 2.35 (s, 3 H, H9), 3.35 (dd, J=19.32, 18.10 Hz, 2 H, H5), 3.71 - 3.82 (m, 3 H, H7), 6.29 (br. s., 1 H, H3). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 22.99 (s, C7), 30.43 (s, C9), 30.80 (s, C5), 51.43 - 52.15 (m, C2), 52.75 (s, C8), 169.93 (s, C1), 170.64 (s, C6), 195.29 (s, C4). HRMS m/z (ESI+): Found 243.0526 [M+Na]+; $C_8H_{12}{}^2H\ O_4N^{23}Na^{32}S$ requires 243.0520. IR (v_{max} cm^{-1}) 3387 (N-H), 1731 (C=O ester), 1701 (C=O thioester), 1634 (C=O, amide).

Synthesis of methyl 2-(acetylamino)-3-sulfanyl(2-^{2}H)propanoate, AcNHdCysOMe



Method A:

Methyl 2-(acetylamino)-3-(acetylsulfanyl)(2-^{2}H)propanoate (AcNHdAccOMe, 100 mg, 0.46 mmol) was dissolved in methanol (3 mL). Potassium carbonate (63 mg, 0.46 mmol) was added. The mixture was stirred at room temperature for 45 minutes. The mixture was treated with 1 N hydrochloric acid solution (2 mL). The organics were extracted twice with dichloromethane (2x3 mL). The organic layers were combined, dried over sodium sulphate and evaporated under reduced pressure. The residue was purified by column chromatography (eluant: dichloromethane:methanol from 0 to 10% methanol) to afford 72 mg (88% yield) of the title compound as clear oil.

Method B:

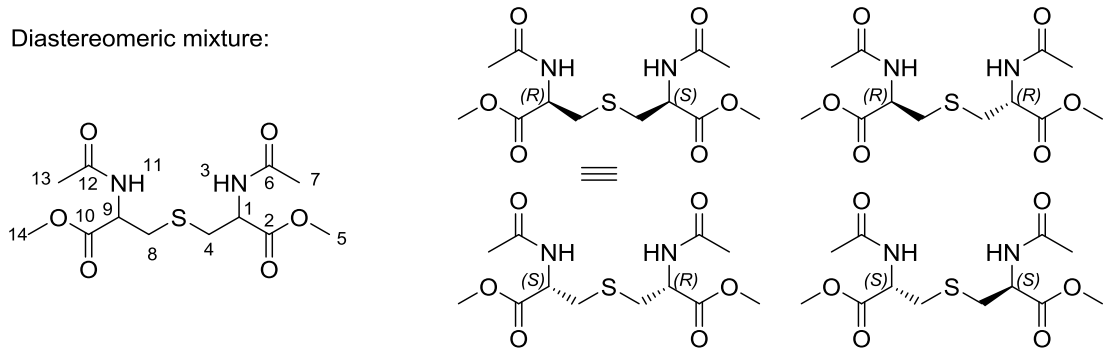
Commercially available sodium thiophosphate hydrate (5 g) was stirred in methanol for 1 h. The solid was filtered and heated up to 100 °C for 1 h to afford 2.6 g of anhydrous sodium

thiophosphate as a white solid, which was stored under argon at room temperature.[203] Anhydrous sodium thiophosphate (144 mg, 0.8 mmol) was solubilised in deuterium oxide (1.5 mL). The pH was adjusted to 9 with a 5 M deuterium oxide solution of deuterium chloride. The pH was then adjusted to pH 8.7 with a 1 M deuterium oxide solution of deuterium chloride. The volume was adjusted to 2 mL to afford a 0.4 M solution of sodium thiophosphate. AcNHDhaOMe (75 mg, 0.52 mmol) was added to the solution. The mixture was stirred at room temperature for 2 hours. The reaction was monitored by ^{1}H NMR. After full conversion to AcNHdpCysOMe (not isolated), the pH of the solution was adjusted to pH 2-3 using a 6 M hydrochloric solution. The mixture was then shaken at 37 °C at 600 rpm for 4 hours. The reaction was monitored by ^{1}H NMR. After full conversion to AcNHCysOMe-d1, the mixture was extracted three times with dichloromethane (2 mL). The organic layers were combined, dried over sodium sulphate and evaporated under reduced pressure to afford 56 mg of the title compound as a clear oil.

^{1}H NMR (400 MHz, D_2O pH 8.7) δ ppm 2.01 (s, 3 H), 3.10 (d, J=12.23 Hz, 2 H), 3.73 (s, 3 H). ^{31}P NMR (162 MHz, D_2O pH 8.7) δ ppm 16.02 (s, 1 P). ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 1.37 (t, J=8.93 Hz, 1 H, H9), 2.04 (s, 3 H, H8), 2.97 (dd, J=8.93 Hz, 1.59 Hz, 2 H), 3.77 (s, 3 H, H7), 6.58 (br. s., 1 H, H3). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 22.97 (s, C7), 26.64 (s, C5), 52.69 (s, C8), 52.98 - 53.65 (m, C2), 169.96 (s, C1), 170.54 (s, C6). LRMS ESI 257 [M-H]-. HRMS m/z (ESI+): Found 201.04120 [M+Na]+; $C_6H_{10}{}^{2}HO_3N^{23}Na^{32}S$ requires 201.04146. IR (v_{max} cm^{-1}) 3293 (NH, amide), 1740 (C=O, ester), 1652 (C=O, amide).

8.3.2. Formation of AcNHLanOMe from the reaction of AcNHDhaOMe with Na_2S

Synthesis of dimethyl 3,3'-thiobis(2-acetamidopropanoate), AcNHLanOMe

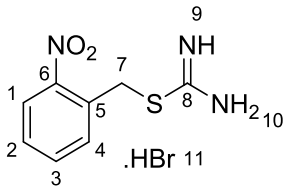


Methyl 2-(acetylamino)prop-2-enoate (AcNHDhaOMe, 50 mg, 0.35 mmol) in water (1 mL) was treated with sodium sulphide (27 mg, 0.35 mmol). The mixture was stirred at room temperature for 30 minutes. The reaction mixture was extracted twice with dichloromethane (1 mL). The organic layers were combined, dried over sodium sulfate and evaporated under reduced pressure to afford 33 mg (60% yield) of the title compound as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.07 - 2.11 (m, 6 H, H7, H13), 2.90 - 3.10 (m, 4 H, H4, H8), 3.76 - 3.79 (m, 6 H, H5, H14), 4.77 - 4.86 (m, 2 H, H1, H9), 6.48 - 6.75 (m, 2 H, H3, H11). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 22.96 (s, C7, C13), 23.00 (s, C7, C13), 34.76 (s, C4, C8), 35.33 (s, C4, C8), 51.95 (s, C5, C14), 52.09 (s, C5, C14), 52.77 (s, C1, C9), 170.34 (s, C2, C10), 170.40 (s, C2, C10), 170.83 (s, C6, C12), 170.96 (s, C6, C12). LRMS ESI 343 [M+Na]+. HRMS m/z (ESI+): Found 343.09314 [M+Na]+; $C_{12}H_{20}O_6N_2{}^{23}Na^{32}S$ requires 343.09343. IR (v_{max} cm^{-1}), 3302 (NH, amide), 1742 (C=O, ester), 1647 (C=O, amide), 1542 (C=O, amide).

8.3.3. Formation of AcNHdNbcOMe from AcNHDhaOMe with a photolabile thiol

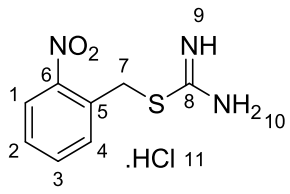
2-nitrobenzyl carbamimidothioate hydrobromide[264]



1-(bromomethyl)-2-nitrobenzene (1.00 g, 4.63 mmol) and carbamimidothioic acid (387 mg, 5.09 mmol) was stirred under reflux in ethanol (10 mL) for 4 hours. The mixture was evaporated under reduced pressure. The residue was triturated in ethanol. The solid was filtered and washed with ethanol to afford 461 mg (34% yield) of 2-nitrobenzyl carbamimidothioate hydrobromide as a yellowish solid.

^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 4.78 (s, 2 H, H7), 7.60 - 7.70 (m, 1 H, H2), 7.72 - 7.87 (m, 2 H, H3, H4), 8.13 (d, *J*=7.93 Hz, 1 H, H1), 9.19 (br. s., 4 H, H9, H10, H11). ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 32.06 (s, C7), 125.63 (s, C4), 129.94 (s, C2), 130.39 (s, 3C), 131.89 (s, C1), 134.44 (s, C1), 147.94 (s, C6), 168.52 (s, C8). ES-API MS 212 [M+H]+.

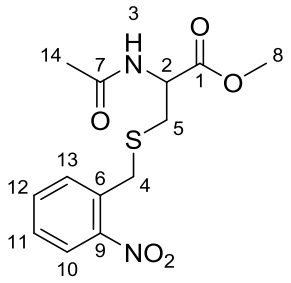
2-nitrobenzyl carbamimidothioate hydrochloride[265]



1-(chloromethyl)-2-nitrobenzene (5.00 g, 29.14 mmol) and carbamimidothioic acid (2.44 g, 32.05 mmol) was stirred under reflux in ethanol (50 mL) for 4 hours. The mixture was evaporated under reduced pressure. The residue was triturated by sonication in acetonitrile (50 mL). The solid was filtered and washed with acetonitrile to afford 6.48 g (90% yield) of the title compound as a light grey solid.

^{1}H NMR (400 MHz, DMSO-d_6) δ ppm: 4.85 (s, 2 H, H7), 7.64 (t, *J*=7.63 Hz, 1 H, H2), 7.73 - 7.90 (m, 2 H, H3, H4), 8.13 (d, *J*=8.09 Hz, 1 H, H1), 9.46 (br. s., 4 H, H9, H10, H11). ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 31.81 (s, C7), 125.63 (s, C4), 129.88 (s, C2), 130.71 (s, C3), 132.03 (s, C1), 134.41 (s, C5), 147.87 (s, C6), 168.72 (s, C8). ES-API MS 212 [M+H]+.

Methyl 2-(acetylamino)-3-[(2-nitrobenzyl)sulfanyl]propanoate, AcNHNbcOMe

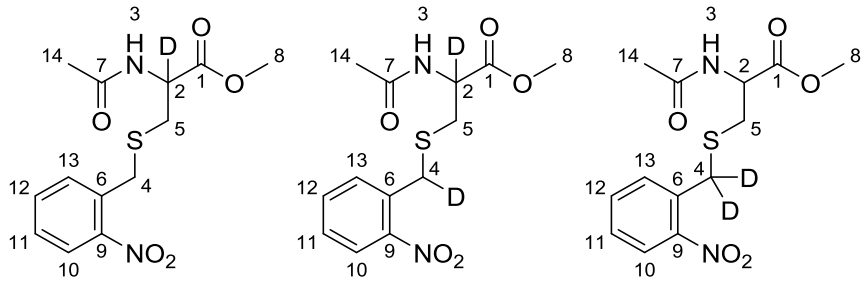


To a solution of potassium carbonate (96 mg, 0.70 mmol) in water (1 mL) was added 2-nitrobenzyl carbamimidothioate hydrochloride (86 mg, 0.35 mmol) followed by methyl 2-(acetylamino)prop-2-enoate (AcNHDhaOMe, 50 mg, 0.35 mmol) in water (0.5 mL). The suspension was shaken at 800 rpm at 37 °C for 2.5 h . The mixture was partitioned between water (1.5 mL) and ethyl acetate (3 mL). The organic layer was separated. The aqueous layer was extracted with ethyl acetate (3 mL). The organic layers were combined, washed with brine (3 mL) and evaporated under reduced pressure. The residue was purified by silica gel column chromatography (eluant: dichloromethane/methanol from 0 to 5% methanol). Evaporation of the fractions containing the desired fractions afforded 42 mg (38% yield) of the title compound as a clear oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.05 (s, 3 H, H14), 2.88 - 3.00 (m, 2 H, H5), 3.76 (s, 3 H, H8), 4.04 - 4.15 (m, 2 H, H4), 4.80 (dt, *J*=7.58, 5.14 Hz, 1 H, H2), 6.35 (d, *J*=7.09 Hz, 1 H, H3), 7.39 - 7.49 (m, 2 H, H10, H11), 7.54 - 7.62 (m, 1 H, H12), 8.00 (dd, *J*=8.44, 1.34 Hz, 1 H, H13). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 23.04 (s, C14), 34.09 (s, C5), 34.15 (s, C4), 51.87 (s, C2), 52.77 (s, C8), 125.52 (s, C13), 128.53 (s, C11), 132.02 (s, C12), 133.19 (s, C10), 133.34 (s, C6), 148.53 (s, C9), 170.01 (s, C1), 171.00 (s, C14). HRMS m/z (ESI+): Found 335.06696 [M+Na]+; $C_{13}H_{16}O_5N_2{}^{23}Na{}^{32}S$ requires 335.06721. IR (v_{max} cm^{-1}) 3307

(NH, amide), 1744 (C=O ester), 1650 (C=O, amide), 1523 (C-N=O, nitro), 1342 (C-NO_2, nitro).

Methyl 2-(acetylamino)-3-[(2-nitrobenzyl)sulfanyl](1-^2H_x), AcNHd$_x$NbcOMe



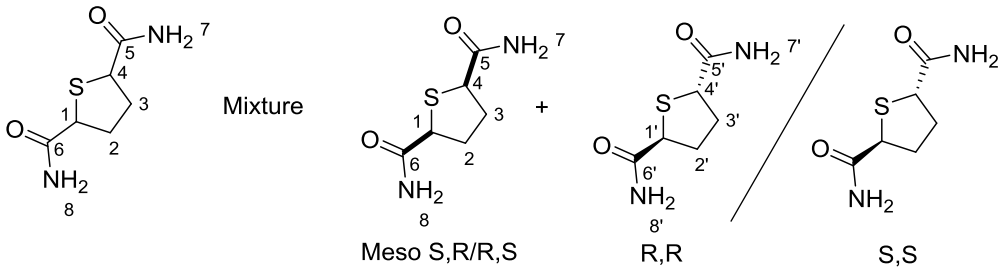
To a solution of potassium carbonate (96 mg, 0.70 mmol) in deuterium oxide (1 mL) was added 2-nitrobenzyl carbamimidothioate hydrochloride (86 mg, 0.35 mmol) followed by methyl 2-(acetylamino)prop-2-enoate (50 mg, 0.35 mmol) in water (0.5 mL). The suspension was shaken at 800 rpm at 37 °C for 2 h 30. The mixture was partitioned between water (1.5 mL) and ethyl acetate (3 mL). The organic layer was separated. The aqueous layer was extracted with ethyl acetate (3 mL). The organic layers were combined, washed with brine (3 mL) and evaporated under reduced pressure. The residue was purified by silica gel column chromatography (eluant: dichloromethane/methanol from 0 to 5% methanol). Evaporation of the fractions containing the desired fractions afforded 42 mg (38% yield) of the title compound as a clear oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.05 (s, 3 H, H14), 2.92 (dd, J=22.25, 8.31 Hz, 2 H, H5), 3.75 (s, 3 H, H8), 4.00 - 4.16 (m, 1 H, H4), 6.37 (br. s., 1 H, H3), 7.40 - 7.49 (m, 2 H, H10, H11), 7.58 (td, J=7.58, 1.22 Hz, 1 H, H12), 8.00 (d, J=8.31 Hz, 1 H, H13). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 23.01 (s, C14), 33.84 - 34.02 (m, C5), 34.11 (s, C4), 51.20 - 52.03 (m, C2), 52.74 (s, C8), 125.51 (s, C13), 128.52 (s, C11), 132.01 (s, C12), 133.19 (s, C10), 133.23 - 133.38 (m, C6), 148.49 (s, C9), 170.03 (s, C1), 170.99 (s, C7). HRMS m/z (ESI+): Mixture 25% Found 314.09177 [M+H]+; $C_{13}H_{16}{}^2HO_5N_2{}^{32}S$ requires 314.09155/ 28% Found

315.09784 [M+H]+; $C_{13}H_{15}{}^{2}H_{2}O_{5}N_{2}{}^{32}S$ requires 315.09782/ 47% Found 316.10403 [M+H]+; $C_{13}H_{14}{}^{2}H_{3}O_{5}N_{2}{}^{32}S$ requires 316.10410. IR (v_{max} cm^{-1}) 3307 (NH, amide), 1741 (C=O ester), 1650 (C=O, amide), 1523 (C-N=O, nitro), 1342 (C-NO_2, nitro).

8.3.4. Formation of THTCD

Tetrahydrothiophene-2,5-dicarboxamide (THTCD)



2,5-dibromohexanediamide (100 mg, 0.33 mmol) suspended in ethanol (1 mL) was treated with anhydrous sodium sulphide (99 mg, 0.41 mmol) and potassium iodide (6.6 mg, 0.039 mmol). Water (66 µL) was added and the mixture was stirred at room temperature for 6 hours. The reaction mixture was filtered. The filtrate was evaporated under reduced pressure to afford 63 mg (quant.) of a 6:4 mixture of both diastereoisomers (S, R meso and RR/SS) tetrahydrothiophene-2,5-dicarboxamide as a white solid.

^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 1.88-2.27 (m, 4 H, H2, H3, H2', H3'), 3.77-3.95 (m, 2 H, H1, H4, H1', H4'), 6.95 (br.s, 1/3 H, H7), 7.15 (br.s, 2/3 H, H7'), 7.43 (br.s, 1/3 H, H8), 7.67 (br.s, 2/3 H, H8'). ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 33.3 (s, C2, C3), 34.2 (s, C2', C3'), 49.4 (s, C1, C4), 50.4 (s, C1', C4'), 173.3 (C5, C6), 174.4 (C5', C6'). ES-API MS 191 [M+Na]+.]+. HRMS m/z (ESI+): Found 197.03538 [M+Na]+; $C_6H_{10}O_2N_2{}^{23}Na^{32}S$ requires 197.03552. IR (v_{max} cm^{-1}) 1669 (C=O, amide), 1358, 1158, 1071, 944, 858.

8.4. Mechanistic study by the detection of THTCD deuteration pattern with various substrate and solvent

8.4.1. Detection method of THTCD

Small molecule LCMS

LCMS-MS was performed on a Thermo Scientific Ultimate 3000 and Thermo Scientific Q-Exactive-S. Two different LC methods were used: Method A, The column used was ACE Ultracore Super C18 column. (2.5 µm, 2x50 mm) Column temperature was 30 °Cwhile sampler temperature was 8 °C. The elution condition was setup as follow: Flow rate 0.4 mL/min, Mobile Phase A: water + 0.1% Formic acid, Mobile phase B: Acetonitrile, isocratic 99.5% A for 10 minutes followed by gradient to 0.1%A for 0.9 minutes, isocratic for 1.1 minutes, then gradient to 99.5% A. Method B: The column used was Waters Atlantis T3 column. (3 µm, 2.1x100) Column temperature was 20 °C while sampler temperature was 8 °C. The elution condition was setup as follow: Mobile Phase A: water + 0.1% Formic acid, Mobile phase B: Acetonitrile, isocratic 99% A for 7 minutes followed by gradient to 0%A for 5 minutes, isocratic for 1 minutes, then gradient for 0.1 minutes to 99% A followed by isocratic for 1.9 minutes.

Full scan MS was collected from 50 to 750 Th with a resolution setting of 17500. MS/MS product ion spectra were collected with normalised collision energy of 35 for five m/z ratios, corresponding to the un-deuterated, mono-, di-, tri- and tetra-deuterated THTCD.

Mass [m/z]	Formula [M]
175.05357	C6H10N2O2S
176.05985	C6H9DN2O2S

177.06613 C6H8D2N2O2S

179.07868 C6H6D4N2O2S

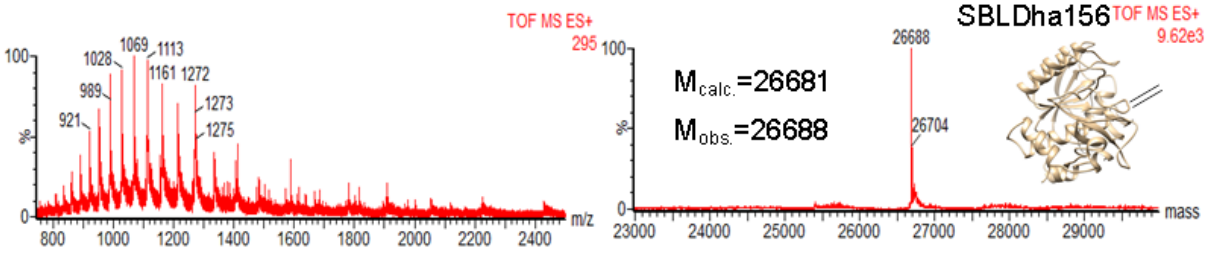
178.07241 C6H7D3N2O2S

Protein LCMS

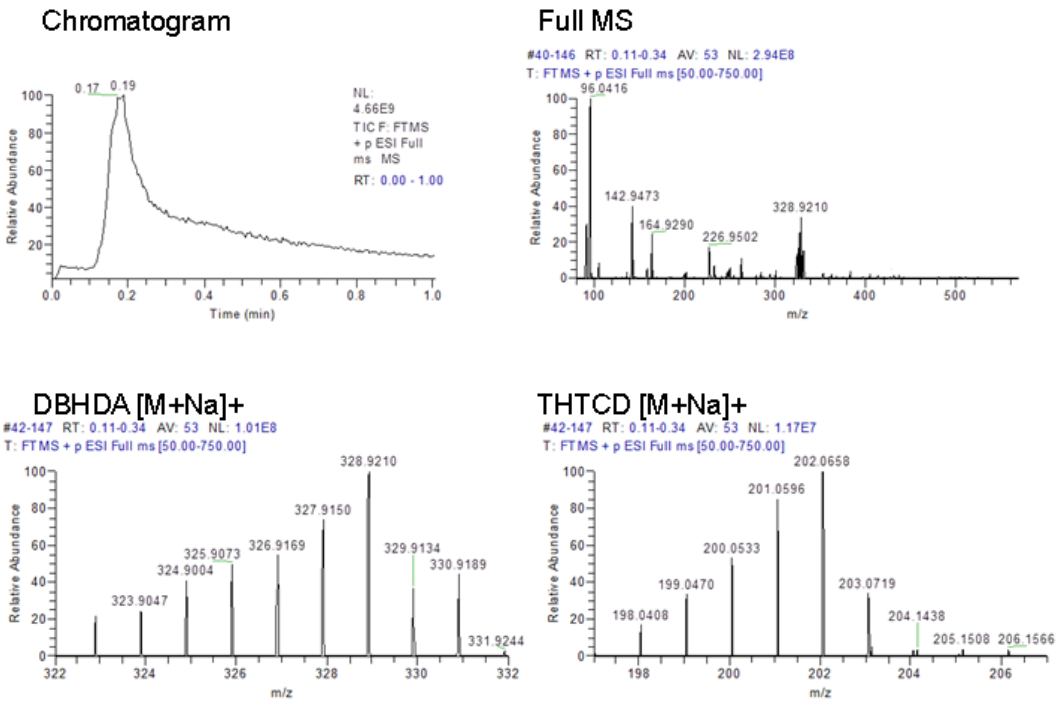
See Chapter 7. 7.2.1. and this chapter 8.1.1.

8.4.2. Reaction SBLC156 and DBHDA in deuterated buffer

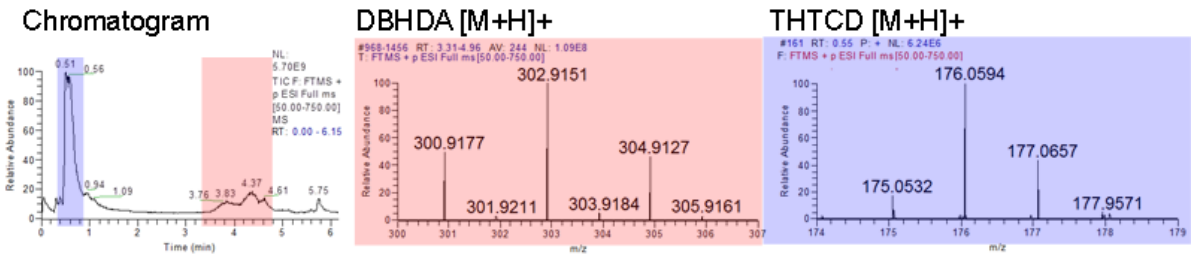
Lyophilised SBLC156 (5 mg) in deuterated 50 mM sodium phosphate buffer pH 8 (reaction buffer) was treated with a 1 M solution of DTT in reaction buffer. The mixture was shaken at 600 rpm at room temperature. The mixture was desalted using a PD-Miditrap™ G25 preequilibrated with reaction buffer to afford 1.5 mL of a 2 mg/mL solution. The solution was treated with DBHDA in solution in DMF (200 Eq. in 1 µL of DMF). The mixture was shaken at 600 rpm at room temperature for 30 minutes and at 37 °C. The reaction was monitored by LCMS-ESI-TOF. Full conversion to perdeuterated SBLDha156 ($M_{obs.}$=26688, $M_{calc.}$=26681) was observed after 2 hours.



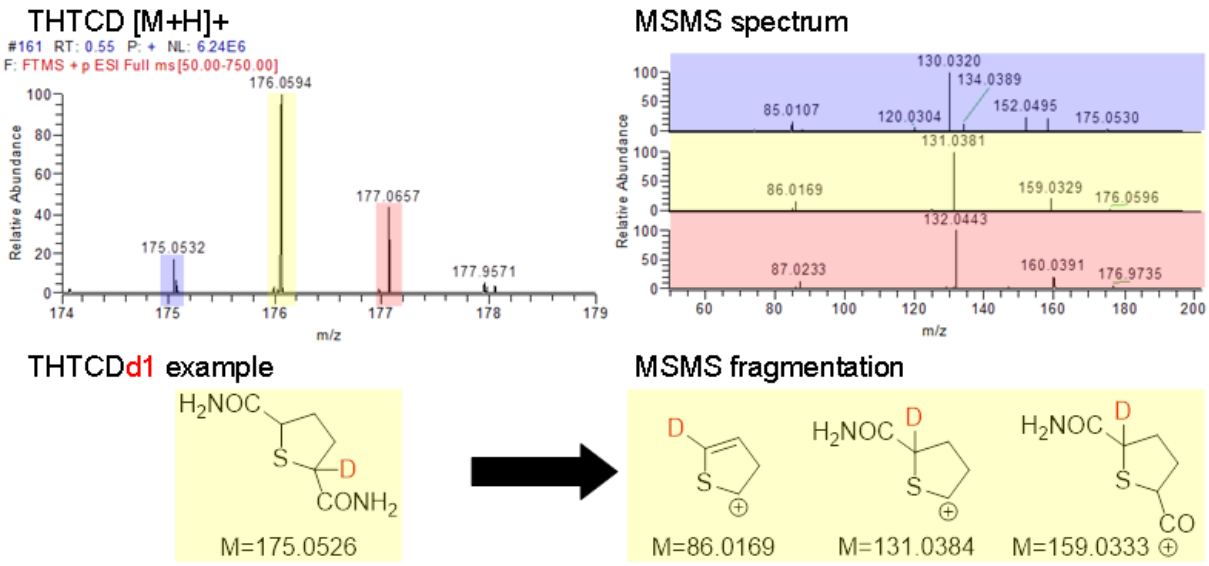
The mixture was then analysed by flow injection MS using Q-exactive.



The spectra could not distinguish between perdeuterated compounds and non-perdeuterated compounds. The addition of LC allowed separation of small molecule compounds and protonation of the exchangeable hydrogens. Excess DBHDA was also detected, and was used as a control for its potential deuteration (Non-deuterated DBHDA, HRMS m/z (ESI+): Found 300.9178 [M+H]+; $C_8H_{11}O_2N_2{}^{79}Br_2$ requires 300.9182).



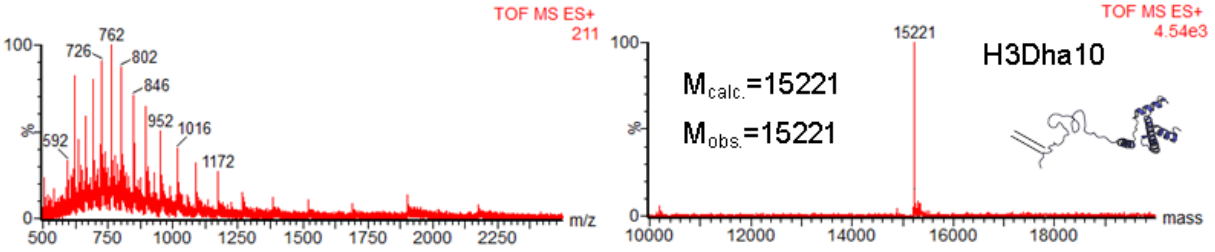
MSMS analysis confirmed that the deuteration occurred on Cγ of THTCD.



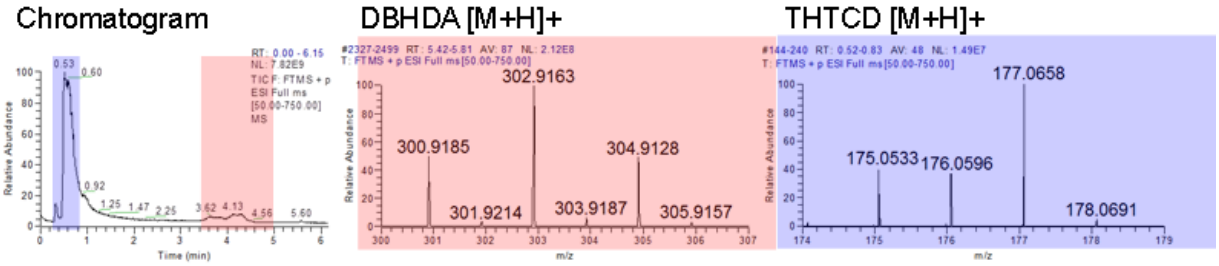
8.4.3. Reaction of H3C10 with DBHDA in deuterated buffer

A solution of lyophilised H3CX was prepared at a concentration of 2.0 mg/mL in deuterium oxide. A 250 µL aliquot of the protein solution was diluted with 250 µL of pH 8.0 sodium phosphate buffer (50 mM). To this diluted protein solution was then added DTT (10 mg). The solution was shaken at room temperature for 15 minutes to reduce any contaminant disulfide. After this time, the protein solution was passed through a PD Minitrap G25, previously equilibrated with deuterium oxide and eluted with 1.0 mL of deuterium oxide. The histone solution was stored on ice until needed. A stock solution of DBHDA was prepared by dissolving 84 mg in 564 µL of DMF. A 250 µL aliquot of the reduced protein was diluted with 250 µL of 50 mM sodium phosphate pH 8.0 and warmed to room temperature. Once the protein was at room temperature, a 50 µL aliquot of the DBHDA solution was added. The reaction was vortexed and then shaken at room temperature for 30 minutes and then 37 ºC for 60 minutes. The reaction mixture was then cooled to room temperature and precipitated DBHDA was removed by centrifugation (2 min, 14K g). The reaction was monitored by

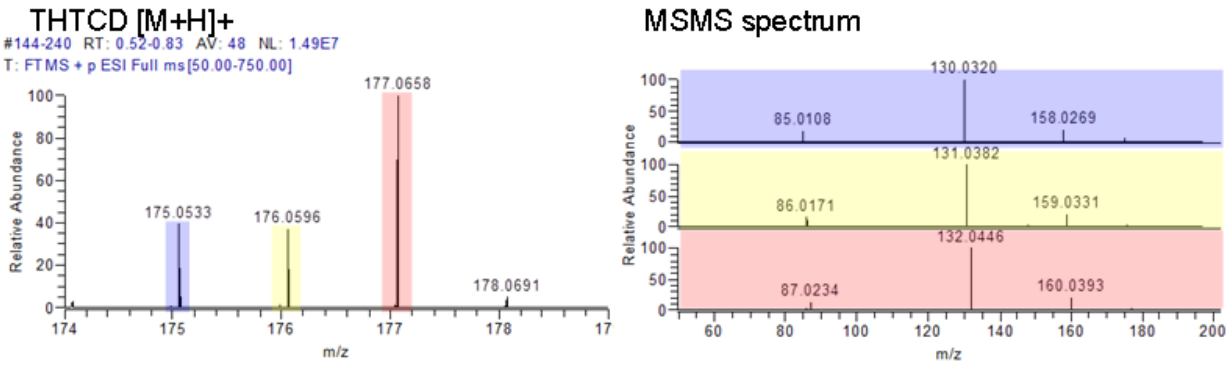
LCMS-ESI-TOF. Full conversion to perdeuterated H3Dha10 ($M_{obs.}$=26688, $M_{calc.}$=26681) was observed after 2 hours.



The small molecules (containing excess DBHDA) were then analysed by LCMS-MS using Q-exactive.



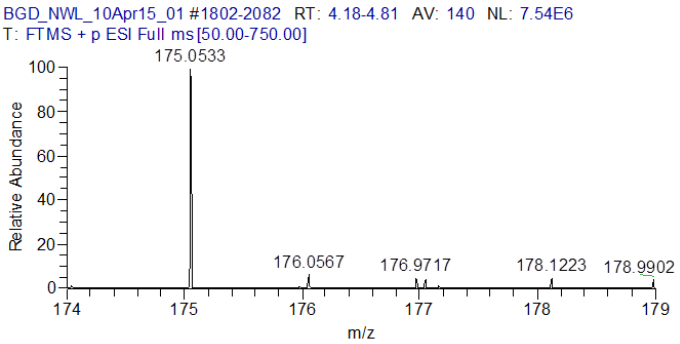
MSMS analysis confirmed that the deuteration occurred on Cγ of THTCD.



8.4.4 .Reaction of SBLdC156 and DBHDA in non-deuterated buffer

SBLdC156 in 50 mM sodium phosphate buffer pH 8 (reaction buffer) was treated with a 1 M solution of DTT in reaction buffer. The mixture was shaken at 600 rpm at room temperature. The mixture was desalted using a PD-Miditrap™ G25 preequilibrated with reaction buffer to

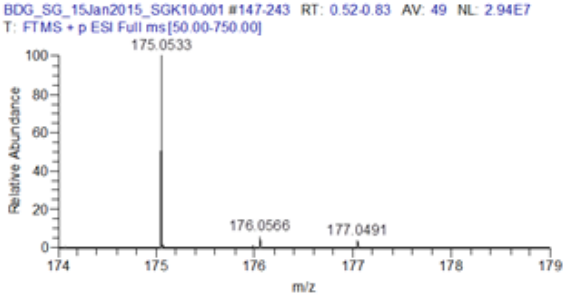
afford 1.5 mL of a 1 mg/mL solution. The solution was treated with DBHDA in solution in DMF (200 Eq. in 1 µL of DMF). The mixture was shaken at 600 rpm at room temperature for 30 minutes and at 37 °C. The reaction was monitored by LCMS-ESI-TOF and analysed by LCMS-MS using Q-exactive.



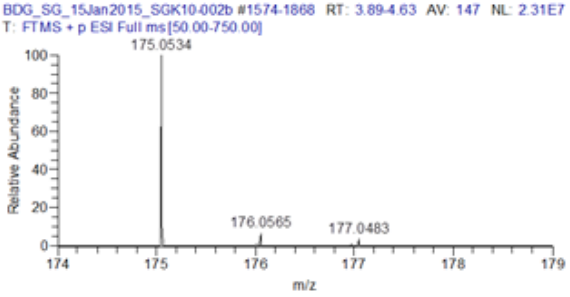
8.4.5. H3dCX and DBHDA in water

A solution of H3dCX was treated buffer exchanged to 50 mM sodium phosphate buffer pH 8. The concentration was adjusted to 0.5 mg/mL. The mixture was treated with a stock solution of DBHDA in DMF (final 200 Eq in 5% DMF.) The reaction was vortexed and then shaken at room temperature for 30 minutes and then 37 ºC for 60 minutes. The reaction mixture was then cooled to room temperature. The reaction was monitored by LCMS and analysed by LCMS-MS.

Reaction with H3dC10

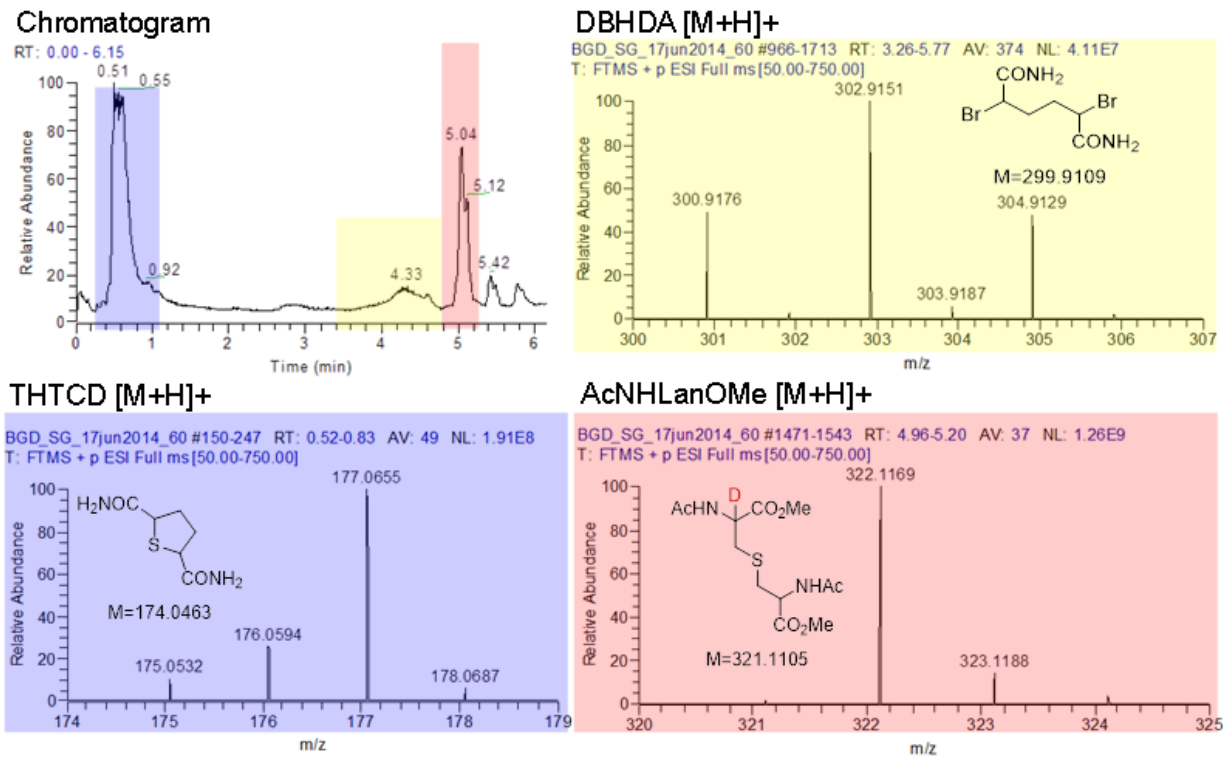


Reaction with H3dC26

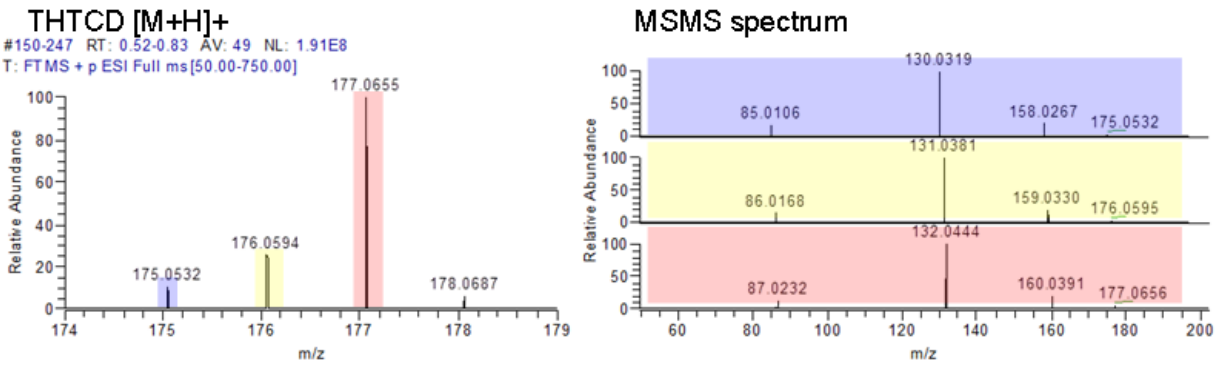


8.4.6. AcNHCysOMe and DBHDA in D2O

Methyl 2-(acetylamino)-3-sulfanylpropanoate (AcNHdCysOMe, 6 mg, 0.03 mmol) in 50 mM sodium phosphate buffer pH 8 (1 mL) was treated with a solution of DBHDA (0.40 mmol) in *N*-*N*-dimethylformamide (100 µL). The reaction mixture was shaken at 600 rpm at room temperature for 30 minutes, followed by shaking at 37 °C for 3 hours to afford mainly methyl S-(2-acetamido-3-methoxy-3-oxopropyl)-N-acetylcysteinate-2-^{2}H (AcNHdLanOMe) and tetrahydrothiophene-2,5-dicarboxamide (THTCD), tetrahydrothiophene-2-^{2}H-2,5-dicarboxamide (THTCD-d1) and tetrahydrothiophene-2,5-$^{2}H_2$-2,5-dicarboxamide (THTCD-d2). The reaction mixture (containing excess DBHDA) was analysed by LCMS-MS.

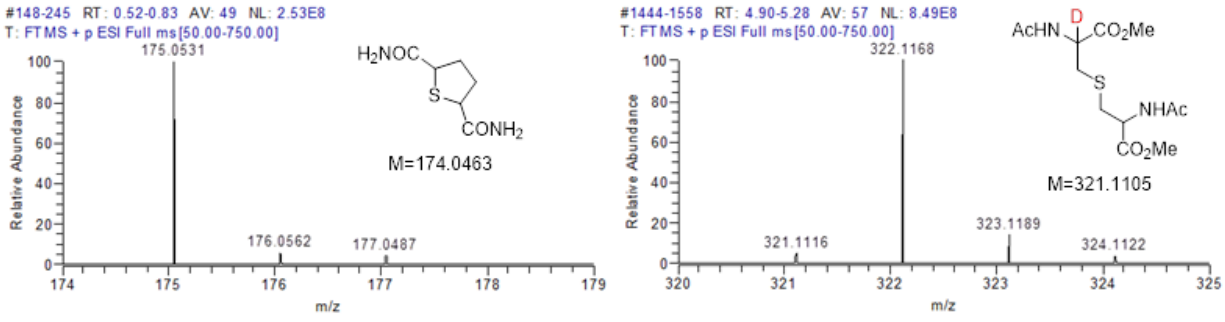


MSMS analysis confirmed that the deuteration occurred on Cγ of THTCD.



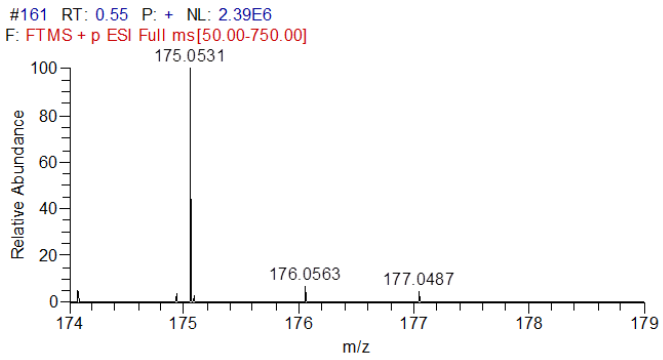
8.4.7. AcNHdCysOMe and DBHDA in water

Methyl 2-(acetylamino)-3-sulfanyl(2-^{2}H)propanoate (AcNHdCysOMe, 6 mg, 0.03 mmol) in 50 mM sodium phosphate buffer pH 8 (1 mL) was treated with a solution of DBHDA (0.40 mmol) in N-N-dimethylformamide (100 µL). The reaction mixture was shaken at 600 rpm at room temperature for 30 minutes, followed by shaking at 37 °C for 3 hours to afford mainly methyl S-(2-acetamido-3-methoxy-3-oxopropyl)-N-acetylcysteinate-2-d (AcNHdLanOMe) and tetrahydrothiophene-2,5-dicarboxamide (THTCD). The reaction mixtures were analysed by LCMS-MS.



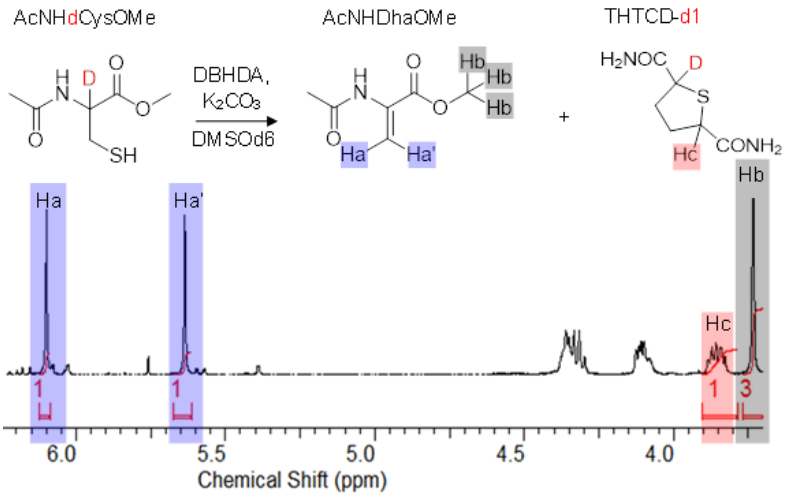
8.4.8. Control experiment with THTCD in deuterated buffer

THTCD (2 mg) in deuterated 50 mM sodium phosphate pH 8 buffer was shaken at 600 rpm at 37 °C for 24 hours. The mixture was analysed by LCMS.



8.4.9. Reaction of AcNHdCysOMe with DBHDA in DMSO-d6

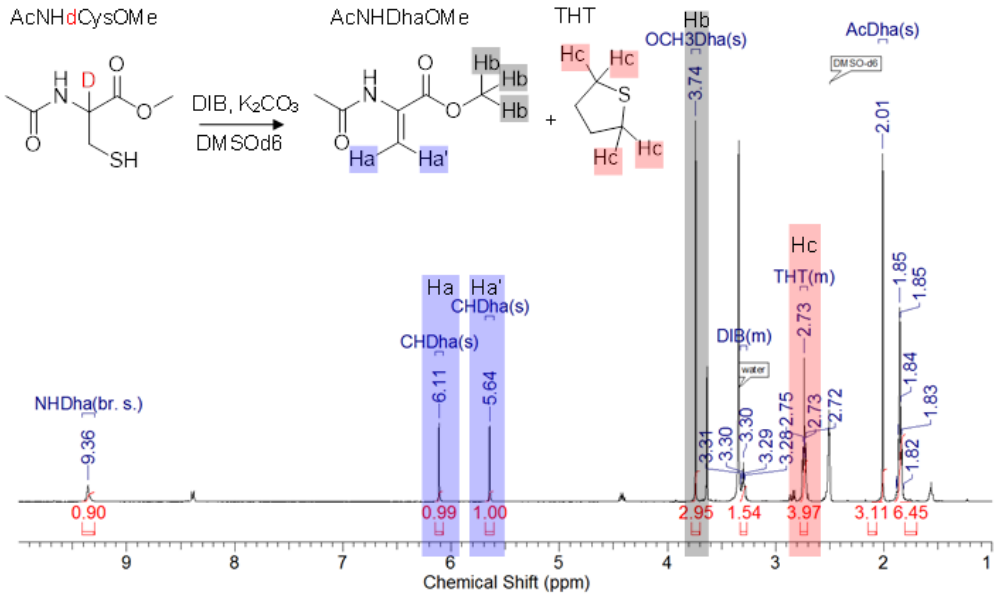
Methyl 2-(acetylamino)-3-sulfanyl(2-^{2}H)propanoate (AcNHdCysOMe, 21 mg, 0.12 mmol), dry 2,5-dibromohexane diiamide (DBHDA, 0.12 mmol) and dry potassium carbonate (50 mg, 0.36 mmol) were stirred at room temperature for 1 hour. The reactions were then stirred at 37 °C to afford mainly methyl 2-(acetylamino)prop-2-enoate (AcNHDhaOMe) and tetrahydrothiophene-2-d-2,5-dicarboxamide (THTCD-d1). The reaction was analysed by ^{1}H NMR.



8.4.10. Reaction of AcNHdCysOMe with DIB in DMSO-d6

Methyl 2-(acetylamino)-3-sulfanyl(2-^{2}H)propanoate (AcNHdCysOMe, 21 mg, 0.12 mmol), 1,4-diiodobutane (DIB, 0.12 mmol) and dry potassium carbonate (50 mg, 0.36 mmol) were

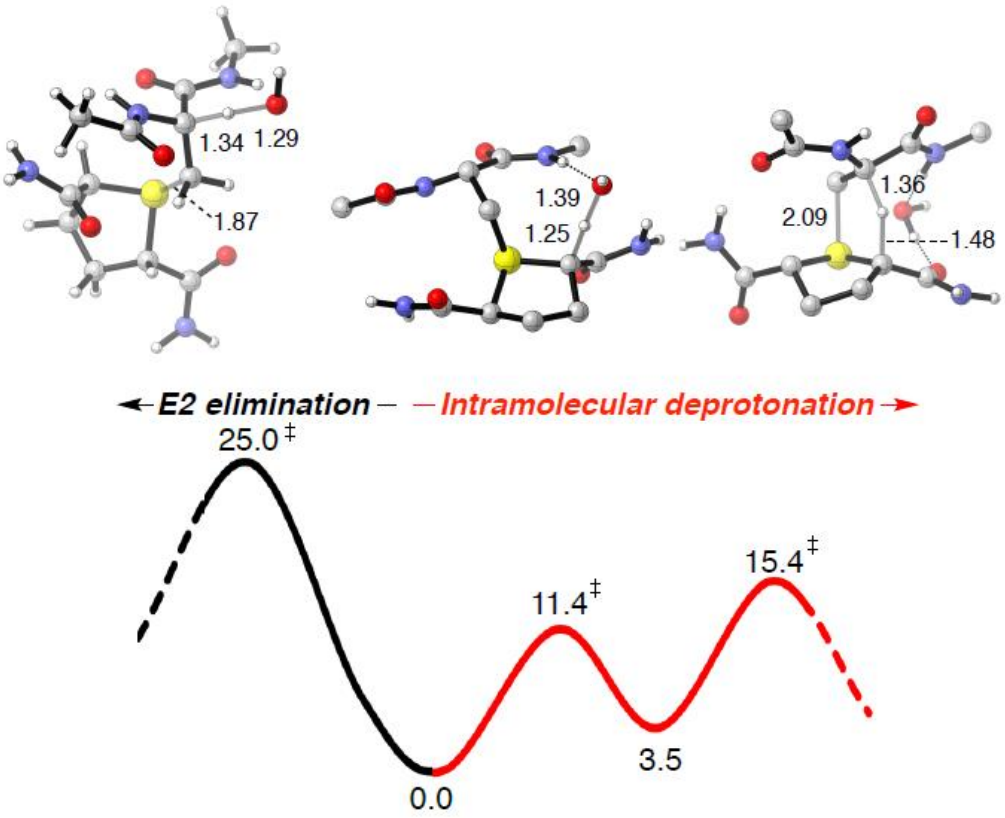
stirred at room temperature for 1 hour. The reactions were then stirred at 37 °C to afford mainly methyl 2-(acetylamino)prop-2-enoate (AcNHDhaOMe) and tetrahydrothiophene (THT). The reaction was analysed by ^{1}H NMR.



8.4.11. DFT calculation of elimination mechanism[‡‡‡‡‡‡‡]

Density functional theory (DFT) calculations were performed with *Gaussian09* rev D.01.[4] The hybrid meta-generalized gradient (GGA) M06-2X exchange-correlation functional[5] was used for all geometry optimizations, a fine integration grid and the diffuse/polarized 6-311++G(d,p) basis set for all elements. A conductor-like polarizable continuum model (CPCM) of DMSO solvation was used during all optimizations.[6] All stationary points were verified as either minima or saddle points by the presence of zero or a single imaginary harmonic vibrational frequency. Relative free energies were evaluated at 298.15 K applying a haptic translational entropy correction to adjust from a standard-state of 1 atm to 1 mol/l, while a quasi-harmonic approximation was used in which the treatment of vibrational entropies switches from a rigid-rotor harmonic oscillator to a free rotor at frequencies below 100 cm^{-1} as described previously. [7]

[‡‡‡‡‡‡‡] In collaboration with Dr. Robert Simion and Prof. Robert Paton, Oxford.



CPCM-M06-2X/6-311++G(d,p) computed free energy profile comparing E2-elimination with the intramolecular deprotonation pathway: TS structures are shown above with key forming/breaking distances in Å. Grel(298.15K) quoted in kcal/mol.

Cartesian Coordinates.

Absolute values (Hartrees) for SCF energy, zero-point vibrational energy (ZPE), enthalpy and free energy (at 298K) for all structures. Geometries of TS structures below.

Structure	E/au	ZPE/au	H/au	G/au
Intramolecular pathway:				
Starting complex	-1463.796873	0.361196	-1463.40982	-1463.483209
Deprotonation TS	-1463.773321	0.355842	-1463.391546	-1463.464988
Sulfonium ylid	-1463.790198	0.361038	-1463.402348	-1463.477693
Intramolecular deprot. TS	-1463.766147	0.355729	-1463.383743	-1463.458739
***E2*-elimination pathway:**				
Starting complex	-1463.785345	0.360164	-1463.398482	-1463.47388
E2-elimination TS	-1463.750013	0.354901	-1463.368585	-1463.44344

BIBLIOGRAPHY

Chapter 9: Bibliography

References

1. Groban, E. S.; Narayanan, A.; Jacobson, M. P. Conformational changes in protein loops and helices induced by post-translational phosphorylation. *PLoS Comput. Biol.* **2006***, 2*, 238-250.

2. Davis, F. Phosphorylation at the Interface. *Structure.* **2011***, 19*, 1726-1727.

3. Nishi, H.; Hashimoto, K.; Panchenko, A. Phosphorylation in Protein-Protein Binding: Effect on Stability and Function. *Structure.* **2011***, 19*, 1807-1815.

4. Nishi, H.; Fong, J. H.; Chang, C.; Teichmann, S. A.; Panchenko, A. R. Regulation of protein-protein binding by coupling between phosphorylation and intrinsic disorder: analysis of human protein complexes. *Mol. BioSyst.* **2013***, 9*, 1620-1626.

5. Nishi, H.; Shaytan, A.; Panchenko, A. R. Physicochemical mechanisms of protein regulation by phosphorylation. *Frontiers in Genetics.* **2014***, 5*, 1-10

6. Pearlman, S.; Serber, Z.; Ferrell Jr., J. A Mechanism for the Evolution of Phosphorylation Sites. *Cell.* **2011***, 147*, 934-946.

7. Xiang, S.; Gapsys, V.; Kim, H.; Bessonov, S.; Hsiao, H.; Möhlmann, S.; Klaukien, V.; Ficner, R.; Becker, S.; Urlaub, H.; Lührmann, R.; de Groot, B.; Zweckstetter, M. Phosphorylation Drives a Dynamic Switch in Serine/Arginine-Rich Proteins. *Structure.* **2013***, 21*, 2162-2174.

8. Tarrant, M. K.; Cole, P. A. The Chemical Biology of Protein Phosphorylation. *Annu. Rev. Biochem.* **2009***, 78*, 797-825.

9. Hunter, T. Tyrosine phosphorylation: thirty years and counting. *Current Opinion in Cell Biology.* **2009***, 21*, 140-146.

10. Meins, M.; Jenö, P.; Müller, D.; Richter, W. J.; Rosenbusch, J. P.; Erni, B. Cysteine phosphorylation of the glucose transporter of Escherichia coli. *Journal of Biological Chemistry.* **1993***, 268*, 11604-11609.

11. Sun, F.; Ding, Y.; Ji, Q.; Liang, Z.; Deng, X.; Wong, C.; Yi, C.; Zhang, L.; Xie, S.; Alvarez, S.; Hicks, L.; Luo, C.; Jiang, H.; Lan, L.; He, C. Protein cysteine phosphorylation of SarA/MgrA family transcriptional regulators mediates bacterial virulence and antibiotic resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2012***, 109*, 15461-15466.

12. Pannifer, A. D. B.; Flint, A. J.; Tonks, N. K.; Barford, D. Visualization of the cysteinyl-phosphate intermediate of a protein-tyrosine phosphatase by x-ray crystallography. *J. Biol. Chem.* **1998***, 273*, 10454-10462.

13. Stock, A. M.; Robinson, V. L.; Goudreau, P. N. Two-Component Signal Transduction. *Annu. Rev. Biochem.* **2000***, 69*, 183-215.

14. Matthews, H. R. Protein kinases and phosphatases that act on histidine, lysine, or arginine residues in eukaryotic proteins: A possible regulator of the mitogen-activated protein kinase cascade. *Pharmacology & Therapeutics.* **1995***, 67*, 323-350.

15. Fuhrmann, J.; Schmidt, A.; Spiess, S.; Lehner, A.; Turgay, K.; Mechtler, K.; Charpentier, E.; Clausen, T. McsB Is a Protein Arginine Kinase That Phosphorylates and Inhibits the Heat-Shock Regulator CtsR. *Science.* **2009***, 324*, 1323.

16. Elsholz, A. K. W.; Turgay, K.; Michalik, S.; Hessling, B.; Gronau, K.; Oertel, D.; Mäder, U.; Bernhardt, J.; Becher, D.; Hecker, M.; Gerth, U. Global impact of protein arginine phosphorylation on the physiology of Bacillus subtilis. *Proceedings of the National Academy of Sciences.* **2012***, 109*, 7451-7456.

17. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science.* **2002***, 298*, 1912.

18. Lindberg, R. A.; Quinn, A. M.; Hunter, T. Dual-specificity protein kinases: will any hydroxyl do? *Trends in Biochemical Sciences.* **1992***, 17*, 114-119.

19. Crews, C.; Alessandrini, A.; Erikson, R. The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product. *Science.* **1992***, 258*, 478-480.

20. Roskoski, R., Jr. MEK1/2 dual-specificity protein kinases: Structure and regulation. *Biochem. Biophys. Res. Commun.* **2012***, 417*, 5-10.

21. Kinoshita, E.; Kinoshita-Kikuta, E.; Kubota, Y.; Takekawa, M.; Koike, T. A Phos-tag SDS-PAGE method that effectively uses phosphoproteomic data for profiling the phosphorylation dynamics of MEK1. *Proteomics.* **2016***, 16*, 1825-1836.

22. Zheng, C. F.; Guan, K. L. Activation of MEK family kinases requires phosphorylation of two conserved Ser/Thr residues. *EMBO J.* **1994***, 13*, 1123-1131.

23. Alessi, D. R.; Saito, Y.; Campbell, D. G.; Cohen, P.; Sithanandam, G.; Rapp, U.; Ashworth, A.; Marshall, C. J.; Cowley, S. Identification of the sites in MAP kinase kinase-1 phosphorylated by p74raf-1. *EMBO J.* **1994***, 13*, 1610-1619.

24. Yan, M.; Templeton, D. J. Identification of 2 serine residues of MEK-1 that are differentially phosphorylated during activation by raf and MEK kinase. *Journal of Biological Chemistry.* **1994***, 269*, 19067-19073.

25. Alessandrini, A.; Greulich, H.; Huang, W.; Erikson, R. L. Mek1 phosphorylation site mutants activate Raf-1 in NIH 3T3 cells. *J. Biol. Chem.* **1996***, 271*, 31612-31618.

26. Gardner, A. M.; Vaillancourt, R. R.; Lange-Carter, C. A.; Johnson, G. L. MEK-1 phosphorylation by MEK kinase, Raf, and mitogen-activated protein kinase: analysis of phosphopeptides and regulation of activity. *Mol. Biol. Cell.* **1994***, 5*, 193-201.

27. Resing, K. A.; Mansour, S. J.; Hermann, A. S.; Johnson, R. S.; Candia, J. M.; Fukasawa, K.; Vande Woude, G. F.; Ahn, N. G. Determination of v-Mos-Catalyzed Phosphorylation Sites and Autophosphorylation Sites on MAP Kinase Kinase by ESI/MS. *Biochemistry (N. Y.).* **1995***, 34*, 2610-2620.

28. Cowley, S.; Paterson, H.; Kemp, P.; Marshall, C. J. Activation of MAP kinase kinase is necessary and sufficient for PC12 differentiation and for transformation of NIH 3T3 cells. *Cell (Cambridge, Mass.).* **1994***, 77*, 841-52.

29. Caunt, C. J.; Sale, M. J.; Smith, P. D.; Cook, S. J. MEK1 and MEK2 inhibitors and cancer therapy: the long and winding road. *Nat Rev Cancer.* **2015***, 15*, 577-592.

30. Abe, H.; Kikuchi, S.; Hayakawa, K.; Iida, T.; Nagahashi, N.; Maeda, K.; Sakamoto, J.; Matsumoto, N.; Miura, T.; Matsumura, K.; Seki, N.; Inaba, T.; Kawasaki, H.; Yamaguchi, T.; Kakefuda, R.; Nanayama, T.; Kurachi, H.; Hori, Y.; Yoshida, T.; Kakegawa, J.; Watanabe, Y.; Gilmartin, A. G.; Richter, M. C.; Moss, K. G.; Laquerre, S. G. Discovery of a Highly Potent and Selective MEK Inhibitor: GSK1120212 (JTP-74057 DMSO Solvate). *ACS Med. Chem. Lett.* **2011***, 2*, 320-324.

31. Rice, K. D.; Aay, N.; Anand, N. K.; Blazey, C. M.; Bowles, O. J.; Bussenius, J.; Costanzo, S.; Curtis, J. K.; Defina, S. C.; Dubenko, L.; Engst, S.; Joshi, A. A.; Kennedy, A. R.; Kim, A. I.; Koltun, E. S.; Lougheed, J. C.; Manalo, J. L.; Martini, J.; Nuss, J. M.; Peto, C. J.; Tsang, T. H.; Yu, P.; Johnston, S. Novel Carboxamide-Based Allosteric MEK Inhibitors: Discovery and Optimization Efforts toward XL518 (GDC-0973). *ACS Med. Chem. Lett.* **2012***, 3*, 416-421.

32. Gilmartin, A. G.; Bleam, M. R.; Groy, A.; Moss, K. G.; Minthorn, E. A.; Kulkarni, S. G.; Rominger, C. M.; Erskine, S.; Fisher, K. E.; Yang, J.; Zappacosta, F.; Annan, R.; Sutton, D.; Laquerre, S. G. GSK1120212 (JTP-74057) Is an Inhibitor of MEK Activity and Activation with Favorable Pharmacokinetic Properties for Sustained In Vivo Pathway Inhibition. *Clinical Cancer Research.* **2011***, 17*, 989-1000.

33. Hatzivassiliou, G.; Haling, J. R.; Chen, H.; Song, K.; Price, S.; Heald, R.; Hewitt, J. F. M.; Zak, M.; Peck, A.; Orr, C.; Merchant, M.; Hoeflich, K. P.; Chan, J.; Luoh, S.; Anderson, D. J.; Ludlam, M. J. C.; Wiesmann, C.; Ultsch, M.; Friedman, L. S.; Malek, S.; Belvin, M. Mechanism of MEK inhibition determines efficacy in mutant KRAS- versus BRAF-driven cancers. *Nature (London, U. K.).* **2013***, 501*, 232-236.

34. Fujioka, A.; Terai, K.; Itoh, R. E.; Aoki, K.; Nakamura, T.; Kuroda, S.; Nishida, E.; Matsuda, M. Dynamics of the Ras/ERK MAPK Cascade as Monitored by Fluorescent Probes. *Journal of Biological Chemistry.* **2006***, 281*, 8917-8926.

35. Aoki, K.; Yamada, M.; Kunida, K.; Yasuda, S.; Matsuda, M. Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proceedings of the National Academy of Sciences.* **2011***, 108*, 12675-12680.

36. Aoki, K.; Takahashi, K.; Kaizu, K.; Matsuda, M. A quantitative model of ERK MAP kinase phosphorylation in crowded media. *Sci. Rep.* **2013***, 3*, 1541-1549.

37. Mansour, S. J.; Matten, W. T.; Hermann, A. S.; Candia, J. M.; Rong, S.; Fukasawa, K.; Vande Woude, G. F.; Ahn, N. G. Transformation of mammalian cells by constitutively active MAP kinase kinase. *Science (Washington, D. C.).* **1994***, 265*, 966-70.

38. Brunet, A.; Pages, G.; Pouyssegur, J. Constitutively active mutants of MAP kinase kinase (MEK1) induce growth factor-relaxation and oncogenicity when expressed in fibroblasts. *Oncogene.* **1994***, 9*, 3379-87.

39. Verlhac, M.; Lefebvre, C.; Kubiak, J. Z.; Umbhauer, M.; Rassinier, P.; Colledge, W.; Maro, B. Mos activates MAPK in mouse oocytes through two opposite pathways. *EMBO J.* **2000***, 19*, 6065-6074.

40. Rice, K. D.; Aay, N.; Anand, N. K.; Blazey, C. M.; Bowles, O. J.; Bussenius, J.; Costanzo, S.; Curtis, J. K.; Defina, S. C.; Dubenko, L.; Engst, S.; Joshi, A. A.; Kennedy, A. R.; Kim, A. I.; Koltun, E. S.; Lougheed, J. C.; Manalo, J. L.; Martini, J.; Nuss, J. M.; Peto, C. J.; Tsang, T. H.; Yu, P.; Johnston, S. Novel Carboxamide-Based Allosteric MEK Inhibitors: Discovery and Optimization Efforts toward XL518 (GDC-0973). *ACS Med. Chem. Lett.* **2012***, 3*, 416-421.

41. Lito, P.; Saborowski, A.; Yue, J.; Solomon, M.; Joseph, E.; Gadal, S.; Saborowski, M.; Kastenhuber, E.; Fellmann, C.; Ohara, K.; Morikami, K.; Miura, T.; Lukacs, C.; Ishii, N.; Lowe, S.; Rosen, N. Disruption of CRAF-Mediated MEK Activation Is Required for Effective MEK Inhibition in KRAS Mutant Tumors. *Cancer Cell.* **2014***, 25*, 697-710.

42. Śmiechowski, M. Theoretical pKa prediction of O-phosphoserine in aqueous solution. *Chemical Physics Letters.* **2010***, 501*, 123-129.

43. Elliott, T. S.; Slowey, A.; Ye, Y.; Conway, S. J. The use of phosphate bioisosteres in medicinal chemistry and chemical biology. *Med. Chem. Commun.* **2012***, 3*, 735-751.

44. Smith, C. K.; Carr, D.; Mayhood, T. W.; Jin, W.; Gray, K.; Windsor, W. T. Expression and purification of phosphorylated and non-phosphorylated human MEK1. *Protein Expr. Purif.* **2007***, 52*, 446-456.

45. Sheth, P. R.; Liu, Y.; Hesson, T.; Zhao, J.; Vilenchik, L.; Liu, Y.; Mayhood, T. W.; Le, H. V. Fully activated MEK1 exhibits compromised affinity for binding of allosteric inhibitors U 0126 and PD 0325901. *Biochemistry.* **2011***, 50*, 7964-7976.

46. Smith, C. K.; Windsor, W. T. Thermodynamics of Nucleotide and Non-ATP-Competitive Inhibitor Binding to MEK1 by Circular Dichroism and Isothermal Titration Calorimetry. *Biochemistry (N. Y.).* **2007***, 46*, 1358-1367.

47. Benvenuti, M.; Mangani, S. Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nat. Protocols.* **2007***, 2*, 1633-1651.

48. Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. Designing logical codon reassignment - Expanding the chemistry in biology. *Chem. Sci.* **2015***, 6*, 50-69.

49. Arslan, T.; Mamaev, S. V.; Mamaeva, N. V.; Hecht, S. M. Structurally Modified Firefly Luciferase. Effects of Amino Acid Substitution at Position 286. *J. Am. Chem. Soc.* **1997***, 119*, 10877-10887.

50. Prestwich, G. D.; Xu, Y.; Qian, L.; Gajewiak, J.; Jiang, G. New metabolically stabilized analogues of lysophosphatidic acid: agonists, antagonists and enzyme inhibitors. *Biochem. Soc. Trans.* **2005***, 33*, 1357.

51. Lu, W.; Shen, K.; Cole, P. A. Chemical Dissection of the Effects of Tyrosine Phosphorylation of SHP-2. *Biochemistry (N. Y.).* **2003***, 42*, 5461-5468.

52. Xie, J.; Supekova, L.; Schultz, P. G. A Genetically Encoded Metabolically Stable Analogue of Phosphotyrosine in Escherichia coli. *ACS Chem. Biol.* **2007***, 2*, 474-478.

53. Rothman, D. M.; Petersson, E. J.; Vázquez, M. E.; Brandt, G. S.; Dougherty, D. A.; Imperiali, B. Caged Phosphoproteins. *J. Am. Chem. Soc.* **2005***, 127*, 846-847.

54. Rothman, D. M.; Shults, M. D.; Imperiali, B. Chemical approaches for investigating phosphorylation in signal transduction networks. *Trends Cell Biol.* **2005***, 15*, 502-510.

55. Fan, C.; Ip, K.; Söll, D. Expanding the genetic code of Escherichia coli with phosphotyrosine. *FEBS Lett.* **2016***, 590*, 3040-3047.

56. Luo, X.; Fu, G.; Wang, R. E.; Zhu, X.; Zambaldo, C.; Liu, R.; Liu, T.; Lyu, X.; Du, J.; Xuan, W.; Yao, A.; Reed, S. A.; Kang, M.; Zhang, Y.; Guo, H.; Huang, C.; Yang, P.; Wilson, I. A.; Schultz, P. G.; Wang, F. Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat Chem Biol.* **2017***, advance online publication*.

57. Hoppmann, C.; Wong, A.; Yang, B.; Li, S.; Hunter, T.; Shokat, K. M.; Wang, L. Site-specific incorporation of phosphotyrosine using an expanded genetic code. *Nat Chem Biol.* **2017***, advance online publication*.

58. Sauerwald, A.; Zhu, W.; Major, T. A.; Roy, H.; Palioura, S.; Jahn, D.; Whitman, W. B.; Yates, J. R.; Ibba, M.; SÃ¶ll, D. RNA-Dependent Cysteine Biosynthesis in Archaea. *Science.* **2005***, 307*, 1969.

59. Park, H.; Hohn, M. J.; Umehara, T.; Guo, L.; Osborne, E. M.; Benner, J.; Noren, C. J.; Rinehart, J.; Söll, D. Expanding the Genetic Code of Escherichia coli with Phosphoserine. *Science.* **2011***, 333*, 1151-1154.

60. Heinemann, I. U.; Rovner, A. J.; Aerni, H. R.; Rogulina, S.; Cheng, L.; Olds, W.; Fischer, J. T.; Soll, D.; Isaacs, F. J.; Rinehart, J. Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett.* **2012***, 586*, 3716-3722.

61. Lee, S.; Oh, S.; Yang, A.; Kim, J.; Söll, D.; Lee, D.; Park, H. A Facile Strategy for Selective Incorporation of Phosphoserine into Histones. *Angewandte Chemie International Edition.* **2013***, 52*, 5771-5775.

62. Rogerson, D. T.; Sachdeva, A.; Wang, K.; Haq, T.; Kazlauskaite, A.; Hancock, S. M.; Huguenin-Dezot, N.; Muqit, M. M. K.; Fry, A. M.; Bayliss, R.; Chin, J. W. Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat Chem Biol.* **2015***, 11*, 496-503.

63. Oza, J. P.; Aerni, H. R.; Pirman, N. L.; Barber, K. W.; ter Haar, C. M.; Rogulina, S.; Amrofell, M. B.; Isaacs, F. J.; Rinehart, J.; Jewett, M. C. Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat Commun.* **2015***, 6*.

64. Pirman, N. L.; Barber, K. W.; Aerni, H. R.; Ma, N. J.; Haimovich, A. D.; Rogulina, S.; Isaacs, F. J.; Rinehart, J. A flexible codon in genomically recoded Escherichia coli permits programmable protein phosphorylation. *Nat Commun.* **2015***, 6*.

65. Chen, Z.; Cole, P. A. Synthetic approaches to protein phosphorylation. *Curr. Opin. Chem. Biol.* **2015***, 28*, 115-122.

66. Gururaja, T. L.; Levine, M. J. Solid-phase synthesis and characterization of human salivary statherin: a tyrosine-rich phosphoprotein inhibitor of calcium phosphate precipitation. *Pept. Res.* **1996***, 9*, 283-289.

67. Jbara, M.; Maity, S. K.; Morgan, M.; Wolberger, C.; Brik, A. Chemical Synthesis of Phosphorylated Histone H2A at Tyr57 Reveals Insight into the Inhibition Mode of the SAGA Deubiquitinating Module. *Angewandte Chemie International Edition.* **2016***, 55*, 4972-4976.

68. Hejjaoui, M.; Butterfield, S.; Fauvet, B.; Vercruysse, F.; Cui, J.; Dikiy, I.; Prudent, M.; Olschewski, D.; Zhang, Y.; Eliezer, D.; Lashuel, H. A. Elucidating the Role of C-Terminal Post-Translational Modifications Using Protein Semisynthesis Strategies: Î±-Synuclein Phosphorylation at Tyrosine 125. *J. Am. Chem. Soc.* **2012***, 134*, 5196-5210.

69. Bondalapati, S.; Mansour, W.; Nakasone, M. A.; Maity, S. K.; Glickman, M. H.; Brik, A. Chemical Synthesis of Phosphorylated Ubiquitin and Diubiquitin Exposes Positional Sensitivities of E1-E2 Enzymes and Deubiquitinases. *Chemistry – A European Journal.* **2015***, 21*, 7360-7364.

70. Muir, T. W. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **2003***, 72*, 249-289.

71. Lu, W.; Gong, D.; Bar-Sagi, D.; Cole, P. A. Site-Specific Incorporation of a Phosphotyrosine Mimetic Reveals a Role for Tyrosine Phosphorylation of SHP-2 in Cell Signaling. *Mol. Cell.* **2001***, 8*, 759-769.

72. Zhang, Z.; Shen, K.; Lu, W.; Cole, P. A. The Role of C-terminal Tyrosine Phosphorylation in the Regulation of SHP-1 Explored via Expressed Protein Ligation. *Journal of Biological Chemistry.* **2003***, 278*, 4668-4674.

73. Shogren-Knaak, M. A.; Fry, C. J.; Peterson, C. L. A Native Peptide Ligation Strategy for Deciphering Nucleosomal Histone Modifications. *Journal of Biological Chemistry.* **2003***, 278*, 15744-15748.

74. Chiang, K. P.; Jensen, M. S.; McGinty, R. K.; Muir, T. W. A Semisynthetic Strategy to Generate Phosphorylated and Acetylated Histone H2B. *ChemBioChem.* **2009***, 10*, 2182-2187.

75. Hahn, M. E.; Muir, T. W. Photocontrol of Smad2, a Multiphosphorylated Cell-Signaling Protein, through Caging of Activating Phosphoserines. *Angewandte Chemie International Edition.* **2004***, 43*, 5800-5803.

76. Broncel, M.; Krause, E.; Schwarzer, D.; Hackenberger, C. P. R. The Alzheimer's Disease Related Tau Protein as a New Target for Chemical Protein Engineering. *Chemistry – A European Journal.* **2012***, 18*, 2488-2492.

77. Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: A general method for protein engineering. *Proceedings of the National Academy of Sciences.* **1998***, 95*, 6705-6710.

78. Cotton, G. J.; Muir, T. W. Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed protein ligation. *Chemistry & Biology.* **2000***, 7*, 253-261.

79. Hølys-Larsen, K. B.; Jensen, K. J. Solid-Phase Synthesis of Phosphopeptides. In *Peptide Synthesis and Applications* Jensen, J. K., Tofteng Shelton, P. and Pedersen, L. S., Eds.; Humana Press: Totowa, NJ, 2013; pp 191-199.

80. Rothman, D. M.; Vazquez, M. E.; Vogel, E. M.; Imperiali, B. Caged Phospho-Amino Acid Building Blocks for Solid-Phase Peptide Synthesis. *J. Org. Chem.* **2003***, 68*, 6795-6798.

81. Eberhard, H.; Seitz, O. N→O-Acyl shift in Fmoc-based synthesis of phosphopeptides. *Org. Biomol. Chem.* **2008***, 6*, 1349-1355.

82. Pirrung, M. C.; James, K. D.; Rana, V. S. Thiophosphorylation of Histidine. *J. Org. Chem.* **2000***, 65*, 8448-8453.

83. Ruman, T.; Długopolska, K.; Jurkiewicz, A.; Rut, D.; Frączyk, T.; Cieśla, J.; Leś, A.; Szewczuk, Z.; Rode, W. Thiophosphorylation of free amino acids and enzyme protein by thiophosphoramidate ions. *Bioorg. Chem.* **2010***, 38*, 74-80.

84. Wilk, P.; Jarmuła, A.; Ruman, T.; Banaszak, K.; Rypniewski, W.; Cieśla, J.; Dowierciał, A.; Rode, W. Crystal structure of phosphoramide-phosphorylated thymidylate synthase reveals pSer127, reflecting probably pHis to pSer phosphotransfer. *Bioorg. Chem.* **2014***, 52*, 44-49.

85. Serwa, R.; Wilkening, I.; Del Signore, G.; Mühlberg, M.; Claußnitzer, I.; Weise, C.; Gerrits, M.; Hackenberger, C. Chemoselective Staudinger-Phosphite Reaction of Azides for the Phosphorylation of Proteins. *Angewandte Chemie International Edition.* **2009***, 48*, 8234-8239.

86. Bertran-Vicente, J.; Serwa, R. A.; Schümann, M.; Schmieder, P.; Krause, E.; Hackenberger, C. P. R. Site-Specifically Phosphorylated Lysine Peptides. *J. Am. Chem. Soc.* **2014***, 136*, 13622-13628.

87. Bertran-Vicente, J.; Schumann, M.; Schmieder, P.; Krause, E.; Hackenberger, C. P. R. Direct access to site-specifically phosphorylated-lysine peptides from a solid-support. *Org. Biomol. Chem.* **2015***, 13*, 6839-6843.

88. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proceedings of the National Academy of Sciences.* **2002***, 99*, 19-24.

89. Link, A. J.; Vink, M. K. S.; Tirrell, D. A. Presentation and Detection of Azide Functionality in Bacterial Cell Surface Proteins. *J. Am. Chem. Soc.* **2004***, 126*, 10598-10602.

90. Miseta, A.; Csutora, P. Relationship Between the Occurrence of Cysteine in Proteins and the Complexity of Organisms. *Mol. Biol. Evol.* **2000***, 17*, 1232-1239.

91. Lamoureaux, T.; Lee, D. Chemical activation of MEK1 - a redox trigger for evaluating the effects of phosphorylation. *Chemical Communications.* **2011***, 47*, 8623-8625.

92. Bertran-Vicente, J.; Penkert, M.; Nieto-Garcia, O.; Jeckelmann, J.; Schmieder, P.; Krause, E.; Hackenberger, C. P. R. Chemoselective synthesis and analysis of naturally occurring phosphorylated cysteine peptides. *Nature Communications.* **2016***, 7*, 12703.

93. Gunnoo, S. B.; Davis, B. G.; Finney, H. M.; Baker, T. S.; Lawson, A. D.; Anthony, D. C. Creation of a gated antibody as a conditionally functional synthetic protein. *Nat Commun.* **2014***, 5*, 1-8.

94. Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. Conversion of Cysteine into Dehydroalanine Enables Access to Synthetic Histones Bearing Diverse Post-Translational Modifications. *Angewandte Chemie International Edition.* **2012***, 51*, 1835-1839.

95. Rowan, F. C.; Richards, M.; Bibby, R. A.; Thompson, A.; Bayliss, R.; Blagg, J. Insights into Aurora-A Kinase Activation Using Unnatural Amino Acids Incorporated by Chemical Modification. *ACS Chem. Biol.* **2013***, 8*, 2184-2191.

96. Chooi, K. P.; Galan, S. R. G.; Raj, R.; McCullagh, J.; Mohammed, S.; Jones, L. H.; Davis, B. G. Synthetic Phosphorylation of p38α Recapitulates Protein Kinase Activity. *J. Am. Chem. Soc.* **2014***, 136*, 1698-1701.

97. Meledin, R.; Mali, S. M.; Singh, S. K.; Brik, A. Protein ubiquitination via dehydroalanine: development and insights into the diastereoselective 1,4-addition step. *Org. Biomol. Chem.* **2016***, 14*, 4817-4823.

98. Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J. L.; Wiewiora, R.; Ng, W.; Raj, R.; Faulkner, S.; Vallée, M. R. J.; Phanumartwiwath, A.; Coleman, O. D.; Thézénas, M.; Khan, M.; Galan, S. R. G.; Lercher, L.; Schombs, M. W.; Gerstberger, S.; Palm-Espling, M. E.; Baldwin, A. J.; Kessler, B. M.; Claridge, T. D. W.; Mohammed, S.; Davis, B. G. Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science.* **2016**, *354*, 6312, aag1465.

99. Chatterjee, C.; Paul, M.; Xie, L.; van, d. D. Biosynthesis and Mode of Action of Lantibiotics. *Chem. Rev.* **2005***, 105*, 633-684.

100. Buchman, G. W.; Banerjee, S.; Hansen, J. N. Structure, expression, and evolution of a gene encoding the precursor of nisin, a small protein antibiotic. *Journal of Biological Chemistry.* **1988***, 263*, 16260-16266.

101. Garg, N.; Salazar-Ocampo, L. M. A.; van der Donk, W. A. In vitro activity of the nisin dehydratase NisB. *Proceedings of the National Academy of Sciences.* **2013***, 110*, 7258-7263.

102. Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van, d. D.; Nair, S. K. Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature.* **2015***, 517*, 509-512.

103. You, Y. O.; Levengood, M. R.; Ihnken, L. A. F.; Knowlton, A. K.; van, d. D. Lacticin 481 Synthetase as a General Serine/Threonine Kinase. *ACS Chem. Biol.* **2009***, 4*, 379-385.

104. Masters, P. In vivo decomposition of phosphoserine and serine in noncollagenous protein from human dentin. *Calcif. Tissue Int.* **1985***, 37*, 236-241.

105. Wang, Z.; Lyons, B.; Truscott, R. J. W.; Schey, K. L. Human protein aging: modification and crosslinking through dehydroalanine and dehydrobutyrine intermediates. *Aging Cell.* **2014***, 13*, 226-234.

106. Li, H.; Xu, H.; Zhou, Y.; Zhang, J.; Long, C.; Li, S.; Chen, S.; Zhou, J.; Shao, F. The Phosphothreonine Lyase Activity of a Bacterial Type III Effector Family. *Science.* **2007***, 315*, 1000-1003.

107. Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; Chen, P. R. Palladium-triggered deprotection chemistry for protein activation in living cells. *Nat Chem.* **2014***, 6*, 352-361.

108. Byford, M. F. Rapid and selective modification of phosphoserine residues catalysed by Ba2+ ions for their detection during peptide microsequencing. *Biochem. J.* **1991***, 280*, 261-265.

109. Yang, A.; Ha, S.; Ahn, J.; Kim, R.; Kim, S.; Lee, Y.; Kim, J.; Söll, D.; Lee, H.; Park, H. A chemical biology route to site-specific authentic protein modifications. *Science.* **2016**.

110. Strumeyer, D. H.; White, W. N.; Koshland, D. E. Role of serine in chymotrypsin action. conversion of the active serine to dehydroalanine. *Proc. Natl. Acad. Sci. U. S. A.* **1963***, 50*, 931-935.

111. Neet, K. E.; Koshland, D. E. The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation". *Proceedings of the National Academy of Sciences.* **1966***, 56*, 1606-1611.

112. Miller, M. J. Isourea-mediated preparation of dehydro amino acids. *J. Org. Chem.* **1980***, 45*, 3131-3132.

113. Gilis, D.; Massar, S.; Cerf, N. J.; Rooman, M. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. *Genome Biol.* **2001***, 2*, research0049.1.

114. Itzkovitz, S.; Alon, U. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Research.* **2007***, 17*, 405-412.

115. Rich, D. H.; Tam, J.; Mathiaparanam, P.; Grant, J. A.; Mabuni, C. General synthesis of didehydroamino-acids and peptides. *J. Chem. Soc. , Chem. Commun.* **1974**, 897-898.

116. Holmes, T. J., Jr.; Lawton, R. G. Cysteine modification and cleavage of proteins with 2-methyl-N1-benzenesulfonyl-N4-bromoacetylquinonediimide. *J. Am. Chem. Soc.* **1977***, 99*, 1984-1986.

117. Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci.* **2011***, 2*, 1666-1676.

118. Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. Facile Conversion of Cysteine and Alkyl Cysteines to Dehydroalanine on Protein Surfaces: Versatile and Switchable Access to Functionalized Proteins. *J. Am. Chem. Soc.* **2008***, 130*, 5052-5053.

119. Younis, I. R.; Elliott, M.; Peer, C. J.; Cooper, A. J. L.; Pinto, J. T.; Konat, G. W.; Kraszpulski, M.; Petros, W. P.; Callery, P. S. Dehydroalanine Analog of Glutathione: An Electrophilic Busulfan Metabolite That Binds to Human Glutathione S-Transferase A1-1. *Journal of Pharmacology and Experimental Therapeutics.* **2008***, 327*, 770-776.

120. Okamoto, R.; Souma, S.; Kajihara, Y. Efficient substitution reaction from cysteine to the serine residue of glycosylated polypeptide: repetitive peptide segment ligation strategy and the synthesis of glycosylated tetracontapeptide having acid labile sialyl-TN antigens. *J. Org. Chem.* **2009***, 74*, 2494-2501.

121. Nathani, R. I.; Moody, P.; Chudasama, V.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S. A novel approach to the site-selective dual labelling of a protein via chemoselective cysteine modification. *Chem. Sci.* **2013***, 4*, 3455-3458.

122. Haj-Yahya, N.; Hemantha, H. P.; Meledin, R.; Bondalapati, S.; Seenaiah, M.; Brik, A. Dehydroalanine-Based Diubiquitin Activity Probes. *Org. Lett.* **2014***, 16*, 540-543.

123. Lercher, L.; Raj, R.; Patel, N. A.; Price, J.; Mohammed, S.; Robinson, C. V.; Schofield, C. J.; Davis, B. G. Generation of a synthetic GlcNAcylated nucleosome reveals regulation of stability by H2A-Thr101 GlcNAcylation. *Nature Communications.* **2015***, 6*, 7978.

124. Raj, R.; Lercher, L.; Mohammed, S.; Davis, B. G. Synthetic Nucleosomes Reveal that GlcNAcylation Modulates Direct Interaction with the FACT Complex. *Angewandte Chemie.* **2016***, 128*, 9064-9068.

125. Timms, N.; Windle, C. L.; Polyakova, A.; Ault, J. R.; Trinh, C. H.; Pearson, A. R.; Nelson, A.; Berry, A. Structural Insights into the Recovery of Aldolase Activity in N-Acetylneuraminic Acid Lyase by Replacement of the Catalytically Active Lysine with ?-Thialysine by Using a Chemical Mutagenesis Strategy. *ChemBioChem.* **2013***, 14*, 474-481.

126. Hashimoto, K.; Sakai, M.; Okuno, T.; Shirahama, H. small beta]-Phenylselenoalanine as a dehydroalanine precursor-efficient synthesis of alternariolide (AM-toxin I). *Chem. Commun.* **1996**, 1139-1140.

127. Okeley, N. M.; Zhu, Y.; van, d. D. Facile Chemoselective Synthesis of Dehydroalanine-Containing Peptides. *Org. Lett.* **2000***, 2*, 3603-3606.

128. Seebeck, F. P.; Szostak, J. W. Ribosomal Synthesis of Dehydroalanine-Containing Peptides. *J. Am. Chem. Soc.* **2006***, 128*, 7150-7151.

129. Wang, J.; Schiller, S. M.; Schultz, P. G. A biosynthetic route to dehydroalanine-containing proteins. *Angew. Chem. , Int. Ed.* **2007***, 46*, 6849-6851.

130. Guo, J.; Wang, J.; Lee, J.; Schultz, P. Site-Specific Incorporation of Methyl- and Acetyl-Lysine Analogues into Recombinant Proteins. *Angewandte Chemie International Edition.* **2008***, 47*, 6399-6401.

131. Varlamova, E. G.; Novoselov, S. V. Methods to biosynthesize mammalian selenocysteine-containing proteins in vitro. *Mol. Biol. (N. Y.).* **2016***, 50*, 37-42.

132. Rowan, F.; Richards, M.; Blagg, J.; Widya, M.; Bayliss, R. Diverse functionalization of aurora-a kinase at specified surface and buried sites by native chemical modification. *PLoS One.* **2014***, 9*, e103935.

133. Steen, H.; Jebanathirajah, J. A.; Rush, J.; Morrice, N.; Kirschner, M. W. Phosphorylation analysis by mass spectrometry. Myths, facts, and the consequences for qualitative and quantitative measurements. *Mol. Cell. Proteomics.* **2006***, 5*, 172-181.

134. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protocols.* **2007***, 1*, 2856-2860.

135. Gundry, R. L.; White, M. Y.; Murray, C. I.; Kane, L. A.; Fu, Q.; Stanley, B. A.; Van Eyk, J.,E. Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow. *Current protocols in molecular biology / edited by Frederick M.Ausubel ...[et al.].* **2009***, CHAPTER*, Unit10.25-Unit10.25.

136. Buchowiecka, A. K. Puzzling over protein cysteine phosphorylation - assessment of proteomic tools for S-phosphorylation profiling. *Analyst.* **2014***, 139*, 4118-4123.

137. Boersema, P. J.; Mohammed, S.; Heck, A. J. R. Phosphopeptide fragmentation and analysis by mass spectrometry. *Journal of Mass Spectrometry.* **2009***, 44*, 861-878.

138. Palumbo, A. M.; Tepe, J. J.; Reid, G. E. Mechanistic Insights into the Multistage Gas-Phase Fragmentation Behavior of Phosphoserine- and Phosphothreonine-Containing Peptides. *J. Proteome Res.* **2008***, 7*, 771-779.

139. Kim, M.; Pandey, A. Electron Transfer Dissociation Mass Spectrometry in Proteomics. *Proteomics.* **2012***, 12*, 530-542.

140. Favata, M. F.; Horiuchi, K. Y.; Manos, E. J.; Daulerio, A. J.; Stradley, D. A.; Feeser, W. S.; Van Dyk, D. E.; Pitts, W. J.; Earl, R. A.; Hobbs, F.; Copeland, R. A.; Magolda, R. L.; Scherle, P. A.; Trzaskos, J. M. Identification of a Novel Inhibitor of Mitogen-activated Protein Kinase Kinase. *Journal of Biological Chemistry.* **1998***, 273*, 18623-18632.

141. VanScyoc, W. S.; Holdgate, G. A.; Sullivan, J. E.; Ward, W. H. J. Enzyme Kinetics and Binding Studies on Inhibitors of MEK Protein Kinase. *Biochemistry (N. Y.).* **2008***, 47*, 5017-5027.

142. Kinoshita, E.; Kinoshita-Kikuta, E.; Takiyama, K.; Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics.* **2006***, 5*, 749-757.

143. Dai, P.; Zhang, C.; Welborn, M.; Shepherd, J. J.; Zhu, T.; Van Voorhis, T.; Pentelute, B. L. Salt Effect Accelerates Site-Selective Cysteine Bioconjugation. *ACS Cent. Sci.* **2016**, 2(9): 637–646.

144. Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. π-Clamp-mediated cysteine conjugation. *Nat Chem.* **2016***, 8*, 120-128.

145. Willwacher, J.; Raj, R.; Mohammed, S.; Davis, B. G. Selective Metal-Site-Guided Arylation of Proteins. *J. Am. Chem. Soc.* **2016***, 138*, 8678-8681.

146. Morrison, P. M.; Foley, P. J.; Warriner, S. L.; Webb, M. E. Chemical generation and modification of peptides containing multiple dehydroalanines. *Chem. Commun.* **2015***, 51*, 13470-13473.

147. Lee, B.; Richards, F. M. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **1971***, 55*, 379-IN4.

149. Liu, Q.; Sabnis, Y.; Zhao, Z.; Zhang, T.; Buhrlage, S.; Jones, L.; Gray, N. Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chem. Biol.* **2013***, 20*, 146-159.

150. Wagle, N.; Emery, C.; Berger, M. F.; Davis, M. J.; Sawyer, A.; Pochanard, P.; Kehoe, S. M.; Johannessen, C. M.; MacConaill, L. E.; Hahn, W. C.; Meyerson, M.; Garraway, L. A. Dissecting Therapeutic Resistance to RAF Inhibition in Melanoma by Tumor Genomic Profiling. *Journal of Clinical Oncology.* **2011***, 29*, 3085-3096.

151. Zhao, Kexin Ju, YoungJun Li, Shuangshuang Altaany, Zaid Wang,Rui Yang, Guangdong S-sulfhydration of MEK1 leads to PARP-1 activation and DNA damage repair. *EMBO Rep.* **2014***, 15*, 792-800.

152. Ellman, G. L. Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **1959***, 82*, 70-77.

153. Collier, H. B. A note on the molar absorptivity of reduced Ellman's reagent, 3-carboxylato-4-nitrothiophenolate. *Anal. Biochem.* **1973***, 56*, 310-311.

154. Riddles, P. W.; Blakeley, R. L.; Zerner, B. [8] Reassessment of Ellman's reagent. In *Methods in Enzymology*Academic Press: 1983; pp 49-60.

155. Linke, P.; Amaning, K.; Maschberger, M.; Vallee, F.; Steier, V.; Baaske, P.; Duhr, S.; Breitsprecher, D.; Rak, A. An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery. *Journal of Biomolecular Screening.* **2016***, 21*, 414-421.

157. Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protocols.* **2007***, 2*, 2212-2221.

158. Xu, S.; Robbins, D.; Frost, J.; Dang, A.; Lange-Carter, C.; Cobb, M. H. MEKK1 phosphorylates MEK1 and MEK2 but does not cause activation of mitogen-activated protein kinase. *Proc. Natl. Acad. Sci. U. S. A.* **1995***, 92*, 6808-6812.

159. Rominger, C. M.; Schaber, M. D.; Yang, J.; Gontarek, R. R.; Weaver, K. L.; Broderick, T.; Carter, L.; Copeland, R. A.; May, E. W. An intrinsic ATPase activity of phospho-MEK-1 uncoupled from downstream ERK phosphorylation. *Arch. Biochem. Biophys.* **2007***, 464*, 130-137.

160. Wagner, S. A.; Beli, P.; Weinert, B. T.; Nielsen, M. L.; Cox, J.; Mann, M.; Choudhary, C. A Proteome-wide, Quantitative Survey of In Vivo Ubiquitylation Sites Reveals Widespread Regulatory Roles. *Molecular & Cellular Proteomics.* **2011***, 10*.

161. Yeung, F.; Ramsey, C. S.; Popko-Scibor, A.; Allison, D. F.; Gray, L. G.; Shin, M.; Kumar, M.; Li, D.; McCubrey, J. A.; Mayo, M. W. Regulation of the mitogen-activated protein kinase kinase (MEK)-1 by NAD+-dependent deacetylases. *Oncogene.* **2015***, 34*, 798-804.

162. Hell, C. Ueber eine neue Bromirungsmethode organischer Säuren. *Berichte der deutschen chemischen Gesellschaft.* **1881***, 14*, 891-893.

163. Volhard, J. 4) Ueber Darstellung α-bromirter Säuren. *Justus Liebigs Ann. Chem.* **1887***, 242*, 141-163.

164. Zelinsky, N. Ueber eine bequeme Darstellungsweise von α-Brompropionsäureester. *Berichte der deutschen chemischen Gesellschaft.* **1887***, 20*, 2026-2026.

165. Harpp, D. N.; Bao, L. Q.; Coyle, C.; Gleason, J. G.; Horovitch, S. 2-Bromohexanoyl chloride. *Org. Synth.* **1976***, 55*, 27-31.

166. Lees, W. J.; Singh, R.; Whitesides, G. M. meso-2,5-Dimercapto-N,N,N',N'-tetramethyladipamide: a readily available, kinetically rapid reagent for the reduction of disulfides in aqueous solution. *J. Org. Chem.* **1991***, 56*, 7328-7331.

167. Palmer, D. E.; Pattaroni, C.; Nunami, K.; Chadha, R. K.; Goodman, M.; Wakamiya, T.; Fukase, K.; Horimoto, S.; Kitazawa, M. Effects of dehydroalanine on peptide conformations. *J. Am. Chem. Soc.* **1992***, 114*, 5634-5642.

168. Nathani, R.; Moody, P.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S. Bioconjugation of Green Fluorescent Protein via an Unexpectedly Stable Cyclic Sulfonium Intermediate. *ChemBioChem.* **2012***, 13*, 1283-1285.

169. Ingold, C. K.; Jessop, J. A.; Kuriyan, K. I.; Mandour, A. M. M. Influence of poles and polar linkings on the course pursued by elimination reactions. XVIII. Thermal decomposition of sulfonium hydroxides. *J. Chem. Soc.* **1933**, 533-537.

170. Johnson, A. W. Ylid Chemistry (Organic Chemistry; A Series of Monographs, Vol. 7). InAcademic: 1966; pp 304-367.

171. Franzen, V.; Mertz, C. Mechanism of the Hofmann-elimination with sulfonium salts. *Chem. Ber.* **1960***, 93*, 2819-2824.

172. Franzen, V.; Schmidt, H. J. The mechanism of the Hofmann elimination with sulfonium salts. II. *Chem. Ber.* **1961***, 94*, 2937-2942.

173. Weygand, F.; Daniel, H. Fragmentation of S-methylthiolanium iodide with phenyllithium to ethylene and methyl vinyl sulfide. *Chem. Ber.* **1961***, 94*, 3145-3146.

174. Cristol, S. J.; Stermitz, F. R. Mechanisms of elimination reactions. XXII. Some cis-and trans-2-phenylcyclohexyl derivatives. Hofmann elimination. *J. Am. Chem. Soc.* **1960***, 82*, 4692-4699.

175. Banait, N. S.; Jencks, W. P. Elimination reactions: experimental confirmation of the predicted elimination of (.beta.-cyanoethyl)sulfonium ions through a concerted, E2 mechanism. *J. Am. Chem. Soc.* **1990***, 112*, 6950-6958.

176. Grob, C. A.; Schiess, P. W. Heterolytic Fragmentation. A Class of Organic Reactions. *Angewandte Chemie International Edition in English.* **1967***, 6*, 1-15.

177. Cowburn, D.; Muir, T. W. [3] - Segmental Isotopic Labeling Using Expressed Protein Ligation. *Meth. Enzymol.* **2001***, 339*, 41-54.

178. Liu, D.; Xu, R.; Cowburn, D. Chapter 8 Segmental Isotopic Labeling of Proteins for Nuclear Magnetic Resonance. *Meth. Enzymol.* **2009***, 462*, 151-175.

179. Cremeens, M. E.; Zimmermann, J.; Yu, W.; Dawson, P. E.; Romesberg, F. E. Direct Observation of Structural Heterogeneity in a Î²-Sheet. *J. Am. Chem. Soc.* **2009***, 131*, 5726-5727.

180. Krawczyk, B.; Ensle, P.; Müller, W. M.; Süssmuth, R. D. Deuterium Labeled Peptides Give Insights into the Directionality of Class III Lantibiotic Synthetase LabKC. *J. Am. Chem. Soc.* **2012***, 134*, 9922-9925.

181. Mukherjee, S.; van, d. D. Mechanistic Studies on the Substrate-Tolerant Lanthipeptide Synthetase ProcM. *J. Am. Chem. Soc.* **2014***, 136*, 10450-10459.

182. Markley, J. L.; Putter, I.; Jardetzky, O. High-Resolution Nuclear Magnetic Resonance Spectra of Selectively Deuterated Staphylococcal Nuclease. *Science.* **1968***, 161*, 1249-1251.

183 Feeney, J.; Roberts, G. C. K.; Birdsall, B.; Griffiths, D. V.; King, R. W.; Scudder, P.; Burgen, A. ^{1}H Nuclear Magnetic Resonance Studies of the Tyrosine Residues of Selectively Deuterated Lactobacillus casei Dihydrofolate Reductase. *Proc R Soc Lond B Biol Sci.* **1977***, 196*, 267-290.

184. Feeney, J.; Birdsall, B.; Ostler, G.; Carr, M. D.; Kairi, M. A novel method of preparing totally α-deuterated amino acids for selective incorporation into proteins: Application to assignment of 1H resonances of valine residues in dihydrofolate reductase. *FEBS Lett.* **1990***, 272*, 197-199.

185. Lian, L.; Middleton, D. A. Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Prog Nucl Magn Reson Spectrosc.* **2001***, 39*, 171-190.

186. Chin, J. K.; Jimenez, R.; Romesberg, F. E. Direct Observation of Protein Vibrations by Selective Incorporation of Spectroscopically Observable Carbon-Deuterium Bonds in Cytochrome c. *J. Am. Chem. Soc.* **2001***, 123*, 2426-2427.

187. Groff, D.; Thielges, M.; Cellitti, S.; Schultz, P.; Romesberg, F. Efforts Toward the Direct Experimental Characterization of Enzyme Microenvironments: Tyrosine100 in Dihydrofolate Reductase. *Angewandte Chemie International Edition.* **2009***, 48*, 3478-3481.

189. Axelsson, B. S.; O'Toole, K. J.; Spencer, P. A.; Young, D. W. A versatile synthesis of stereospecificaily labelled D-amino acids and related enzyme inhibitors. *J. Chem. Soc. , Chem. Commun.* **1991**, 1085-1086.

190. Strauss, E.; Begley, T. P. Stereochemical studies on phosphopantothenoylcysteine decarboxylase from Escherichia coli. *Bioorg. Med. Chem. Lett.* **2003***, 13*, 339-342.

191. Khatik, G. L.; Kumar, R.; Chakraborti, A. K. Catalyst-Free Conjugated Addition of Thiols to α,β-Unsaturated Carbonyl Compounds in Water. *Org. Lett.* **2006***, 8*, 2433-2436.

192. Polgar, L.; Bender, M. L. A New Enzyme Containing a Synthetically Formed Active Site. Thiol-Subtilisin1. *J. Am. Chem. Soc.* **1966***, 88*, 3153-3154.

193. Hermanson, G. T.; Editor. *Bioconjugate Techniques*; Academic: 1995; pp 786.

194. Pan, P.; Bayley, H. Caged cysteine and thiophosphoryl peptides. *FEBS Lett.* **1997***, 405*, 81-85.

195. Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. A Genetically Encoded Photocaged Amino Acid. *J. Am. Chem. Soc.* **2004***, 126*, 14306-14307.

196. Jones, D. H.; Cellitti, S. E.; Hao, X.; Zhang, Q.; Jahnz, M.; Summerer, D.; Schultz, P. G.; Uno, T.; Geierstanger, B. H. Site-specific labeling of proteins with NMR-active unnatural amino acids. *J. Biomol. NMR.* **2010***, 46*, 89-100.

197. Devarie-Baez, N.; Bagdon, P. E.; Peng, B.; Zhao, Y.; Park, C.; Xian, M. Light-Induced Hydrogen Sulfide Release from "Caged" gem-Dithiols. *Org. Lett.* **2013***, 15*, 2786-2789.

198. Nguyen, D. P.; Mahesh, M.; Elsasser, S. J.; Hancock, S. M.; Uttamapinant, C.; Chin, J. W. Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. *J. Am. Chem. Soc.* **2014***, 136*, 2240-2243.

199. Zhao, Y.; Ge, Z.; Cheng, T.; Li, R. Thia-Michael Addition Using Cheap and Odorless S-Alkylisothiouronium Salts as Thiol Equivalents in Water. *Synlett.* **2007***, 2007*, 1529-1532.

200. Smith,Amos B., I.,II; Savinov, S. N.; Manjappara, U. V.; Chaiken, I. M. Peptide-small molecule hybrids via orthogonal deprotection-chemoselective conjugation to cysteine-anchored scaffolds. A model study. *Org. Lett.* **2002***, 4*, 4041-4044.

201. Kim, M. S.; Diamond, S. L. Photocleavage of o-nitrobenzyl ether derivatives for rapid biomedical release applications. *Bioorganic & Medicinal Chemistry Letters.* **2006***, 16*, 4007-4010.

202. Adhvaryu, K. K.; Selker, E. U. Protein phosphatase PP1 is required for normal DNA methylation in Neurospora. *Genes Dev.* **2008***, 22*, 3391-3396.

203. Yasuda, S. K.; Lambert, J. L. Preparation and Properties of Anhydrous Trisodium and Tripotassium Monothiophosphates. *J. Am. Chem. Soc.* **1954***, 76*, 5356-5356.

204. Schurter, B. T.; Koh, S. S.; Chen, D.; Bunick, G. J.; Harp, J. M.; Hanson, B. L.; Henschen-Edman, A.; Mackay, D. R.; Stallcup, M. R.; Aswad, D. W. Methylation of Histone H3 by Coactivator-Associated Arginine Methyltransferase 1. *Biochemistry (N. Y.).* **2001***, 40*, 5747-5756.

205. Cheng, J.; Liu, B.; Zhang, X. Radical Substituent Effects of Î±-Sulfonium Groups. *J. Org. Chem.* **1998***, 63*, 7574-7575.

206. Bordwell, F. G. Equilibrium acidities in dimethyl sulfoxide solution. *Acc. Chem. Res.* **1988***, 21*, 456-463.

207. Zhou, H.; van, d. D. Biomimetic Stereoselective Formation of Methyllanthionine. *Org. Lett.* **2002***, 4*, 1335-1338.

208. Zhu, Y.; Gieselman, M. D.; Zhou, H.; Averin, O.; van, d. D. Biomimetic studies on the mechanism of stereoselective lanthionine formation. *Org. Biomol. Chem.* **2003***, 1*, 3304-3315.

209. Toogood, P. L. Model studies of lantibiotic biogenesis. *Tetrahedron Lett.* **1993***, 34*, 7833-7836.

210. Rippa, M.; Bellini, T.; Signorini, M.; Dallocchio, F. Evidence for multiple pairs of vicinal thiols in some proteins. *J. Biol. Chem.* **1981***, 256*, 451-455.

211. Haccard, O.; Jessus, C.; Cayla, X.; Goris, J.; Merlevede, W.; Ozon, R. In vivo activation of a microtubule-associated protein kinase during meiotic maturation of the Xenopus oocyte. *Eur. J. Biochem.* **1990***, 192*, 633-642.

212. Moran, M. F.; Koch, C. A.; Anderson, D.; Ellis, C.; England, L.; Martin, G. S.; Pawson, T. Src homology region 2 domains direct protein-protein interactions in signal transduction. *Proc. Natl. Acad. Sci. U. S. A.* **1990***, 87*, 8622-8226.

213. Alessi, D. R.; Gomez, N.; Moorhead, G.; Lewis, T.; Keyes, S. M.; Cohen, P. Inactivation of p42 MAP kinase by protein phosphatase 2A and a protein tyrosine phosphatase, but not CL100, in various cell lines. *Curr. Biol.* **1995***, 5*, 283-295.

214. Cho, U. S.; Xu, W. Crystal structure of a protein phosphatase 2A heterotrimeric holoenzyme. *Nature.* **2007***, 445*, 53-57.

215. Ahn, N. G.; Seger, R.; Bratlien, R. L.; Diltz, C. D.; Tonks, N. K.; Krebs, E. G. Multiple components in an epidermal growth factor-stimulated protein kinase cascade. In vitro activation of a myelin basic protein/microtubule-associated protein 2 kinase. *Journal of Biological Chemistry.* **1991***, 266*, 4220-4227.

216. Gomez, N.; Cohen, P. Dissection of the protein kinase cascade by which nerve growth factor activates MAP kinases. *Nature.* **1991***, 353*, 170-173.

217. Seger, R.; Ahn, N. G.; Posada, J.; Munar, E. S.; Jensen, A. M.; Cooper, J. A.; Cobb, M. H.; Krebs, E. G. Purification and characterization of mitogen-activated protein kinase activator(s) from epidermal growth factor-stimulated A431 cells. *Journal of Biological Chemistry.* **1992***, 267*, 14373-14381.

218. Robbins, D. J.; Zhen, E.; Owaki, H.; Vanderbilt, C. A.; Ebert, D.; Geppert, T. D.; Cobb, M. H. Regulation and properties of extracellular signal-regulated protein kinases 1 and 2 in vitro. *Journal of Biological Chemistry.* **1993***, 268*, 5097-5106.

219. Burack, W. R.; Sturgill, T. W. The Activating Dual Phosphorylation of MAPK by MEK Is Nonprocessive. *Biochemistry (N. Y.).* **1997***, 36*, 5929-5933.

220. Delaney, A. M.; Printen, J. A.; Chen, H.; Fauman, E. B.; Dudley, D. T. Identification of a Novel Mitogen-Activated Protein Kinase Kinase Activation Domain Recognized by the Inhibitor PD 184352. *Molecular and Cellular Biology.* **2002***, 22*, 7593-7602.

221. Haling, J.; Sudhamsu, J.; Yen, I.; Sideris, S.; Sandoval, W.; Phung, W.; Bravo, B.; Giannetti, A.; Peck, A.; Masselot, A.; Morales, T.; Smith, D.; Brandhuber, B.; Hymowitz, S.; Malek, S. Structure of the BRAF-MEK Complex Reveals a Kinase Activity Independent Role for BRAF in MAPK Signaling. *Cancer Cell.* **2014***, 26*, 402-413.

222. Brennan, D. F.; Dar, A. C.; Hertz, N. T.; Chao, W. C. H.; Burlingame, A. L.; Shokat, K. M.; Barford, D. A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. *Nature.* **2011***, 472*, 366-369.

223. Barbero, N.; Napione, L.; Visentin, S.; Alvaro, M.; Veglio, A.; Bussolino, F.; Viscardi, G. A transient kinetic study between signaling proteins: the case of the MEK-ERK interaction. *Chem. Sci.* **2011***, 2*, 1804-1809.

225. Ohren, J. F.; Chen, H.; Pavlovsky, A.; Whitehead, C.; Zhang, E.; Kuffa, P.; Yan, C.; McConnell, P.; Spessard, C.; Banotai, C.; Mueller, W. T.; Delaney, A.; Omer, C.; Sebolt-Leopold, J.; Dudley, D. T.; Leung, I. K.; Flamme, C.; Warmus, J.; Kaufman, M.; Barrett, S.; Tecle, H.; Hasemann, C. A. Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nat Struct Mol Biol.* **2004***, 11*, 1192-1197.

226. Spicer, J. A.; Rewcastle, G. W.; Kaufman, M. D.; Black, S. L.; Plummer, M. S.; Denny, W. A.; Quin, J.,3rd; Shahripour, A. B.; Barrett, S. D.; Whitehead, C. E.; Milbank, J. B.; Ohren, J. F.; Gowan, R. C.; Omer, C.; Camp, H. S.; Esmaeil, N.; Moore, K.; Sebolt-Leopold, J. S.; Pryzbranowski, S.; Merriman, R. L.; Ortwine, D. F.; Warmus, J. S.; Flamme, C. M.; Pavlovsky, A. G.; Tecle, H. 4-Anilino-5-Carboxamido-2-Pyridone Derivatives as Noncompetitive Inhibitors of Mitogen-Activated Protein Kinase Kinase. *J. Med. Chem.* **2007***, 50*, 5090-5102.

227. Tecle, H.; Shao, J.; Li, Y.; Kothe, M.; Kazmirski, S.; Penzotti, J.; Ding, Y. H.; Ohren, J.; Moshinsky, D.; Coli, R.; Jhawar, N.; Bora, E.; Jacques-O'Hagan, S.; Wu, J. Beyond the MEK-pocket: can current MEK kinase inhibitors be utilized to synthesize novel type III NCKIs? Does the MEK-pocket exist in kinases other than MEK? *Bioorg. Med. Chem. Lett.* **2009***, 19*, 226-229.

228. Iverson, C.; Larson, G.; Lai, C.; Yeh, L. T.; Dadson, C.; Weingarten, P.; Appleby, T.; Vo, T.; Maderna, A.; Vernier, J. M.; Hamatake, R.; Miner, J. N.; Quart, B. RDEA119/BAY 869766: a potent, selective, allosteric inhibitor of MEK1/2 for the treatment of cancer. *Cancer Res.* **2009***, 69*, 6839-6847.

229. Heald, R. A.; Jackson, P.; Savy, P.; Jones, M.; Gancia, E.; Burton, B.; Newman, R.; Boggs, J.; Chan, E.; Chan, J.; Choo, E.; Merchant, M.; Rudewicz, P.; Ultsch, M.; Wiesmann, C.; Yue, Q.; Belvin, M.; Price, S. Discovery of Novel Allosteric Mitogen-Activated Protein Kinase Kinase (MEK) 1,2 Inhibitors Possessing Bidentate Ser212 Interactions. *J. Med. Chem.* **2012***, 55*, 4594-4604.

230. Hartung, I. V.; Hitchcock, M.; Pühler, F.; Neuhaus, R.; Scholz, A.; Hammer, S.; Petersen, K.; Siemeister, G.; Brittain, D.; Hillig, R. C. Optimization of allosteric MEK inhibitors. Part 1: Venturing into underexplored SAR territories. *Bioorg. Med. Chem. Lett.* **2013***, 23*, 2384-2390.

231. Robarge, K. D.; Lee, W.; Eigenbrot, C.; Ultsch, M.; Wiesmann, C.; Heald, R.; Price, S.; Hewitt, J.; Jackson, P.; Savy, P.; Burton, B.; Choo, E. F.; Pang, J.; Boggs, J.; Yang, A.; Yang, X.; Baumgardner, M. Structure based design of novel 6,5 heterobicyclic mitogen-activated protein kinase kinase (MEK) inhibitors leading to the discovery of imidazo[1,5-a] pyrazine G-479. *Bioorg. Med. Chem. Lett.* **2014**, 4714-4723.

232. Isshiki, Y.; Kohchi, Y.; Iikura, H.; Matsubara, Y.; Asoh, K.; Murata, T.; Kohchi, M.; Mizuguchi, E.; Tsujii, S.; Hattori, K.; Miura, T.; Yoshimura, Y.; Aida, S.; Miwa, M.; Saitoh, R.; Murao, N.; Okabe, H.; Belunis, C.; Janson, C.; Lukacs, C.; Schuck, V.; Shimma, N. Design and synthesis of novel allosteric MEK inhibitor CH4987655 as an orally available anticancer agent. *Bioorg. Med. Chem. Lett.* **2011***, 21*, 1795-1801.

233. Amaning, K.; Lowinski, M.; Vallee, F.; Steier, V.; Marcireau, C.; Ugolini, A.; Delorme, C.; Foucalt, F.; McCort, G.; Derimay, N.; Andouche, C.; Vougier, S.; Llopart, S.; Halland, N.; Rak, A. The use of virtual screening and differential scanning fluorimetry for the rapid identification of fragments active against MEK1. *Bioorg. Med. Chem. Lett.* **2013***, 23*, 3620-3626.

234. Sini, P.; Gürtler, U.; Zahn, S. K.; Baumann, C.; Rudolph, D.; Baumgartinger, R.; Strauss, E.; Haslinger, C.; Tontsch-Grunt, U.; Waizenegger, I. C.; Solca, F.; Bader, G.; Zoephel, A.; Treu, M.; Reiser, U.; Garin-Chesa, P.; Boehmelt, G.; Kraut, N.; Quant, J.; Adolf, G. R. Pharmacological Profile of BI 847325, an Orally Bioavailable, ATP-competitive Inhibitor of MEK and Aurora Kinases. *Mol Cancer Ther.* **2016**. *15*(10):2388-2398

235. Jedrychowski, M. P.; Huttlin, E. L.; Haas, W.; Sowa, M. E.; Rad, R.; Gygi, S. P. Evaluation of HCD- and CID-type Fragmentation Within Their Respective Detection Platforms For Murine Phosphoproteomics. *Molecular & Cellular Proteomics : MCP.* **2011***, 10*, M111.009910.

236. Zhu, Y.; Li, H.; Long, C.; Hu, L.; Xu, H.; Liu, L.; Chen, S.; Wang, D.; Shao, F. Structural Insights into the Enzymatic Mechanism of the Pathogenic MAPK Phosphothreonine Lyase. *Mol. Cell.* **2007***, 28*, 899-913.

237. Langer, M.; Lieber, A.; Retey, J. Histidine Ammonia-Lyase Mutant S143C Is Posttranslationally Converted into Fully Active Wild-Type Enzyme. Evidence for Serine 143 To Be the Precursor of Active Site Dehydroalanine. *Biochemistry (N. Y.).* **1994***, 33*, 14034-14038.

238. Mansour, S. J.; Candia, J. M.; Matsuura, J. E.; Manning, M. C.; Ahn, N. G. Interdependent Domains Controlling the Enzymatic Activity of Mitogen-Activated Protein Kinase Kinase 1 . *Biochemistry (N. Y.).* **1996***, 35*, 15529-15536.

239. Xu, B.; Wilsbacher, J. L.; Collisson, T.; Cobb, M. H. The N-terminal ERK-binding site of MEK1 is required for efficient feedback phosphorylation by ERK2 in vitro and ERK activation in vivo. *J. Biol. Chem.* **1999***, 274*, 34029-34035.

240. Desforges, E.; Grysan, A.; Oget, N.; Sindt, M.; Mieloszynski, J. A quantitative synthesis of β-carboxylated thiolophosphates via a Michael reaction. *Tetrahedron Lett.* **2003***, 44*, 6273-6276.

241. Frey, P. A.; Sammons, R. D. Bond order and charge localization in nucleoside phosphorothioates. *Science.* **1985***, 228*, 541.

242. Gutiérrez-Jiménez, M. I.; Aydillo, C.; Navo, C. D.; Avenoza, A.; Corzana, F.; Jiménez-Osés, G.; Zurbano, M. M.; Busto, J. H.; Peregrina, J. M. Bifunctional Chiral Dehydroalanines for Peptide Coupling and Stereoselective S-Michael Addition. *Org. Lett.* **2016***, 18*, 2796-2799.

243. Wang, G.; Yao, S. Q. Combinatorial Synthesis of a Small-Molecule Library Based on the Vinyl Sulfone Scaffold. *Org. Lett.* **2003***, 5*, 4437-4440.

245. Evans, P.; Johnson, P.; Taylor, R. J. K. The Epoxy-Ramberg–Bäcklund Reaction (ERBR): A Sulfone-Based Method for the Synthesis of Allylic Alcohols. *European Journal of Organic Chemistry.* **2006***, 2006*, 1740-1754.

246. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry.* **2004***, 25*, 1605-1612.

247. Patel, S. B.; Cameron, P. M.; Frantz-Wattley, B.; O'Neill, E.; Becker, J. W.; Scapin, G. Lattice stabilization and enhanced diffraction in human p38α crystals by protein engineering. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics.* **2004***, 1696*, 67-73.

248. Fischmann, T. O.; Smith, C. K.; Mayhood, T. W.; Myers, J. E., Jr.; Reichert, P.; Mannarino, A.; Carr, D.; Zhu, H.; Wong, J.; Yang, R.; Le, H. V.; Madison, V. S. Crystal Structures of MEK1 Binary and Ternary Complexes with Nucleotides and Inhibitors. *Biochemistry.* **2009***, 48*, 2661-2674.

249. Goddette, D. W.; Paech, C.; Yang, S. S.; Mielenz, J. R.; Bystroff, C.; Wilke, M. E.; Fletterick, R. J. The crystal structure of the Bacillus lentus alkaline protease, subtilisin BL, at 1.4 Å resolution. *J. Mol. Biol.* **1992***, 228*, 580-595.

250. Chaikuad, A.; M, C. T.; Zimmer, J.; Liang, Y.; Gray, N. S.; Tarsounas, M.; Knapp, S. A unique inhibitor binding site in ERK1/2 is associated with slow binding kinetics. *Nat Chem Biol.* **2014***, 10*, 853-860.

251. Kelley, L. A.; Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protocols.* **2009***, 4*, 363-371.

252. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990***, 215*, 403-410.

253. Hubbard, S. J.; Campbell, S. F.; Thornton, J. M. Molecular recognition. *J. Mol. Biol.* **1991***, 220*, 507-530.

254. Gundry, R. L.; White, M. Y.; Murray, C. I.; Kane, L. A.; Fu, Q.; Stanley, B. A.; Van Eyk, J.,E. Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow. *Current protocols in molecular biology / edited by Frederick M.Ausubel ...[et al.].* **2009***, CHAPTER*, Unit10.25-Unit10.25.

255. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech.* **2008***, 26*, 1367-1372.

256. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011***, 10*, 1794-1805.

257. Cimmperman, P.; Baranauskienė, L.; Jachimovičiūtė, S.; Jachno, J.; Torresan, J.; Michailovienė, V.; Matulienė, J.; Sereikaitė, J.; Bumelis, V.; Matulis, D. A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. *Biophys. J.* **2008***, 95*, 3222-3231.

258. Nielsen, G.; Jonker, H. R. A.; Vajpai, N.; Grzesiek, S.; Schwalbe, H. Kinase in Motion: Insights into the Dynamic Nature of p38α by High-Pressure NMR Spectroscopic Studies. *ChemBioChem.* **2013***, 14*, 1799-1806.

259. Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. Spontaneous α-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **1999***, 267*, 169-184.

260.`Kim, K. M.; Yi, E. C.; Baker, D.; Zhang, K. Y. J. Post-translational modification of the N-terminal His tag interferes with the crystallization of the wild-type and mutant SH3 domains from chicken src tyrosine kinase. *Acta Crystallogr. , Sect. D: Biol. Crystallogr.* **2001***, D57*, 759-762.

261.`Harpp, D. N.; Smith, R. A.; Steliou, K. Studies toward cyclic trisulfides. Trisulfide polymers and sulfur extrusion. *J. Org. Chem.* **1981***, 46*, 2072-9.

262.`Crofts, P. C.; Kosolapoff, G. M. The preparation and the determination of apparent dissociation constants of some substituted aliphatic phosphonic acids. *J. Am. Chem. Soc.* **1953***, 75*, 5738-5740.

263. Flynn, R. M.; Burton, D. J. Synthetic and mechanistic aspects of halo-F-methylphosphonates. *J. Fluorine Chem.* **2011***, 132*, 815-828.

264. Mamalis, P. 924. Reaction of some substituted methylthiomethyl-benzenes and -naphthalenes with methyl iodide. *J. Chem. Soc.* **1960**, 4747-4753.

265. Teotino, U. M.; Cignarella, G. Synthesis of Derivatives of a New Heterocyclic Ring. 1,3-Dihydro-2,3,5-benzothiadiazepine 2,2-Dioxide. *J. Am. Chem. Soc.* **1959***, 81*, 4935-4939.

Appendix A: Naccess Analysis

p38α (PDB 1R3C)

Residue	Number	All		Total side chain		Main chain		Non polar		All-polar	
		ABS	REL	ABS	**REL**	ABS	REL	ABS	REL	ABS	REL
GLU	4	217.4	126.2	161.39	**119.8**	56.02	149.3	88.15	146.2	129.25	115.4
ARG	5	116.98	49	89.55	**44.5**	27.43	73.1	55.99	72	60.98	37.9
PRO	6	71.04	52.2	70.1	**58.5**	0.94	5.8	70.27	58.1	0.77	5.1
THR	7	81.18	58.3	75.33	**74.1**	5.85	15.6	38.88	51.3	42.3	66.6
PHE	8	28.36	14.2	6.48	**3.9**	21.88	61.9	6.49	3.9	21.87	63.9
TYR	9	65.36	30.7	65.01	**36.7**	0.35	1	62.55	45.8	2.82	3.7
ARG	10	144.99	60.7	116.54	**57.9**	28.45	75.8	51.19	65.8	93.8	58.3
GLN	11	62.75	35.2	61.94	**43.9**	0.81	2.1	26.97	51.7	35.77	28.3
GLU	12	130.66	75.9	105.69	**78.4**	24.97	66.6	35.25	58.5	95.41	85.2
LEU	13	23.7	13.3	23.7	**16.8**	0	0	23.7	16.7	0	0
ASN	14	45.35	31.5	27.92	**26.3**	17.43	46.2	14.11	30.5	31.24	32
LYS	15	106.12	52.8	68.2	**41.8**	37.92	101.1	68.68	58.9	37.44	44.4
THR	16	43.19	31	43.18	**42.5**	0.02	0	26.51	35	16.68	26.2
ILE	17	83.85	47.9	75.55	**54.8**	8.3	22.3	75.55	54.3	8.3	23.1
TRP	18	1.03	0.4	0.28	**0.1**	0.75	2	0.28	0.1	0.75	1.3
GLU	19	39.1	22.7	39.1	**29**	0	0	4.51	7.5	34.58	30.9
VAL	20	1.17	0.8	1.17	**1**	0	0	1.17	1	0	0
PRO	21	3.84	2.8	3.84	**3.2**	0	0	3.84	3.2	0	0
GLU	22	96.93	56.3	92.31	**68.5**	4.63	12.3	28.86	47.9	68.08	60.8
ARG	23	50.61	21.2	47.16	**23.4**	3.44	9.2	25.51	32.8	25.1	15.6
TYR	24	0.45	0.2	0.45	**0.3**	0	0	0.02	0	0.43	0.6
GLN	25	66.45	37.2	66.45	**47.1**	0	0	19.91	38.1	46.54	36.9
ASN	26	103.99	72.2	94.31	**88.8**	9.67	25.7	35.27	76.3	68.71	70.3
LEU	27	15.62	8.7	0.82	**0.6**	14.8	39.5	0.82	0.6	14.8	40.7
SER	28	55.55	47.7	55.55	**71.1**	0	0	26.21	54	29.34	43.2
PRO	29	69.96	51.4	51.6	**43**	18.36	113.1	52.93	43.8	17.02	112.1
VAL	30	91.06	60.1	57.67	**50.5**	33.39	89.9	62.05	53.7	29.02	80.7
GLY	31	33.68	42	28.94	**89.5**	4.74	9.9	30.52	81.3	3.16	7.4
SER	32	78.26	67.2	42.28	**54.1**	35.98	93.7	12.29	25.3	65.97	97.1

GLY	33	35.32	44.1	19.17	**59.3**	16.15	33.8	24.18	64.4	11.13	26.2
ALA	34	107.04	99.2	81.43	**117.3**	25.61	66.5	82.62	115.7	24.42	66.8
TYR	35	97.92	46	93.92	**52.9**	4	11.3	65.03	47.6	32.88	43.1
GLY	36	1.18	1.5	0.29	**0.9**	0.89	1.9	1.18	3.1	0	0
SER	37	17.53	15	17.07	**21.9**	0.45	1.2	0.91	1.9	16.62	24.5
VAL	38	21.16	14	21.16	**18.5**	0	0	21.16	18.3	0	0
CYS	39	0	0	0	**0**	0	0	0	0	0	0
ALA	40	8.53	7.9	8.38	**12.1**	0.14	0.4	8.53	11.9	0	0
ALA	41	0	0	0	**0**	0	0	0	0	0	0
PHE	42	67.73	34	67.73	**41.3**	0	0	67.73	41	0	0
ASP	43	1.05	0.7	1.05	**1**	0	0	0	0	1.05	1.2
THR	44	75.63	54.3	49.44	**48.6**	26.19	69.7	41.18	54.4	34.45	54.2
LYS	45	133.22	66.3	107.08	**65.6**	26.14	69.7	62.12	53.3	71.1	84.4
THR	46	63.6	45.7	35.04	**34.5**	28.56	76	38.55	50.9	25.05	39.4
GLY	47	46.12	57.6	21.12	**65.3**	25	52.3	21.97	58.5	24.15	56.7
LEU	48	63.04	35.3	63.04	**44.7**	0	0	63.04	44.3	0	0
ARG	49	123.82	51.9	117.66	**58.5**	6.17	16.4	56.81	73	67.01	41.6
VAL	50	0.05	0	0.05	**0**	0	0	0.05	0	0	0
ALA	51	18.93	17.5	18.32	**26.4**	0.61	1.6	18.34	25.7	0.59	1.6
VAL	52	0.03	0	0.03	**0**	0	0	0.03	0	0	0
LYS	53	36.09	18	36.09	**22.1**	0	0	24.48	21	11.61	13.8
LYS	54	21.4	10.7	20.19	**12.4**	1.21	3.2	14.08	12.1	7.31	8.7
LEU	55	5.17	2.9	3.96	**2.8**	1.21	3.2	4.1	2.9	1.06	2.9
SER	56	52.56	45.1	48.51	**62.1**	4.05	10.6	45.7	94.1	6.87	10.1
ARG	57	151.24	63.3	144.17	**71.6**	7.07	18.8	62.89	80.8	88.35	54.9
PRO	58	6.78	5	6.78	**5.7**	0	0	6.78	5.6	0	0
PHE	59	17.7	8.9	8.64	**5.3**	9.06	25.6	8.7	5.3	9	26.3
GLN	60	76.86	43.1	52.2	**37**	24.66	65.7	6.89	13.2	69.97	55.4
SER	61	34.56	29.7	34.56	**44.3**	0	0	33.29	68.6	1.27	1.9
ILE	62	72.44	41.4	70.03	**50.8**	2.41	6.5	70.03	50.3	2.41	6.7
ILE	63	131.01	74.8	129.5	**93.9**	1.5	4	129.52	93.1	1.49	4.1
HIS	64	72.74	39.8	72.3	**49.2**	0.44	1.2	50.76	52.3	21.98	25.6
ALA	65	0	0	0	**0**	0	0	0	0	0	0
LYS	66	37.12	18.5	37.12	**22.7**	0	0	14.23	12.2	22.89	27.2

ARG	67	34.75	14.6	34.75	**17.3**	0	0	19.23	24.7	15.53	9.6
THR	68	3.52	2.5	2.99	**2.9**	0.53	1.4	0.41	0.5	3.11	4.9
TYR	69	8.33	3.9	8.33	**4.7**	0	0	0.35	0.3	7.97	10.5
ARG	70	46.36	19.4	46.36	**23**	0	0	16.92	21.8	29.44	18.3
GLU	71	6.71	3.9	6.71	**5**	0	0	0.94	1.6	5.77	5.2
LEU	72	1.02	0.6	0.97	**0.7**	0.05	0.1	0.97	0.7	0.05	0.1
ARG	73	38.04	15.9	36.63	**18.2**	1.41	3.7	5.04	6.5	33	20.5
LEU	74	14.23	8	14.14	**10**	0.09	0.2	14.23	10	0	0
LEU	75	19.3	10.8	18.58	**13.2**	0.72	1.9	18.58	13.1	0.72	2
LYS	76	40.14	20	30.46	**18.7**	9.68	25.8	28.35	24.3	11.8	14
HIS	77	36.12	19.8	32.3	**22**	3.82	10.7	14.73	15.2	21.39	25
MET	78	2.49	1.3	0.66	**0.4**	1.83	4.9	0.66	0.4	1.83	5
LYS	79	138.31	68.9	117.93	**72.2**	20.38	54.3	76.04	65.2	62.27	73.9
HIS	80	17.91	9.8	7	**4.8**	10.9	30.5	10.07	10.4	7.84	9.1
GLU	81	118.49	68.8	110.75	**82.2**	7.74	20.6	53.54	88.8	64.95	58
ASN	82	8.66	6	8.66	**8.1**	0	0	0	0	8.66	8.9
VAL	83	3.51	2.3	0	**0**	3.51	9.4	0	0	3.51	9.8
ILE	84	18.75	10.7	18.75	**13.6**	0	0	18.75	13.5	0	0
GLY	85	18.2	22.7	13.32	**41.2**	4.88	10.2	15.38	40.9	2.83	6.6
LEU	86	11.35	6.4	0.89	**0.6**	10.46	27.9	1.48	1	9.87	27.2
LEU	87	20.18	11.3	15.24	**10.8**	4.94	13.2	15.24	10.7	4.94	13.6
ASP	88	17.57	12.5	17.57	**17.1**	0	0	1.67	3.4	15.91	17.5
VAL	89	7.11	4.7	0	**0**	7.11	19.1	0	0	7.11	19.8
PHE	90	14.13	7.1	14.13	**8.6**	0	0	14.13	8.6	0	0
THR	91	12.89	9.3	5.79	**5.7**	7.09	18.9	5.63	7.4	7.25	11.4
PRO	92	22.09	16.2	18.75	**15.6**	3.34	20.6	20.14	16.7	1.95	12.8
ALA	93	11.23	10.4	3.36	**4.8**	7.87	20.4	3.36	4.7	7.87	21.5
ARG	94	157.54	66	131.81	**65.5**	25.74	68.6	49.79	64	107.76	66.9
SER	95	36.76	31.6	36.76	**47.1**	0	0	29.92	61.6	6.84	10.1
LEU	96	62.71	35.1	58.29	**41.3**	4.42	11.8	58.29	41	4.42	12.2
GLU	97	153.64	89.2	126.69	**94**	26.96	71.9	82.45	136.8	71.19	63.6
GLU	98	66.17	38.4	60.94	**45.2**	5.23	13.9	38.91	64.5	27.25	24.3
PHE	99	0.1	0	0	**0**	0.1	0.3	0	0	0.1	0.3
ASN	100	90.01	62.5	84.76	**79.8**	5.24	13.9	24.51	53	65.5	67

ASP	101	44.36	31.6	42.99	**41.9**	1.38	3.7	0.87	1.8	43.49	47.7
VAL	102	0	0	0	**0**	0	0	0	0	0	0
TYR	103	1.75	0.8	1.75	**1**	0	0	0	0	1.75	2.3
LEU	104	4.48	2.5	3.5	**2.5**	0.98	2.6	3.54	2.5	0.94	2.6
VAL	105	0.49	0.3	0.49	**0.4**	0	0	0.49	0.4	0	0
THR	106	15.51	11.1	15.51	**15.2**	0	0	4.64	6.1	10.87	17.1
HIS	107	52.61	28.8	48.51	**33**	4.1	11.4	35.18	36.2	17.43	20.3
LEU	108	56.28	31.5	44.47	**31.5**	11.81	31.5	44.47	31.3	11.81	32.5
MET	109	52.78	27.2	34.44	**22**	18.34	48.9	34.47	21.8	18.31	50.4
GLY	110	74.07	92.5	37.95	**117.4**	36.12	75.6	40.12	106.8	33.95	79.8
ALA	111	36.6	33.9	34.05	**49.1**	2.55	6.6	35.82	50.2	0.77	2.1
ASP	112	26.97	19.2	25.36	**24.7**	1.6	4.3	7.07	14.4	19.89	21.8
LEU	113	2.12	1.2	2.12	**1.5**	0	0	2.12	1.5	0	0
ASN	114	46.69	32.4	46.69	**43.9**	0	0	2.85	6.2	43.84	44.9
ASN	115	91.45	63.5	66.57	**62.7**	24.87	66	20.58	44.5	70.87	72.5
ILE	116	87.5	50	67.85	**49.2**	19.65	52.9	68.27	49.1	19.24	53.5
VAL	117	12.2	8.1	10.03	**8.8**	2.17	5.9	10.03	8.7	2.17	6
LYS	118	110.66	55.1	102.8	**63**	7.86	21	68.02	58.3	42.64	50.6
CYS	119	97.32	72.5	74.5	**77**	22.81	60.8	74.83	76.4	22.49	61.9
GLN	120	106.28	59.5	106.05	**75.2**	0.23	0.6	21.26	40.7	85.02	67.3
LYS	121	155.01	77.2	123.67	**75.7**	31.34	83.5	74.89	64.2	80.12	95.1
LEU	122	16.72	9.4	13.3	**9.4**	3.42	9.1	14.16	9.9	2.56	7
THR	123	95.17	68.3	94.58	**93**	0.59	1.6	87.36	115.4	7.81	12.3
ASP	124	40.78	29	32.16	**31.3**	8.62	22.9	28.43	57.7	12.34	13.5
ASP	125	88.85	63.3	86.63	**84.4**	2.22	5.9	30.17	61.3	58.69	64.4
HIS	126	66.97	36.6	66.36	**45.1**	0.61	1.7	51.99	53.5	14.98	17.5
VAL	127	0	0	0	**0**	0	0	0	0	0	0
GLN	128	18.21	10.2	18.21	**12.9**	0	0	0.68	1.3	17.53	13.9
PHE	129	52.43	26.3	52.43	**31.9**	0	0	52.43	31.7	0	0
LEU	130	0.42	0.2	0.42	**0.3**	0	0	0.42	0.3	0	0
ILE	131	0	0	0	**0**	0	0	0	0	0	0
TYR	132	15.38	7.2	15.38	**8.7**	0	0	12.17	8.9	3.22	4.2
GLN	133	8.5	4.8	8.5	**6**	0	0	4.97	9.5	3.53	2.8
ILE	134	0	0	0	**0**	0	0	0	0	0	0

LEU	135	0	0	0	**0**	0	0	0	0	0	0
ARG	136	13.41	5.6	13.22	**6.6**	0.19	0.5	5.11	6.6	8.29	5.2
GLY	137	0	0	0	**0**	0	0	0	0	0	0
LEU	138	0	0	0	**0**	0	0	0	0	0	0
LYS	139	40.98	20.4	39.56	**24.2**	1.42	3.8	19.29	16.5	21.69	25.7
TYR	140	9.64	4.5	9.64	**5.4**	0	0	5.98	4.4	3.66	4.8
ILE	141	0.08	0	0.08	**0.1**	0	0	0.08	0.1	0	0
HIS	142	12.07	6.6	6.9	**4.7**	5.17	14.4	7.33	7.5	4.74	5.5
SER	143	46.91	40.3	23.98	**30.7**	22.94	59.7	24.16	49.8	22.75	33.5
ALA	144	2.03	1.9	1.38	**2**	0.65	1.7	1.74	2.4	0.29	0.8
ASP	145	107.74	76.7	98.05	**95.5**	9.69	25.7	30.19	61.3	77.55	85.1
ILE	146	6.4	3.7	6.24	**4.5**	0.16	0.4	6.31	4.5	0.09	0.3
ILE	147	26.19	15	21.79	**15.8**	4.4	11.8	21.79	15.7	4.4	12.2
HIS	148	0	0	0	**0**	0	0	0	0	0	0
ARG	149	92.92	38.9	88.23	**43.8**	4.7	12.5	37.26	47.9	55.67	34.6
ASP	150	14.15	10.1	13.94	**13.6**	0.21	0.6	4.37	8.9	9.78	10.7
LEU	151	2.5	1.4	0	**0**	2.5	6.7	0	0	2.5	6.9
LYS	152	24.73	12.3	24.73	**15.1**	0	0	14.98	12.8	9.75	11.6
PRO	153	4.93	3.6	4.93	**4.1**	0	0	4.93	4.1	0	0
SER	154	47.68	40.9	36.43	**46.6**	11.25	29.3	29.57	60.9	18.11	26.7
ASN	155	4.72	3.3	4.72	**4.4**	0	0	0	0	4.72	4.8
LEU	156	0.01	0	0	**0**	0.01	0	0.01	0	0	0
ALA	157	3.43	3.2	3.07	**4.4**	0.35	0.9	3.15	4.4	0.27	0.7
VAL	158	11.47	7.6	8.57	**7.5**	2.9	7.8	8.69	7.5	2.79	7.7
ASN	159	59.48	41.3	50.22	**47.3**	9.26	24.6	28.3	61.2	31.19	31.9
GLU	160	177.67	103.1	148.98	**110.6**	28.69	76.5	70.19	116.4	107.48	96
ASP	161	103.41	73.7	91.13	**88.7**	12.27	32.6	30.72	62.4	72.69	79.7
SER	162	22.12	19	22	**28.2**	0.12	0.3	7.56	15.6	14.56	21.4
GLU	163	89.87	52.2	85.97	**63.8**	3.9	10.4	33.02	54.8	56.85	50.8
LEU	164	1.34	0.7	0.03	**0**	1.31	3.5	0.15	0.1	1.19	3.3
LYS	165	44.46	22.1	44.27	**27.1**	0.19	0.5	21.66	18.6	22.8	27.1
ILE	166	0	0	0	**0**	0	0	0	0	0	0
LEU	167	24.11	13.5	24.11	**17.1**	0	0	24.11	16.9	0	0
ASP	168	89.12	63.5	83.18	**81**	5.94	15.8	22.95	46.6	66.17	72.6

PHE	169	12.62	6.3	4.52	**2.8**	8.1	22.9	4.52	2.7	8.1	23.7
GLY	170	26.3	32.8	13.32	**41.2**	12.98	27.2	15.75	41.9	10.56	24.8
LEU	171	29.33	16.4	17.61	**12.5**	11.72	31.2	20.38	14.3	8.95	24.7
ALA	172	59.99	55.6	43.56	**62.8**	16.43	42.6	45.72	64.1	14.26	39
ARG	173	105.35	44.1	75.96	**37.7**	29.39	78.4	75.96	97.6	29.39	18.3
HIS	174	58.38	31.9	39.41	**26.8**	18.97	53	20.86	21.5	37.52	43.8
THR	175	78.71	56.5	77.62	**76.3**	1.09	2.9	52.3	69.1	26.41	41.6
ASP	176	124.97	89	121.82	**118.6**	3.15	8.3	45.9	93.2	79.07	86.7
ASP	177	113.62	80.9	106.12	**103.3**	7.5	19.9	35.43	72	78.19	85.8
GLU	178	77.59	45	75.17	**55.8**	2.42	6.4	38.01	63	39.58	35.4
MET	179	21.05	10.8	21.05	**13.4**	0	0	21.05	13.3	0	0
THR	180	81.57	58.6	72.99	**71.8**	8.58	22.8	60.31	79.7	21.26	33.5
GLY	181	16.54	20.6	16.23	**50.2**	0.31	0.6	16.23	43.2	0.31	0.7
TYR	182	134.28	63.1	112.99	**63.7**	21.29	60.2	79.97	58.6	54.31	71.2
VAL	183	42.83	28.3	42.83	**37.5**	0	0	42.83	37.1	0	0
ALA	184	46.29	42.9	40.91	**58.9**	5.38	14	40.91	57.3	5.38	14.7
THR	185	36.71	26.4	35.1	**34.5**	1.62	4.3	35.85	47.4	0.86	1.4
ARG	186	0	0	0	**0**	0	0	0	0	0	0
TRP	187	28.28	11.3	27.61	**13.1**	0.67	1.7	27.61	14.6	0.67	1.1
TYR	188	8.57	4	8.32	**4.7**	0.25	0.7	4.61	3.4	3.97	5.2
ARG	189	39.32	16.5	39.32	**19.5**	0	0	6.47	8.3	32.85	20.4
ALA	190	0.07	0.1	0	**0**	0.07	0.2	0	0	0.07	0.2
PRO	191	0	0	0	**0**	0	0	0	0	0	0
GLU	192	2.1	1.2	2.1	**1.6**	0	0	1.98	3.3	0.12	0.1
ILE	193	24.42	13.9	15.71	**11.4**	8.71	23.4	15.71	11.3	8.71	24.2
MET	194	12.25	6.3	2.68	**1.7**	9.56	25.5	2.68	1.7	9.56	26.3
LEU	195	44.66	25	35.19	**24.9**	9.47	25.3	35.62	25	9.04	24.9
ASN	196	121.12	84.1	109.84	**103.4**	11.28	29.9	34.41	74.4	86.71	88.7
TRP	197	3.65	1.5	3.51	**1.7**	0.14	0.4	1.09	0.6	2.56	4.3
MET	198	34.71	17.9	25.5	**16.3**	9.21	24.6	29.82	18.9	4.89	13.5
HIS	199	176.38	96.4	162.49	**110.5**	13.89	38.8	109.12	112.3	67.27	78.5
TYR	200	86.97	40.9	68.96	**38.9**	18	50.9	48.63	35.6	38.33	50.3
ASN	201	59.42	41.3	59.42	**55.9**	0	0	32.8	71	26.62	27.2
GLN	202	61.78	34.6	60.64	**43**	1.14	3	20.98	40.2	40.8	32.3

THR	203	14.88	10.7	14.88	**14.6**	0	0	14.54	19.2	0.34	0.5
VAL	204	1.57	1	1.01	**0.9**	0.56	1.5	1.13	1	0.44	1.2
ASP	205	0.65	0.5	0.65	**0.6**	0	0	0	0	0.65	0.7
ILE	206	1.98	1.1	1.98	**1.4**	0	0	1.98	1.4	0	0
TRP	207	0	0	0	**0**	0	0	0	0	0	0
SER	208	6.37	5.5	6.37	**8.1**	0	0	0	0	6.37	9.4
VAL	209	0	0	0	**0**	0	0	0	0	0	0
GLY	210	0	0	0	**0**	0	0	0	0	0	0
CYS	211	1.08	0.8	1.07	**1.1**	0.01	0	1.07	1.1	0.01	0
ILE	212	0.2	0.1	0.2	**0.1**	0	0	0.2	0.1	0	0
MET	213	0	0	0	**0**	0	0	0	0	0	0
ALA	214	0.69	0.6	0.58	**0.8**	0.11	0.3	0.69	1	0	0
GLU	215	23.25	13.5	22.84	**17**	0.4	1.1	3.96	6.6	19.28	17.2
LEU	216	1.05	0.6	1.05	**0.7**	0	0	1.05	0.7	0	0
LEU	217	19.15	10.7	10.81	**7.7**	8.34	22.2	11.15	7.8	8	22
THR	218	51.5	37	36.98	**36.4**	14.53	38.7	35.08	46.3	16.43	25.8
GLY	219	7.38	9.2	2.47	**7.6**	4.91	10.3	3.06	8.2	4.32	10.2
ARG	220	156.47	65.5	155.72	**77.4**	0.75	2	64.48	82.9	91.99	57.1
THR	221	34.93	25.1	20.79	**20.4**	14.14	37.6	3.38	4.5	31.55	49.6
LEU	222	32.47	18.2	18.25	**12.9**	14.22	37.9	18.25	12.8	14.22	39.2
PHE	223	3.54	1.8	3.05	**1.9**	0.49	1.4	3.05	1.8	0.49	1.4
PRO	224	40.93	30.1	36.22	**30.2**	4.72	29.1	36.22	29.9	4.72	31
GLY	225	0	0	0	**0**	0	0	0	0	0	0
THR	226	73.69	52.9	66.11	**65**	7.59	20.2	55.89	73.8	17.8	28
ASP	227	60.4	43	60.4	**58.8**	0	0	32.27	65.5	28.12	30.9
HIS	228	53	29	52.56	**35.7**	0.44	1.2	28.81	29.7	24.19	28.2
ILE	229	67.29	38.4	67.12	**48.7**	0.17	0.5	67.12	48.2	0.17	0.5
ASP	230	44.36	31.6	44.02	**42.9**	0.34	0.9	10.11	20.5	34.25	37.6
GLN	231	0.15	0.1	0.15	**0.1**	0	0	0	0	0.15	0.1
LEU	232	0.32	0.2	0.32	**0.2**	0	0	0.32	0.2	0	0
LYS	233	105.06	52.3	101.72	**62.3**	3.33	8.9	57.78	49.6	47.28	56.1
LEU	234	63.23	35.4	59.49	**42.2**	3.74	10	59.49	41.8	3.74	10.3
ILE	235	0	0	0	**0**	0	0	0	0	0	0
LEU	236	3.05	1.7	3.05	**2.2**	0	0	3.05	2.1	0	0

ARG	237	143.17	60	142.9	**71**	0.28	0.7	34.83	44.8	108.34	67.3
LEU	238	9.57	5.4	9.57	**6.8**	0	0	9.57	6.7	0	0
VAL	239	17.56	11.6	3.3	**2.9**	14.27	38.4	3.3	2.9	14.27	39.7
GLY	240	2.95	3.7	2.61	**8.1**	0.34	0.7	2.77	7.4	0.18	0.4
THR	241	25.93	18.6	7.87	**7.7**	18.06	48.1	5.06	6.7	20.87	32.8
PRO	242	9.11	6.7	2.02	**1.7**	7.1	43.7	2.02	1.7	7.1	46.7
GLY	243	43.28	54	36.25	**112.1**	7.04	14.7	36.25	96.5	7.04	16.5
ALA	244	76.42	70.8	66.92	**96.4**	9.5	24.7	66.92	93.8	9.5	26
GLU	245	82.3	47.8	79.78	**59.2**	2.52	6.7	37.09	61.5	45.21	40.4
LEU	246	9.4	5.3	8.68	**6.2**	0.72	1.9	8.68	6.1	0.72	2
LEU	247	35.95	20.1	35.82	**25.4**	0.14	0.4	35.95	25.3	0	0
LYS	248	144.95	72.2	121.5	**74.4**	23.45	62.5	89.21	76.5	55.73	66.2
LYS	249	66.1	32.9	56.47	**34.6**	9.62	25.7	52.28	44.8	13.81	16.4
ILE	250	0.67	0.4	0	**0**	0.67	1.8	0.38	0.3	0.3	0.8
SER	251	78.32	67.2	59.65	**76.4**	18.67	48.6	48.76	100.4	29.56	43.5
SER	252	37.57	32.2	37.08	**47.5**	0.48	1.3	29.81	61.4	7.75	11.4
GLU	253	125.34	72.8	118.89	**88.2**	6.45	17.2	42.31	70.2	83.03	74.2
SER	254	87.05	74.7	80.63	**103.2**	6.42	16.7	42.04	86.6	45.01	66.2
ALA	255	15.93	14.8	14.9	**21.5**	1.03	2.7	14.9	20.9	1.03	2.8
ARG	256	63.17	26.5	63.17	**31.4**	0	0	13.78	17.7	49.39	30.7
ASN	257	98.31	68.3	95.46	**89.8**	2.85	7.6	13.79	29.8	84.52	86.5
TYR	258	73.29	34.4	73.29	**41.3**	0	0	69.95	51.2	3.34	4.4
ILE	259	7.53	4.3	4.07	**2.9**	3.46	9.3	4.07	2.9	3.46	9.6
GLN	260	115.46	64.7	89.43	**63.4**	26.03	69.4	32.97	63.1	82.49	65.3
SER	261	79.95	68.6	51.48	**65.9**	28.48	74.2	48.59	100.1	31.36	46.2
LEU	262	38.35	21.5	29.78	**21.1**	8.57	22.9	30.68	21.6	7.68	21.1
THR	263	115.22	82.7	106.32	**104.5**	8.89	23.7	71.58	94.5	43.64	68.7
GLN	264	152.21	85.3	132.77	**94.2**	19.44	51.8	43.57	83.4	108.64	86
MET	265	55.13	28.4	47.27	**30.2**	7.85	20.9	50.39	31.9	4.74	13
PRO	266	112.49	82.6	110.61	**92.3**	1.87	11.5	110.61	91.5	1.87	12.3
LYS	267	99.48	49.5	74.33	**45.5**	25.15	67	65.96	56.6	33.52	39.8
MET	268	80.64	41.5	79.48	**50.7**	1.16	3.1	80.2	50.8	0.44	1.2
ASN	269	86.05	59.8	79.41	**74.7**	6.63	17.6	27.23	58.9	58.82	60.2
PHE	270	1.13	0.6	0.58	**0.4**	0.54	1.5	0.58	0.4	0.54	1.6

ALA	271	57.07	52.9	47.28	**68.1**	9.79	25.4	47.28	66.2	9.79	26.8
ASN	272	106.15	73.7	90.75	**85.4**	15.4	40.9	25.75	55.7	80.4	82.3
VAL	273	47.51	31.4	31.36	**27.4**	16.15	43.5	31.36	27.2	16.15	44.9
PHE	274	7.31	3.7	5.48	**3.3**	1.83	5.2	6.18	3.7	1.13	3.3
ILE	275	148.43	84.8	120.73	**87.5**	27.7	74.5	120.73	86.8	27.7	77
GLY	276	80.38	100.3	39.13	**121**	41.25	86.4	43.04	114.6	37.34	87.8
ALA	277	25.77	23.9	14.91	**21.5**	10.86	28.2	15.2	21.3	10.58	28.9
ASN	278	42.6	29.6	37.42	**35.2**	5.18	13.7	15.67	33.9	26.93	27.6
PRO	279	107.4	78.9	101.17	**84.4**	6.23	38.4	102.62	84.8	4.79	31.5
LEU	280	73.98	41.4	73.98	**52.4**	0	0	73.98	52	0	0
ALA	281	0	0	0	**0**	0	0	0	0	0	0
VAL	282	14.28	9.4	14.28	**12.5**	0	0	14.28	12.4	0	0
ASP	283	47.59	33.9	47.54	**46.3**	0.05	0.1	18.74	38.1	28.84	31.6
LEU	284	0	0	0	**0**	0	0	0	0	0	0
LEU	285	0	0	0	**0**	0	0	0	0	0	0
GLU	286	53.64	31.1	46.81	**34.7**	6.83	18.2	13.57	22.5	40.07	35.8
LYS	287	85.83	42.7	85.43	**52.3**	0.4	1.1	57.47	49.3	28.36	33.7
MET	288	0	0	0	**0**	0	0	0	0	0	0
LEU	289	3.88	2.2	0	**0**	3.88	10.3	0	0	3.88	10.7
VAL	290	38.95	25.7	38.95	**34.1**	0	0	38.95	33.7	0	0
LEU	291	15.74	8.8	10.89	**7.7**	4.85	12.9	10.89	7.7	4.85	13.4
ASP	292	12.07	8.6	11.28	**11**	0.79	2.1	8.64	17.5	3.43	3.8
SER	293	9.5	8.2	4.71	**6**	4.79	12.5	0.52	1.1	8.98	13.2
ASP	294	99.73	71	71.23	**69.4**	28.5	75.6	31.09	63.1	68.64	75.3
LYS	295	119.63	59.6	100.41	**61.5**	19.23	51.3	71.77	61.6	47.86	56.8
ARG	296	11.26	4.7	0.09	**0**	11.18	29.8	0.13	0.2	11.13	6.9
ILE	297	24.07	13.7	22.78	**16.5**	1.3	3.5	22.81	16.4	1.26	3.5
THR	298	45.82	32.9	45.82	**45.1**	0	0	44.01	58.1	1.81	2.8
ALA	299	2.04	1.9	1.35	**1.9**	0.69	1.8	1.35	1.9	0.69	1.9
ALA	300	53.3	49.4	40.98	**59**	12.32	32	42.21	59.1	11.09	30.3
GLN	301	104.02	58.3	99.41	**70.5**	4.6	12.3	45.38	86.9	58.63	46.4
ALA	302	0	0	0	**0**	0	0	0	0	0	0
LEU	303	1.52	0.9	1.52	**1.1**	0	0	1.52	1.1	0	0
ALA	304	68.29	63.3	41.49	**59.8**	26.8	69.5	43.4	60.8	24.89	68.1

HIS	305	24.77	13.5	23.24	**15.8**	1.53	4.3	21.16	21.8	3.61	4.2
ALA	306	58.39	54.1	48.93	**70.5**	9.45	24.5	48.93	68.6	9.45	25.8
TYR	307	3.8	1.8	0.51	**0.3**	3.3	9.3	0.51	0.4	3.3	4.3
PHE	308	0	0	0	**0**	0	0	0	0	0	0
ALA	309	80.57	74.6	58.63	**84.5**	21.93	56.9	59.55	83.4	21.02	57.5
GLN	310	126.4	70.8	100.33	**71.2**	26.07	69.5	38	72.8	88.39	70
TYR	311	63.91	30	51.78	**29.2**	12.13	34.3	37.38	27.4	26.53	34.8
HIS	312	60.54	33.1	40.34	**27.4**	20.19	56.4	30.41	31.3	30.13	35.1
ASP	313	59.61	42.5	59.61	**58**	0	0	21.74	44.2	37.86	41.5
PRO	314	97.89	71.9	79.01	**65.9**	18.88	116.3	79.14	65.4	18.75	123.4
ASP	315	138.46	98.6	106.62	**103.8**	31.85	84.5	39.65	80.5	98.81	108.4
ASP	316	72.27	51.5	72.27	**70.4**	0	0	16.66	33.8	55.61	61
GLU	317	7.51	4.4	2.65	**2**	4.86	12.9	1.57	2.6	5.94	5.3
PRO	318	64.76	47.6	64.36	**53.7**	0.4	2.5	64.36	53.2	0.4	2.7
VAL	319	90.78	59.9	75.19	**65.8**	15.59	42	75.19	65.1	15.59	43.3
ALA	320	14.74	13.7	0.93	**1.3**	13.81	35.8	2.4	3.4	12.34	33.7
ASP	321	105.23	75	101.57	**98.9**	3.66	9.7	26.5	53.8	78.73	86.4
PRO	322	123.64	90.8	110.28	**92**	13.37	82.4	110.64	91.5	13	85.6
TYR	323	18.26	8.6	2.97	**1.7**	15.29	43.2	4.04	3	14.22	18.6
ASP	324	84.11	59.9	80.31	**78.2**	3.8	10.1	44.86	91.1	39.25	43.1
GLN	325	62.02	34.7	53.86	**38.2**	8.17	21.8	26.71	51.2	35.31	28
SER	326	60.11	51.6	55.4	**70.9**	4.7	12.2	51	105	9.11	13.4
PHE	327	1.39	0.7	1.18	**0.7**	0.21	0.6	1.18	0.7	0.21	0.6
GLU	328	60.5	35.1	46.18	**34.3**	14.33	38.2	33.67	55.8	26.83	24
SER	329	106.28	91.2	79.77	**102.1**	26.51	69	47.14	97.1	59.14	87
ARG	330	104.48	43.8	96.82	**48.1**	7.65	20.4	47.31	60.8	57.17	35.5
ASP	331	140.5	100.1	117.13	**114.1**	23.37	62	46.02	93.5	94.48	103.6
LEU	332	34.85	19.5	31.18	**22.1**	3.68	9.8	31.56	22.2	3.29	9.1
LEU	333	116.96	65.5	116.96	**82.9**	0	0	116.96	82.2	0	0
ILE	334	57.03	32.6	54.57	**39.6**	2.47	6.6	54.57	39.2	2.47	6.9
ASP	335	101.17	72.1	95.31	**92.8**	5.85	15.5	37.4	76	63.77	70
GLU	336	76.19	44.2	75.42	**56**	0.77	2	29.73	49.3	46.45	41.5
TRP	337	1.37	0.6	1.37	**0.6**	0	0	0.2	0.1	1.17	2
LYS	338	33.92	16.9	33.92	**20.8**	0	0	7.31	6.3	26.61	31.6

SER	339	35.82	30.7	35.76	**45.8**	0.05	0.1	16.72	34.4	19.09	28.1
LEU	340	54.57	30.5	53.49	**37.9**	1.08	2.9	53.49	37.6	1.08	3
THR	341	0	0	0	**0**	0	0	0	0	0	0
TYR	342	28.96	13.6	28.96	**16.3**	0	0	16.45	12.1	12.51	16.4
ASP	343	70.14	50	69.42	**67.6**	0.72	1.9	30.66	62.3	39.48	43.3
GLU	344	37.17	21.6	37.17	**27.6**	0	0	2.75	4.6	34.43	30.7
VAL	345	10.58	7	5.61	**4.9**	4.97	13.4	5.61	4.9	4.97	13.8
ILE	346	111.45	63.6	85.38	**61.9**	26.07	70.2	85.58	61.5	25.88	71.9
SER	347	64.05	55	53.77	**68.8**	10.29	26.8	48.71	100.3	15.34	22.6
PHE	348	28.24	14.2	11.98	**7.3**	16.26	46	11.98	7.2	16.26	47.5
VAL	349	110.43	72.9	101.21	**88.6**	9.23	24.8	101.21	87.6	9.23	25.7
PRO	350	73.35	53.9	63.61	**53.1**	9.73	60	63.61	52.6	9.73	64.1
PRO	351	48.77	35.8	35.47	**29.6**	13.3	82	35.59	29.4	13.18	86.8
PRO	352	185.95	136.6	124.41	**103.8**	61.54	379.2	157.79	130.5	28.16	185.4
Absolute sum overall chain											
Total		16794.7		4418.1		2376.6		9524.9		7269.9	

MEK1 (PDB 3W8Q)

Residues	Number	All		Total side chain		Main chain		Non polar		All polar	
		ABS	REL	ABS	**REL**	ABS	REL	ABS	REL	ABS	REL
GLU	39	116.56	67.7	70.15	**52.1**	46.41	123.7	39.79	66	76.77	68.6
LEU	40	147.04	82.3	126.21	**89.4**	20.83	55.5	126.6	89	20.44	56.3
GLU	41	150.95	87.6	132.04	**98**	18.91	50.4	73.52	121.9	77.43	69.2
LEU	42	61.52	34.4	35.08	**24.9**	26.44	70.5	37.18	26.1	24.34	67
ASP	43	80.89	57.6	79.18	**77.1**	1.71	4.5	51.66	104.9	29.23	32.1
GLU	44	109.4	63.5	102.05	**75.7**	7.35	19.6	72.34	120	37.06	33.1
GLN	45	105.8	59.3	101.48	**72**	4.32	11.5	45.44	87	60.36	47.8
GLN	46	41.84	23.4	41.41	**29.4**	0.43	1.2	16.05	30.7	25.79	20.4
ARG	47	75.47	31.6	75.47	**37.5**	0	0	39.56	50.8	35.92	22.3
LYS	48	117.98	58.8	116	**71**	1.97	5.3	69.43	59.6	48.55	57.6
ARG	49	78.14	32.7	70.5	**35**	7.63	20.4	30.75	39.5	47.39	29.4
LEU	50	21.56	12.1	20.38	**14.4**	1.19	3.2	20.38	14.3	1.19	3.3
GLU	51	83.86	48.7	82.91	**61.5**	0.95	2.5	41.02	68	42.84	38.3

ALA	52	49.25	45.6	46.94	**67.6**	2.32	6	46.94	65.8	2.32	6.3
PHE	53	28.88	14.5	28.39	**17.3**	0.49	1.4	28.39	17.2	0.49	1.4
LEU	54	65.6	36.7	65.41	**46.4**	0.19	0.5	65.41	46	0.19	0.5
THR	55	71.96	51.7	69.33	**68.2**	2.63	7	47.11	62.2	24.85	39.1
GLN	56	64.33	36	63.7	**45.2**	0.62	1.7	33.62	64.4	30.7	24.3
LYS	57	26.25	13.1	20.34	**12.5**	5.91	15.8	16.68	14.3	9.57	11.4
GLN	58	132.8	74.4	107.54	**76.3**	25.26	67.3	54.78	104.9	78.01	61.8
LYS	59	157.06	78.2	137.22	**84**	19.84	52.9	105.99	90.9	51.06	60.6
VAL	60	12.73	8.4	3.08	**2.7**	9.65	26	3.91	3.4	8.81	24.5
GLY	61	54.85	68.5	36.99	**114.4**	17.86	37.4	40.86	108.8	13.99	32.9
GLU	62	150.53	87.4	135.6	**100.6**	14.93	39.8	61.4	101.8	89.13	79.6
LEU	63	20.94	11.7	12.36	**8.8**	8.58	22.9	12.36	8.7	8.58	23.6
LYS	64	98.22	48.9	98.22	**60.1**	0	0	73.52	63.1	24.7	29.3
ASP	65	65.51	46.7	62.36	**60.7**	3.15	8.4	16.99	34.5	48.51	53.2
ASP	66	118.53	84.4	97.4	**94.8**	21.13	56	49.28	100.1	69.24	76
ASP	67	22.22	15.8	22.15	**21.6**	0.07	0.2	0.64	1.3	21.58	23.7
PHE	68	23.27	11.7	1.13	**0.7**	22.15	62.6	1.62	1	21.66	63.3
GLU	69	76.88	44.6	76.88	**57.1**	0	0	38.67	64.1	38.21	34.1
LYS	70	136.59	68	103.46	**63.4**	33.13	88.3	65.1	55.8	71.49	84.9
ILE	71	91.6	52.3	64.38	**46.7**	27.22	73.3	65.22	46.9	26.37	73.3
SER	72	51.38	44.1	51.17	**65.5**	0.21	0.6	41.01	84.5	10.37	15.3
GLU	73	104.79	60.8	78.73	**58.4**	26.06	69.5	31.21	51.8	73.58	65.7
LEU	74	50.67	28.4	39.34	**27.9**	11.33	30.2	41.78	29.4	8.89	24.5
GLY	75	22.28	27.8	21.11	**65.3**	1.17	2.4	22.04	58.7	0.24	0.6
ALA	76	84.94	78.7	48.38	**69.7**	36.56	94.9	49.28	69	35.66	97.5
GLY	77	30.04	37.5	14.69	**45.4**	15.35	32.1	17.3	46.1	12.74	29.9
ASN	78	143.37	99.6	109.14	**102.7**	34.23	90.8	36.03	77.9 1	7.34	109.8
GLY	79	27.62	34.5	15.35	**47.5**	12.27	25.7	19.72	52.5	7.9	18.6
GLY	80	14.41	18	11.25	**34.8**	3.16	6.6	11.35	30.2	3.06	7.2
VAL	81	40.22	26.6	39.8	**34.8**	0.42	1.1	39.8	34.5	0.42	1.2
VAL	82	22.19	14.7	22.19	**19.4**	0	0	22.19	19.2	0	0
PHE	83	25.7	12.9	25.7	**15.7**	0	0	25.7	15.6	0	0
LYS	84	53.51	26.6	53.51	**32.8**	0	0	12.26	10.5	41.25	49
VAL	85	2.96	2	2.96	**2.6**	0	0	2.96	2.6	0	0

SER	86	20.65	17.7	20.65	**26.4**	0	0	14.6	30.1	6.05	8.9
HIS	87	0.98	0.5	0.98	**0.7**	0	0	0.16	0.2	0.82	1
LYS	88	123.65	61.6	105.97	**64.9**	17.69	47.1	61.32	52.6	62.33	74
PRO	89	87.28	64.1	58.19	**48.5**	29.1	179.3	60.51	50	26.77	176.3
SER	90	50.75	43.6	22.71	**29.1**	28.04	73	22.08	45.5	28.67	42.2
GLY	91	50.54	63.1	28.29	**87.5**	22.24	46.6	31.58	84.1	18.96	44.5
LEU	92	43.57	24.4	43.46	**30.8**	0.11	0.3	43.46	30.5	0.11	0.3
VAL	93	63.55	42	51.34	**44.9**	12.21	32.9	51.34	44.5	12.21	33.9
MET	94	0	0	0	**0**	0	0	0	0	0	0
ALA	95	7.33	6.8	7.33	**10.6**	0	0	7.33	10.3	0	0
ARG	96	17.09	7.2	17.09	**8.5**	0	0	1.96	2.5	15.13	9.4
LYS	97	22.47	11.2	22.47	**13.8**	0	0	18.2	15.6	4.28	5.1
LEU	98	26.19	14.7	26.19	**18.6**	0	0	26.19	18.4	0	0
ILE	99	3.7	2.1	3.7	**2.7**	0	0	3.7	2.7	0	0
HIS	100	118.16	64.6	96.44	**65.6**	21.72	60.7	68.5	70.5	49.66	57.9
LEU	101	38.83	21.7	36.4	**25.8**	2.42	6.5	36.98	26	1.85	5.1
GLU	102	102.03	59.2	95.7	**71**	6.33	16.9	71.6	118.8	30.43	27.2
ILE	103	2.49	1.4	2.49	**1.8**	0	0	2.49	1.8	0	0
LYS	104	112.07	55.8	112.07	**68.6**	0	0	89.54	76.8	22.53	26.7
PRO	105	98.44	72.3	84.56	**70.5**	13.88	85.5	86.74	71.7	11.7	77
ALA	106	79.71	73.8	64.41	**92.8**	15.3	39.7	64.45	90.3	15.26	41.7
ILE	107	22.72	13	22.72	**16.5**	0	0	22.72	16.3	0	0
ARG	108	74.35	31.1	74.35	**36.9**	0	0	25.33	32.6	49.02	30.5
ASN	109	84.7	58.8	84.7	**79.7**	0	0	22.82	49.4	61.88	63.3
GLN	110	81.35	45.6	71.1	**50.4**	10.25	27.3	35.9	68.8	45.45	36
ILE	111	7.07	4	7.01	**5.1**	0.05	0.1	7.01	5	0.05	0.2
ILE	112	42.23	24.1	42.23	**30.6**	0	0	42.23	30.3	0	0
ARG	113	149.62	62.7	148.82	**73.9**	0.81	2.1	67.73	87.1	81.89	50.9
GLU	114	54.77	31.8	54.77	**40.7**	0	0	26.69	44.3	28.09	25.1
LEU	115	1.53	0.9	1.53	**1.1**	0	0	1.53	1.1	0	0
GLN	116	77.22	43.3	77.22	**54.8**	0	0	34.31	65.7	42.9	34
VAL	117	88.8	58.6	85.43	**74.8**	3.38	9.1	85.43	74	3.38	9.4
LEU	118	58.46	32.7	47.82	**33.9**	10.65	28.4	47.82	33.6	10.65	29.3
HIS	119	16.65	9.1	12.73	**8.7**	3.92	10.9	9.76	10	6.89	8

GLU	120	44.82	26	35.01	**26**	9.81	26.2	19.02	31.5	25.8	23
CYS	121	22.71	16.9	5.3	**5.5**	17.41	46.4	8.52	8.7	14.2	39.1
ASN	122	59.82	41.6	52.23	**49.2**	7.59	20.1	32.35	70	27.47	28.1
SER	123	6.04	5.2	3.11	**4**	2.93	7.6	1.94	4	4.11	6
PRO	124	17.43	12.8	15.39	**12.8**	2.04	12.6	15.39	12.7	2.04	13.4
TYR	125	34.12	16	34.03	**19.2**	0.08	0.2	26.31	19.3	7.8	10.2
ILE	126	10.17	5.8	6.07	**4.4**	4.1	11	6.07	4.4	4.1	11.4
VAL	127	4.19	2.8	1.03	**0.9**	3.16	8.5	1.03	0.9	3.16	8.8
GLY	128	9.95	12.4	9.86	**30.5**	0.08	0.2	9.92	26.4	0.03	0.1
PHE	129	2.79	1.4	2.06	**1.3**	0.73	2.1	2.06	1.2	0.73	2.1
TYR	130	18.35	8.6	7.18	**4**	11.17	31.6	0.6	0.4	17.75	23.3
GLY	131	9.43	11.8	9.43	**29.2**	0	0	9.43	25.1	0	0
ALA	132	19.29	17.9	0.67	**1**	18.62	48.3	0.88	1.2	18.41	50.3
PHE	133	27.4	13.7	26.35	**16.1**	1.05	3	27.3	16.5	0.1	0.3
TYR	134	32.32	15.2	10.19	**5.7**	22.13	62.5	10.04	7.4	22.28	29.2
SER	135	32.27	27.7	31.5	**40.3**	0.77	2	29.45	60.7	2.82	4.1
ASP	136	131.43	93.6	109.13	**106.3**	22.3	59.2	51.89	105.4	79.54	87.3
GLY	137	19.81	24.7	11.81	**36.5**	8	16.8	11.81	31.5	8	18.8
GLU	138	38.32	22.2	37.99	**28.2**	0.33	0.9	3.09	5.1	35.23	31.5
ILE	139	3.07	1.8	3.07	**2.2**	0	0	3.07	2.2	0	0
SER	140	0.68	0.6	0.68	**0.9**	0	0	0.31	0.6	0.37	0.5
ILE	141	0.06	0	0.06	**0**	0	0	0.06	0	0	0
CYS	142	0	0	0	**0**	0	0	0	0	0	0
MET	143	10.62	5.5	10.62	**6.8**	0	0	10.62	6.7	0	0
GLU	144	32.31	18.8	30.89	**22.9**	1.42	3.8	12.22	20.3	20.08	17.9
HIS	145	40.21	22	23.31	**15.9**	16.89	47.2	10.98	11.3	29.23	34.1
MET	146	7.45	3.8	0.65	**0.4**	6.8	18.1	0.65	0.4	6.8	18.7
ASP	147	54.18	38.6	39.69	**38.7**	14.49	38.4	21.45	43.6	32.73	35.9
GLY	148	5.25	6.6	0	**0**	5.25	11	1.81	4.8	3.45	8.1
GLY	149	15.94	19.9	15.03	**46.5**	0.91	1.9	15.03	40	0.91	2.1
SER	150	12.49	10.7	11.79	**15.1**	0.69	1.8	2.91	6	9.58	14.1
LEU	151	0	0	0	**0**	0	0	0	0	0	0
ASP	152	43.9	31.3	41.73	**40.6**	2.17	5.7	19.74	40.1	24.15	26.5
GLN	153	62.88	35.2	59.53	**42.2**	3.35	8.9	34.59	66.2	28.3	22.4

VAL	154	1.67	1.1	1.64	**1.4**	0.02	0.1	1.64	1.4	0.02	0.1
LEU	155	17.7	9.9	16.57	**11.7**	1.13	3	16.67	11.7	1.03	2.8
LYS	156	155.43	77.4	131.65	**80.6**	23.78	63.4	108.22	92.8	47.2	56
LYS	157	131.72	65.6	107.43	**65.8**	24.29	64.7	66.61	57.1	65.11	77.3
ALA	158	37.72	34.9	11.19	**16.1**	26.54	68.9	14.52	20.3	23.2	63.4
GLY	159	43.59	54.4	36.52	**112.9**	7.08	14.8	37.69	100.4	5.91	13.9
ARG	160	91.67	38.4	82.28	**40.9**	9.39	25	19.26	24.8	72.41	45
ILE	161	0.86	0.5	0	**0**	0.86	2.3	0.09	0.1	0.77	2.1
PRO	162	60.02	44.1	59.94	**50**	0.08	0.5	59.94	49.6	0.08	0.5
GLU	163	25.86	15	22.89	**17**	2.98	7.9	17.19	28.5	8.67	7.7
GLN	164	87.24	48.9	85.81	**60.9**	1.42	3.8	28.91	55.4	58.33	46.2
ILE	165	5.74	3.3	5.74	**4.2**	0	0	5.74	4.1	0	0
LEU	166	0	0	0	**0**	0	0	0	0	0	0
GLY	167	0	0	0	**0**	0	0	0	0	0	0
LYS	168	38.79	19.3	37.01	**22.7**	1.77	4.7	25.95	22.3	12.83	15.2
VAL	169	0.46	0.3	0.46	**0.4**	0	0	0.46	0.4	0	0
SER	170	0.2	0.2	0.2	**0.2**	0	0	0.2	0.4	0	0
ILE	171	14.07	8	14.07	**10.2**	0	0	14.07	10.1	0	0
ALA	172	4.86	4.5	3.75	**5.4**	1.11	2.9	4.67	6.5	0.19	0.5
VAL	173	1.92	1.3	1.92	**1.7**	0	0	1.92	1.7	0	0
ILE	174	0.15	0.1	0.15	**0.1**	0	0	0.15	0.1	0	0
LYS	175	37.8	18.8	34.15	**20.9**	3.65	9.7	14.29	12.3	23.51	27.9
GLY	176	0	0	0	**0**	0	0	0	0	0	0
LEU	177	0	0	0	**0**	0	0	0	0	0	0
THR	178	46.19	33.2	46.19	**45.4**	0	0	41.1	54.3	5.09	8
TYR	179	40.07	18.8	40	**22.6**	0.07	0.2	31.95	23.4	8.12	10.6
LEU	180	6.62	3.7	6.62	**4.7**	0	0	6.62	4.7	0	0
ARG	181	67.12	28.1	48.75	**24.2**	18.37	49	14	18	53.12	33
GLU	182	121.16	70.3	92.07	**68.3**	29.09	77.6	39.33	65.2	81.84	73.1
LYS	183	139.64	69.5	111.97	**68.6**	27.67	73.8	68.12	58.4	71.52	84.9
HIS	184	58.79	32.1	27.93	**19**	30.86	86.2	21.67	22.3	37.12	43.3
LYS	185	42.48	21.2	17.63	**10.8**	24.85	66.2	15.07	12.9	27.41	32.5
ILE	186	55.44	31.7	31.74	**23**	23.7	63.8	31.74	22.8	23.7	65.9
MET	187	61.44	31.6	61.26	**39.1**	0.18	0.5	61.26	38.8	0.18	0.5

HIS	188	106.6	58.3	93.39	**63.5**	13.22	36.9	73.77	75.9	32.83	38.3
ARG	189	16.6	7	16.6	**8.2**	0	0	4.01	5.1	12.59	7.8
ASP	190	53.06	37.8	47.94	**46.7**	5.13	13.6	28.05	57	25.02	27.4
VAL	191	0.28	0.2	0.2	**0.2**	0.08	0.2	0.2	0.2	0.08	0.2
LYS	192	109.79	54.7	109.79	**67.2**	0	0	87.2	74.8	22.59	26.8
PRO	193	8.22	6	8.22	**6.9**	0	0	8.22	6.8	0	0
SER	194	49.01	42.1	38.76	**49.6**	10.24	26.7	26.13	53.8	22.87	33.7
ASN	195	10.2	7.1	10.2	**9.6**	0	0	0	0	10.2	10.4
ILE	196	2.63	1.5	2.63	**1.9**	0	0	2.63	1.9	0	0
LEU	197	26.11	14.6	26.11	**18.5**	0	0	26.11	18.3	0	0
VAL	198	0	0	0	**0**	0	0	0	0	0	0
ASN	199	9.4	6.5	9.4	**8.8**	0	0	3.55	7.7	5.85	6
SER	200	37.75	32.4	25.2	**32.3**	12.54	32.7	7.95	16.4	29.8	43.9
ARG	201	150.63	63.1	130.5	**64.8**	20.12	53.6	51.16	65.8	99.47	61.8
GLY	202	19.78	24.7	2.98	**9.2**	16.8	35.2	4.55	12.1	15.23	35.8
GLU	203	31.71	18.4	31.71	**23.5**	0	0	20.25	33.6	11.46	10.2
ILE	204	17.55	10	11.97	**8.7**	5.59	15	11.97	8.6	5.59	15.5
LYS	205	16.3	8.1	15.88	**9.7**	0.42	1.1	1.58	1.4	14.72	17.5
LEU	206	1.23	0.7	0.82	**0.6**	0.41	1.1	0.82	0.6	0.41	1.1
CYS	207	22.17	16.5	21.67	**22.4**	0.5	1.3	21.67	22.1	0.5	1.4
ASP	208	11.51	8.2	11.04	**10.7**	0.47	1.3	4.8	9.7	6.71	7.4
PHE	209	23.71	11.9	19.6	**11.9**	4.11	11.6	19.6	11.9	4.11	12
GLY	210	40.29	50.3	27.86	**86.2**	12.43	26	28.62	76.2	11.67	27.4
VAL	211	2.89	1.9	0.89	**0.8**	2	5.4	0.89	0.8	2	5.6
SER	212	15.83	13.6	13.1	**16.8**	2.73	7.1	12.37	25.5	3.46	5.1
GLY	213	47.77	59.6	37.75	**116.8**	10.02	21	42.35	112.8	5.42	12.7
GLN	214	22.21	12.4	21.18	**15**	1.03	2.8	14.32	27.4	7.89	6.3
LEU	215	0.35	0.2	0.35	**0.2**	0	0	0.35	0.2	0	0
ILE	216	90.02	51.4	89.88	**65.1**	0.14	0.4	89.88	64.6	0.14	0.4
ASP	217	73.19	52.1	72.83	**70.9**	0.36	1	25.45	51.7	47.74	52.4
SER	218	9.53	8.2	6.95	**8.9**	2.59	6.7	6.58	13.6	2.96	4.4
MET	219	27.24	14	16.73	**10.7**	10.52	28	18.51	11.7	8.74	24.1
ALA	220	68.87	63.8	49.65	**71.5**	19.22	49.9	50.15	70.3	18.72	51.2
ASN	221	60.73	42.2	46.26	**43.5**	14.47	38.4	23.9	51.7	36.82	37.7

SER	222	43.35	37.2	30	**38.4**	13.34	34.7	29.51	60.8	13.84	20.4
PHE	223	87.48	43.9	63.65	**38.8**	23.84	67.4	66.11	40	21.38	62.5
VAL	224	136.18	89.9	125.99	**110.2**	10.19	27.4	126.36	109.4	9.82	27.3
GLY	225	50.11	62.6	39.22	**121.3**	10.89	22.8	39.54	105.3	10.57	24.8
THR	226	144.07	103.4	115.08	**113.2**	28.99	77.2	86.8	114.6	57.27	90.1
ARG	227	144.91	60.7	129.03	**64.1**	15.87	42.3	37.83	48.6 1	7.08	66.5
SER	228	72.39	62.1	40.59	**52**	31.8	82.8	36.62	75.4	35.77	52.6
TYR	229	169.08	79.5	137.19	**77.3**	31.89	90.1	122.79	90	46.29	60.7
MET	230	48.51	25	34.09	**21.8**	14.42	38.5	35.26	22.3	13.25	36.5
SER	231	37.74	32.4	35.67	**45.7**	2.07	5.4	16.07	33.1	21.67	31.9
PRO	232	98.95	72.7	98.24	**81.9**	0.71	4.4	98.24	81.2	0.71	4.7
GLU	233	100	58.1	92.67	**68.8**	7.33	19.5	44.27	73.4	55.73	49.8
ARG	234	97.03	40.6	92.39	**45.9**	4.64	12.4	19.27	24.8	77.76	48.3
LEU	235	104.42	58.5	74.68	**52.9**	29.74	79.3	75.41	53	29.01	79.9
GLN	236	140.26	78.6	128.83	**91.4**	11.43	30.5	55.36	106	84.9	67.2
GLY	237	4.82	6	0.06	**0.2**	4.76	10	0.27	0.7	4.55	10.7
THR	238	102.51	73.6	90.86	**89.3**	11.65	31	72.18	95.3	30.33	47.7
HIS	239	123.97	67.8	96.43	**65.6**	27.54	76.9	63.23	65.1	60.74	70.8
TYR	240	36.6	17.2	36.31	**20.5**	0.29	0.8	29.97	22	6.63	8.7
SER	241	41.04	35.2	39.12	**50.1**	1.92	5	25.95	53.4	15.1	22.2
VAL	242	18.78	12.4	17.95	**15.7**	0.83	2.2	17.95	15.5	0.83	2.3
GLN	243	31.56	17.7	31.56	**22.4**	0	0	6.03	11.6	25.52	20.2
SER	244	37.75	32.4	36.52	**46.7**	1.24	3.2	24.38	50.2	13.38	19.7
ASP	245	0.23	0.2	0.23	**0.2**	0	0	0	0	0.23	0.2
ILE	246	0	0	0	**0**	0	0	0	0	0	0
TRP	247	36.54	14.7	36.54	**17.3**	0	0	25.98	13.7	10.56	17.7
SER	248	12.98	11.1	12.09	**15.5**	0.89	2.3	4.19	8.6	8.8	12.9
MET	249	3.18	1.6	3.18	**2**	0	0	3.18	2	0	0
GLY	250	0	0	0	**0**	0	0	0	0	0	0
LEU	251	21.46	12	21.46	**15.2**	0	0	21.46	15.1	0	0
SER	252	5.77	5	5.74	**7.4**	0.03	0.1	1.66	3.4	4.11	6.1
LEU	253	2.21	1.2	2.21	**1.6**	0	0	2.21	1.6	0	0
VAL	254	1.99	1.3	1.99	**1.7**	0	0	1.99	1.7	0	0
GLU	255	36.43	21.1	36.43	**27**	0	0	0.43	0.7	36	32.2

MET	256	0.6	0.3	0.6	**0.4**	0	0	0.6	0.4	0	0
ALA	257	0.85	0.8	0	**0**	0.85	2.2	0	0	0.85	2.3
VAL	258	5.54	3.7	3.64	**3.2**	1.89	5.1	3.64	3.2	1.89	5.3
GLY	259	28.69	35.8	2.13	**6.6**	26.56	55.6	4.17	11.1	24.52	57.6
ARG	260	63.79	26.7	63.79	**31.7**	0	0	40.73	52.4	23.06	14.3
TYR	261	88.79	41.7	80.06	**45.1**	8.72	24.7	71.4	52.3	17.39	22.8
PRO	262	3.14	2.3	2.76	**2.3**	0.38	2.3	2.76	2.3	0.38	2.5
ILE	263	6.68	3.8	1.93	**1.4**	4.76	12.8	1.93	1.4	4.76	13.2
PRO	264	62.24	45.7	62.24	**51.9**	0	0	62.24	51.5	0	0
PRO	265	59.38	43.6	46.08	**38.4**	13.31	82	46.13	38.1	13.25	87.2
PRO	266	24.72	18.2	16.3	**13.6**	8.41	51.8	16.53	13.7	8.19	53.9
ASP	267	90.74	64.6	89.26	**86.9**	1.48	3.9	53.93	109.5	36.81	40.4
ALA	268	77.74	72	68.23	**98.3**	9.5	24.7	68.82	96.4	8.91	24.4
LYS	269	131.68	65.6	127.19	**77.9**	4.49	12	81.33	69.8	50.36	59.8
GLU	270	62.23	36.1	62.23	**46.2**	0	0	38.63	64.1	23.61	21.1
LEU	271	28.04	15.7	28.04	**19.9**	0	0	28.04	19.7	0	0
GLU	272	82.2	47.7	78.48	**58.2**	3.72	9.9	24.16	40.1	58.04	51.8
LEU	273	119.72	67	95.89	**67.9**	23.83	63.5	97.7	68.7	22.02	60.6
MET	274	47.43	24.4	35.63	**22.7**	11.8	31.5	35.63	22.6	11.8	32.5
PHE	275	32.36	16.2	8.25	**5**	24.11	68.2	12.81	7.8	19.55	57.1
GLY	276	55.54	69.3	36.01	**111.4**	19.53	40.9	37.77	100.6	17.77	41.8
CYS	277	107.05	79.7	90.47	**93.5**	16.58	44.2	92.77	94.7	14.28	39.3
GLN	278	94.78	53.1	62.44	**44.3**	32.34	86.2	33.12	63.4	61.67	48.8
ARG	305	139.42	58.4	80.11	**39.8**	59.31	158.1	81.34	104.5	58.08	36.1
PRO	306	123.08	90.4	106.37	**88.7**	16.71	103	107.97	89.3	15.11	99.5
PRO	307	120.04	88.2	105.85	**88.3**	14.2	87.5	105.85	87.5	14.19	93.4
MET	308	56.42	29.1	34.27	**21.9**	22.15	59.1	34.27	21.7	22.15	61
ALA	309	61.65	57.1	60.47	**87.1**	1.18	3	61.32	85.9	0.33	0.9
ILE	310	113.49	64.8	110.18	**79.9**	3.32	8.9	110.18	79.2	3.32	9.2
PHE	311	159.67	80	158.98	**96.9**	0.69	1.9	158.98	96.2	0.69	2
GLU	312	84.56	49.1	82.41	**61.2**	2.15	5.7	37.27	61.8	47.29	42.2
LEU	313	35.39	19.8	35.39	**25.1**	0	0	35.39	24.9	0	0
LEU	314	94.25	52.8	92.21	**65.3**	2.04	5.4	93.19	65.5	1.06	2.9
ASP	315	59.5	42.4	58.55	**57**	0.95	2.5	25.49	51.8	34.02	37.3

TYR	316	50.27	23.6	49.1	**27.7**	1.18	3.3	34.78	25.5	15.49	20.3
ILE	317	48.09	27.5	45.44	**32.9**	2.65	7.1	45.44	32.7	2.65	7.4
VAL	318	105.78	69.8	96.78	**84.7**	9	24.2	96.78	83.8	9	25
ASN	319	96.39	67	71.41	**67.2**	24.98	66.3	24.58	53.2	71.81	73.5
GLU	320	69.04	40.1	66.46	**49.3**	2.59	6.9	13.53	22.4	55.52	49.6
PRO	321	95.49	70.1	91.65	**76.4**	3.84	23.6	91.65	75.8	3.84	25.3
PRO	322	47	34.5	41.91	**35**	5.09	31.4	42.04	34.8	4.96	32.6
PRO	323	8.94	6.6	7.63	**6.4**	1.31	8.1	7.63	6.3	1.31	8.6
LYS	324	138.04	68.7	120.13	**73.6**	17.91	47.8	78.49	67.3	59.54	70.7
LEU	325	8.36	4.7	0	**0**	8.36	22.3	0.49	0.3	7.87	21.7
PRO	326	12.52	9.2	12.41	**10.4**	0.1	0.6	12.41	10.3	0.1	0.7
SER	327	58.19	50	56.87	**72.8**	1.32	3.4	38.52	79.3	19.67	28.9
GLY	328	68.08	85	34.1	**105.5**	33.98	71.1	37.04	98.6	31.04	73
VAL	329	52.33	34.6	31.77	**27.8**	20.55	55.3	33.16	28.7	19.16	53.3
PHE	330	18.18	9.1	14.25	**8.7**	3.93	11.1	15.08	9.1	3.1	9
SER	331	33.37	28.6	30.01	**38.4**	3.36	8.8	30.01	61.8	3.36	5
LEU	332	137.91	77.2	130.41	**92.4**	7.51	20	131.54	92.4	6.37	17.5
GLU	333	58.49	34	58.49	**43.4**	0	0	33.7	55.9	24.78	22.1
PHE	334	0	0	0	**0**	0	0	0	0	0	0
GLN	335	24.65	13.8	24.58	**17.4**	0.07	0.2	7	13.4	17.65	14
ASP	336	54.15	38.6	53.91	**52.5**	0.24	0.6	19.64	39.9	34.5	37.9
PHE	337	0	0	0	**0**	0	0	0	0	0	0
VAL	338	0.08	0.1	0.08	**0.1**	0	0	0.08	0.1	0	0
ASN	339	62.3	43.3	59.75	**56.2**	2.56	6.8	22.32	48.3	39.98	40.9
LYS	340	62.43	31.1	62.43	**38.2**	0	0	24.77	21.2	37.66	44.7
CYS	341	0	0	0	**0**	0	0	0	0	0	0
LEU	342	1.05	0.6	0.83	**0.6**	0.22	0.6	1.05	0.7	0	0
ILE	343	49.11	28	49.11	**35.6**	0	0	49.11	35.3	0	0
LYS	344	74.86	37.3	53.38	**32.7**	21.47	57.2	50.96	43.7	23.89	28.4
ASN	345	72.14	50.1	71.37	**67.2**	0.77	2	29.45	63.7	42.69	43.7
PRO	346	69.45	51	67.62	**56.4**	1.84	11.3	68.67	56.8	0.78	5.2
ALA	347	90.39	83.7	61.52	**88.6**	28.87	74.9	61.87	86.7	28.52	78
GLU	348	120.5	70	106.08	**78.7**	14.41	38.4	42.76	70.9	77.74	69.4
ARG	349	22.57	9.5	17.66	**8.8**	4.92	13.1	0.05	0.1	22.52	14

ALA	350	0.49	0.5	0	**0**	0.49	1.3	0	0	0.49	1.3
ASP	351	35.87	25.6	35.43	**34.5**	0.45	1.2	17.07	34.7	18.81	20.6
LEU	352	16.3	9.1	14.59	**10.3**	1.71	4.5	15.29	10.7	1.02	2.8
LYS	353	158.23	78.8	154.42	**94.6**	3.81	10.2	115.07	98.7	43.15	51.2
GLN	354	74.77	41.9	73.26	**52**	1.51	4	32.16	61.6	42.62	33.7
LEU	355	0	0	0	**0**	0	0	0	0	0	0
MET	356	25.39	13.1	25.36	**16.2**	0.03	0.1	25.36	16.1	0.03	0.1
VAL	357	81.29	53.7	75.67	**66.2**	5.62	15.1	75.7	65.6	5.6	15.6
HIS	358	20.2	11	20.2	**13.7**	0	0	18.43	19	1.77	2.1
ALA	359	50.17	46.5	42.84	**61.7**	7.33	19	43.13	60.4	7.03	19.2
PHE	360	0	0	0	**0**	0	0	0	0	0	0
ILE	361	0.66	0.4	0.66	**0.5**	0	0	0.66	0.5	0	0
LYS	362	117.24	58.4	109.78	**67.2**	7.46	19.9	91.71	78.7	25.53	30.3
ARG	363	70.67	29.6	69.81	**34.7**	0.86	2.3	61.01	78.4	9.66	6
SER	364	0.03	0	0.03	**0**	0	0	0.03	0.1	0	0
ASP	365	67.02	47.7	47.97	**46.7**	19.06	50.5	14.15	28.7	52.87	58
ALA	366	83.57	77.4	64.96	**93.6**	18.61	48.3	65.84	92.2	17.73	48.5
GLU	367	39.5	22.9	35.55	**26.4**	3.94	10.5	27.26	45.2	12.24	10.9
GLU	368	178.48	103.6	147.1	**109.2**	31.37	83.6	69.97	116.1 1	8.5	96.9
VAL	369	33.24	21.9	9.29	**8.1**	23.95	64.4	10.75	9.3	22.49	62.5
ASP	370	102.81	73.2	90.65	**88.3**	12.16	32.2	47.94	97.4	54.87	60.2
PHE	371	13.67	6.9	12.25	**7.5**	1.42	4	12.25	7.4	1.42	4.2
ALA	372	30.08	27.9	27.65	**39.8**	2.42	6.3	28.55	40	1.52	4.2
GLY	373	32.46	40.5	26.09	**80.7**	6.37	13.3	30.99	82.5	1.47	3.4
TRP	374	54.56	21.9	52.19	**24.7**	2.38	6.2	39	20.6	15.56	26.1
LEU	375	0	0	0	**0**	0	0	0	0	0	0
CYS	376	23.4	17.4	21.18	**21.9**	2.22	5.9	22.54	23	0.86	2.4
SER	377	94.6	81.2	72.86	**93.3**	21.74	56.6	42.26	87.1	52.34	77
THR	378	29.74	21.4	20.25	**19.9**	9.49	25.3	13.93	18.4	15.81	24.9
ILE	379	37.47	21.4	19.47	**14.1**	18	48.4	19.93	14.3	17.54	48.7
GLY	380	57.68	72	32.38	**100.1**	25.31	53	33.17	88.3	24.51	57.6
LEU	381	60.46	33.8	47.32	**33.5**	13.14	35	47.77	33.6	12.69	34.9
ASN	382	183.48	127.5	123.43	**116.2**	60.05	159.3	73.68	159.4 1	9.79	112.4
Absolute sums overall chain											

Total		16468.7		4023.1		2445.6		80.8		87.9	

MEK1+ATP (3W8Q)

Residues	Number	All		Total side chain		Main chain		Non polar		All polar	
		ABS	REL	ABS	**REL**	ABS	REL	ABS	REL	ABS	REL
GLU	39	116.56	67.7	70.15	**52.1**	46.41	123.7	39.79	66	76.77	68.6
LEU	40	147.04	82.3	126.21	**89.4**	20.83	55.5	126.6	89	20.44	56.3
GLU	41	150.95	87.6	132.04	**98**	18.91	50.4	73.52	121.9	77.43	69.2
LEU	42	61.52	34.4	35.08	**24.9**	26.44	70.5	37.18	26.1	24.34	67
ASP	43	80.89	57.6	79.18	**77.1**	1.71	4.5	51.66	104.9	29.23	32.1
GLU	44	109.4	63.5	102.05	**75.7**	7.35	19.6	72.34	120	37.06	33.1
GLN	45	105.8	59.3	101.48	**72**	4.32	11.5	45.44	87	60.36	47.8
GLN	46	41.84	23.4	41.41	**29.4**	0.43	1.2	16.05	30.7	25.79	20.4
ARG	47	75.47	31.6	75.47	**37.5**	0	0	39.56	50.8	35.92	22.3
LYS	48	117.98	58.8	116	**71**	1.97	5.3	69.43	59.6	48.55	57.6
ARG	49	78.14	32.7	70.5	**35**	7.63	20.4	30.75	39.5	47.39	29.4
LEU	50	21.56	12.1	20.38	**14.4**	1.19	3.2	20.38	14.3	1.19	3.3
GLU	51	83.86	48.7	82.91	**61.5**	0.95	2.5	41.02	68	42.84	38.3
ALA	52	49.25	45.6	46.94	**67.6**	2.32	6	46.94	65.8	2.32	6.3
PHE	53	28.88	14.5	28.39	**17.3**	0.49	1.4	28.39	17.2	0.49	1.4
LEU	54	65.6	36.7	65.41	**46.4**	0.19	0.5	65.41	46	0.19	0.5
THR	55	71.96	51.7	69.33	**68.2**	2.63	7	47.11	62.2	24.85	39.1
GLN	56	64.33	36	63.7	**45.2**	0.62	1.7	33.62	64.4	30.7	24.3
LYS	57	26.25	13.1	20.34	**12.5**	5.91	15.8	16.68	14.3	9.57	11.4
GLN	58	132.8	74.4	107.54	**76.3**	25.26	67.3	54.78	104.9	78.01	61.8
LYS	59	157.06	78.2	137.22	**84**	19.84	52.9	105.99	90.9	51.06	60.6
VAL	60	12.73	8.4	3.08	**2.7**	9.65	26	3.91	3.4	8.81	24.5
GLY	61	54.85	68.5	36.99	**114.4**	17.86	37.4	40.86	108.8	13.99	32.9
GLU	62	150.53	87.4	135.6	**100.6**	14.93	39.8	61.4	101.8	89.13	79.6
LEU	63	20.94	11.7	12.36	**8.8**	8.58	22.9	12.36	8.7	8.58	23.6
LYS	64	98.22	48.9	98.22	**60.1**	0	0	73.52	63.1	24.7	29.3
ASP	65	65.51	46.7	62.36	**60.7**	3.15	8.4	16.99	34.5	48.51	53.2
ASP	66	118.53	84.4	97.4	**94.8**	21.13	56	49.28	100.1	69.24	76

ASP	67	22.22	15.8	22.15	**21.6**	0.07	0.2	0.64	1.3	21.58	23.7
PHE	68	23.27	11.7	1.13	**0.7**	22.15	62.6	1.62	1	21.66	63.3
GLU	69	76.88	44.6	76.88	**57.1**	0	0	38.67	64.1	38.21	34.1
LYS	70	136.59	68	103.46	**63.4**	33.13	88.3	65.1	55.8	71.49	84.9
ILE	71	91.6	52.3	64.38	**46.7**	27.22	73.3	65.22	46.9	26.37	73.3
SER	72	51.38	44.1	51.17	**65.5**	0.21	0.6	41.01	84.5	10.37	15.3
GLU	73	104.79	60.8	78.73	**58.4**	26.06	69.5	31.21	51.8	73.58	65.7
LEU	74	32.07	18	23.2	**16.4**	8.87	23.7	24.17	17	7.9	21.7
GLY	75	11.91	14.9	11.91	**36.8**	0	0	11.91	31.7	0	0
ALA	76	71.4	66.1	48.38	**69.7**	23.02	59.7	48.38	67.8	23.02	62.9
GLY	77	14.76	18.4	0	**0**	14.76	30.9	2.02	5.4	12.74	29.9
ASN	78	121.76	84.6	97.3	**91.6**	24.47	64.9	24.19	52.3	97.57	99.9
GLY	79	25.41	31.7	15.35	**47.5**	10.06	21.1	19.72	52.5	5.7	13.4
GLY	80	13.83	17.3	11.25	**34.8**	2.58	5.4	11.34	30.2	2.49	5.9
VAL	81	40.22	26.6	39.8	**34.8**	0.42	1.1	39.8	34.5	0.42	1.2
VAL	82	0	0	0	**0**	0	0	0	0	0	0
PHE	83	25.7	12.9	25.7	**15.7**	0	0	25.7	15.6	0	0
LYS	84	53.51	26.6	53.51	**32.8**	0	0	12.26	10.5	41.25	49
VAL	85	2.96	2	2.96	**2.6**	0	0	2.96	2.6	0	0
SER	86	20.65	17.7	20.65	**26.4**	0	0	14.6	30.1	6.05	8.9
HIS	87	0.98	0.5	0.98	**0.7**	0	0	0.16	0.2	0.82	1
LYS	88	123.65	61.6	105.97	**64.9**	17.69	47.1	61.32	52.6	62.33	74
PRO	89	87.28	64.1	58.19	**48.5**	29.1	179.3	60.51	50	26.77	176.3
SER	90	50.75	43.6	22.71	**29.1**	28.04	73	22.08	45.5	28.67	42.2
GLY	91	50.54	63.1	28.29	**87.5**	22.24	46.6	31.58	84.1	18.96	44.5
LEU	92	43.57	24.4	43.46	**30.8**	0.11	0.3	43.46	30.5	0.11	0.3
VAL	93	63.55	42	51.34	**44.9**	12.21	32.9	51.34	44.5	12.21	33.9
MET	94	0	0	0	**0**	0	0	0	0	0	0
ALA	95	0	0	0	**0**	0	0	0	0	0	0
ARG	96	17.09	7.2	17.09	**8.5**	0	0	1.96	2.5	15.13	9.4
LYS	97	0.09	0	0.09	**0.1**	0	0	0.03	0	0.06	0.1
LEU	98	26.19	14.7	26.19	**18.6**	0	0	26.19	18.4	0	0
ILE	99	1.33	0.8	1.33	**1**	0	0	1.33	1	0	0
HIS	100	118.16	64.6	96.44	**65.6**	21.72	60.7	68.5	70.5	49.66	57.9

LEU	101	38.83	21.7	36.4	**25.8**	2.42	6.5	36.98	26	1.85	5.1
GLU	102	102.03	59.2	95.7	**71**	6.33	16.9	71.6	118.8	30.43	27.2
ILE	103	2.49	1.4	2.49	**1.8**	0	0	2.49	1.8	0	0
LYS	104	112.07	55.8	112.07	**68.6**	0	0	89.54	76.8	22.53	26.7
PRO	105	98.44	72.3	84.56	**70.5**	13.88	85.5	86.74	71.7	11.7	77
ALA	106	79.71	73.8	64.41	**92.8**	15.3	39.7	64.45	90.3	15.26	41.7
ILE	107	22.72	13	22.72	**16.5**	0	0	22.72	16.3	0	0
ARG	108	74.35	31.1	74.35	**36.9**	0	0	25.33	32.6	49.02	30.5
ASN	109	84.7	58.8	84.7	**79.7**	0	0	22.82	49.4	61.88	63.3
GLN	110	81.35	45.6	71.1	**50.4**	10.25	27.3	35.9	68.8	45.45	36
ILE	111	7.07	4	7.01	**5.1**	0.05	0.1	7.01	5	0.05	0.2
ILE	112	42.23	24.1	42.23	**30.6**	0	0	42.23	30.3	0	0
ARG	113	149.62	62.7	148.82	**73.9**	0.81	2.1	67.73	87.1	81.89	50.9
GLU	114	54.77	31.8	54.77	**40.7**	0	0	26.69	44.3	28.09	25.1
LEU	115	1.53	0.9	1.53	**1.1**	0	0	1.53	1.1	0	0
GLN	116	77.22	43.3	77.22	**54.8**	0	0	34.31	65.7	42.9	34
VAL	117	88.8	58.6	85.43	**74.8**	3.38	9.1	85.43	74	3.38	9.4
LEU	118	58.46	32.7	47.82	**33.9**	10.65	28.4	47.82	33.6	10.65	29.3
HIS	119	16.65	9.1	12.73	**8.7**	3.92	10.9	9.76	10	6.89	8
GLU	120	44.82	26	35.01	**26**	9.81	26.2	19.02	31.5	25.8	23
CYS	121	22.71	16.9	5.3	**5.5**	17.41	46.4	8.52	8.7	14.2	39.1
ASN	122	59.82	41.6	52.23	**49.2**	7.59	20.1	32.35	70	27.47	28.1
SER	123	6.04	5.2	3.11	**4**	2.93	7.6	1.94	4	4.11	6
PRO	124	17.43	12.8	15.39	**12.8**	2.04	12.6	15.39	12.7	2.04	13.4
TYR	125	34.12	16	34.03	**19.2**	0.08	0.2	26.31	19.3	7.8	10.2
ILE	126	10.17	5.8	6.07	**4.4**	4.1	11	6.07	4.4	4.1	11.4
VAL	127	3.84	2.5	0.68	**0.6**	3.16	8.5	0.68	0.6	3.16	8.8
GLY	128	9.95	12.4	9.86	**30.5**	0.08	0.2	9.92	26.4	0.03	0.1
PHE	129	2.79	1.4	2.06	**1.3**	0.73	2.1	2.06	1.2	0.73	2.1
TYR	130	18.35	8.6	7.18	**4**	11.17	31.6	0.6	0.4	17.75	23.3
GLY	131	9.43	11.8	9.43	**29.2**	0	0	9.43	25.1	0	0
ALA	132	19.29	17.9	0.67	**1**	18.62	48.3	0.88	1.2	18.41	50.3
PHE	133	27.4	13.7	26.35	**16.1**	1.05	3	27.3	16.5	0.1	0.3
TYR	134	32.32	15.2	10.19	**5.7**	22.13	62.5	10.04	7.4	22.28	29.2

SER	135	32.27	27.7	31.5	**40.3**	0.77	2	29.45	60.7	2.82	4.1
ASP	136	131.43	93.6	109.13	**106.3**	22.3	59.2	51.89	105.4	79.54	87.3
GLY	137	19.81	24.7	11.81	**36.5**	8	16.8	11.81	31.5	8	18.8
GLU	138	38.32	22.2	37.99	**28.2**	0.33	0.9	3.09	5.1	35.23	31.5
ILE	139	3.07	1.8	3.07	**2.2**	0	0	3.07	2.2	0	0
SER	140	0.68	0.6	0.68	**0.9**	0	0	0.31	0.6	0.37	0.5
ILE	141	0.06	0	0.06	**0**	0	0	0.06	0	0	0
CYS	142	0	0	0	**0**	0	0	0	0	0	0
MET	143	0.87	0.4	0.87	**0.6**	0	0	0.87	0.6	0	0
GLU	144	30.89	17.9	30.89	**22.9**	0	0	12.22	20.3	18.67	16.7
HIS	145	39.79	21.8	22.9	**15.6**	16.89	47.2	10.56	10.9	29.23	34.1
MET	146	0.46	0.2	0.01	**0**	0.45	1.2	0.01	0	0.45	1.2
ASP	147	54.18	38.6	39.69	**38.7**	14.49	38.4	21.45	43.6	32.73	35.9
GLY	148	5.25	6.6	0	**0**	5.25	11	1.81	4.8	3.45	8.1
GLY	149	11.77	14.7	10.85	**33.6**	0.91	1.9	10.85	28.9	0.91	2.1
SER	150	0	0	0	**0**	0	0	0	0	0	0
LEU	151	0	0	0	**0**	0	0	0	0	0	0
ASP	152	42.6	30.3	40.43	**39.4**	2.17	5.7	19.74	40.1	22.86	25.1
GLN	153	59.28	33.2	55.93	**39.7**	3.35	8.9	34.59	66.2	24.69	19.6
VAL	154	1.67	1.1	1.64	**1.4**	0.02	0.1	1.64	1.4	0.02	0.1
LEU	155	17.7	9.9	16.57	**11.7**	1.13	3	16.67	11.7	1.03	2.8
LYS	156	155.43	77.4	131.65	**80.6**	23.78	63.4	108.22	92.8	47.2	56
LYS	157	131.72	65.6	107.43	**65.8**	24.29	64.7	66.61	57.1	65.11	77.3
ALA	158	37.72	34.9	11.19	**16.1**	26.54	68.9	14.52	20.3	23.2	63.4
GLY	159	43.59	54.4	36.52	**112.9**	7.08	14.8	37.69	100.4	5.91	13.9
ARG	160	91.67	38.4	82.28	**40.9**	9.39	25	19.26	24.8	72.41	45
ILE	161	0.86	0.5	0	**0**	0.86	2.3	0.09	0.1	0.77	2.1
PRO	162	60.02	44.1	59.94	**50**	0.08	0.5	59.94	49.6	0.08	0.5
GLU	163	25.86	15	22.89	**17**	2.98	7.9	17.19	28.5	8.67	7.7
GLN	164	87.24	48.9	85.81	**60.9**	1.42	3.8	28.91	55.4	58.33	46.2
ILE	165	5.74	3.3	5.74	**4.2**	0	0	5.74	4.1	0	0
LEU	166	0	0	0	**0**	0	0	0	0	0	0
GLY	167	0	0	0	**0**	0	0	0	0	0	0
LYS	168	38.79	19.3	37.01	**22.7**	1.77	4.7	25.95	22.3	12.83	15.2

VAL	169	0.46	0.3	0.46	**0.4**	0	0	0.46	0.4	0	0
SER	170	0.2	0.2	0.2	**0.2**	0	0	0.2	0.4	0	0
ILE	171	14.07	8	14.07	**10.2**	0	0	14.07	10.1	0	0
ALA	172	4.86	4.5	3.75	**5.4**	1.11	2.9	4.67	6.5	0.19	0.5
VAL	173	1.92	1.3	1.92	**1.7**	0	0	1.92	1.7	0	0
ILE	174	0.15	0.1	0.15	**0.1**	0	0	0.15	0.1	0	0
LYS	175	37.8	18.8	34.15	**20.9**	3.65	9.7	14.29	12.3	23.51	27.9
GLY	176	0	0	0	**0**	0	0	0	0	0	0
LEU	177	0	0	0	**0**	0	0	0	0	0	0
THR	178	46.19	33.2	46.19	**45.4**	0	0	41.1	54.3	5.09	8
TYR	179	40.07	18.8	40	**22.6**	0.07	0.2	31.95	23.4	8.12	10.6
LEU	180	6.62	3.7	6.62	**4.7**	0	0	6.62	4.7	0	0
ARG	181	67.12	28.1	48.75	**24.2**	18.37	49	14	18	53.12	33
GLU	182	121.16	70.3	92.07	**68.3**	29.09	77.6	39.33	65.2	81.84	73.1
LYS	183	139.64	69.5	111.97	**68.6**	27.67	73.8	68.12	58.4	71.52	84.9
HIS	184	58.79	32.1	27.93	**19**	30.86	86.2	21.67	22.3	37.12	43.3
LYS	185	42.48	21.2	17.63	**10.8**	24.85	66.2	15.07	12.9	27.41	32.5
ILE	186	55.44	31.7	31.74	**23**	23.7	63.8	31.74	22.8	23.7	65.9
MET	187	61.44	31.6	61.26	**39.1**	0.18	0.5	61.26	38.8	0.18	0.5
HIS	188	106.6	58.3	93.39	**63.5**	13.22	36.9	73.77	75.9	32.83	38.3
ARG	189	16.6	7	16.6	**8.2**	0	0	4.01	5.1	12.59	7.8
ASP	190	53.06	37.8	47.94	**46.7**	5.13	13.6	28.05	57	25.02	27.4
VAL	191	0.28	0.2	0.2	**0.2**	0.08	0.2	0.2	0.2	0.08	0.2
LYS	192	109.79	54.7	109.79	**67.2**	0	0	87.2	74.8	22.59	26.8
PRO	193	8.22	6	8.22	**6.9**	0	0	8.22	6.8	0	0
SER	194	29.6	25.4	29.33	**37.5**	0.27	0.7	19.48	40.1	10.12	14.9
ASN	195	0.93	0.6	0.93	**0.9**	0	0	0	0	0.93	0.9
ILE	196	2.63	1.5	2.63	**1.9**	0	0	2.63	1.9	0	0
LEU	197	0	0	0	**0**	0	0	0	0	0	0
VAL	198	0	0	0	**0**	0	0	0	0	0	0
ASN	199	9.4	6.5	9.4	**8.8**	0	0	3.55	7.7	5.85	6
SER	200	37.75	32.4	25.2	**32.3**	12.54	32.7	7.95	16.4	29.8	43.9
ARG	201	150.63	63.1	130.5	**64.8**	20.12	53.6	51.16	65.8	99.47	61.8
GLY	202	19.78	24.7	2.98	**9.2**	16.8	35.2	4.55	12.1	15.23	35.8

GLU	203	31.71	18.4	31.71	**23.5**	0	0	20.25	33.6	11.46	10.2
ILE	204	17.55	10	11.97	**8.7**	5.59	15	11.97	8.6	5.59	15.5
LYS	205	16.3	8.1	15.88	**9.7**	0.42	1.1	1.58	1.4	14.72	17.5
LEU	206	1.23	0.7	0.82	**0.6**	0.41	1.1	0.82	0.6	0.41	1.1
CYS	207	0.5	0.4	0	**0**	0.5	1.3	0	0	0.5	1.4
ASP	208	11.04	7.9	11.04	**10.7**	0	0	4.8	9.7	6.24	6.8
PHE	209	13.24	6.6	9.13	**5.6**	4.11	11.6	9.13	5.5	4.11	12
GLY	210	26.76	33.4	23.63	**73.1**	3.13	6.6	24.39	65	2.37	5.6
VAL	211	0	0	0	**0**	0	0	0	0	0	0
SER	212	11.82	10.1	10.87	**13.9**	0.95	2.5	11.17	23	0.64	0.9
GLY	213	47.77	59.6	37.75	**116.8**	10.02	21	42.35	112.8	5.42	12.7
GLN	214	22.21	12.4	21.18	**15**	1.03	2.8	14.32	27.4	7.89	6.3
LEU	215	0.35	0.2	0.35	**0.2**	0	0	0.35	0.2	0	0
ILE	216	90.02	51.4	89.88	**65.1**	0.14	0.4	89.88	64.6	0.14	0.4
ASP	217	73.19	52.1	72.83	**70.9**	0.36	1	25.45	51.7	47.74	52.4
SER	218	9.53	8.2	6.95	**8.9**	2.59	6.7	6.58	13.6	2.96	4.4
MET	219	27.24	14	16.73	**10.7**	10.52	28	18.51	11.7	8.74	24.1
ALA	220	68.87	63.8	49.65	**71.5**	19.22	49.9	50.15	70.3	18.72	51.2
ASN	221	60.73	42.2	46.26	**43.5**	14.47	38.4	23.9	51.7	36.82	37.7
SER	222	43.35	37.2	30	**38.4**	13.34	34.7	29.51	60.8	13.84	20.4
PHE	223	87.48	43.9	63.65	**38.8**	23.84	67.4	66.11	40	21.38	62.5
VAL	224	136.18	89.9	125.99	**110.2**	10.19	27.4	126.36	109.4	9.82	27.3
GLY	225	50.11	62.6	39.22	**121.3**	10.89	22.8	39.54	105.3	10.57	24.8
THR	226	144.07	103.4	115.08	**113.2**	28.99	77.2	86.8	114.6	57.27	90.1
ARG	227	144.91	60.7	129.03	**64.1**	15.87	42.3	37.83	48.6	107.08	66.5
SER	228	72.39	62.1	40.59	**52**	31.8	82.8	36.62	75.4	35.77	52.6
TYR	229	169.08	79.5	137.19	**77.3**	31.89	90.1	122.79	90	46.29	60.7
MET	230	48.51	25	34.09	**21.8**	14.42	38.5	35.26	22.3	13.25	36.5
SER	231	37.74	32.4	35.67	**45.7**	2.07	5.4	16.07	33.1	21.67	31.9
PRO	232	98.95	72.7	98.24	**81.9**	0.71	4.4	98.24	81.2	0.71	4.7
GLU	233	100	58.1	92.67	**68.8**	7.33	19.5	44.27	73.4	55.73	49.8
ARG	234	97.03	40.6	92.39	**45.9**	4.64	12.4	19.27	24.8	77.76	48.3
LEU	235	104.42	58.5	74.68	**52.9**	29.74	79.3	75.41	53	29.01	79.9
GLN	236	140.26	78.6	128.83	**91.4**	11.43	30.5	55.36	106	84.9	67.2

GLY	237	4.82	6	0.06	**0.2**	4.76	10	0.27	0.7	4.55	10.7
THR	238	102.51	73.6	90.86	**89.3**	11.65	31	72.18	95.3	30.33	47.7
HIS	239	123.97	67.8	96.43	**65.6**	27.54	76.9	63.23	65.1	60.74	70.8
TYR	240	36.6	17.2	36.31	**20.5**	0.29	0.8	29.97	22	6.63	8.7
SER	241	41.04	35.2	39.12	**50.1**	1.92	5	25.95	53.4	15.1	22.2
VAL	242	18.78	12.4	17.95	**15.7**	0.83	2.2	17.95	15.5	0.83	2.3
GLN	243	31.56	17.7	31.56	**22.4**	0	0	6.03	11.6	25.52	20.2
SER	244	37.75	32.4	36.52	**46.7**	1.24	3.2	24.38	50.2	13.38	19.7
ASP	245	0.23	0.2	0.23	**0.2**	0	0	0	0	0.23	0.2
ILE	246	0	0	0	**0**	0	0	0	0	0	0
TRP	247	36.54	14.7	36.54	**17.3**	0	0	25.98	13.7	10.56	17.7
SER	248	12.98	11.1	12.09	**15.5**	0.89	2.3	4.19	8.6	8.8	12.9
MET	249	3.18	1.6	3.18	**2**	0	0	3.18	2	0	0
GLY	250	0	0	0	**0**	0	0	0	0	0	0
LEU	251	21.46	12	21.46	**15.2**	0	0	21.46	15.1	0	0
SER	252	5.77	5	5.74	**7.4**	0.03	0.1	1.66	3.4	4.11	6.1
LEU	253	2.21	1.2	2.21	**1.6**	0	0	2.21	1.6	0	0
VAL	254	1.99	1.3	1.99	**1.7**	0	0	1.99	1.7	0	0
GLU	255	36.43	21.1	36.43	**27**	0	0	0.43	0.7	36	32.2
MET	256	0.6	0.3	0.6	**0.4**	0	0	0.6	0.4	0	0
ALA	257	0.85	0.8	0	**0**	0.85	2.2	0	0	0.85	2.3
VAL	258	5.54	3.7	3.64	**3.2**	1.89	5.1	3.64	3.2	1.89	5.3
GLY	259	28.69	35.8	2.13	**6.6**	26.56	55.6	4.17	11.1	24.52	57.6
ARG	260	63.79	26.7	63.79	**31.7**	0	0	40.73	52.4	23.06	14.3
TYR	261	88.79	41.7	80.06	**45.1**	8.72	24.7	71.4	52.3	17.39	22.8
PRO	262	3.14	2.3	2.76	**2.3**	0.38	2.3	2.76	2.3	0.38	2.5
ILE	263	6.68	3.8	1.93	**1.4**	4.76	12.8	1.93	1.4	4.76	13.2
PRO	264	62.24	45.7	62.24	**51.9**	0	0	62.24	51.5	0	0
PRO	265	59.38	43.6	46.08	**38.4**	13.31	82	46.13	38.1	13.25	87.2
PRO	266	24.72	18.2	16.3	**13.6**	8.41	51.8	16.53	13.7	8.19	53.9
ASP	267	90.74	64.6	89.26	**86.9**	1.48	3.9	53.93	109.5	36.81	40.4
ALA	268	77.74	72	68.23	**98.3**	9.5	24.7	68.82	96.4	8.91	24.4
LYS	269	131.68	65.6	127.19	**77.9**	4.49	12	81.33	69.8	50.36	59.8
GLU	270	62.23	36.1	62.23	**46.2**	0	0	38.63	64.1	23.61	21.1

LEU	271	28.04	15.7	28.04	**19.9**	0	0	28.04	19.7	0	0
GLU	272	82.2	47.7	78.48	**58.2**	3.72	9.9	24.16	40.1	58.04	51.8
LEU	273	119.72	67	95.89	**67.9**	23.83	63.5	97.7	68.7	22.02	60.6
MET	274	47.43	24.4	35.63	**22.7**	11.8	31.5	35.63	22.6	11.8	32.5
PHE	275	32.36	16.2	8.25	**5**	24.11	68.2	12.81	7.8	19.55	57.1
GLY	276	55.54	69.3	36.01	**111.4**	19.53	40.9	37.77	100.6	17.77	41.8
CYS	277	107.05	79.7	90.47	**93.5**	16.58	44.2	92.77	94.7	14.28	39.3
GLN	278	94.78	53.1	62.44	**44.3**	32.34	86.2	33.12	63.4	61.67	48.8
ARG	305	139.42	58.4	80.11	**39.8**	59.31	158.1	81.34	104.5	58.08	36.1
PRO	306	123.08	90.4	106.37	**88.7**	16.71	103	107.97	89.3	15.11	99.5
PRO	307	120.04	88.2	105.85	**88.3**	14.2	87.5	105.85	87.5	14.19	93.4
MET	308	56.42	29.1	34.27	**21.9**	22.15	59.1	34.27	21.7	22.15	61
ALA	309	61.65	57.1	60.47	**87.1**	1.18	3	61.32	85.9	0.33	0.9
ILE	310	113.49	64.8	110.18	**79.9**	3.32	8.9	110.18	79.2	3.32	9.2
PHE	311	159.67	80	158.98	**96.9**	0.69	1.9	158.98	96.2	0.69	2
GLU	312	84.56	49.1	82.41	**61.2**	2.15	5.7	37.27	61.8	47.29	42.2
LEU	313	35.39	19.8	35.39	**25.1**	0	0	35.39	24.9	0	0
LEU	314	94.25	52.8	92.21	**65.3**	2.04	5.4	93.19	65.5	1.06	2.9
ASP	315	59.5	42.4	58.55	**57**	0.95	2.5	25.49	51.8	34.02	37.3
TYR	316	50.27	23.6	49.1	**27.7**	1.18	3.3	34.78	25.5	15.49	20.3
ILE	317	48.09	27.5	45.44	**32.9**	2.65	7.1	45.44	32.7	2.65	7.4
VAL	318	105.78	69.8	96.78	**84.7**	9	24.2	96.78	83.8	9	25
ASN	319	96.39	67	71.41	**67.2**	24.98	66.3	24.58	53.2	71.81	73.5
GLU	320	69.04	40.1	66.46	**49.3**	2.59	6.9	13.53	22.4	55.52	49.6
PRO	321	95.49	70.1	91.65	**76.4**	3.84	23.6	91.65	75.8	3.84	25.3
PRO	322	47	34.5	41.91	**35**	5.09	31.4	42.04	34.8	4.96	32.6
PRO	323	8.94	6.6	7.63	**6.4**	1.31	8.1	7.63	6.3	1.31	8.6
LYS	324	138.04	68.7	120.13	**73.6**	17.91	47.8	78.49	67.3	59.54	70.7
LEU	325	8.36	4.7	0	**0**	8.36	22.3	0.49	0.3	7.87	21.7
PRO	326	12.52	9.2	12.41	**10.4**	0.1	0.6	12.41	10.3	0.1	0.7
SER	327	58.19	50	56.87	**72.8**	1.32	3.4	38.52	79.3	19.67	28.9
GLY	328	68.08	85	34.1	**105.5**	33.98	71.1	37.04	98.6	31.04	73
VAL	329	52.33	34.6	31.77	**27.8**	20.55	55.3	33.16	28.7	19.16	53.3
PHE	330	18.18	9.1	14.25	**8.7**	3.93	11.1	15.08	9.1	3.1	9

SER	331	33.37	28.6	30.01	**38.4**	3.36	8.8	30.01	61.8	3.36	5
LEU	332	137.91	77.2	130.41	**92.4**	7.51	20	131.54	92.4	6.37	17.5
GLU	333	58.49	34	58.49	**43.4**	0	0	33.7	55.9	24.78	22.1
PHE	334	0	0	0	**0**	0	0	0	0	0	0
GLN	335	24.65	13.8	24.58	**17.4**	0.07	0.2	7	13.4	17.65	14
ASP	336	54.15	38.6	53.91	**52.5**	0.24	0.6	19.64	39.9	34.5	37.9
PHE	337	0	0	0	**0**	0	0	0	0	0	0
VAL	338	0.08	0.1	0.08	**0.1**	0	0	0.08	0.1	0	0
ASN	339	62.3	43.3	59.75	**56.2**	2.56	6.8	22.32	48.3	39.98	40.9
LYS	340	62.43	31.1	62.43	**38.2**	0	0	24.77	21.2	37.66	44.7
CYS	341	0	0	0	**0**	0	0	0	0	0	0
LEU	342	1.05	0.6	0.83	**0.6**	0.22	0.6	1.05	0.7	0	0
ILE	343	49.11	28	49.11	**35.6**	0	0	49.11	35.3	0	0
LYS	344	74.86	37.3	53.38	**32.7**	21.47	57.2	50.96	43.7	23.89	28.4
ASN	345	72.14	50.1	71.37	**67.2**	0.77	2	29.45	63.7	42.69	43.7
PRO	346	69.45	51	67.62	**56.4**	1.84	11.3	68.67	56.8	0.78	5.2
ALA	347	90.39	83.7	61.52	**88.6**	28.87	74.9	61.87	86.7	28.52	78
GLU	348	120.5	70	106.08	**78.7**	14.41	38.4	42.76	70.9	77.74	69.4
ARG	349	22.57	9.5	17.66	**8.8**	4.92	13.1	0.05	0.1	22.52	14
ALA	350	0.49	0.5	0	**0**	0.49	1.3	0	0	0.49	1.3
ASP	351	35.87	25.6	35.43	**34.5**	0.45	1.2	17.07	34.7	18.81	20.6
LEU	352	16.3	9.1	14.59	**10.3**	1.71	4.5	15.29	10.7	1.02	2.8
LYS	353	158.23	78.8	154.42	**94.6**	3.81	10.2	115.07	98.7	43.15	51.2
GLN	354	74.77	41.9	73.26	**52**	1.51	4	32.16	61.6	42.62	33.7
LEU	355	0	0	0	**0**	0	0	0	0	0	0
MET	356	25.39	13.1	25.36	**16.2**	0.03	0.1	25.36	16.1	0.03	0.1
VAL	357	81.29	53.7	75.67	**66.2**	5.62	15.1	75.7	65.6	5.6	15.6
HIS	358	20.2	11	20.2	**13.7**	0	0	18.43	19	1.77	2.1
ALA	359	50.17	46.5	42.84	**61.7**	7.33	19	43.13	60.4	7.03	19.2
PHE	360	0	0	0	**0**	0	0	0	0	0	0
ILE	361	0.66	0.4	0.66	**0.5**	0	0	0.66	0.5	0	0
LYS	362	117.24	58.4	109.78	**67.2**	7.46	19.9	91.71	78.7	25.53	30.3
ARG	363	70.67	29.6	69.81	**34.7**	0.86	2.3	61.01	78.4	9.66	6
SER	364	0.03	0	0.03	**0**	0	0	0.03	0.1	0	0

ASP	365	67.02	47.7	47.97	**46.7**	19.06	50.5	14.15	28.7	52.87	58
ALA	366	83.57	77.4	64.96	**93.6**	18.61	48.3	65.84	92.2	17.73	48.5
GLU	367	39.5	22.9	35.55	**26.4**	3.94	10.5	27.26	45.2	12.24	10.9
GLU	368	178.48	103.6	147.1	**109.2**	31.37	83.6	69.97	116.1	108.5	96.9
VAL	369	33.24	21.9	9.29	**8.1**	23.95	64.4	10.75	9.3	22.49	62.5
ASP	370	102.81	73.2	90.65	**88.3**	12.16	32.2	47.94	97.4	54.87	60.2
PHE	371	13.67	6.9	12.25	**7.5**	1.42	4	12.25	7.4	1.42	4.2
ALA	372	30.08	27.9	27.65	**39.8**	2.42	6.3	28.55	40	1.52	4.2
GLY	373	32.46	40.5	26.09	**80.7**	6.37	13.3	30.99	82.5	1.47	3.4
TRP	374	54.56	21.9	52.19	**24.7**	2.38	6.2	39	20.6	15.56	26.1
LEU	375	0	0	0	**0**	0	0	0	0	0	0
CYS	376	23.4	17.4	21.18	**21.9**	2.22	5.9	22.54	23	0.86	2.4
SER	377	94.6	81.2	72.86	**93.3**	21.74	56.6	42.26	87.1	52.34	77
THR	378	29.74	21.4	20.25	**19.9**	9.49	25.3	13.93	18.4	15.81	24.9
ILE	379	37.47	21.4	19.47	**14.1**	18	48.4	19.93	14.3	17.54	48.7
GLY	380	57.68	72	32.38	**100.1**	25.31	53	33.17	88.3	24.51	57.6
LEU	381	60.46	33.8	47.32	**33.5**	13.14	35	47.77	33.6	12.69	34.9
ASN	382	183.48	127.5	123.43	**116.2**	60.05	159.3	73.68	159.4	109.79	112.4
AGS	401	32.17	-99.9	32.17	**-99.9**	0	-99.9	4.19	-99.9	27.98	-99.9
Absolute sums overall chain											
Total		16216.1	1	3832.8		2383.3		9889.7		6326.4	

MEK1+ATP-Mg (3EQD)

Residues	Number	All		Total side chain		Main chain		Non polar		All polar	
		ABS	REL	ABS	**REL**	ABS	REL	ABS	REL	ABS	REL
GLU	39	90.33	52.4	63.16	**46.9**	27.17	72.4	23.5	39	66.84	59.7
LEU	40	47.72	26.7	35.54	**25.2**	12.17	32.5	37.85	26.6	9.86	27.2
GLU	41	161.27	93.6	145.82	**108.2**	15.45	41.2	69.73	115.7	91.54	81.8
LEU	42	72.88	40.8	41.52	**29.4**	31.37	83.6	42.39	29.8	30.49	84
ASP	43	91.54	65.2	90.18	**87.8**	1.36	3.6	21.04	42.7	70.5	77.3
GLU	44	160.56	93.2	150.77	**111.9**	9.78	26.1	71.46	118.5	89.09	79.6
GLN	45	88.49	49.6	86.48	**61.3**	2.01	5.4	21.97	42.1	66.52	52.7
GLN	46	35.15	19.7	33.91	**24.1**	1.24	3.3	13.21	25.3	21.94	17.4

ARG	47	158.55	66.4	157.44	**78.2**	1.11	3	41.02	52.7	117.53	73
LYS	48	115.97	57.7	113.55	**69.5**	2.42	6.4	84.76	72.7	31.21	37
ARG	49	47.95	20.1	41.66	**20.7**	6.29	16.8	14.57	18.7	33.38	20.7
LEU	50	31.17	17.4	30.33	**21.5**	0.83	2.2	30.33	21.3	0.83	2.3
GLU	51	96.2	55.8	94.76	**70.3**	1.43	3.8	47.19	78.3	49	43.8
ALA	52	50.66	46.9	49.9	**71.9**	0.75	2	49.9	69.9	0.75	2.1
PHE	53	26.35	13.2	25.78	**15.7**	0.57	1.6	25.78	15.6	0.57	1.7
LEU	54	66.42	37.2	65.39	**46.3**	1.03	2.8	65.39	45.9	1.03	2.8
THR	55	64.84	46.6	64.37	**63.3**	0.47	1.3	46.09	60.9	18.75	29.5
GLN	56	80.22	44.9	79.77	**56.6**	0.45	1.2	19.74	37.8	60.48	47.9
LYS	57	31.92	15.9	24.35	**14.9**	7.58	20.2	19.32	16.6	12.61	15
GLN	58	143.92	80.6	117.57	**83.4**	26.35	70.2	52.27	100.1	91.65	72.6
LYS	59	158.53	78.9	135.47	**83**	23.06	61.5	89.34	76.6	69.19	82.1
VAL	60	16.68	11	2.46	**2.2**	14.22	38.3	2.58	2.2	14.1	39.2
GLY	61	57.4	71.7	37.22	**115.1**	20.18	42.2	41.94	111.7	15.46	36.3
GLU	62	135.8	78.8	119.09	**88.4**	16.71	44.6	60.83	100.9	74.98	67
LEU	63	19.01	10.6	5.12	**3.6**	13.89	37	5.12	3.6	13.89	38.2
LYS	64	108.63	54.1	108.63	**66.5**	0	0	82.73	71	25.9	30.7
ASP	65	11.21	8	4.65	**4.5**	6.55	17.4	4.65	9.4	6.55	7.2
ASP	66	90.4	64.4	82.53	**80.4**	7.88	20.9	42.43	86.2	47.98	52.6
ASP	67	37.12	26.4	37.12	**36.1**	0	0	1.22	2.5	35.9	39.4
PHE	68	20.73	10.4	1.17	**0.7**	19.55	55.3	1.52	0.9	19.21	56.1
GLU	69	90.25	52.4	90.25	**67**	0	0	36.15	60	54.1	48.3
LYS	70	124.24	61.9	95.96	**58.8**	28.28	75.4	64.8	55.6	59.44	70.6
ILE	71	76.63	43.8	51.66	**37.4**	24.97	67.2	52.49	37.7	24.13	67.1
SER	72	48	41.2	48	**61.4**	0	0	37.59	77.4	10.41	15.3
GLU	73	97.06	56.3	68.99	**51.2**	28.07	74.8	26.27	43.6	70.79	63.2
LEU	74	39.01	21.8	29.17	**20.7**	9.83	26.2	30.82	21.7	8.19	22.5
GLY	75	18.4	23	18.33	**56.7**	0.08	0.2	18.33	48.8	0.08	0.2
ALA	76	60.85	56.4	33.44	**48.2**	27.41	71.1	35	49	25.85	70.7
GLY	77	34.08	42.5	11.53	**35.6**	22.56	47.2	14.98	39.9	19.1	44.9
ASN	78	76.69	53.3	67.07	**63.1**	9.61	25.5	6.13	13.3	70.55	72.2
GLY	79	16.03	20	1.53	**4.7**	14.49	30.3	5.19	13.8	10.84	25.5
GLY	80	9.56	11.9	5.47	**16.9**	4.08	8.6	6.01	16	3.55	8.3

VAL	81	46.92	31	46.92	**41.1**	0	0	46.92	40.6	0	0
VAL	82	0	0	0	**0**	0	0	0	0	0	0
PHE	83	30.79	15.4	30.79	**18.8**	0	0	30.79	18.6	0	0
LYS	84	55.17	27.5	55.17	**33.8**	0	0	11.69	10	43.47	51.6
VAL	85	3.57	2.4	3.57	**3.1**	0	0	3.57	3.1	0	0
SER	86	19.11	16.4	19.11	**24.5**	0	0	6.11	12.6	13	19.1
HIS	87	2.23	1.2	2.23	**1.5**	0	0	1.94	2	0.29	0.3
LYS	88	110.23	54.9	90.95	**55.7**	19.28	51.4	67.15	57.6	43.08	51.1
PRO	89	88.74	65.2	56.99	**47.5**	31.75	195.6	59.3	49	29.44	193.8
SER	90	43.42	37.3	18	**23**	25.42	66.2	20.08	41.4	23.34	34.3
GLY	91	54.1	67.5	28.23	**87.3**	25.87	54.2	30.95	82.4	23.16	54.4
LEU	92	33	18.5	33	**23.4**	0	0	33	23.2	0	0
VAL	93	64.09	42.3	49.31	**43.1**	14.78	39.8	49.31	42.7	14.78	41.1
MET	94	0	0	0	**0**	0	0	0	0	0	0
ALA	95	0	0	0	**0**	0	0	0	0	0	0
ARG	96	15.63	6.5	15.63	**7.8**	0	0	1.11	1.4	14.52	9
LYS	97	20.42	10.2	20.42	**12.5**	0	0	7.75	6.6	12.67	15
LEU	98	26.89	15.1	24.26	**17.2**	2.63	7	24.28	17.1	2.62	7.2
ILE	99	2.8	1.6	2.8	**2**	0	0	2.8	2	0	0
HIS	100	100.83	55.1	96.88	**65.9**	3.95	11	74.24	76.4	26.59	31
LEU	101	7.39	4.1	6.96	**4.9**	0.43	1.1	6.96	4.9	0.43	1.2
GLU	102	133.32	77.4	107.04	**79.4**	26.29	70.1	60.87	101	72.45	64.7
ILE	103	33.96	19.4	21.47	**15.6**	12.49	33.6	26.08	18.7	7.88	21.9
LYS	104	125.28	62.4	124.12	**76**	1.16	3.1	75.04	64.4	50.23	59.6
PRO	105	74.99	55.1	74.96	**62.5**	0.03	0.2	74.99	62	0	0
ALA	106	74.37	68.9	66.32	**95.6**	8.04	20.9	68.47	95.9	5.9	16.1
ILE	107	36.47	20.8	36.26	**26.3**	0.2	0.5	36.26	26.1	0.2	0.6
ARG	108	48.02	20.1	48.02	**23.9**	0	0	7.76	10	40.26	25
ASN	109	76.56	53.2	73.84	**69.5**	2.72	7.2	20.33	44	56.23	57.5
GLN	110	58.82	33	58.82	**41.7**	0	0	23.4	44.8	35.42	28.1
ILE	111	0.59	0.3	0.59	**0.4**	0	0	0.59	0.4	0	0
ILE	112	32.92	18.8	32.92	**23.9**	0	0	32.92	23.7	0	0
ARG	113	150.01	62.8	138.6	**68.9**	11.41	30.4	74.03	95.2	75.98	47.2
GLU	114	46.81	27.2	45.1	**33.5**	1.7	4.5	18.74	31.1	28.07	25.1

LEU	115	5.84	3.3	5.84	**4.1**	0	0	5.84	4.1	0	0
GLN	116	66.58	37.3	63.47	**45**	3.1	8.3	42.68	81.7	23.89	18.9
VAL	117	41.61	27.5	39.81	**34.8**	1.8	4.9	39.81	34.5	1.8	5
LEU	118	24.23	13.6	23.06	**16.3**	1.18	3.1	23.06	16.2	1.18	3.2
HIS	119	62.54	34.2	53.1	**36.1**	9.44	26.4	40.94	42.1	21.6	25.2
GLU	120	105.17	61.1	97.29	**72.2**	7.88	21	44.21	73.3	60.96	54.4
CYS	121	5.23	3.9	0.71	**0.7**	4.52	12.1	0.71	0.7	4.52	12.4
ASN	122	48.68	33.8	38.7	**36.4**	9.99	26.5	10.54	22.8	38.15	39
SER	123	7.77	6.7	3.02	**3.9**	4.75	12.4	2.59	5.3	5.18	7.6
PRO	124	17.32	12.7	14.82	**12.4**	2.49	15.4	14.82	12.3	2.49	16.4
TYR	125	31.41	14.8	31.41	**17.7**	0	0	21.03	15.4	10.38	13.6
ILE	126	5.7	3.3	2.27	**1.6**	3.43	9.2	2.27	1.6	3.43	9.5
VAL	127	10.54	7	2.54	**2.2**	7.99	21.5	2.56	2.2	7.98	22.2
GLY	128	9.59	12	9.26	**28.7**	0.33	0.7	9.3	24.8	0.29	0.7
PHE	129	11.17	5.6	8.11	**4.9**	3.06	8.6	8.11	4.9	3.06	8.9
TYR	130	17.88	8.4	10.04	**5.7**	7.84	22.2	0.2	0.1	17.67	23.2
GLY	131	5.77	7.2	5.77	**17.8**	0	0	5.77	15.4	0	0
ALA	132	12.39	11.5	0	**0**	12.39	32.1	0	0	12.39	33.9
PHE	133	22.46	11.3	22.46	**13.7**	0	0	22.46	13.6	0	0
TYR	134	54.57	25.6	31.74	**17.9**	22.83	64.5	22.51	16.5	32.06	42
SER	135	19.6	16.8	19.48	**24.9**	0.12	0.3	14.08	29	5.52	8.1
ASP	136	113.03	80.5	91.28	**88.9**	21.75	57.7	37.45	76.1	75.58	82.9
GLY	137	8.69	10.8	5.73	**17.7**	2.95	6.2	5.73	15.3	2.95	6.9
GLU	138	36.93	21.4	35.7	**26.5**	1.23	3.3	8.01	13.3	28.92	25.8
ILE	139	2.16	1.2	2.16	**1.6**	0	0	2.16	1.5	0	0
SER	140	0.13	0.1	0.13	**0.2**	0	0	0	0	0.13	0.2
ILE	141	19.21	11	19.21	**13.9**	0	0	19.21	13.8	0	0
CYS	142	0	0	0	**0**	0	0	0	0	0	0
MET	143	13.13	6.8	13.13	**8.4**	0	0	13.13	8.3	0	0
GLU	144	29.36	17	29.36	**21.8**	0	0	13.14	21.8	16.22	14.5
HIS	145	39.71	21.7	23.38	**15.9**	16.33	45.6	10.4	10.7	29.31	34.2
MET	146	1.01	0.5	0	**0**	1.01	2.7	0	0	1.01	2.8
ASP	147	55.51	39.5	39.54	**38.5**	15.98	42.4	25.17	51.1	30.35	33.3
GLY	148	5.76	7.2	0	**0**	5.76	12.1	2.19	5.8	3.57	8.4

GLY	149	11.32	14.1	10.55	**32.6**	0.78	1.6	10.55	28.1	0.78	1.8
SER	150	0	0	0	**0**	0	0	0	0	0	0
LEU	151	0.92	0.5	0.92	**0.7**	0	0	0.92	0.6	0	0
ASP	152	40.82	29.1	39.16	**38.1**	1.67	4.4	18.57	37.7	22.25	24.4
GLN	153	68.93	38.6	62.07	**44**	6.86	18.3	35.84	68.6	33.09	26.2
VAL	154	0.49	0.3	0.49	**0.4**	0	0	0.49	0.4	0	0
LEU	155	14.69	8.2	14.09	**10**	0.6	1.6	14.09	9.9	0.6	1.7
LYS	156	171.7	85.5	144.93	**88.8**	26.77	71.4	117.51	100.8	54.19	64.3
LYS	157	127.52	63.5	103.78	**63.6**	23.74	63.3	74.08	63.6	53.44	63.4
ALA	158	33.89	31.4	7.11	**10.2**	26.78	69.5	8.37	11.7	25.52	69.8
GLY	159	44.9	56.1	36.6	**113.2**	8.3	17.4	39.55	105.3	5.35	12.6
ARG	160	101.33	42.4	91.71	**45.6**	9.61	25.6	22.65	29.1	78.68	48.9
ILE	161	2.26	1.3	0.51	**0.4**	1.75	4.7	0.94	0.7	1.32	3.7
PRO	162	58.18	42.7	57.62	**48.1**	0.56	3.5	57.62	47.6	0.56	3.7
GLU	163	27.88	16.2	22.38	**16.6**	5.5	14.7	17.39	28.8	10.49	9.4
GLN	164	91.69	51.4	90.55	**64.2**	1.15	3.1	24.39	46.7	67.3	53.3
ILE	165	6.48	3.7	6.48	**4.7**	0	0	6.48	4.7	0	0
LEU	166	0	0	0	**0**	0	0	0	0	0	0
GLY	167	0	0	0	**0**	0	0	0	0	0	0
LYS	168	37.02	18.4	34.6	**21.2**	2.42	6.4	24.55	21.1	12.46	14.8
VAL	169	0.58	0.4	0.58	**0.5**	0	0	0.58	0.5	0	0
SER	170	0	0	0	**0**	0	0	0	0	0	0
ILE	171	12.55	7.2	12.55	**9.1**	0	0	12.55	9	0	0
ALA	172	4.59	4.3	3.87	**5.6**	0.72	1.9	4.3	6	0.29	0.8
VAL	173	0.68	0.5	0.68	**0.6**	0	0	0.68	0.6	0	0
ILE	174	0	0	0	**0**	0	0	0	0	0	0
LYS	175	29.59	14.7	29.59	**18.1**	0	0	1.52	1.3	28.07	33.3
GLY	176	0	0	0	**0**	0	0	0	0	0	0
LEU	177	1.48	0.8	0.68	**0.5**	0.8	2.1	0.68	0.5	0.8	2.2
THR	178	4.42	3.2	4.42	**4.4**	0	0	4.17	5.5	0.25	0.4
TYR	179	11.38	5.3	11.38	**6.4**	0	0	6.01	4.4	5.37	7
LEU	180	0.11	0.1	0.11	**0.1**	0	0	0.11	0.1	0	0
ARG	181	93.12	39	87.96	**43.7**	5.16	13.8	30.15	38.8	62.97	39.1
GLU	182	95.67	55.5	73.17	**54.3**	22.5	60	24.89	41.3	70.78	63.2

LYS	183	125.87	62.7	101.18	**62**	24.7	65.8	58.22	49.9	67.65	80.3
HIS	184	38.71	21.2	24.22	**16.5**	14.49	40.5	18.68	19.2	20.03	23.4
LYS	185	188.34	93.8	159.44	**97.6**	28.89	77	113	96.9	75.33	89.4
ILE	186	20.91	11.9	19.91	**14.4**	1	2.7	20.85	15	0.06	0.2
MET	187	18.58	9.6	17.38	**11.1**	1.2	3.2	17.38	11	1.2	3.3
HIS	188	0.05	0	0.03	**0**	0.02	0.1	0	0	0.05	0.1
ARG	189	91.44	38.3	83.57	**41.5**	7.87	21	28.47	36.6	62.96	39.1
ASP	190	46.18	32.9	46.18	**45**	0	0	22.56	45.8	23.62	25.9
VAL	191	1.98	1.3	0.38	**0.3**	1.6	4.3	0.38	0.3	1.6	4.4
LYS	192	7.27	3.6	7.27	**4.5**	0	0	4.47	3.8	2.8	3.3
PRO	193	3.81	2.8	3.81	**3.2**	0	0	3.81	3.1	0	0
SER	194	11.52	9.9	10.89	**13.9**	0.63	1.6	9.74	20.1	1.78	2.6
ASN	195	0.31	0.2	0.31	**0.3**	0	0	0.31	0.7	0	0
ILE	196	1.44	0.8	1.44	**1**	0	0	1.44	1	0	0
LEU	197	0	0	0	**0**	0	0	0	0	0	0
VAL	198	0	0	0	**0**	0	0	0	0	0	0
ASN	199	9.01	6.3	9.01	**8.5**	0	0	3.43	7.4	5.59	5.7
SER	200	44.67	38.3	31.48	**40.3**	13.2	34.4	16.82	34.7	27.85	41
ARG	201	140.26	58.7	125.64	**62.4**	14.62	39	51.68	66.4	88.58	55
GLY	202	23.77	29.7	4.97	**15.4**	18.8	39.3	6.72	17.9	17.05	40.1
GLU	203	26.19	15.2	26.19	**19.4**	0	0	14.98	24.9	11.21	10
ILE	204	12.9	7.4	8.79	**6.4**	4.11	11.1	8.79	6.3	4.11	11.4
LYS	205	15.83	7.9	15.73	**9.6**	0.1	0.3	0.6	0.5	15.23	18.1
LEU	206	1.97	1.1	1.91	**1.4**	0.06	0.2	1.91	1.3	0.06	0.2
CYS	207	4.24	3.2	4.24	**4.4**	0	0	4.24	4.3	0	0
ASP	208	47.3	33.7	33.47	**32.6**	13.83	36.7	23.43	47.6	23.87	26.2
PHE	209	14.86	7.5	7.07	**4.3**	7.79	22	7.95	4.8	6.91	20.2
GLY	210	26.78	33.4	17.21	**53.2**	9.57	20	20.2	53.8	6.58	15.5
VAL	211	37.1	24.5	11.65	**10.2**	25.45	68.5	15.12	13.1	21.98	61.1
SER	212	3.36	2.9	0.62	**0.8**	2.74	7.1	0.62	1.3	2.74	4
GLY	213	43.19	53.9	33.34	**103.1**	9.86	20.6	34.74	92.5	8.45	19.9
GLN	214	37.08	20.8	34.37	**24.4**	2.71	7.2	1.48	2.8	35.6	28.2
LEU	215	13.61	7.6	13.61	**9.6**	0	0	13.61	9.6	0	0
ILE	216	86.77	49.5	85.66	**62.1**	1.11	3	85.66	61.6	1.11	3.1

ASP	217	90.93	64.8	83.05	**80.9**	7.89	20.9	27.43	55.7	63.5	69.7
SER	218	12.08	10.4	11.86	**15.2**	0.22	0.6	7.38	15.2	4.7	6.9
MET	219	37.78	19.5	37.55	**24**	0.23	0.6	37.78	23.9	0	0
ALA	220	64.55	59.8	43.41	**62.5**	21.14	54.9	44.31	62.1	20.24	55.3
ASN	221	108.78	75.6	84.17	**79.2**	24.62	65.3	32.93	71.2	75.86	77.6
SER	222	95.19	81.7	77.79	**99.6**	17.4	45.3	46.58	95.9	48.61	71.5
PHE	223	34.56	17.3	27.14	**16.5**	7.42	21	27.93	16.9	6.63	19.4
VAL	224	99.69	65.8	94.55	**82.7**	5.14	13.8	94.55	81.9	5.14	14.3
GLY	225	33.05	41.3	12.2	**37.7**	20.85	43.7	14.67	39.1	18.38	43.2
THR	226	54.85	39.4	47.29	**46.5**	7.56	20.1	37.64	49.7	17.22	27.1
ARG	227	102.43	42.9	101.54	**50.5**	0.89	2.4	33.01	42.4	69.42	43.1
SER	228	11.61	10	11.61	**14.9**	0	0	5.74	11.8	5.87	8.6
TYR	229	29.87	14	28.76	**16.2**	1.11	3.1	15.63	11.4	14.25	18.7
MET	230	31.74	16.3	31.63	**20.2**	0.1	0.3	31.63	20	0.1	0.3
SER	231	0.41	0.3	0	**0**	0.41	1.1	0	0	0.41	0.6
PRO	232	6.37	4.7	6.32	**5.3**	0.05	0.3	6.37	5.3	0	0
GLU	233	6	3.5	5.56	**4.1**	0.44	1.2	5.56	9.2	0.44	0.4
ARG	234	44.7	18.7	33.8	**16.8**	10.9	29	3.48	4.5	41.21	25.6
LEU	235	33.21	18.6	4.68	**3.3**	28.53	76.1	5.84	4.1	27.37	75.4
GLN	236	107.48	60.2	85.85	**60.9**	21.63	57.7	31.15	59.6	76.34	60.5
GLY	237	77.26	96.5	39.04	**120.8**	38.21	80	41.2	109.7	36.05	84.7
THR	238	78.15	56.1	67.59	**66.5**	10.56	28.1	49	64.7	29.15	45.9
HIS	239	83.48	45.6	78.6	**53.4**	4.88	13.6	53.83	55.4	29.64	34.6
TYR	240	85.81	40.3	59.3	**33.4**	26.51	74.9	51.75	37.9	34.07	44.7
SER	241	41.39	35.5	41.39	**53**	0	0	38.7	79.7	2.7	4
VAL	242	17.95	11.9	15.93	**13.9**	2.01	5.4	15.93	13.8	2.01	5.6
GLN	243	18.65	10.4	18.65	**13.2**	0	0	14.49	27.7	4.16	3.3
SER	244	2.5	2.1	0.87	**1.1**	1.63	4.3	1.95	4	0.55	0.8
ASP	245	1.44	1	1.44	**1.4**	0	0	0	0	1.44	1.6
ILE	246	0	0	0	**0**	0	0	0	0	0	0
TRP	247	1.09	0.4	1.09	**0.5**	0	0	1.09	0.6	0	0
SER	248	8.62	7.4	8.62	**11**	0	0	0	0	8.62	12.7
MET	249	1.57	0.8	1.57	**1**	0	0	1.57	1	0	0
GLY	250	0	0	0	**0**	0	0	0	0	0	0

LEU	251	0	0	0	**0**	0	0	0	0	0	0
SER	252	0.7	0.6	0.7	**0.9**	0	0	0.7	1.4	0	0
LEU	253	1.55	0.9	1.55	**1.1**	0	0	1.55	1.1	0	0
VAL	254	1.05	0.7	1.05	**0.9**	0	0	1.05	0.9	0	0
GLU	255	22.21	12.9	22.21	**16.5**	0	0	0	0	22.21	19.8
MET	256	0.39	0.2	0.39	**0.3**	0	0	0.39	0.2	0	0
ALA	257	0.52	0.5	0	**0**	0.52	1.3	0	0	0.52	1.4
VAL	258	7.4	4.9	3.25	**2.8**	4.15	11.2	3.25	2.8	4.15	11.5
GLY	259	30.49	38.1	2.49	**7.7**	27.99	58.6	3.21	8.6	27.27	64.1
ARG	260	62.01	26	62.01	**30.8**	0	0	32.3	41.5	29.71	18.5
TYR	261	45.7	21.5	37.69	**21.2**	8.02	22.7	37.69	27.6	8.02	10.5
PRO	262	2.86	2.1	2.61	**2.2**	0.25	1.5	2.61	2.2	0.25	1.6
ILE	263	9.7	5.5	1.84	**1.3**	7.86	21.1	1.84	1.3	7.86	21.8
PRO	264	49.1	36.1	49.1	**41**	0	0	49.1	40.6	0	0
PRO	265	65.02	47.8	55.79	**46.5**	9.23	56.9	55.79	46.1	9.23	60.8
PRO	266	28.53	21	15.56	**13**	12.97	79.9	15.56	12.9	12.97	85.4
ASP	267	85.72	61.1	85.72	**83.5**	0	0	24.59	49.9	61.13	67.1
ALA	268	86.73	80.3	68.9	**99.3**	17.82	46.2	69.87	97.9	16.86	46.1
LYS	269	154.65	77	154.6	**94.7**	0.05	0.1	106.15	91.1	48.49	57.6
GLU	270	53.97	31.3	53.97	**40.1**	0	0	24.96	41.4	29.01	25.9
LEU	271	48.43	27.1	47.52	**33.7**	0.91	2.4	47.75	33.6	0.68	1.9
GLU	272	100.41	58.3	96.15	**71.4**	4.26	11.3	29.15	48.3	71.26	63.7
LEU	273	145.87	81.7	117.68	**83.4**	28.19	75.2	118.41	83.2	27.46	75.6
MET	274	42.87	22.1	36.23	**23.1**	6.64	17.7	36.49	23.1	6.38	17.6
PHE	275	43.28	21.7	13.88	**8.5**	29.39	83.1	18.7	11.3	24.57	71.8
GLY	276	65.52	81.8	31.2	**96.5**	34.32	71.8	34.2	91.1	31.32	73.6
CYS	277	60.31	44.9	60.31	**62.3**	0	0	60.31	61.6	0	0
PRO	307	186.67	137.1	142.04	**118.5**	44.63	275	144.32	119.3	42.35	278.8
MET	308	76.93	39.6	45.46	**29**	31.47	83.9	45.57	28.9	31.36	86.3
ALA	309	31.54	29.2	31.42	**45.3**	0.12	0.3	31.42	44	0.12	0.3
ILE	310	90.74	51.8	85.38	**61.9**	5.36	14.4	85.38	61.4	5.36	14.9
PHE	311	143.43	71.9	137.08	**83.5**	6.35	18	138.15	83.6	5.28	15.4
GLU	312	75.62	43.9	75.11	**55.7**	0.51	1.4	31.51	52.3	44.1	39.4
LEU	313	19.56	10.9	19.56	**13.9**	0	0	19.56	13.7	0	0

LEU	314	32.87	18.4	31.63	**22.4**	1.23	3.3	32.87	23.1	0	0
ASP	315	82.84	59	82.27	**80.1**	0.57	1.5	27.06	55	55.78	61.2
TYR	316	49.4	23.2	49.09	**27.7**	0.31	0.9	35.06	25.7	14.33	18.8
ILE	317	4.39	2.5	3.25	**2.4**	1.14	3.1	3.25	2.3	1.14	3.2
VAL	318	27.7	18.3	17.23	**15.1**	10.47	28.2	17.23	14.9	10.47	29.1
ASN	319	91.95	63.9	67.03	**63.1**	24.92	66.1	25.89	56	66.05	67.6
GLU	320	68.17	39.6	65.79	**48.8**	2.38	6.3	14.53	24.1	53.64	47.9
PRO	321	94.9	69.7	89.46	**74.6**	5.44	33.5	89.49	74	5.41	35.6
PRO	322	44.7	32.8	34.46	**28.7**	10.24	63.1	34.46	28.5	10.24	67.4
PRO	323	10.17	7.5	8.65	**7.2**	1.52	9.3	8.65	7.2	1.52	10
LYS	324	143.97	71.7	128.42	**78.6**	15.56	41.5	85.41	73.3	58.56	69.5
LEU	325	6.29	3.5	0	**0**	6.29	16.8	0.15	0.1	6.14	16.9
PRO	326	22.34	16.4	22.34	**18.6**	0.01	0	22.34	18.5	0.01	0
SER	327	89.97	77.2	68	**87.1**	21.97	57.2	34.01	70	55.97	82.4
GLY	328	79.73	99.5	40.53	**125.4**	39.2	82.1	43.34	115.4	36.39	85.5
VAL	329	68.73	45.4	40.39	**35.3**	28.34	76.3	41.39	35.8	27.34	76
PHE	330	22.34	11.2	13.17	**8**	9.17	25.9	16.51	10	5.84	17
SER	331	38.3	32.9	35.18	**45**	3.12	8.1	35.18	72.5	3.12	4.6
LEU	332	141.45	79.2	132.26	**93.7**	9.19	24.5	133.54	93.8	7.91	21.8
GLU	333	64.59	37.5	64.45	**47.8**	0.13	0.4	29.47	48.9	35.11	31.4
PHE	334	0	0	0	**0**	0	0	0	0	0	0
GLN	335	23.31	13.1	23.31	**16.5**	0	0	7.46	14.3	15.85	12.6
ASP	336	54.64	38.9	54.19	**52.8**	0.46	1.2	23.1	46.9	31.54	34.6
PHE	337	0	0	0	**0**	0	0	0	0	0	0
VAL	338	0.05	0	0.05	**0**	0	0	0.05	0	0	0
ASN	339	67.83	47.1	65.4	**61.6**	2.42	6.4	21.84	47.2	45.99	47.1
LYS	340	80.39	40	80.39	**49.2**	0	0	50.14	43	30.25	35.9
CYS	341	0	0	0	**0**	0	0	0	0	0	0
LEU	342	1.25	0.7	1.09	**0.8**	0.16	0.4	1.25	0.9	0	0
ILE	343	55.85	31.9	55.85	**40.5**	0	0	55.85	40.1	0	0
LYS	344	55.45	27.6	48.5	**29.7**	6.95	18.5	46.21	39.6	9.24	11
ASN	345	64.28	44.7	62.36	**58.7**	1.92	5.1	23.65	51.2	40.63	41.6
PRO	346	22.17	16.3	19.91	**16.6**	2.27	14	20.18	16.7	1.99	13.1
ALA	347	92.41	85.6	63.46	**91.4**	28.95	75.1	65	91.1	27.41	74.9

GLU	348	119.55	69.4	101.34	**75.2**	18.22	48.6	46.59	77.3	72.96	65.2
ARG	349	4.39	1.8	0	**0**	4.39	11.7	0	0	4.39	2.7
ALA	350	5.96	5.5	5.4	**7.8**	0.56	1.5	5.4	7.6	0.56	1.5
ASP	351	54.1	38.5	54.09	**52.7**	0.02	0	18.47	37.5	35.64	39.1
LEU	352	3.54	2	2.74	**1.9**	0.8	2.1	2.82	2	0.72	2
LYS	353	156.29	77.8	153.6	**94.1**	2.69	7.2	105.24	90.3	51.05	60.6
GLN	354	102.68	57.5	97.93	**69.5**	4.76	12.7	34.1	65.3	68.58	54.3
LEU	355	0.86	0.5	0.86	**0.6**	0	0	0.86	0.6	0	0
MET	356	24.38	12.6	18.94	**12.1**	5.45	14.5	18.94	12	5.45	15
VAL	357	92.15	60.9	86.81	**76**	5.35	14.4	87.07	75.4	5.08	14.1
HIS	358	18.94	10.4	18.94	**12.9**	0	0	17.48	18	1.47	1.7
ALA	359	43.83	40.6	37.26	**53.7**	6.58	17.1	37.26	52.2	6.58	18
PHE	360	0	0	0	**0**	0	0	0	0	0	0
ILE	361	1.52	0.9	1.52	**1.1**	0	0	1.52	1.1	0	0
LYS	362	117.17	58.3	110.84	**67.9**	6.33	16.9	88.84	76.2	28.33	33.6
ARG	363	69.99	29.3	68.88	**34.2**	1.11	3	58.44	75.1	11.55	7.2
SER	364	0.04	0	0.04	**0.1**	0	0	0	0	0.04	0.1
ASP	365	76.79	54.7	59.73	**58.2**	17.05	45.2	14.77	30	62.01	68
ALA	366	85.02	78.8	63.69	**91.8**	21.33	55.3	65.35	91.6	19.67	53.8
GLU	367	33.69	19.6	31.54	**23.4**	2.16	5.8	25	41.5	8.7	7.8
GLU	368	172.81	100.3	137.17	**101.8**	35.64	95	73.05	121.2	99.76	89.1
VAL	369	34.29	22.6	13.32	**11.7**	20.97	56.4	15.22	13.2	19.06	53
ASP	370	100.3	71.4	90.77	**88.4**	9.54	25.3	48.8	99.1	51.51	56.5
PHE	371	10.66	5.3	8.03	**4.9**	2.63	7.4	8.03	4.9	2.63	7.7
ALA	372	23.92	22.2	21.93	**31.6**	1.99	5.2	22.41	31.4	1.52	4.1
GLY	373	30.26	37.8	22.02	**68.1**	8.24	17.3	26.47	70.5	3.8	8.9
TRP	374	49.75	20	47.63	**22.5**	2.12	5.6	37.62	19.8	12.13	20.3
LEU	375	0	0	0	**0**	0	0	0	0	0	0
CYS	376	33.48	24.9	31.91	**33**	1.57	4.2	32.79	33.5	0.69	1.9
SER	377	92.16	79.1	68.02	**87.1**	24.15	62.9	57.08	117.6	35.08	51.6
THR	378	32.41	23.3	21.4	**21**	11.01	29.3	13.53	17.9	18.88	29.7
ILE	379	27.95	16	19.02	**13.8**	8.93	24	19.24	13.8	8.71	24.2
GLY	380	62.97	78.6	31.48	**97.4**	31.49	65.9	34.08	90.8	28.88	67.9
LEU	381	62.19	34.8	43.49	**30.8**	18.7	49.9	45.31	31.8	16.89	46.5

ASN	382	90.2	62.7	80.34	**75.6**	9.85	26.1	80.34	173.8	9.85	10.1
AGS	2	46.88	-99.9	46.88	**-99.9**	0	-99.9	15.04	-99.9	31.85	-99.9
MG	3	1.04	-99.9	1.04	**-99.9**	0	-99.9	1.04	-99.9	0	-99.9
CA	5	52.76	-99.9	52.76	**-99.9**	0	-99.9	0	-99.9	52.76	-99.9
NA	6	48.55	-99.9	48.55	**-99.9**	0	-99.9	0	-99.9	48.55	-99.9
Absolute sums overall chain											
Total		14724.6		2517.7		2206.9		8355.5		6369.1	

MEK1+ADP-Mg (3EQI)

Residues	Number	All		Total side chain		Main chain		Non polar		All polar	
		ABS	REL	ABS	**REL**	ABS	REL	ABS	REL	ABS	REL
GLU	39	99.63	57.8	70.96	**52.7**	28.66	76.4	29.37	48.7	70.26	62.8
LEU	40	44.05	24.7	32.5	**23**	11.55	30.8	35	24.6	9.05	24.9
GLU	41	164.35	95.4	147.62	**109.6**	16.74	44.6	71.6	118.8	92.75	82.8
LEU	42	76.07	42.6	44.25	**31.4**	31.81	84.8	45.63	32.1	30.44	83.8
ASP	43	92.46	65.9	89.48	**87.1**	2.99	7.9	19.96	40.5	72.5	79.5
GLU	44	158.55	92	149.93	**111.3**	8.62	23	70.67	117.2	87.88	78.5
GLN	45	87.9	49.2	86.19	**61.1**	1.71	4.6	18.29	35	69.61	55.1
GLN	46	38.05	21.3	37.09	**26.3**	0.96	2.6	17.97	34.4	20.09	15.9
ARG	47	128.65	53.9	128.65	**63.9**	0	0	51.91	66.7	76.74	47.7
LYS	48	117.83	58.7	115.25	**70.6**	2.59	6.9	84.97	72.9	32.87	39
ARG	49	54.42	22.8	47.66	**23.7**	6.75	18	16.04	20.6	38.38	23.8
LEU	50	26.68	14.9	26.08	**18.5**	0.6	1.6	26.08	18.3	0.6	1.7
GLU	51	90.72	52.7	89.73	**66.6**	0.99	2.6	30.22	50.1	60.5	54
ALA	52	49.03	45.4	48.33	**69.6**	0.7	1.8	48.33	67.7	0.7	1.9
PHE	53	24.88	12.5	24.5	**14.9**	0.39	1.1	24.5	14.8	0.39	1.1
LEU	54	59.65	33.4	58.94	**41.8**	0.7	1.9	59.65	41.9	0	0
THR	55	77.68	55.8	76.5	**75.2**	1.18	3.1	56.42	74.5	21.27	33.5
GLN	56	74.75	41.9	74.14	**52.6**	0.61	1.6	25.27	48.4	49.48	39.2
LYS	57	24.27	12.1	18.98	**11.6**	5.29	14.1	18.49	15.9	5.78	6.9
GLN	58	138.89	77.8	113.32	**80.4**	25.58	68.2	58.72	112.4	80.18	63.5
LYS	59	159.59	79.5	134.66	**82.5**	24.93	66.5	86.9	74.5	72.69	86.3
VAL	60	17.44	11.5	3.42	**3**	14.01	37.7	3.69	3.2	13.75	38.2

GLY	61	56.2	70.2	38.87	**120.2**	17.33	36.3	43.54	116	12.66	29.8
GLU	62	134.89	78.3	113.93	**84.6**	20.96	55.9	53.99	89.5	80.9	72.3
LEU	63	13.17	7.4	6.01	**4.3**	7.16	19.1	6.01	4.2	7.16	19.7
LYS	64	106.48	53	106.48	**65.2**	0	0	77.15	66.2	29.33	34.8
ASP	65	11.79	8.4	5.16	**5**	6.63	17.6	4.91	10	6.88	7.5
ASP	66	93.43	66.6	83.49	**81.3**	9.94	26.4	40.86	83	52.57	57.7
ASP	67	39.93	28.4	39.93	**38.9**	0	0	2	4.1	37.93	41.6
PHE	68	19.18	9.6	0.49	**0.3**	18.69	52.8	1.34	0.8	17.83	52.1
GLU	69	92.43	53.7	92.38	**68.6**	0.05	0.1	32.8	54.4	59.62	53.3
LYS	70	121.96	60.7	89.95	**55.1**	32.01	85.3	64.96	55.7	57	67.7
ILE	71	85.12	48.6	60.1	**43.6**	25.02	67.3	60.73	43.6	24.39	67.8
SER	72	46.83	40.2	46.83	**60**	0	0	36.9	76	9.92	14.6
GLU	73	106.01	61.5	79.48	**59**	26.53	70.7	27.24	45.2	78.76	70.4
LEU	74	37.65	21.1	25.66	**18.2**	11.98	32	27.61	19.4	10.03	27.6
GLY	75	18.8	23.5	17.94	**55.5**	0.87	1.8	18.7	49.8	0.1	0.2
ALA	76	62.39	57.8	43.39	**62.5**	19	49.3	46.53	65.2	15.87	43.4
GLY	77	34.28	42.8	18.16	**56.2**	16.12	33.7	22.1	58.9	12.18	28.6
ASN	78	126.09	87.6	109.81	**103.4**	16.28	43.2	28.84	62.4	97.25	99.5
GLY	79	20.28	25.3	14.76	**45.6**	5.52	11.6	17.72	47.2	2.56	6
GLY	80	9.73	12.1	8.58	**26.5**	1.14	2.4	8.58	22.9	1.14	2.7
VAL	81	48.99	32.3	48.94	**42.8**	0.04	0.1	48.94	42.4	0.04	0.1
VAL	82	0	0	0	**0**	0	0	0	0	0	0
PHE	83	30.81	15.4	30.81	**18.8**	0	0	30.81	18.6	0	0
LYS	84	56.49	28.1	56.49	**34.6**	0	0	12.71	10.9	43.78	52
VAL	85	3.69	2.4	3.69	**3.2**	0	0	3.69	3.2	0	0
SER	86	22.7	19.5	22.7	**29.1**	0	0	9.57	19.7	13.13	19.3
HIS	87	2.25	1.2	2.25	**1.5**	0	0	1.64	1.7	0.61	0.7
LYS	88	103.77	51.7	83.49	**51.1**	20.28	54.1	54.77	47	49	58.2
PRO	89	87.82	64.5	57.32	**47.8**	30.5	187.9	59.33	49.1	28.48	187.5
SER	90	44.52	38.2	16.82	**21.5**	27.7	72.1	17.8	36.7	26.72	39.3
GLY	91	52.42	65.4	27.23	**84.2**	25.19	52.7	29.86	79.5	22.57	53
LEU	92	31.33	17.5	31.33	**22.2**	0	0	31.33	22	0	0
VAL	93	61.86	40.8	48.8	**42.7**	13.06	35.2	48.8	42.3	13.06	36.3
MET	94	0	0	0	**0**	0	0	0	0	0	0

ALA	95	0	0	0	**0**	0	0	0	0	0	0
ARG	96	16.68	7	16.68	**8.3**	0	0	0.9	1.2	15.78	9.8
LYS	97	23.03	11.5	23.03	**14.1**	0	0	7.51	6.4	15.51	18.4
LEU	98	26.56	14.9	23.64	**16.8**	2.92	7.8	23.66	16.6	2.9	8
ILE	99	2.48	1.4	2.48	**1.8**	0	0	2.48	1.8	0	0
HIS	100	103.49	56.6	97.92	**66.6**	5.57	15.6	76.96	79.2	26.54	31
LEU	101	8.89	5	8.89	**6.3**	0	0	8.89	6.3	0	0
GLU	102	135.32	78.6	105.75	**78.5**	29.56	78.8	63.96	106.1	71.36	63.7
ILE	103	31.3	17.9	21.96	**15.9**	9.34	25.1	26.2	18.8	5.1	14.2
LYS	104	129.97	64.7	127.14	**77.9**	2.83	7.5	78.43	67.3	51.54	61.2
PRO	105	82.49	60.6	81.98	**68.4**	0.51	3.1	82.49	68.2	0	0
ALA	106	70.8	65.6	63.43	**91.4**	7.37	19.1	65.1	91.2	5.69	15.6
ILE	107	41.29	23.6	41.22	**29.9**	0.07	0.2	41.22	29.6	0.07	0.2
ARG	108	49.73	20.8	49.73	**24.7**	0	0	5.37	6.9	44.35	27.6
ASN	109	70.65	49.1	67.55	**63.6**	3.1	8.2	11.17	24.2	59.47	60.9
GLN	110	72.41	40.6	72.41	**51.4**	0	0	25.77	49.3	46.64	36.9
ILE	111	0.4	0.2	0.4	**0.3**	0	0	0.4	0.3	0	0
ILE	112	32.08	18.3	32.08	**23.3**	0	0	32.08	23.1	0	0
ARG	113	150.84	63.2	141.02	**70.1**	9.83	26.2	74.95	96.3	75.9	47.2
GLU	114	41.21	23.9	41.21	**30.6**	0	0	15.14	25.1	26.06	23.3
LEU	115	7.99	4.5	7.99	**5.7**	0	0	7.99	5.6	0	0
GLN	116	72.51	40.6	67.01	**47.5**	5.5	14.7	40.86	78.2	31.65	25.1
VAL	117	53.62	35.4	53.43	**46.7**	0.19	0.5	53.43	46.3	0.19	0.5
LEU	118	24.21	13.6	23.55	**16.7**	0.66	1.8	23.55	16.5	0.66	1.8
HIS	119	61.54	33.7	51.4	**34.9**	10.14	28.3	39.13	40.3	22.41	26.1
GLU	120	117.36	68.1	109.49	**81.3**	7.87	21	40.61	67.4	76.75	68.5
CYS	121	3.87	2.9	1.08	**1.1**	2.79	7.4	1.08	1.1	2.79	7.7
ASN	122	48.17	33.5	34.93	**32.9**	13.24	35.1	11.23	24.3	36.93	37.8
SER	123	8.04	6.9	3.28	**4.2**	4.76	12.4	2.47	5.1	5.57	8.2
PRO	124	16.52	12.1	13.76	**11.5**	2.76	17	13.76	11.4	2.76	18.2
TYR	125	33.05	15.5	33	**18.6**	0.05	0.1	21.96	16.1	11.09	14.5
ILE	126	6.12	3.5	2.33	**1.7**	3.79	10.2	2.33	1.7	3.79	10.5
VAL	127	10.73	7.1	2.76	**2.4**	7.97	21.4	2.76	2.4	7.97	22.2
GLY	128	8.47	10.6	8.21	**25.4**	0.26	0.5	8.31	22.1	0.16	0.4

PHE	129	8.5	4.3	7.01	**4.3**	1.49	4.2	7.01	4.2	1.49	4.3
TYR	130	17.84	8.4	9.91	**5.6**	7.92	22.4	0.09	0.1	17.75	23.3
GLY	131	5.44	6.8	5.44	**16.8**	0	0	5.44	14.5	0	0
ALA	132	11.33	10.5	0	**0**	11.33	29.4	0	0	11.33	31
PHE	133	21.85	11	21.85	**13.3**	0	0	21.85	13.2	0	0
TYR	134	51.79	24.3	26.79	**15.1**	25	70.7	19.56	14.3	32.23	42.3
SER	135	18.63	16	18.2	**23.3**	0.44	1.1	12.03	24.8	6.6	9.7
ASP	136	129.6	92.3	105.17	**102.4**	24.43	64.8	45.05	91.5	84.55	92.8
GLY	137	8.85	11	5.91	**18.3**	2.93	6.1	5.91	15.7	2.93	6.9
GLU	138	47.95	27.8	47.45	**35.2**	0.49	1.3	8.84	14.7	39.11	34.9
ILE	139	0.89	0.5	0.89	**0.6**	0	0	0.89	0.6	0	0
SER	140	0.38	0.3	0.38	**0.5**	0	0	0	0	0.38	0.6
ILE	141	17.39	9.9	17.39	**12.6**	0	0	17.39	12.5	0	0
CYS	142	0	0	0	**0**	0	0	0	0	0	0
MET	143	13.41	6.9	13.41	**8.6**	0	0	13.41	8.5	0	0
GLU	144	28.16	16.3	28.16	**20.9**	0	0	9.57	15.9	18.59	16.6
HIS	145	41.93	22.9	24.13	**16.4**	17.8	49.7	9.54	9.8	32.39	37.8
MET	146	0.87	0.5	0	**0**	0.87	2.3	0	0	0.87	2.4
ASP	147	57.02	40.6	39.62	**38.6**	17.39	46.1	24.34	49.4	32.68	35.8
GLY	148	6.2	7.7	0	**0**	6.2	13	2.45	6.5	3.75	8.8
GLY	149	12.33	15.4	11.39	**35.2**	0.94	2	11.39	30.3	0.94	2.2
SER	150	0	0	0	**0**	0	0	0	0	0	0
LEU	151	0.82	0.5	0.82	**0.6**	0	0	0.82	0.6	0	0
ASP	152	37.72	26.9	36.5	**35.5**	1.23	3.3	15.45	31.4	22.27	24.4
GLN	153	66.11	37	60.55	**42.9**	5.56	14.8	33.59	64.3	32.52	25.8
VAL	154	0.29	0.2	0.17	**0.2**	0.12	0.3	0.17	0.1	0.12	0.3
LEU	155	15.84	8.9	14.87	**10.5**	0.98	2.6	14.87	10.4	0.97	2.7
LYS	156	168.52	83.9	142.91	**87.5**	25.61	68.3	114.03	97.8	54.49	64.7
LYS	157	134.34	66.9	109.69	**67.2**	24.65	65.7	69.14	59.3	65.19	77.4
ALA	158	33.93	31.4	7.09	**10.2**	26.84	69.6	9.53	13.4	24.4	66.7
GLY	159	46.88	58.5	37.9	**117.2**	8.98	18.8	40.92	109	5.95	14
ARG	160	97.5	40.8	88.64	**44**	8.86	23.6	24.34	31.3	73.16	45.5
ILE	161	1.81	1	0.35	**0.3**	1.46	3.9	0.5	0.4	1.31	3.6
PRO	162	58.96	43.3	58.54	**48.8**	0.42	2.6	58.54	48.4	0.42	2.8

GLU	163	26.3	15.3	21.43	**15.9**	4.87	13	16.06	26.6	10.24	9.1
GLN	164	89.74	50.3	88.6	**62.8**	1.13	3	26.42	50.6	63.31	50.1
ILE	165	7.09	4.1	7.09	**5.1**	0	0	7.09	5.1	0	0
LEU	166	0	0	0	**0**	0	0	0	0	0	0
GLY	167	0	0	0	**0**	0	0	0	0	0	0
LYS	168	40.04	19.9	37.74	**23.1**	2.3	6.1	22.49	19.3	17.54	20.8
VAL	169	0.8	0.5	0.8	**0.7**	0	0	0.8	0.7	0	0
SER	170	0.11	0.1	0.11	**0.1**	0	0	0.11	0.2	0	0
ILE	171	15	8.6	15	**10.9**	0	0	15	10.8	0	0
ALA	172	4.3	4	3.73	**5.4**	0.57	1.5	4.05	5.7	0.25	0.7
VAL	173	1.06	0.7	1.06	**0.9**	0	0	1.06	0.9	0	0
ILE	174	0.14	0.1	0.14	**0.1**	0	0	0.14	0.1	0	0
LYS	175	33.1	16.5	33.1	**20.3**	0	0	4.03	3.5	29.06	34.5
GLY	176	0	0	0	**0**	0	0	0	0	0	0
LEU	177	2.7	1.5	1.57	**1.1**	1.13	3	1.57	1.1	1.13	3.1
THR	178	12.47	9	12.47	**12.3**	0	0	11.41	15.1	1.06	1.7
TYR	179	12.14	5.7	12.14	**6.8**	0	0	7.15	5.2	4.99	6.5
LEU	180	0.16	0.1	0.16	**0.1**	0	0	0.16	0.1	0	0
ARG	181	122.3	51.2	117	**58.1**	5.31	14.1	17.75	22.8 1	4.56	65
GLU	182	115.66	67.1	91.69	**68.1**	23.96	63.9	26.56	44.1	89.09	79.6
LYS	183	124.68	62.1	99.94	**61.2**	24.74	66	58.5	50.2	66.18	78.6
HIS	184	43.86	24	29.51	**20.1**	14.35	40.1	20.98	21.6	22.89	26.7
LYS	185	186.47	92.9	158.26	**96.9**	28.2	75.2	111.36	95.5	75.1	89.2
ILE	186	21.5	12.3	20.31	**14.7**	1.19	3.2	21.14	15.2	0.36	1
MET	187	25.26	13	23.44	**15**	1.82	4.8	23.44	14.9	1.82	5
HIS	188	0	0	0	**0**	0	0	0	0	0	0
ARG	189	65.22	27.3	57.17	**28.4**	8.05	21.5	25.14	32.3	40.08	24.9
ASP	190	49.32	35.1	49.32	**48**	0	0	22.48	45.6	26.85	29.5
VAL	191	1.65	1.1	0.12	**0.1**	1.52	4.1	0.12	0.1	1.52	4.2
LYS	192	15.27	7.6	15.27	**9.4**	0	0	5.46	4.7	9.81	11.6
PRO	193	3.46	2.5	3.46	**2.9**	0	0	3.46	2.9	0	0
SER	194	11.74	10.1	11.18	**14.3**	0.56	1.5	10.32	21.3	1.42	2.1
ASN	195	0.23	0.2	0.23	**0.2**	0	0	0.14	0.3	0.09	0.1
ILE	196	1.54	0.9	1.54	**1.1**	0	0	1.54	1.1	0	0

LEU	197	0	0	0	**0**	0	0	0	0	0	0
VAL	198	0	0	0	**0**	0	0	0	0	0	0
ASN	199	7.58	5.3	7.58	**7.1**	0	0	3.07	6.6	4.52	4.6
SER	200	43.37	37.2	29.8	**38.2**	13.57	35.3	17.2	35.4	26.17	38.5
ARG	201	129.82	54.4	113.64	**56.5**	16.17	43.1	52.59	67.6	77.22	48
GLY	202	24.15	30.2	6.03	**18.7**	18.12	37.9	8.47	22.6	15.68	36.9
GLU	203	23.75	13.8	23.75	**17.6**	0	0	14.17	23.5	9.58	8.6
ILE	204	13.13	7.5	9.22	**6.7**	3.91	10.5	9.22	6.6	3.91	10.9
LYS	205	14.91	7.4	14.65	**9**	0.25	0.7	0.89	0.8	14.02	16.6
LEU	206	1.62	0.9	1.62	**1.1**	0	0	1.62	1.1	0	0
CYS	207	3.4	2.5	3.4	**3.5**	0	0	3.4	3.5	0	0
ASP	208	46.72	33.3	31.41	**30.6**	15.31	40.6	21.19	43	25.53	28
PHE	209	17.95	9	8.34	**5.1**	9.62	27.2	8.87	5.4	9.08	26.5
GLY	210	23.93	29.9	17.72	**54.8**	6.22	13	19.58	52.1	4.36	10.2
VAL	211	34.37	22.7	11.86	**10.4**	22.51	60.6	15.01	13	19.36	53.8
SER	212	3.25	2.8	0.63	**0.8**	2.62	6.8	0.63	1.3	2.62	3.9
GLY	213	41.95	52.4	32.35	**100.1**	9.6	20.1	34.72	92.5	7.24	17
GLN	214	42.58	23.9	39.2	**27.8**	3.38	9	1.74	3.3	40.84	32.3
LEU	215	13.72	7.7	13.72	**9.7**	0	0	13.72	9.6	0	0
ILE	216	81.83	46.7	80.58	**58.4**	1.25	3.4	80.58	57.9	1.25	3.5
ASP	217	84.71	60.3	77.58	**75.5**	7.13	18.9	33.03	67.1	51.68	56.7
SER	218	10.05	8.6	9.7	**12.4**	0.34	0.9	7.7	15.9	2.35	3.5
MET	219	42.12	21.7	42.02	**26.8**	0.09	0.3	42.12	26.7	0	0
ALA	220	60.88	56.4	47.9	**69**	12.98	33.7	49	68.6	11.88	32.5
ASN	221	117.8	81.8	97.59	**91.9**	20.21	53.6	32.08	69.4	85.73	87.7
SER	222	93.08	79.9	73.18	**93.7**	19.9	51.8	59.78	123.1	33.3	49
PHE	223	27.04	13.6	24.27	**14.8**	2.77	7.8	25.27	15.3	1.77	5.2
VAL	224	109.96	72.6	88.43	**77.4**	21.54	58	89.83	77.8	20.14	56
GLY	225	52.55	65.6	31.99	**98.9**	20.57	43.1	33.66	89.6	18.89	44.4
THR	226	61.03	43.8	49.03	**48.2**	12	31.9	36.32	48	24.71	38.9
ARG	227	100.72	42.2	98.89	**49.1**	1.83	4.9	28.27	36.3	72.44	45
SER	228	11.09	9.5	11.09	**14.2**	0	0	5.85	12.1	5.23	7.7
TYR	229	29.42	13.8	28.4	**16**	1.02	2.9	14.95	11	14.47	19
MET	230	31.22	16.1	31.22	**19.9**	0	0	31.22	19.8	0	0

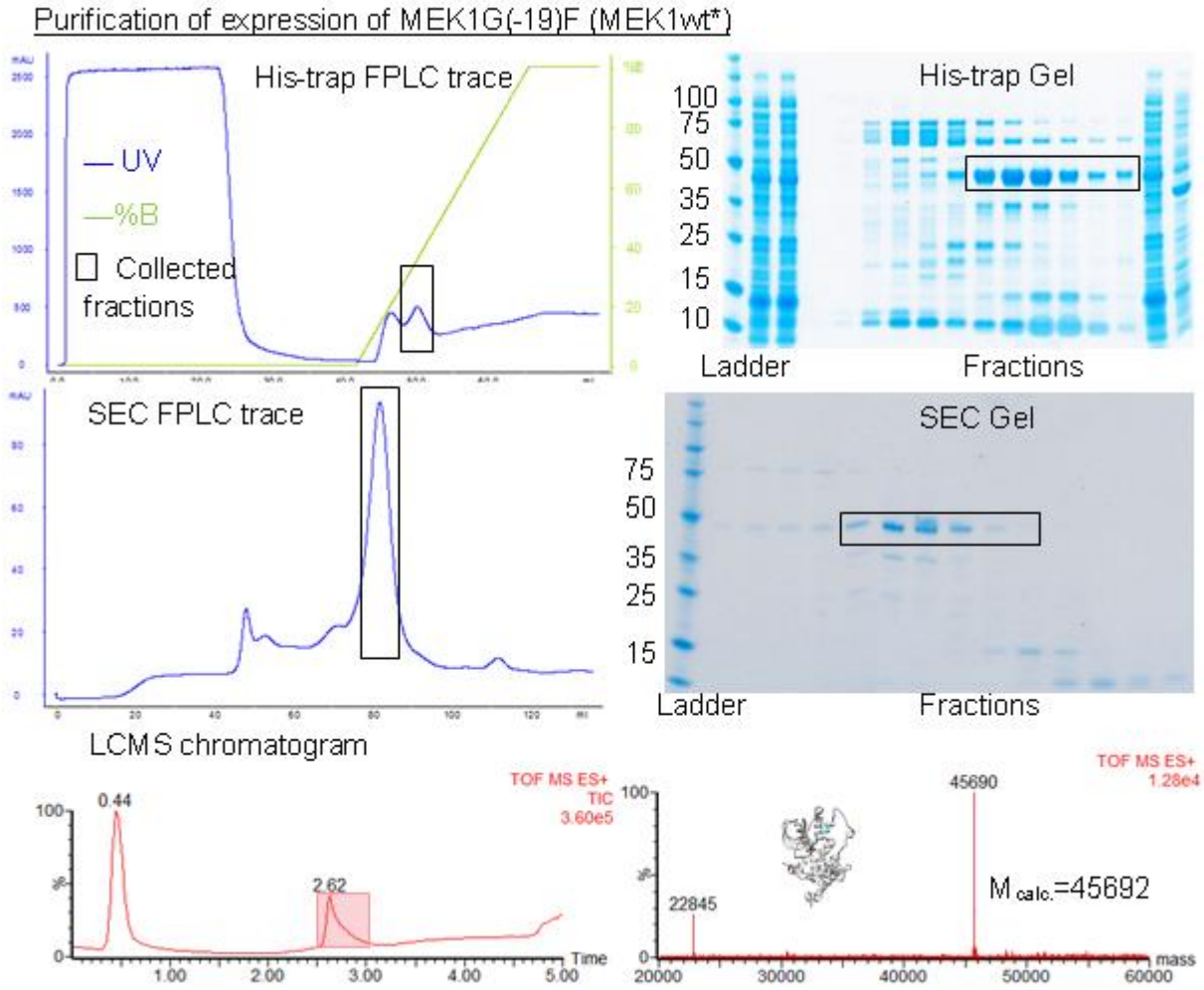
SER	231	0	0	0	**0**	0	0	0	0	0	0
PRO	232	6.69	4.9	6.45	**5.4**	0.24	1.5	6.69	5.5	0	0
GLU	233	6.92	4	6.41	**4.8**	0.5	1.3	6.41	10.6	0.5	0.4
ARG	234	46	19.3	34.78	**17.3**	11.21	29.9	4.76	6.1	41.24	25.6
LEU	235	34.98	19.6	5.59	**4**	29.39	78.3	6.59	4.6	28.38	78.2
GLN	236	114.02	63.9	88.44	**62.7**	25.58	68.2	29.53	56.5	84.5	66.9
GLY	237	73.74	92.1	37.08	**114.7**	36.66	76.7	40.16	107	33.57	78.9
THR	238	87.31	62.7	74.15	**72.9**	13.16	35	66.02	87.2	21.29	33.5
HIS	239	56.25	30.8	55.94	**38**	0.32	0.9	35.01	36	21.25	24.8
TYR	240	118.72	55.8	90.15	**50.8**	28.57	80.7	81.53	59.7	37.19	48.8
SER	241	29.69	25.5	29.69	**38**	0	0	27.27	56.2	2.42	3.6
VAL	242	10.31	6.8	7.96	**7**	2.35	6.3	7.96	6.9	2.35	6.5
GLN	243	17.67	9.9	17.67	**12.5**	0	0	14.42	27.6	3.26	2.6
SER	244	2.65	2.3	1.97	**2.5**	0.68	1.8	0.68	1.4	1.97	2.9
ASP	245	1.83	1.3	1.83	**1.8**	0	0	0	0	1.83	2
ILE	246	0	0	0	**0**	0	0	0	0	0	0
TRP	247	1.12	0.4	1.12	**0.5**	0	0	1.12	0.6	0	0
SER	248	7.74	6.6	7.74	**9.9**	0	0	0	0	7.74	11.4
MET	249	2.04	1.1	2.04	**1.3**	0	0	2.04	1.3	0	0
GLY	250	0	0	0	**0**	0	0	0	0	0	0
LEU	251	0	0	0	**0**	0	0	0	0	0	0
SER	252	0.69	0.6	0.69	**0.9**	0	0	0.69	1.4	0	0
LEU	253	1.58	0.9	1.58	**1.1**	0	0	1.58	1.1	0	0
VAL	254	1.25	0.8	1.25	**1.1**	0	0	1.25	1.1	0	0
GLU	255	21.81	12.7	21.81	**16.2**	0	0	0	0	21.81	19.5
MET	256	0.46	0.2	0.46	**0.3**	0	0	0.46	0.3	0	0
ALA	257	0.56	0.5	0	**0**	0.56	1.4	0	0	0.56	1.5
VAL	258	7.2	4.8	3.49	**3.1**	3.7	10	3.49	3	3.7	10.3
GLY	259	29.12	36.4	1.86	**5.8**	27.26	57.1	1.97	5.2	27.15	63.8
ARG	260	58.34	24.4	58.34	**29**	0	0	30.89	39.7	27.46	17.1
TYR	261	45.46	21.4	35.94	**20.3**	9.53	26.9	35.94	26.3	9.53	12.5
PRO	262	2.94	2.2	2.63	**2.2**	0.31	1.9	2.63	2.2	0.31	2
ILE	263	11.23	6.4	3.5	**2.5**	7.72	20.8	3.5	2.5	7.72	21.5
PRO	264	49.47	36.3	49.47	**41.3**	0	0	49.47	40.9	0	0

PRO	265	68.64	50.4	59.47	**49.6**	9.17	56.5	59.57	49.2	9.08	59.8
PRO	266	27.65	20.3	16.45	**13.7**	11.2	69	17.03	14.1	10.62	69.9
ASP	267	93.38	66.5	92.49	**90.1**	0.89	2.4	27.92	56.7	65.46	71.8
ALA	268	86.83	80.4	72.49	**104.4**	14.34	37.2	73.53	103	13.3	36.4
LYS	269	154.5	76.9	152.41	**93.3**	2.09	5.6	103.56	88.8	50.94	60.5
GLU	270	58.19	33.8	57.99	**43**	0.2	0.5	29.08	48.2	29.11	26
LEU	271	50.74	28.4	49.7	**35.2**	1.04	2.8	50	35.1	0.74	2
GLU	272	94.99	55.1	92.39	**68.6**	2.6	6.9	36.98	61.3	58.01	51.8
LEU	273	141.82	79.4	111.74	**79.2**	30.08	80.2	112.85	79.3	28.96	79.7
MET	274	43.27	22.3	35.24	**22.5**	8.03	21.4	35.51	22.5	7.76	21.4
PHE	275	41.29	20.7	14.91	**9.1**	26.37	74.6	19.57	11.8	21.72	63.4
GLY	276	65.93	82.3	31.05	**96**	34.89	73	33.56	89.4	32.37	76.1
CYS	277	59.19	44.1	59.19	**61.2**	0	0	59.19	60.4	0	0
PRO	307	190.59	140	143.21	**119.4**	47.38	292	146.3	121	44.29	291.6
MET	308	70.23	36.2	41.96	**26.8**	28.27	75.4	42.39	26.9	27.84	76.6
ALA	309	32.3	29.9	32.01	**46.1**	0.29	0.7	32.01	44.8	0.29	0.8
ILE	310	88.3	50.4	82.83	**60**	5.47	14.7	82.83	59.5	5.47	15.2
PHE	311	144.19	72.3	137.14	**83.6**	7.04	19.9	138	83.5	6.19	18.1
GLU	312	76.54	44.4	75.77	**56.2**	0.77	2	31.17	51.7	45.37	40.5
LEU	313	20.19	11.3	20.19	**14.3**	0	0	20.19	14.2	0	0
LEU	314	33.76	18.9	32.72	**23.2**	1.05	2.8	33.76	23.7	0	0
ASP	315	81.74	58.2	81.08	**79**	0.66	1.7	27.49	55.8	54.24	59.5
TYR	316	59.28	27.9	58.97	**33.2**	0.31	0.9	37.89	27.8	21.39	28
ILE	317	4.53	2.6	3.63	**2.6**	0.9	2.4	3.63	2.6	0.9	2.5
VAL	318	29.34	19.4	18.51	**16.2**	10.83	29.2	18.51	16	10.83	30.1
ASN	319	92.59	64.3	67.23	**63.3**	25.36	67.3	24.76	53.6	67.84	69.4
GLU	320	68.22	39.6	65.68	**48.7**	2.55	6.8	13.75	22.8	54.47	48.7
PRO	321	99.98	73.4	92.21	**76.9**	7.76	47.8	92.21	76.2	7.76	51.1
PRO	322	45.04	33.1	33.85	**28.2**	11.19	68.9	33.86	28	11.18	73.6
PRO	323	10.37	7.6	8.52	**7.1**	1.85	11.4	8.52	7	1.85	12.1
LYS	324	138.86	69.2	122.94	**75.3**	15.92	42.4	94.8	81.3	44.06	52.3
LEU	325	4.58	2.6	0	**0**	4.58	12.2	0.06	0	4.51	12.4
PRO	326	17.71	13	17.68	**14.7**	0.03	0.2	17.68	14.6	0.03	0.2
SER	327	90.11	77.3	69.33	**88.8**	20.78	54.1	35.6	73.3	54.51	80.2

GLY	328	78.42	97.9	39.39	**121.8**	39.04	81.7	41.5	110.5	36.92	86.8
VAL	329	77.88	51.4	50.42	**44.1**	27.46	73.9	52	45	25.88	71.9
PHE	330	21.98	11	13.26	**8.1**	8.73	24.7	16.13	9.8	5.86	17.1
SER	331	38.91	33.4	35.97	**46.1**	2.94	7.7	35.97	74.1	2.94	4.3
LEU	332	142.32	79.7	133.27	**94.4**	9.06	24.1	134.44	94.5	7.89	21.7
GLU	333	65.24	37.9	64.98	**48.2**	0.26	0.7	30.06	49.9	35.18	31.4
PHE	334	0	0	0	**0**	0	0	0	0	0	0
GLN	335	22.6	12.7	22.6	**16**	0	0	6.23	11.9	16.37	13
ASP	336	55.17	39.3	54.68	**53.2**	0.49	1.3	22.94	46.6	32.23	35.4
PHE	337	0	0	0	**0**	0	0	0	0	0	0
VAL	338	0	0	0	**0**	0	0	0	0	0	0
ASN	339	69.09	48	66.49	**62.6**	2.6	6.9	21.96	47.5	47.13	48.2
LYS	340	79.04	39.4	79.04	**48.4**	0	0	55.39	47.5	23.65	28.1
CYS	341	0	0	0	**0**	0	0	0	0	0	0
LEU	342	1.45	0.8	1.23	**0.9**	0.22	0.6	1.45	1	0	0
ILE	343	52.34	29.9	52.34	**37.9**	0	0	52.34	37.6	0	0
LYS	344	53.43	26.6	47.53	**29.1**	5.9	15.7	44.69	38.3	8.74	10.4
ASN	345	65.82	45.7	64.13	**60.4**	1.68	4.5	26.24	56.8	39.58	40.5
PRO	346	24.29	17.8	21.57	**18**	2.72	16.7	22.1	18.3	2.19	14.4
ALA	347	89.93	83.3	62.48	**90**	27.45	71.2	64.04	89.7	25.89	70.8
GLU	348	119.62	69.4	102.38	**76**	17.24	46	47.44	78.7	72.18	64.5
ARG	349	5	2.1	0	**0**	5	13.3	0	0	5	3.1
ALA	350	5.76	5.3	5.25	**7.6**	0.51	1.3	5.25	7.4	0.51	1.4
ASP	351	54.21	38.6	54.04	**52.6**	0.17	0.4	20	40.6	34.21	37.5
LEU	352	15.85	8.9	13.77	**9.8**	2.07	5.5	14.39	10.1	1.45	4
LYS	353	158.31	78.8	155.05	**94.9**	3.26	8.7	106.37	91.2	51.94	61.7
GLN	354	103.59	58	98.9	**70.1**	4.69	12.5	36.18	69.3	67.41	53.4
LEU	355	0.79	0.4	0.79	**0.6**	0	0	0.79	0.6	0	0
MET	356	23.34	12	16.89	**10.8**	6.45	17.2	16.89	10.7	6.45	17.8
VAL	357	91.83	60.6	87.36	**76.4**	4.47	12	87.55	75.8	4.28	11.9
HIS	358	19.98	10.9	19.98	**13.6**	0	0	18.3	18.8	1.68	2
ALA	359	43.76	40.5	36.94	**53.2**	6.82	17.7	36.94	51.7	6.82	18.6
PHE	360	0.08	0	0.08	**0.1**	0	0	0.08	0.1	0	0
ILE	361	1.41	0.8	1.41	**1**	0	0	1.41	1	0	0

LYS	362	118.53	59	112.19	**68.7**	6.33	16.9	88.77	76.1	29.76	35.3
ARG	363	68.71	28.8	67.78	**33.7**	0.93	2.5	55.97	71.9	12.75	7.9
SER	364	0.16	0.1	0.16	**0.2**	0	0	0	0	0.16	0.2
ASP	365	74.68	53.2	60.02	**58.4**	14.66	38.9	16.68	33.9	58	63.6
ALA	366	78.79	73	60.67	**87.4**	18.12	47	62.27	87.2	16.52	45.2
GLU	367	34.14	19.8	31.53	**23.4**	2.61	6.9	25.25	41.9	8.88	7.9
GLU	368	171.64	99.6	138.78	**103**	32.86	87.6	71.84	119.2	99.8	89.1
VAL	369	37.49	24.8	14.21	**12.4**	23.28	62.6	16.29	14.1	21.2	58.9
ASP	370	99.96	71.2	91.3	**88.9**	8.66	23	47.88	97.2	52.08	57.1
PHE	371	9.62	4.8	7.56	**4.6**	2.06	5.8	7.56	4.6	2.06	6
ALA	372	23.77	22	21.71	**31.3**	2.06	5.3	22.32	31.3	1.45	4
GLY	373	28.44	35.5	22.34	**69.1**	6.1	12.8	27.03	72	1.41	3.3
TRP	374	52.18	20.9	49.92	**23.6**	2.26	5.9	39.37	20.8	12.8	21.5
LEU	375	0	0	0	**0**	0	0	0	0	0	0
CYS	376	26.52	19.7	26.24	**27.1**	0.27	0.7	26.35	26.9	0.17	0.5
SER	377	96.13	82.5	71.67	**91.8**	24.46	63.7	43.13	88.8	53	78
THR	378	32.49	23.3	21.28	**20.9**	11.21	29.8	13.16	17.4	19.33	30.4
ILE	379	28.62	16.3	17.99	**13**	10.63	28.6	18.8	13.5	9.82	27.3
GLY	380	64.73	80.8	30.58	**94.6**	34.15	71.5	32.63	86.9	32.09	75.4
LEU	381	84.77	47.5	53.17	**37.7**	31.61	84.3	55.23	38.8	29.54	81.3
ASN	382	70.81	49.2	70.37	**66.2**	0.44	1.2	70.37	152.2	0.44	0.4
ADP	2	48.42	-99.9	48.42	-99.9	0	-99.9	8.81	-99.9	39.62	-99.9
MG	3	4.28	-99.9	4.28	-99.9	0	-99.9	4.28	-99.9	0	-99.9
CA	5	52.96	-99.9	52.96	-99.9	0	-99.9	0	-99.9	52.96	-99.9
NA	6	56.16	-99.9	56.16	-99.9	0	-99.9	0	-99.9	56.16	-99.9
Absolute sums overall chain											
Total		14924.8	1	2726.2		2198.6		8470.8	64	54	

Appendix B: Protein Purification, Characterisation and Sequences

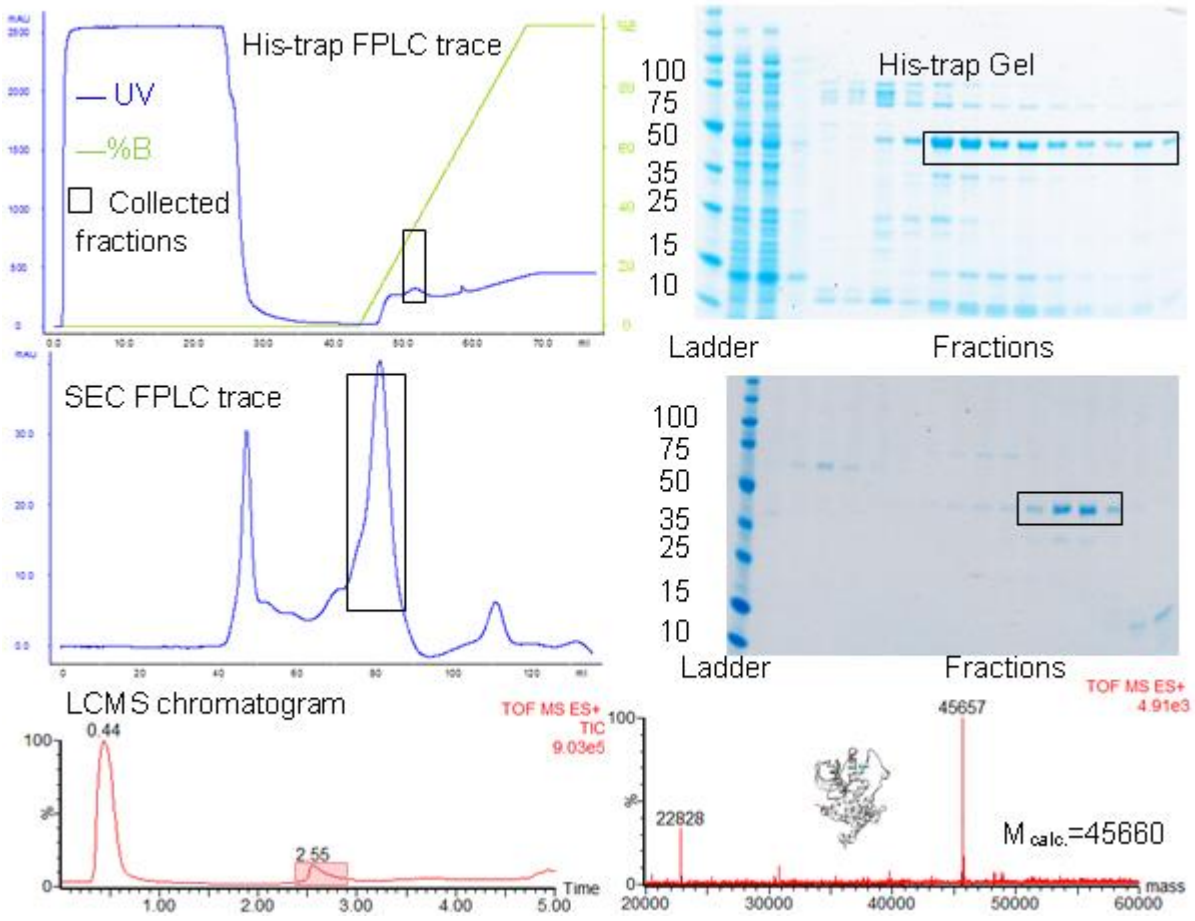


Protein sequence (Thrombin cleavage site in green):

MFSSHHHHHH SSGLVPRGSH

```
1     MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL
51    EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH
101   LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS
151   LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS
201   RGEIKLCDFG VSGQLIDSMA NSFVGTRSYM SPERLQGTHY SVQSDIWSMG
251   LSLVEMAVGR YPIPPPDAKE LELMFGCQVE GDAAETPPRP RTPGRPLSSY
301   GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA
351   DLKQLMVHAF IKRSDAEEVD FAGWLCSTIG LNQPSTPTHA AGV
```

Purification of expression of MEK1G(-19)FC277SC376S (MEK1wt)



Protein sequence:

MFSSHHHHHH SSGLVPRGSH

1 MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL

51 EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH

101 LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS

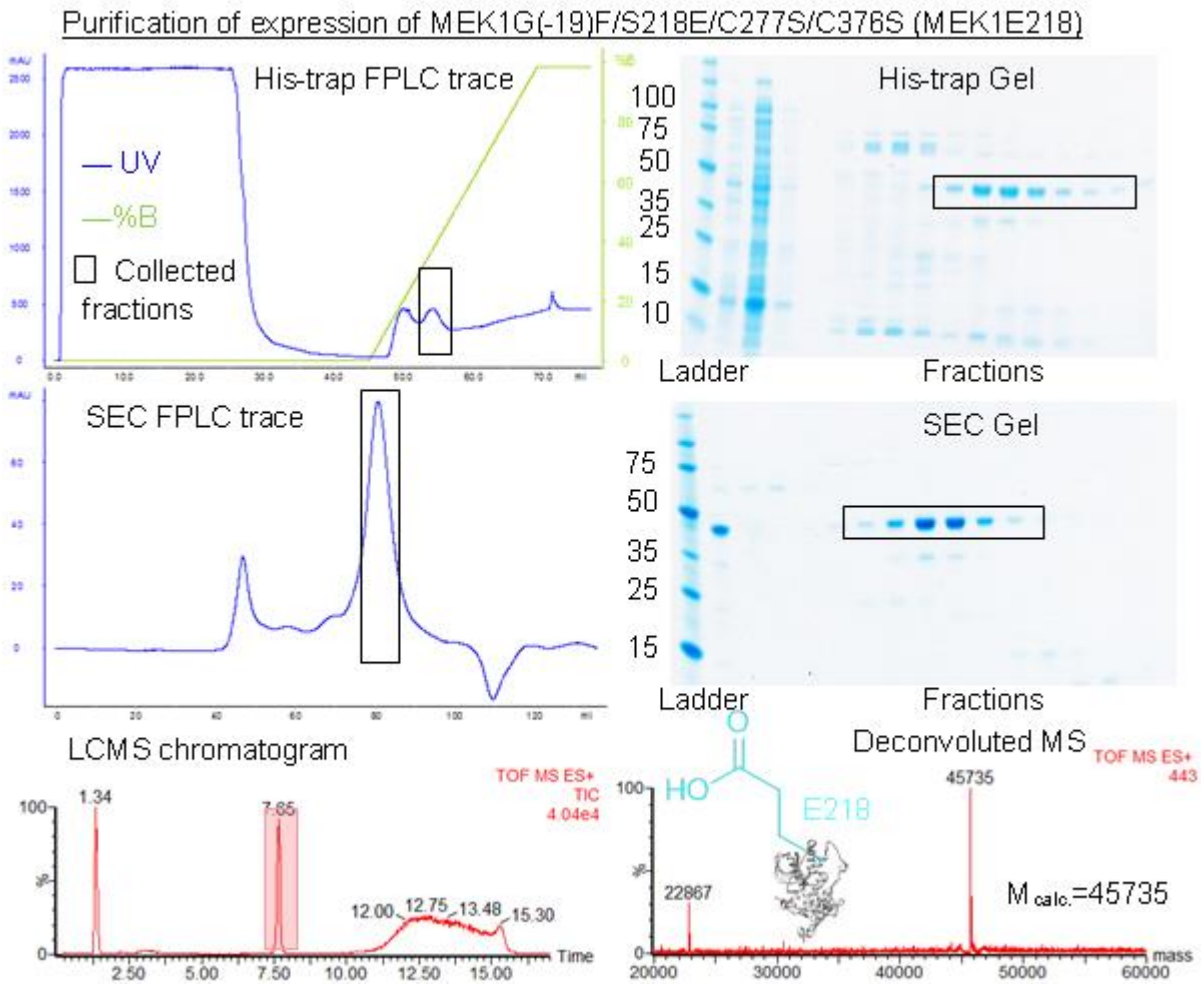
151 LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS

201 RGEIKLCDFG VSGQLIDSMA NSFVGTRSYM SPERLQGTHY SVQSDIWSMG

251 LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY

301 GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA

351 DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNQPSTPTHA AGV



Protein sequence:

MFSSHHHHHH SSGLVPRGSH

1 MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL

51 EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH

101 LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS

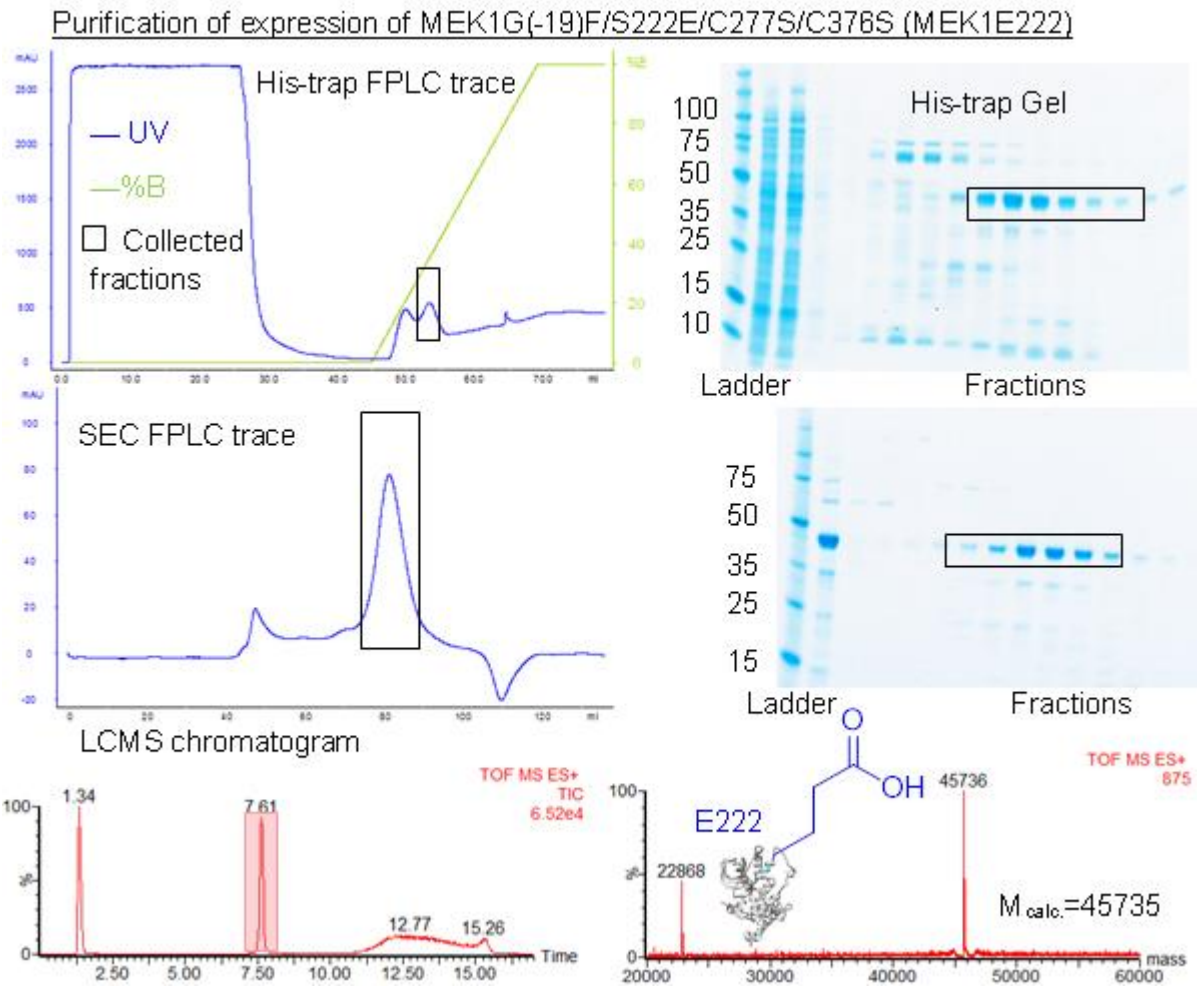
151 LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS

201 RGEIKLCDFG VSGQLIDEMA NSFVGTRSYM SPERLQGTHY SVQSDIWSMG

251 LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY

301 GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA

351 DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNQPSTPTHA AGV



Protein sequence:

MFSSHHHHHH SSGLVPRGSH

1 MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL

51 EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH

101 LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS

151 LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS

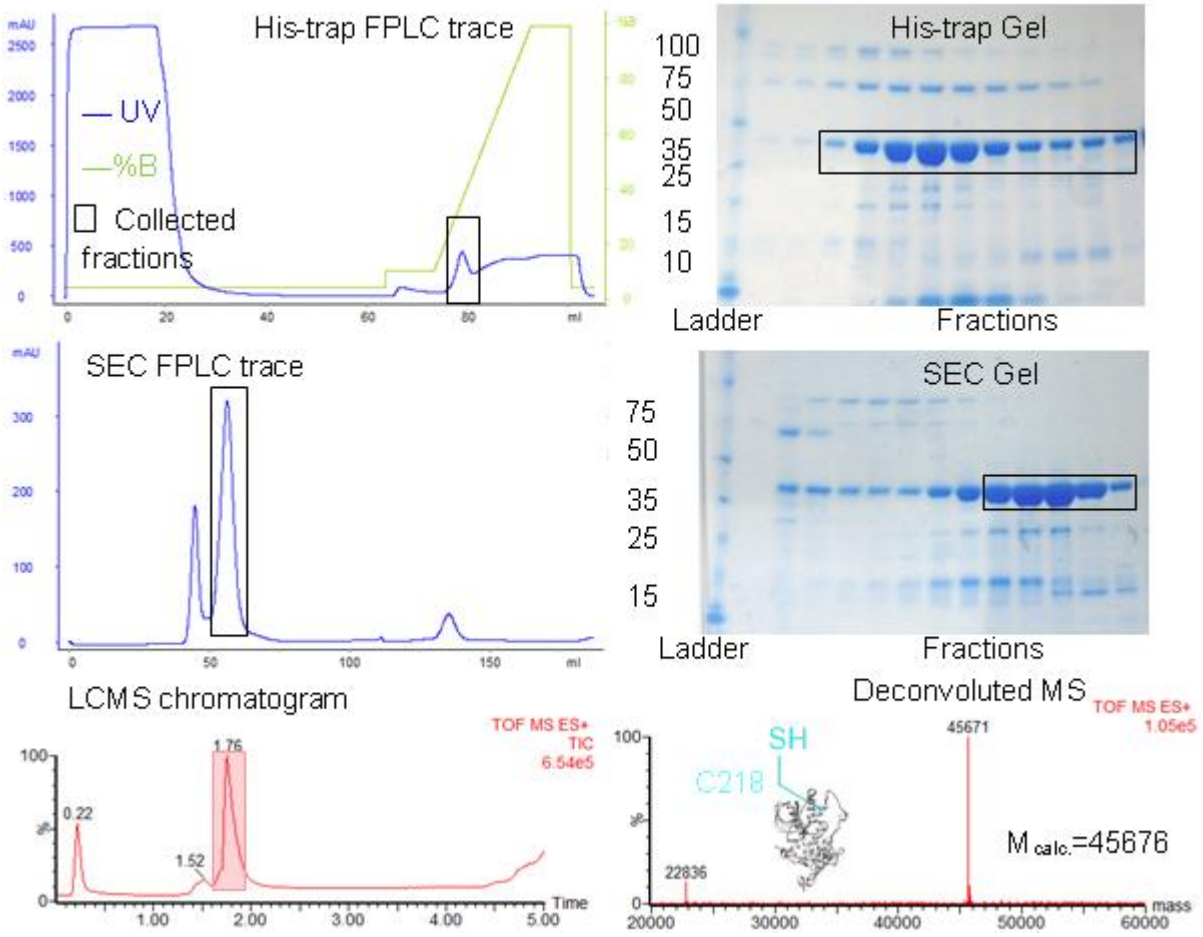
201 RGEIKLCDFG VSGQLIDSMA NEFVGTRSYM SPERLQGTHY SVQSDIWSMG

251 LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY

301 GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA

351 DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNQPSTPTHA AGV

Purification of expression of MEK1G(-19)F/S218C/C277S/C376S (MEK1C218)



Protein sequence:

MFSSHHHHHH SSGLVPRGSH

1 MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL

51 EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH

101 LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS

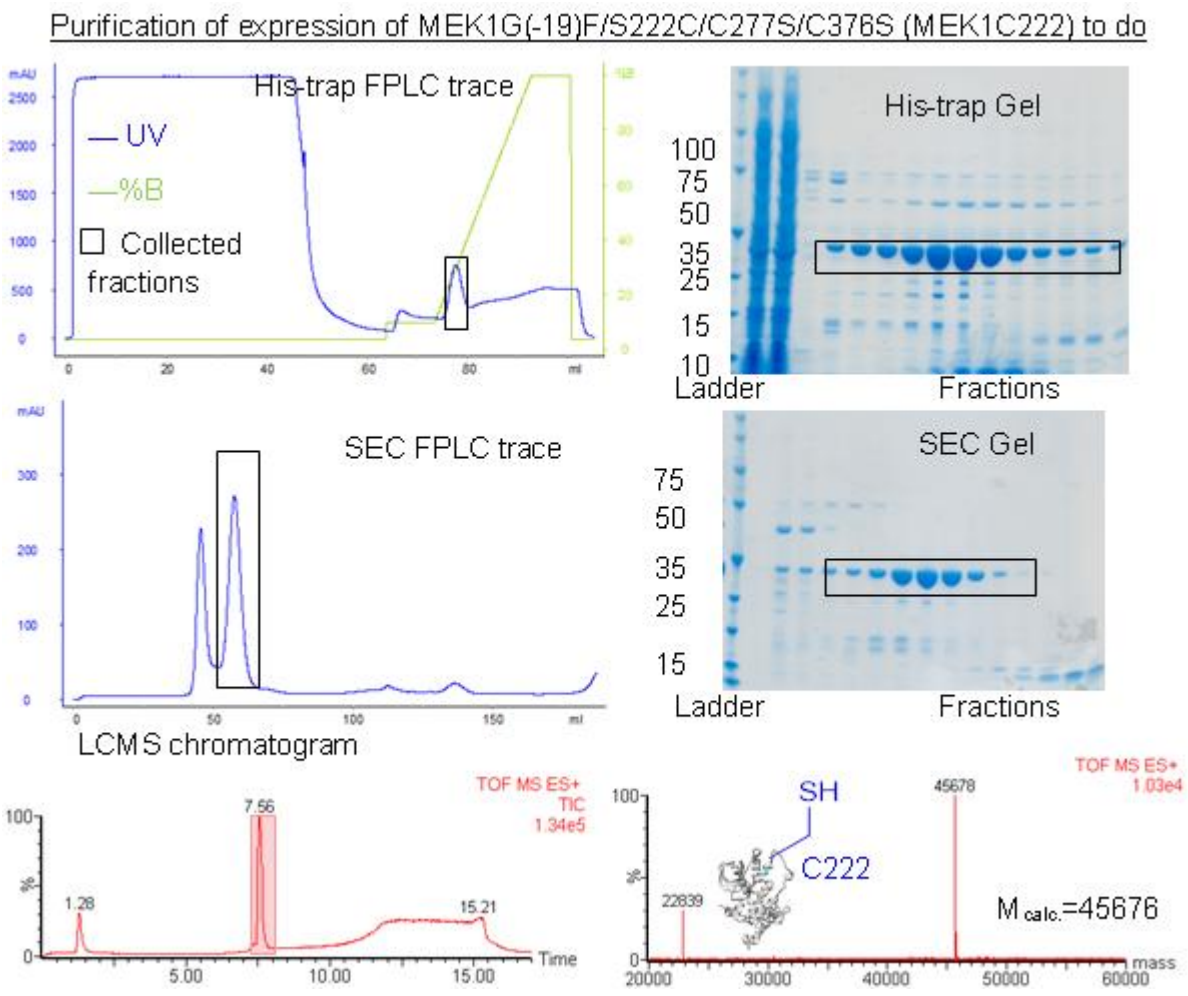
151 LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS

201 RGEIKLCDFG VSGQLIDCMA NSFVGTRSYM SPERLQGTHY SVQSDIWSMG

251 LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY

301 GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA

351 DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNQPSTPTHA AGV



Protein sequence:

MFSSHHHHHH SSGLVPRGSH

1 MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL

51 EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH

101 LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS

151 LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS

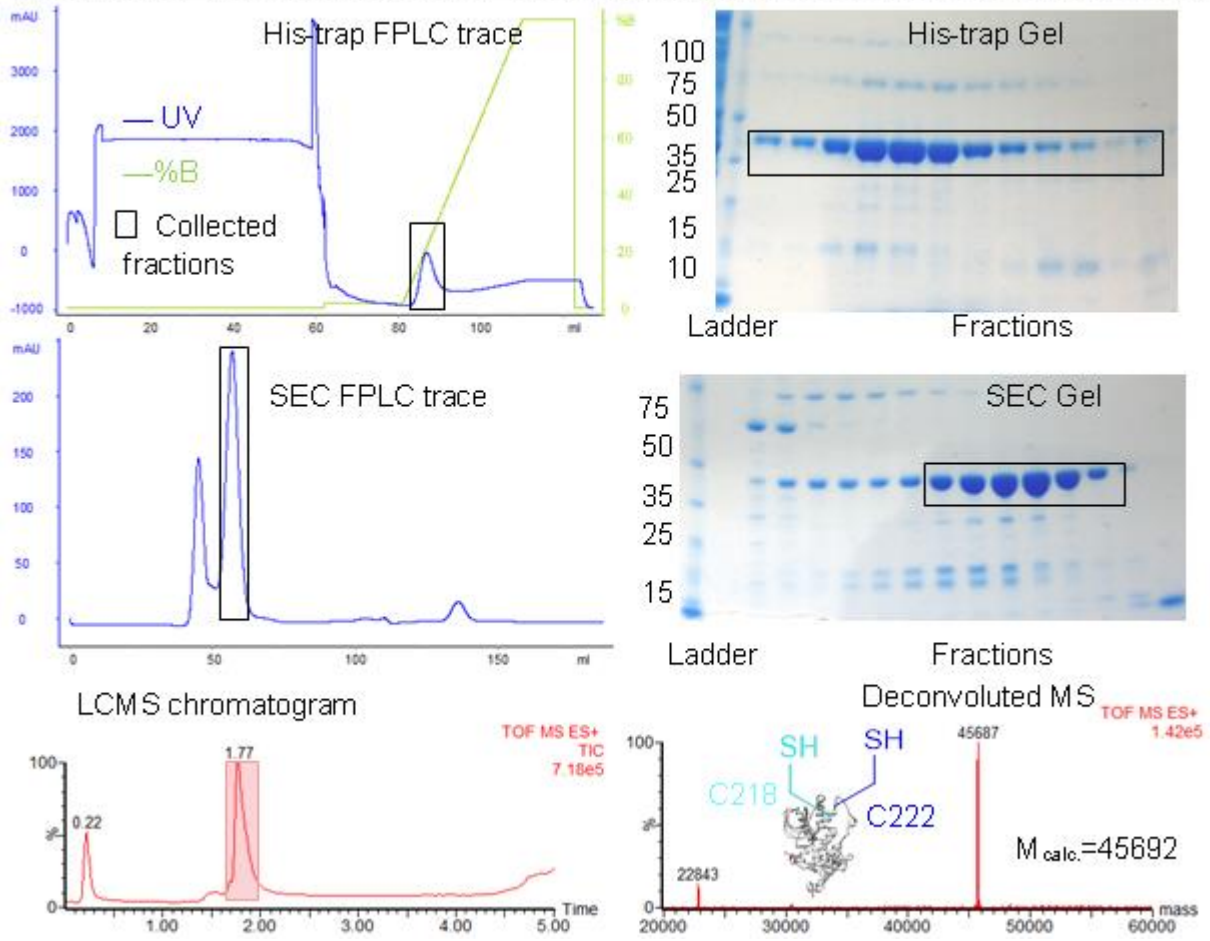
201 RGEIKLCDFG VSGQLIDSMA NCFVGTRSYM SPERLQGTHY SVQSDIWSMG

251 LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY

301 GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA

351 DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNQPSTPTHA AGV

Purification of expression of MEK1G(-19)F/S218C/S222C/C277S/C376S (MEK1C218C222)



Protein sequence:

MFSSHHHHHH SSGLVPRGSH

1 MPKKKPTPIQ LNPAPDGSAV NGTSSAETNL EALQKKLEEL ELDEQQRKRL

51 EAFLTQKQKV GELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH

101 LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS

151 LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS

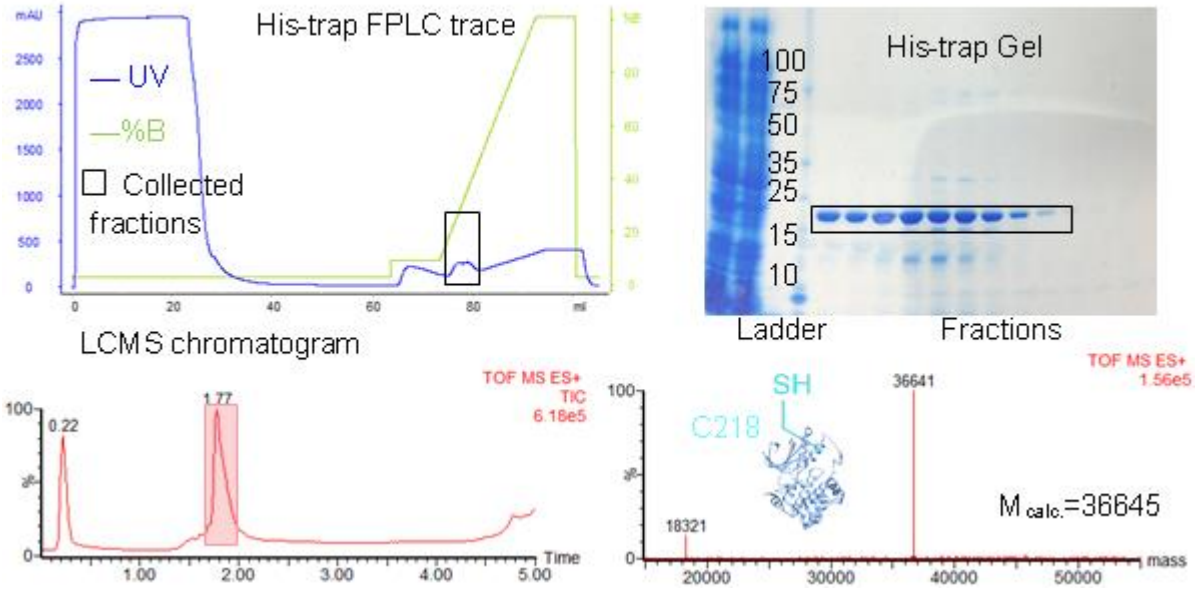
201 RGEIKLCDFG VSGQLIDCMA NCFVGTRSYM SPERLQGTHY SVQSDIWSMG

251 LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY

301 GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA

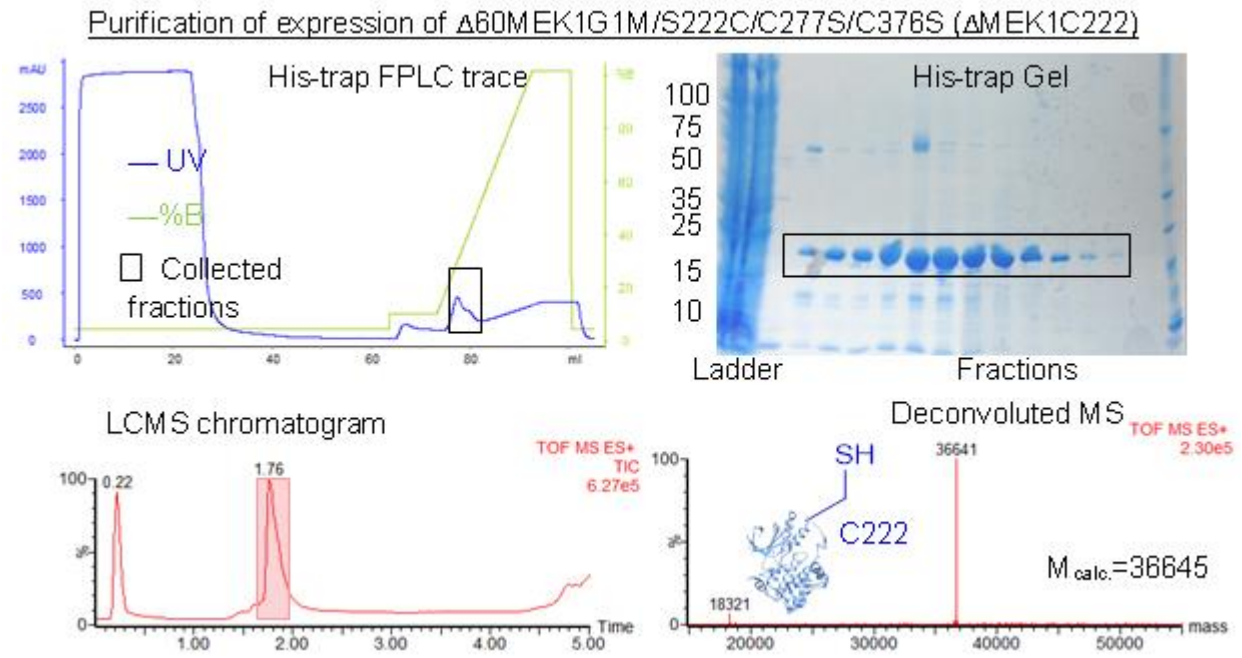
351 DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNQPSTPTHA AGV

Purification of expression of Δ61MEK1S218C/C277S/C376S (ΔMEK1C218)



Protein sequence:

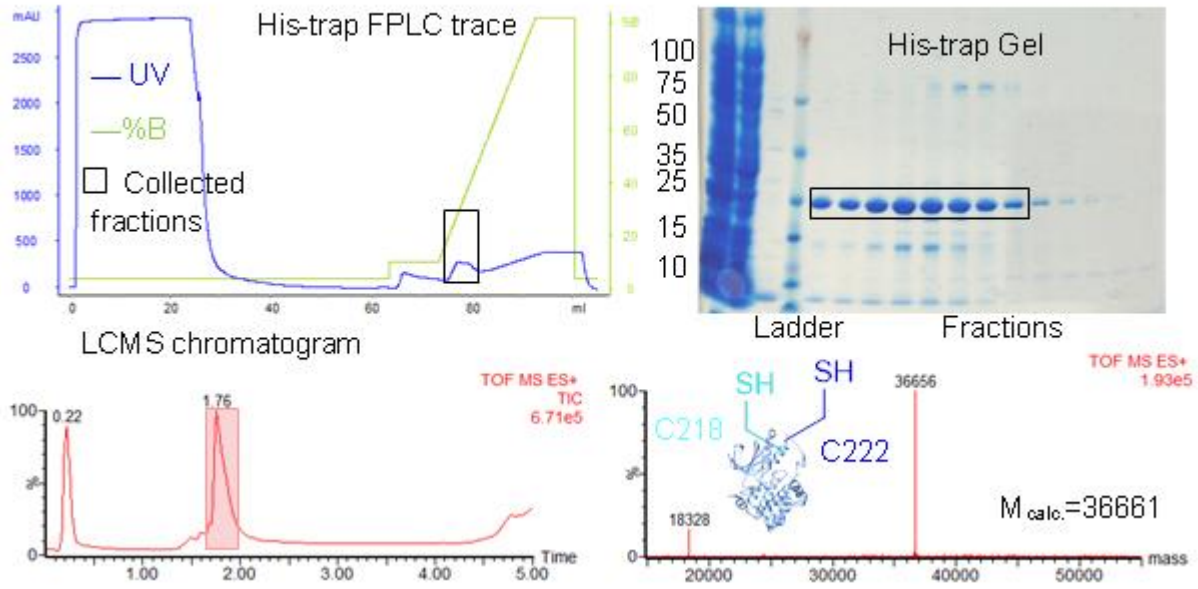
```
1
51                MELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH
101   LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS
151   LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS
201   RGEIKLCDFG VSGQLIDCMA NSFVGTRSYM SPERLQGTHY SVQSDIWSMG
251   LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY
301   GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA
351   DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNAHHHHHH
```



Protein sequence:

```
1
51              MELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH
101   LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS
151   LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS
201   RGEIKLCDFG VSGQLIDSMA NCFVGTRSYM SPERLQGTHY SVQSDIWSMG
251   LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY
301   GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA
351   DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNAHHHHHH
```

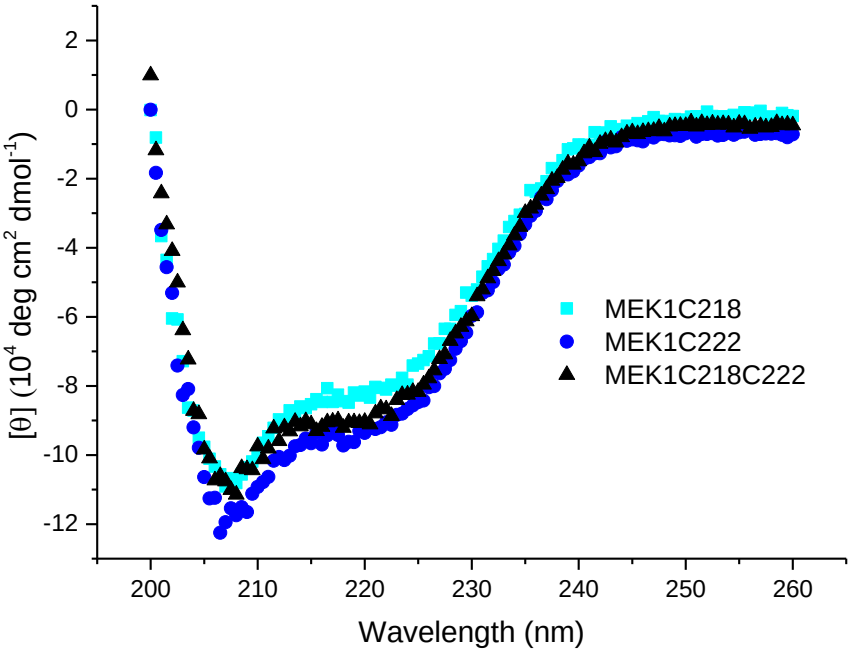
Purification of expression of Δ60MEK1G1M/S218C/S222C/C277S/C376S (ΔMEK1C218C222)



Protein sequence:

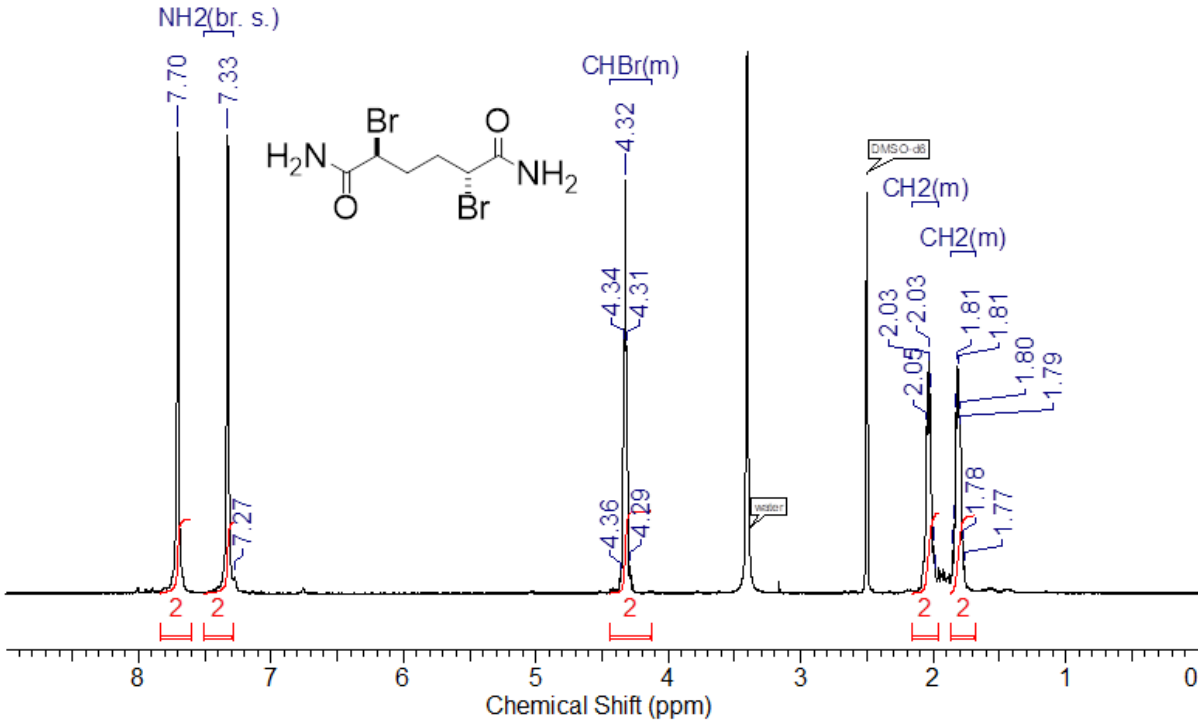
```
1
51                MELKDDDFEK ISELGAGNGG VVFKVSHKPS GLVMARKLIH
101   LEIKPAIRNQ IIRELQVLHE CNSPYIVGFY GAFYSDGEIS ICMEHMDGGS
151   LDQVLKKAGR IPEQILGKVS IAVIKGLTYL REKHKIMHRD VKPSNILVNS
201   RGEIKLCDFG VSGQLIDCMA NCFVGTRSYM SPERLQGTHY SVQSDIWSMG
251   LSLVEMAVGR YPIPPPDAKE LELMFGSQVE GDAAETPPRP RTPGRPLSSY
301   GMDSRPPMAI FELLDYIVNE PPPKLPSGVF SLEFQDFVNK CLIKNPAERA
351   DLKQLMVHAF IKRSDAEEVD FAGWLSSTIG LNAHHHHHH
```

Appendix C: Circular Dichroism of MEK1C218, MEK1C222 and MEK1C218C222

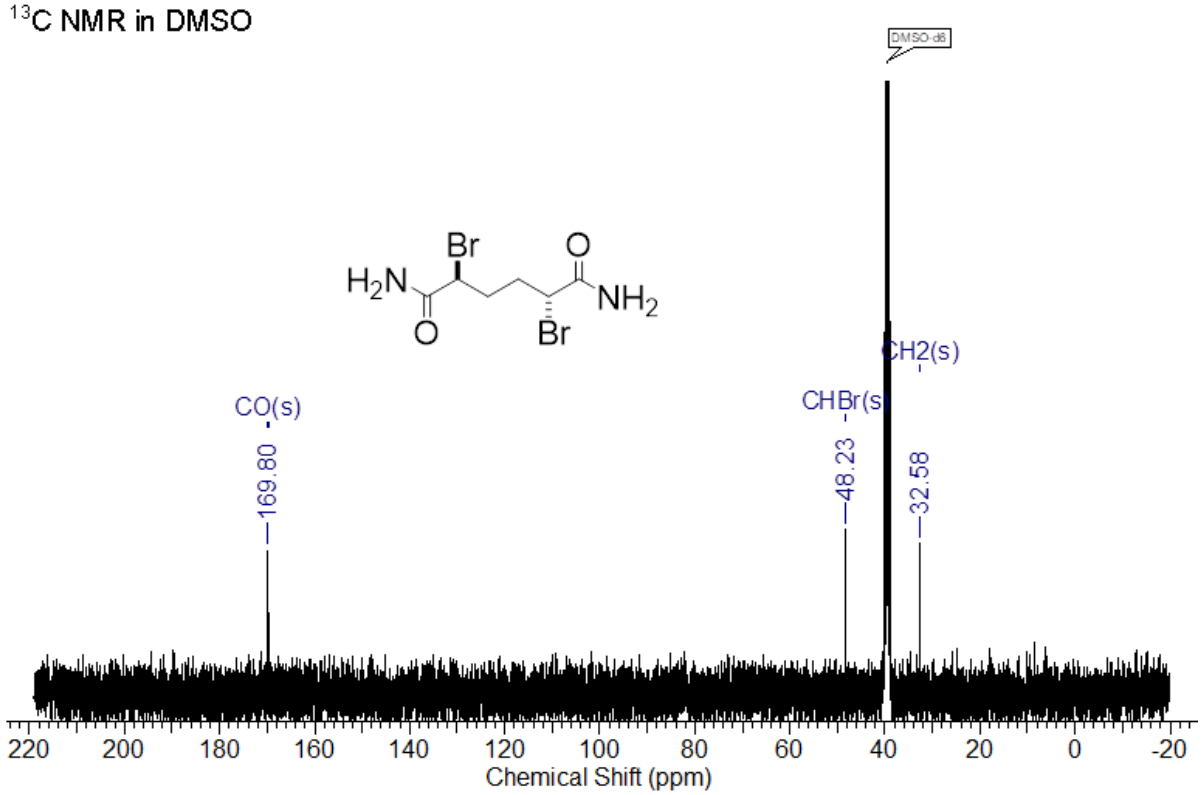


Appendix D: NMR spectra for small molecules used in the Site-selective Activation of MEK1

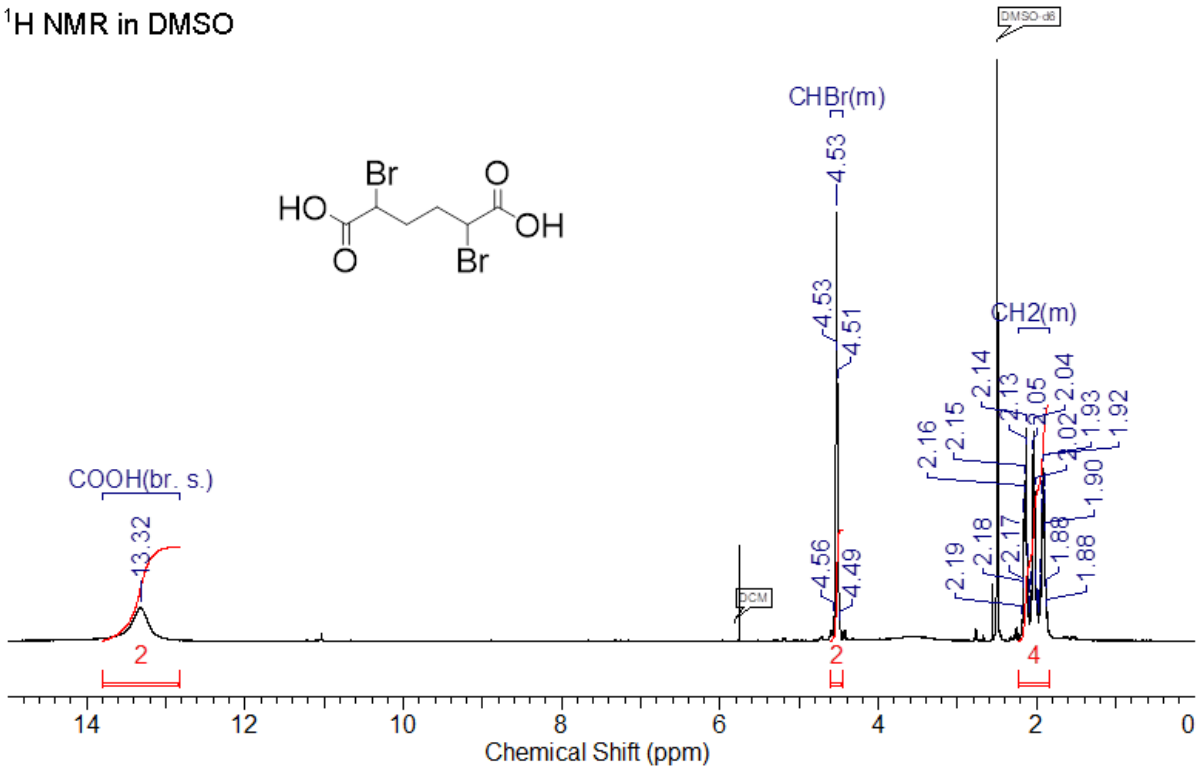
^{1}H NMR in DMSO



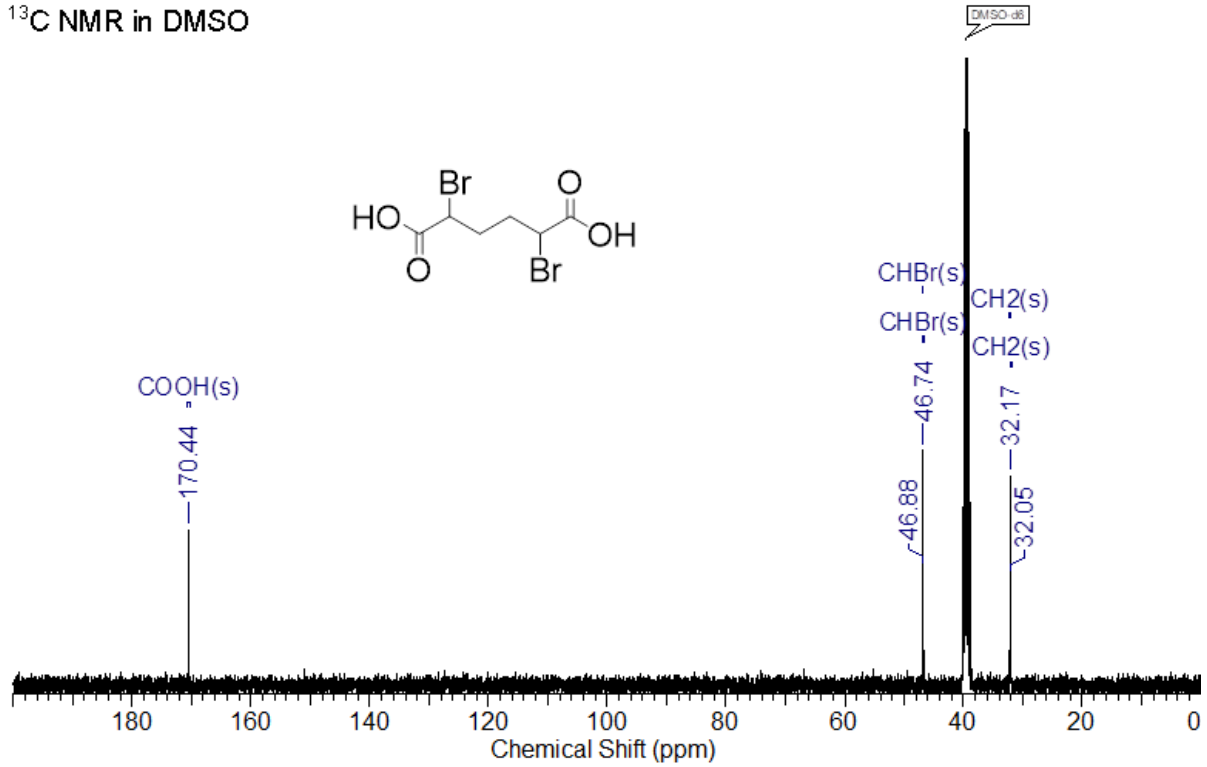
^{13}C NMR in DMSO



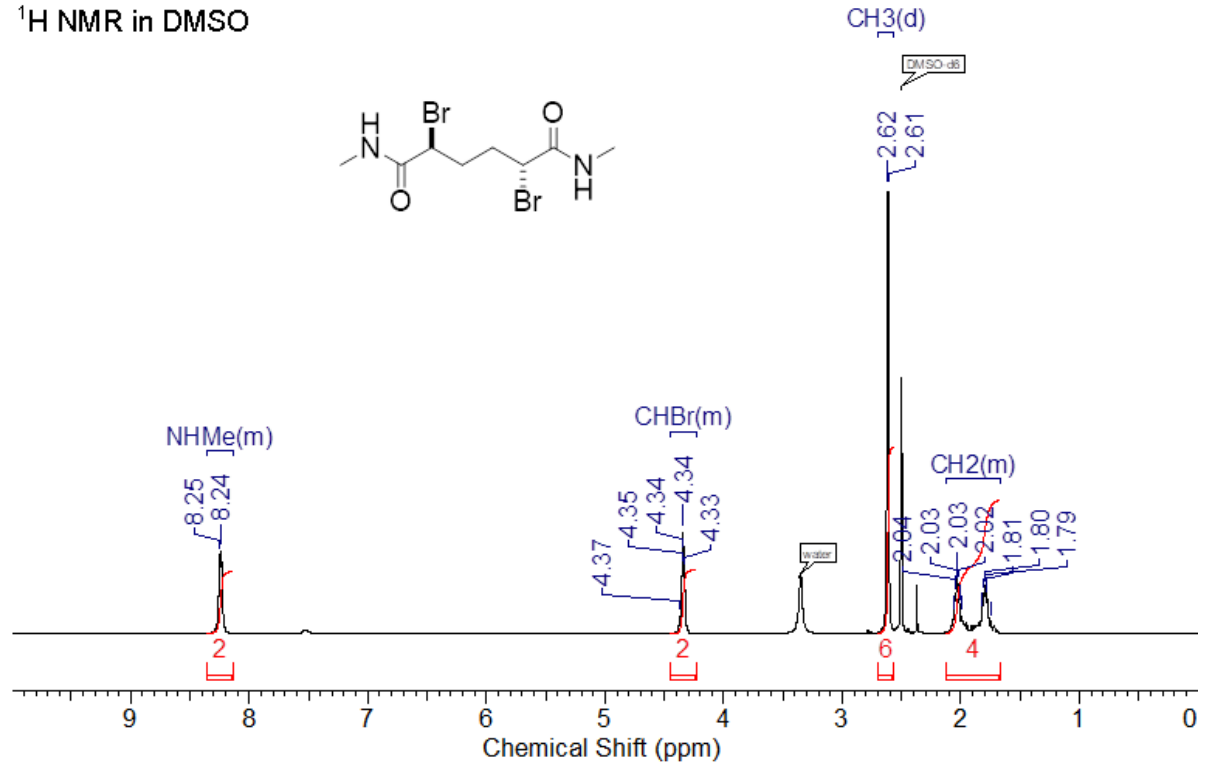
^{1}H NMR in DMSO



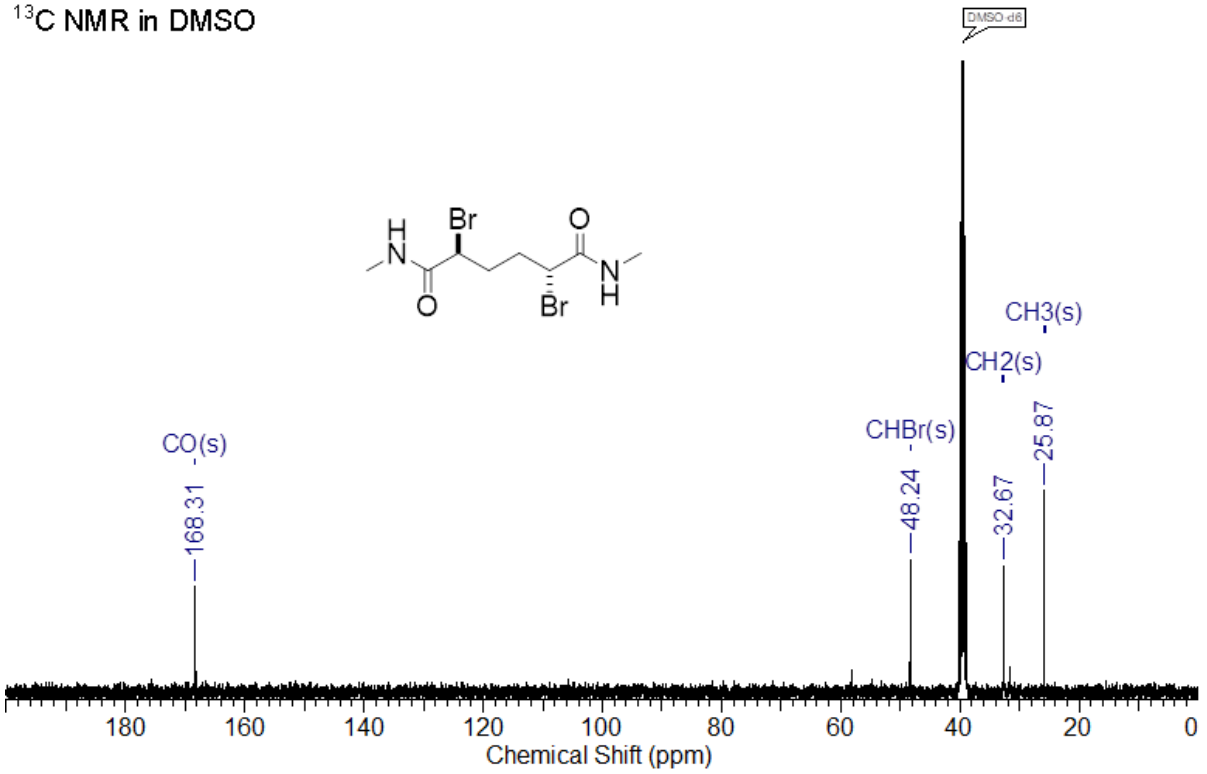
^{13}C NMR in DMSO

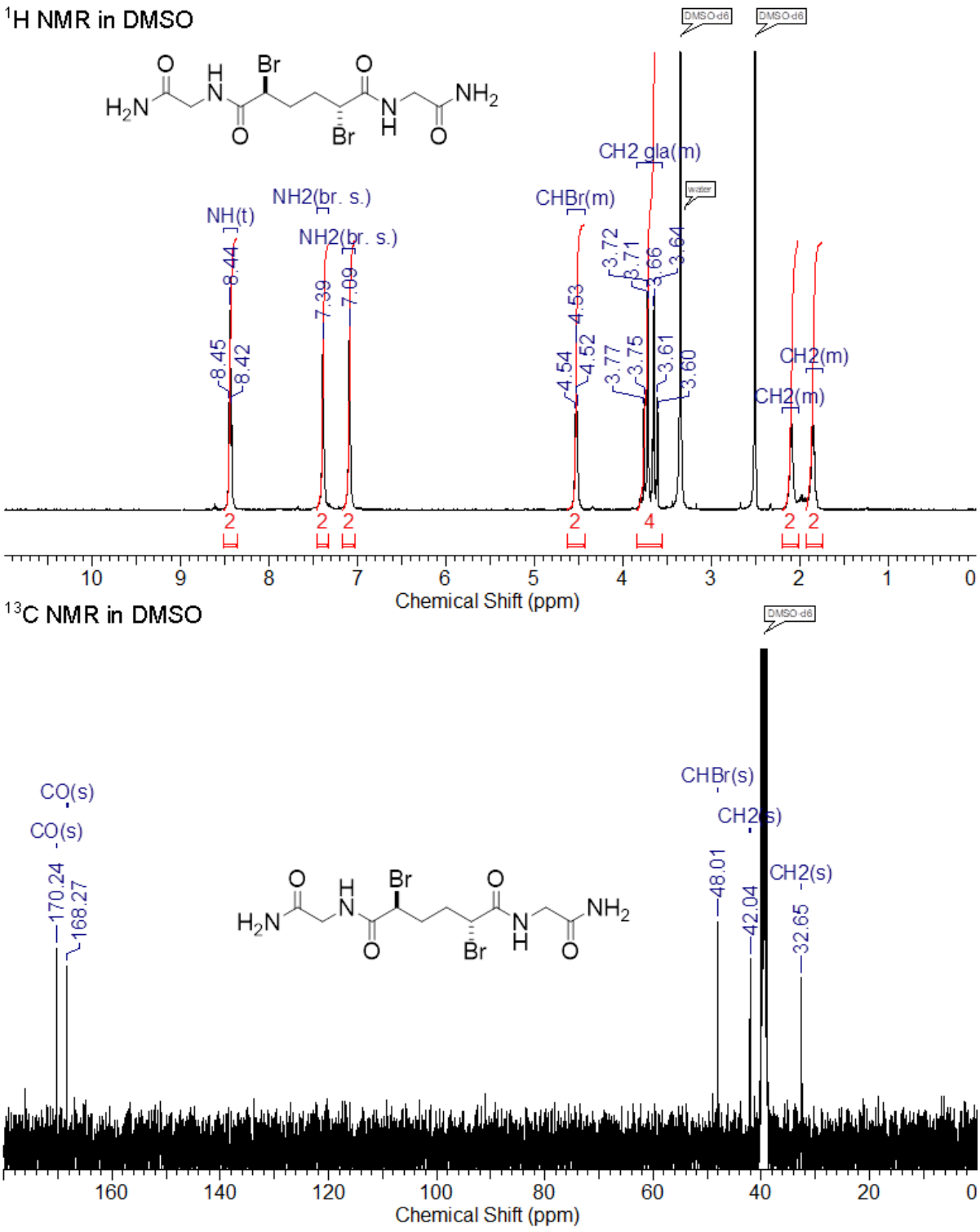


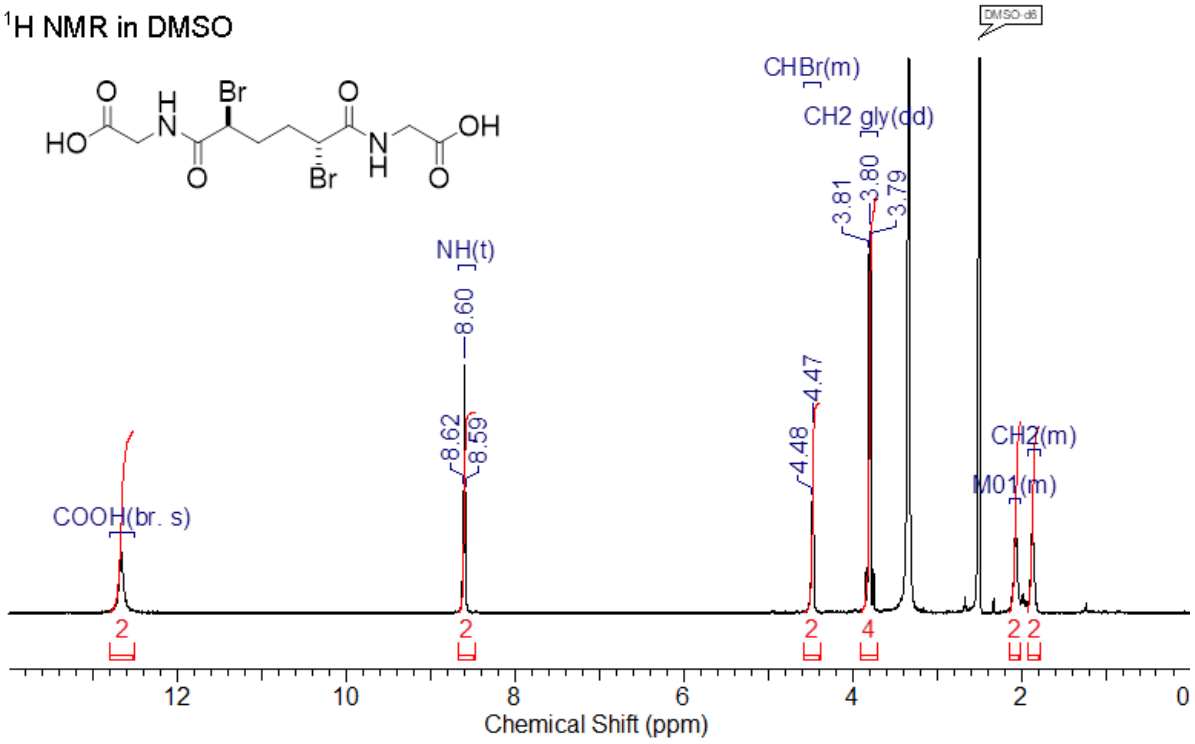
^{1}H NMR in DMSO



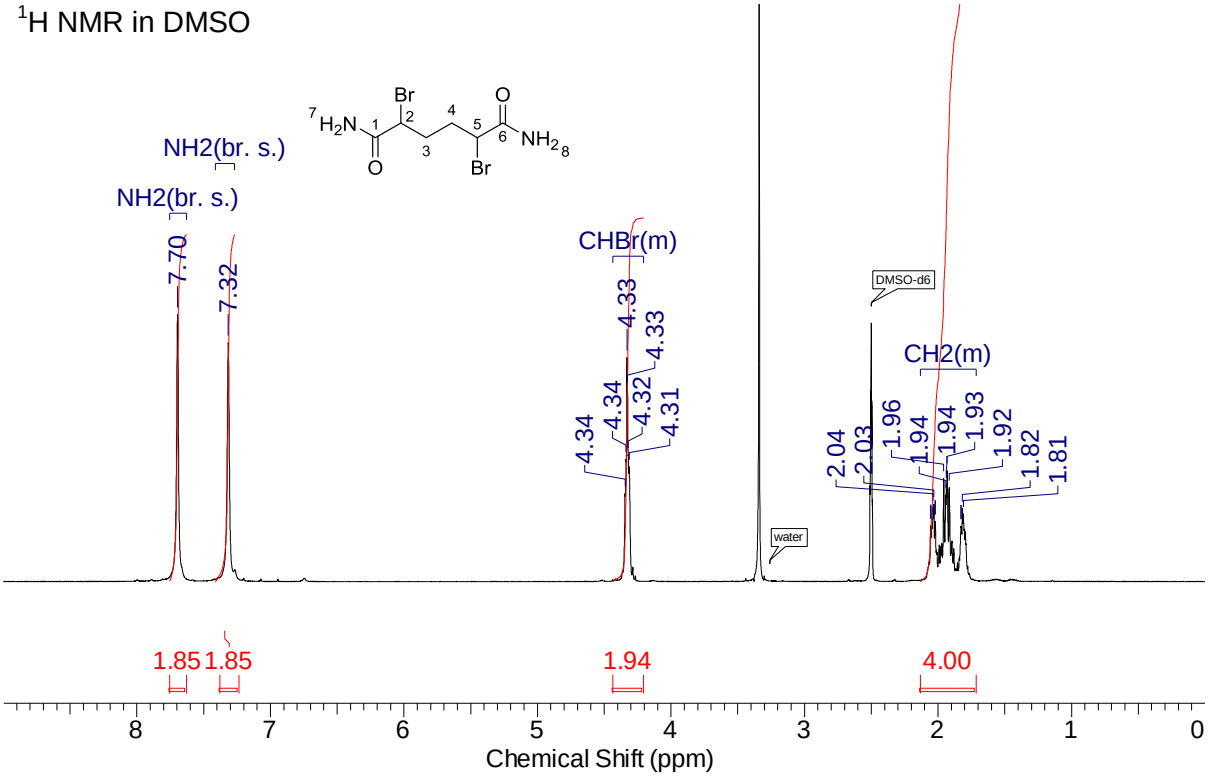
^{13}C NMR in DMSO

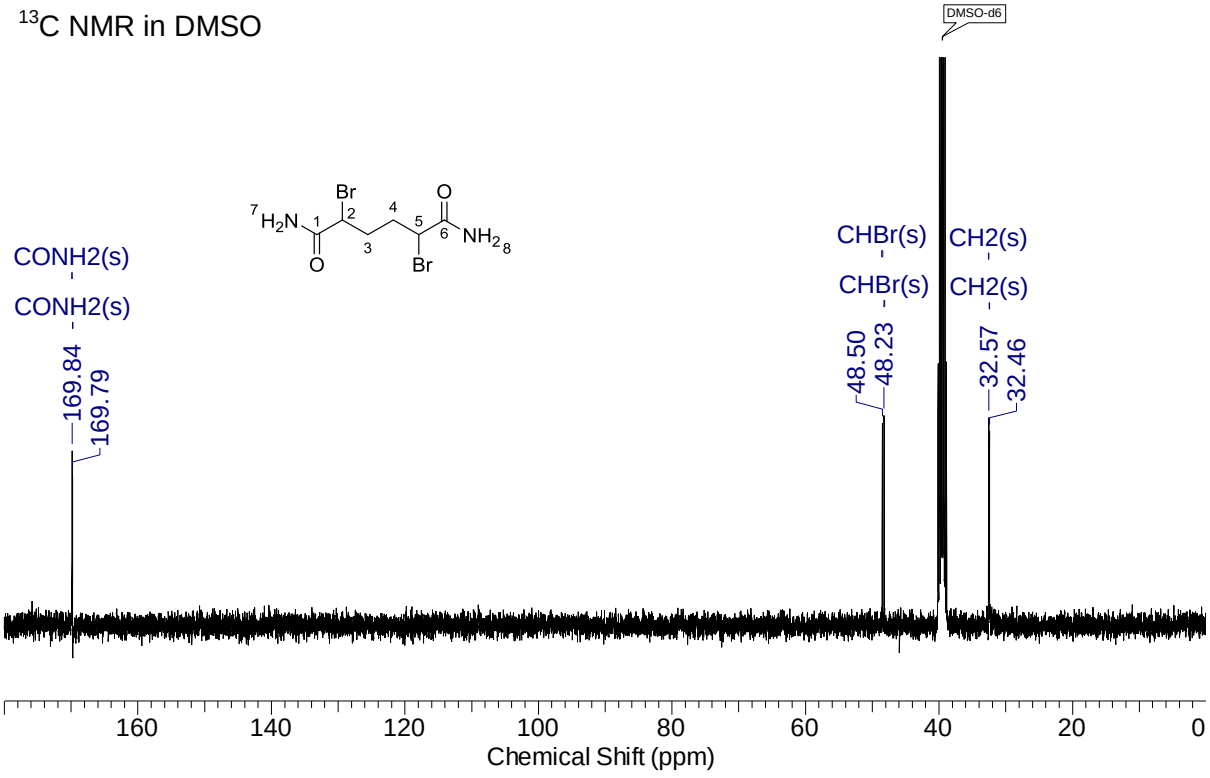



1H NMR in DMSO
DMSO-d6
DMSO-d6
water
NH(t)
NH2(br. s.)
NH2(br. s.)
CHBr(m)
CH2 gla(m)
CH2(m)
CH2(m)
8.45
8.44
8.42
7.39
7.09
4.54
4.53
4.52
3.77
3.75
3.72
3.71
3.66
3.64
3.61
3.60
2
2
2
2
4
2
2
Chemical Shift (ppm)
13C NMR in DMSO
DMSO-d6
CO(s)
CO(s)
170.24
168.27
CHBr(s)
CH2(s)
48.01
42.04
CH2(s)
32.65
Chemical Shift (ppm)

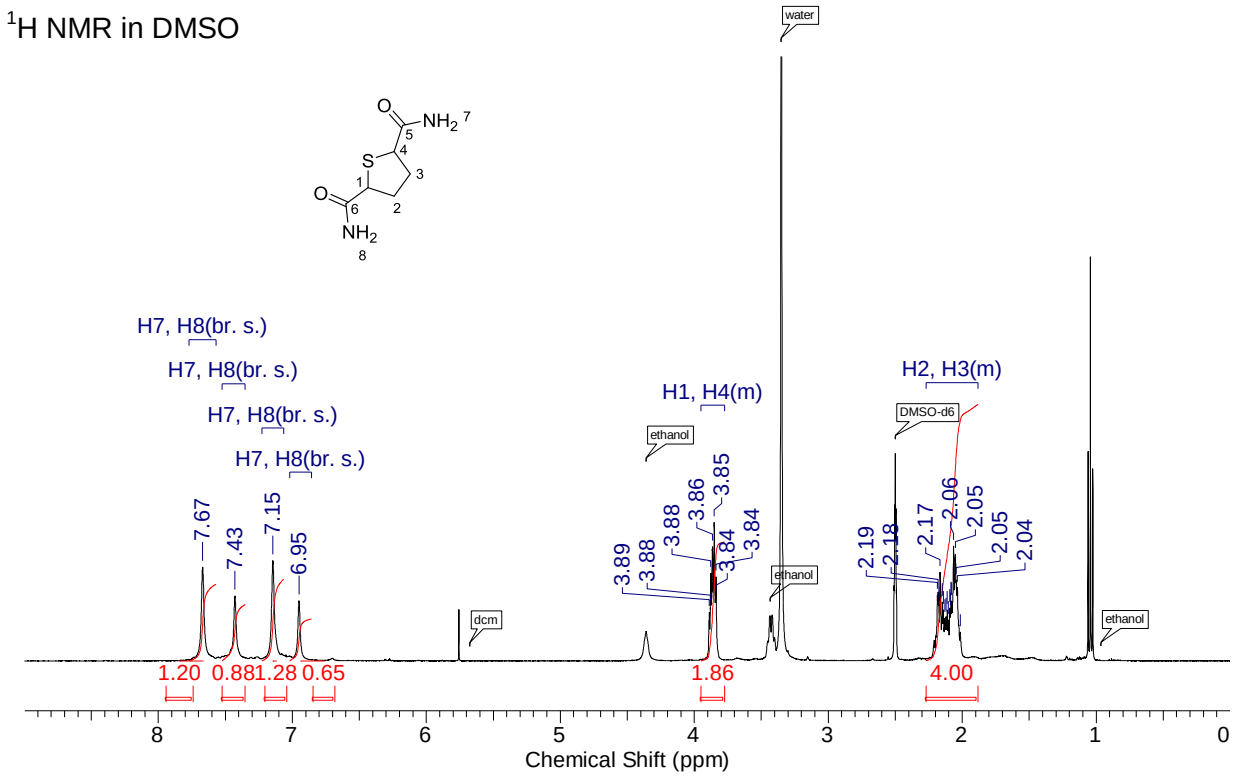

1H NMR in DMSO
DMSO-d6
CHBr(m)
CH2 gly(dd)
NH(t)
COOH(br. s)
CH2(m)
M01(m)
8.60
8.62
8.59
4.47
4.48
3.81
3.80
3.79
2
2
2
4
2
2
12
10
8
6
4
2
0
Chemical Shift (ppm)

Appendix E: NMR spectra for small molecules used in the mechanism elucidation

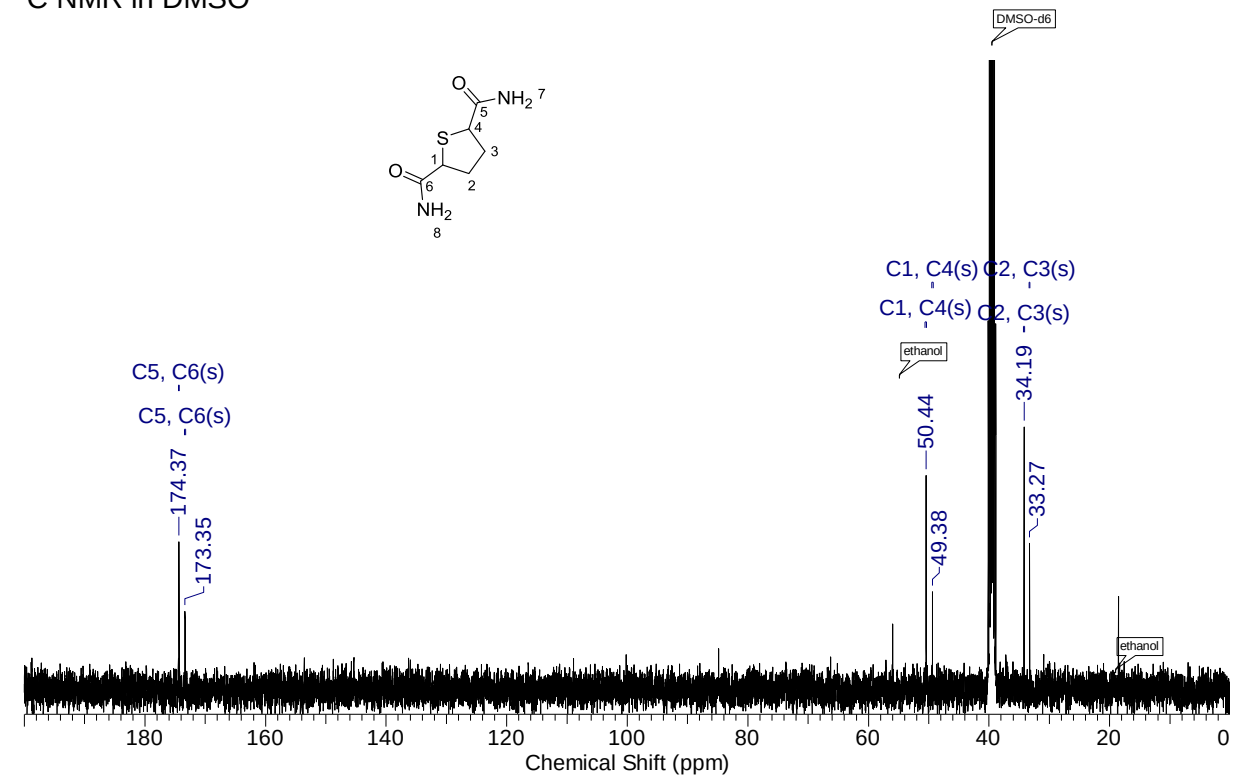




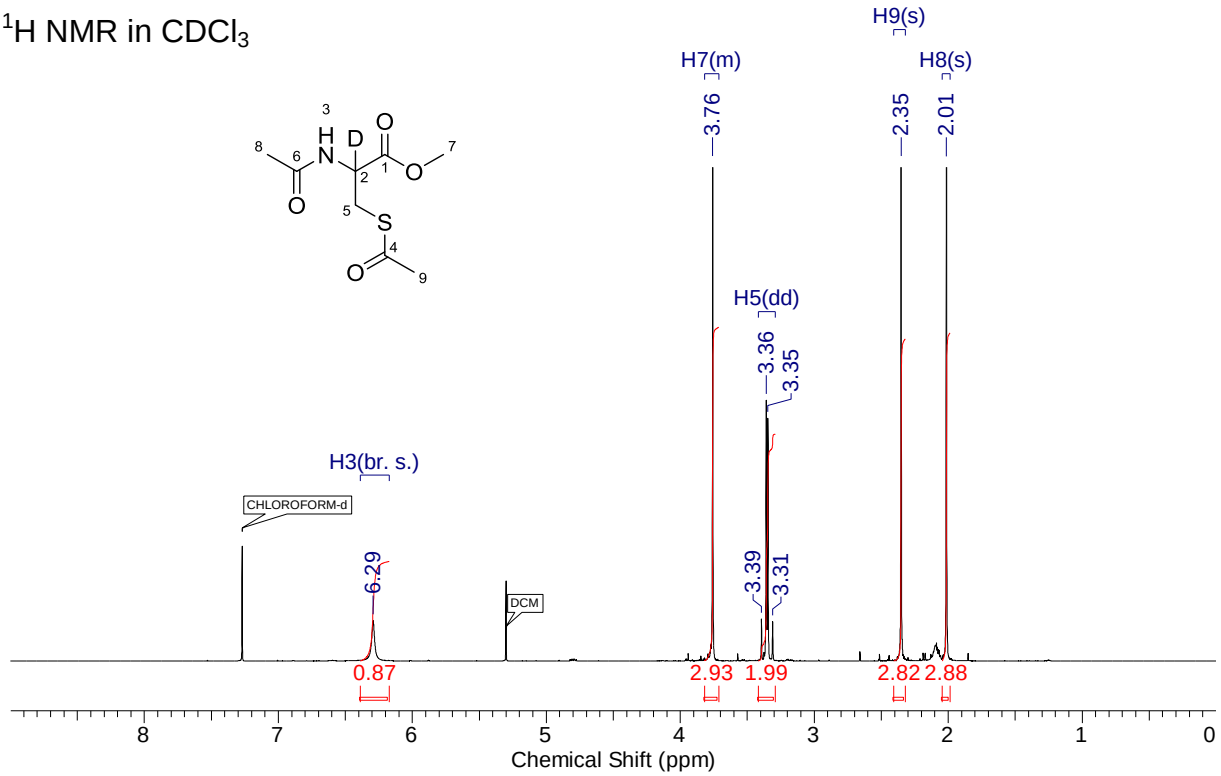
^{1}H NMR in DMSO



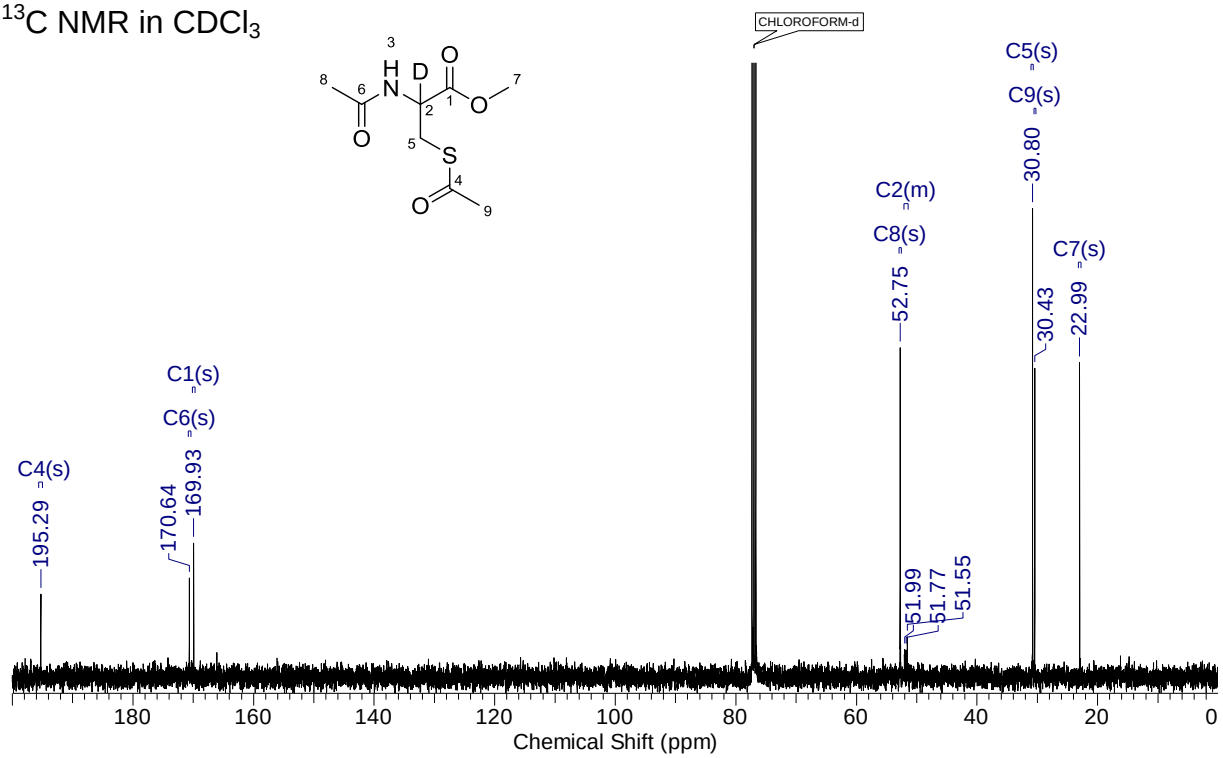
^{13}C NMR in DMSO



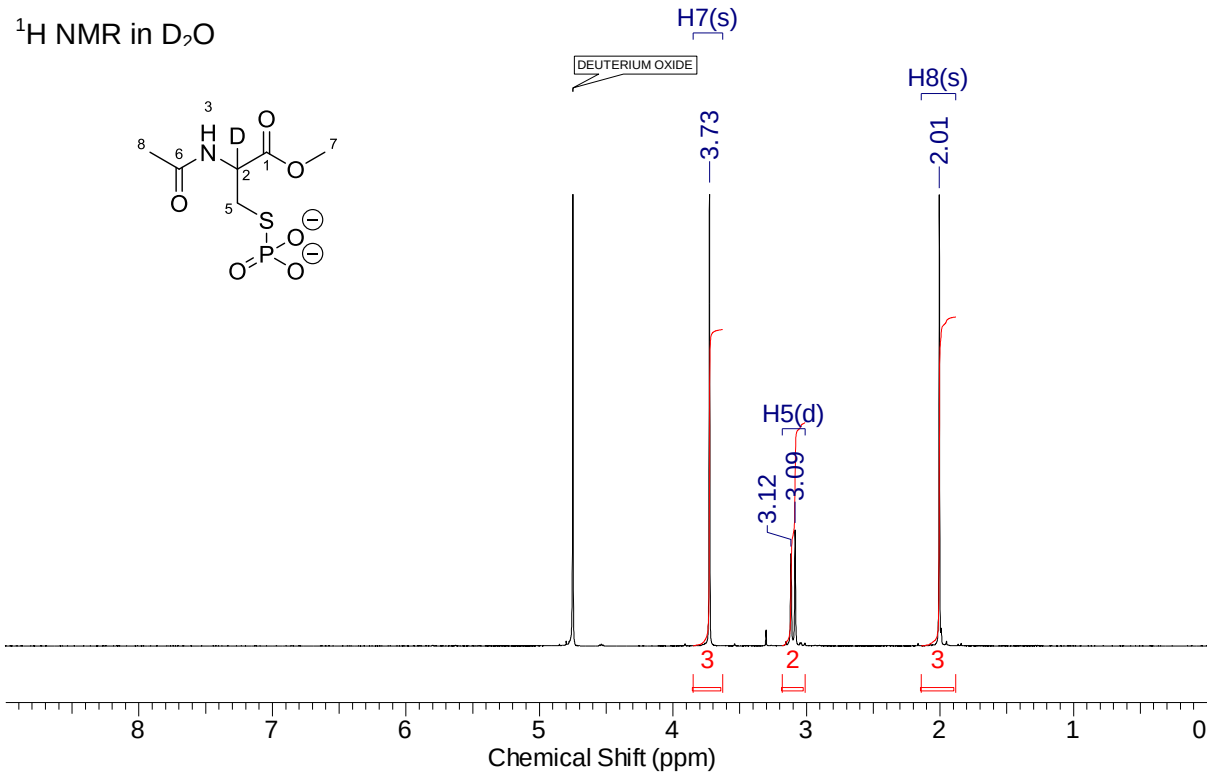
^{1}H NMR in $CDCl_3$



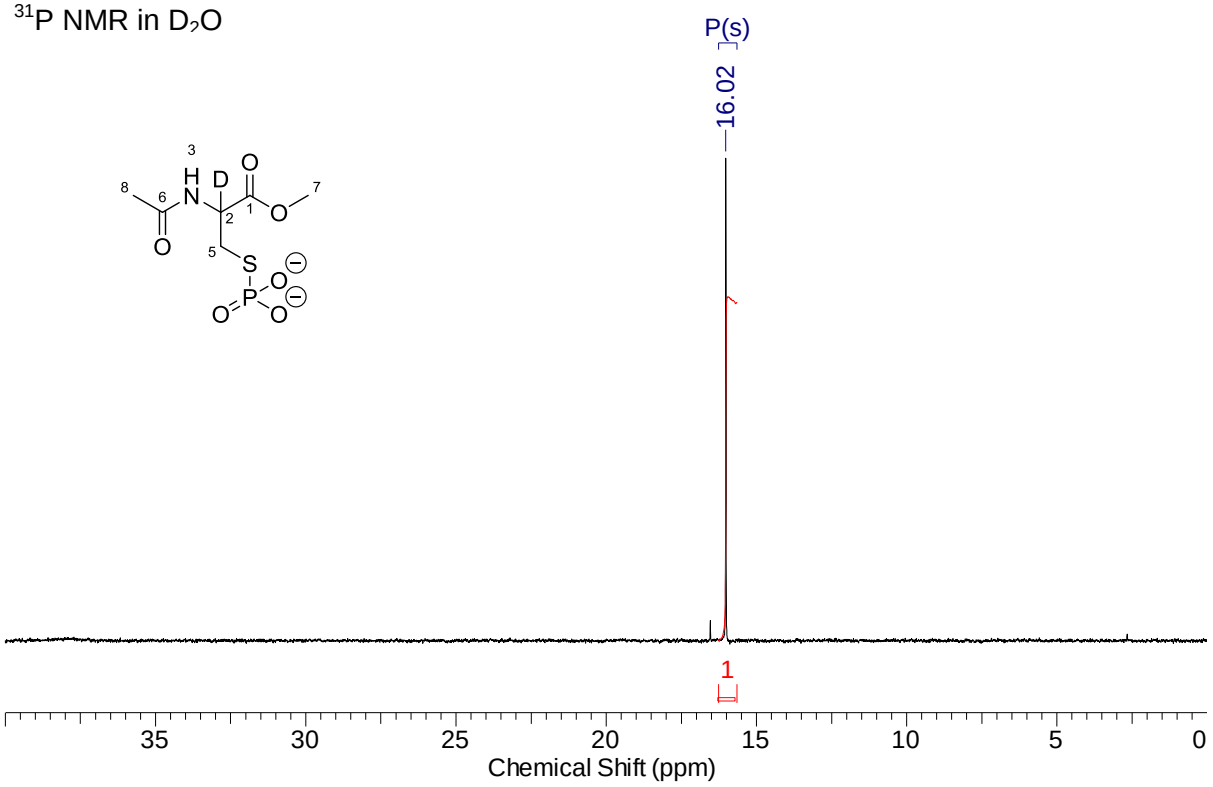
^{13}C NMR in $CDCl_3$



^{1}H NMR in D_2O

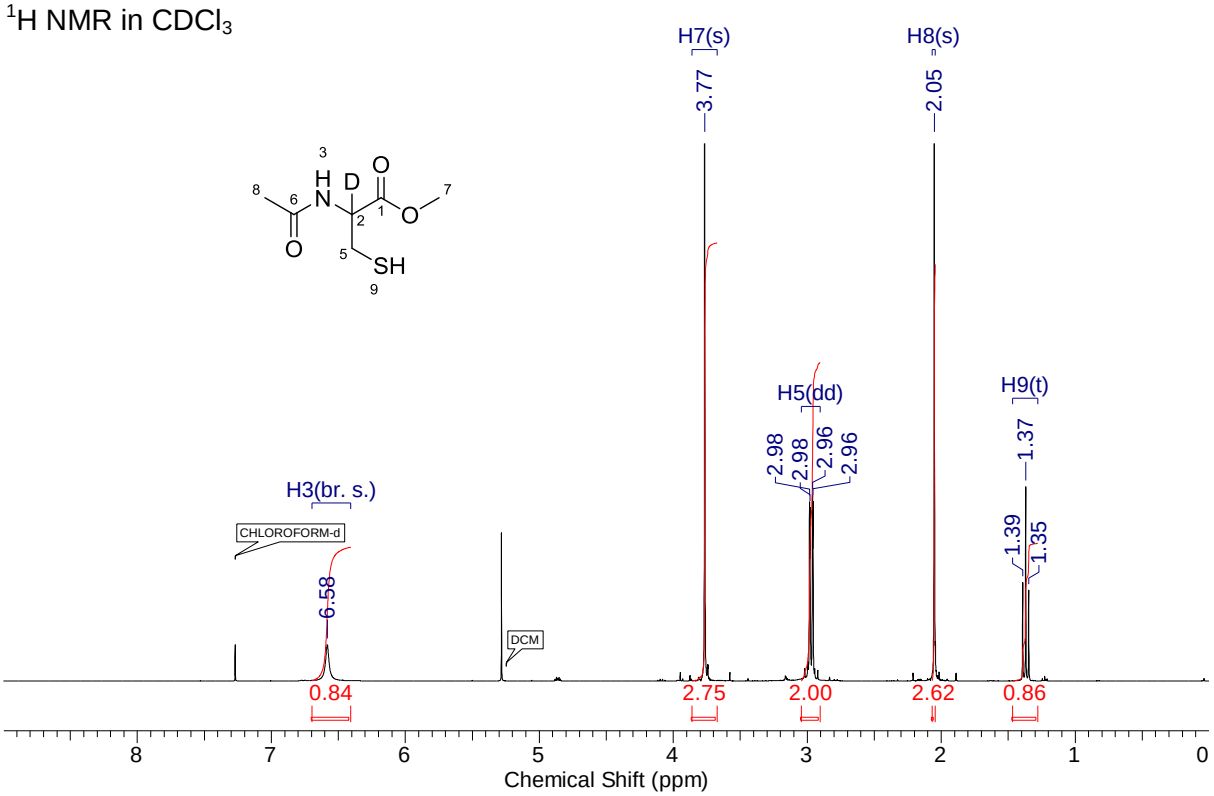


^{31}P NMR in D_2O

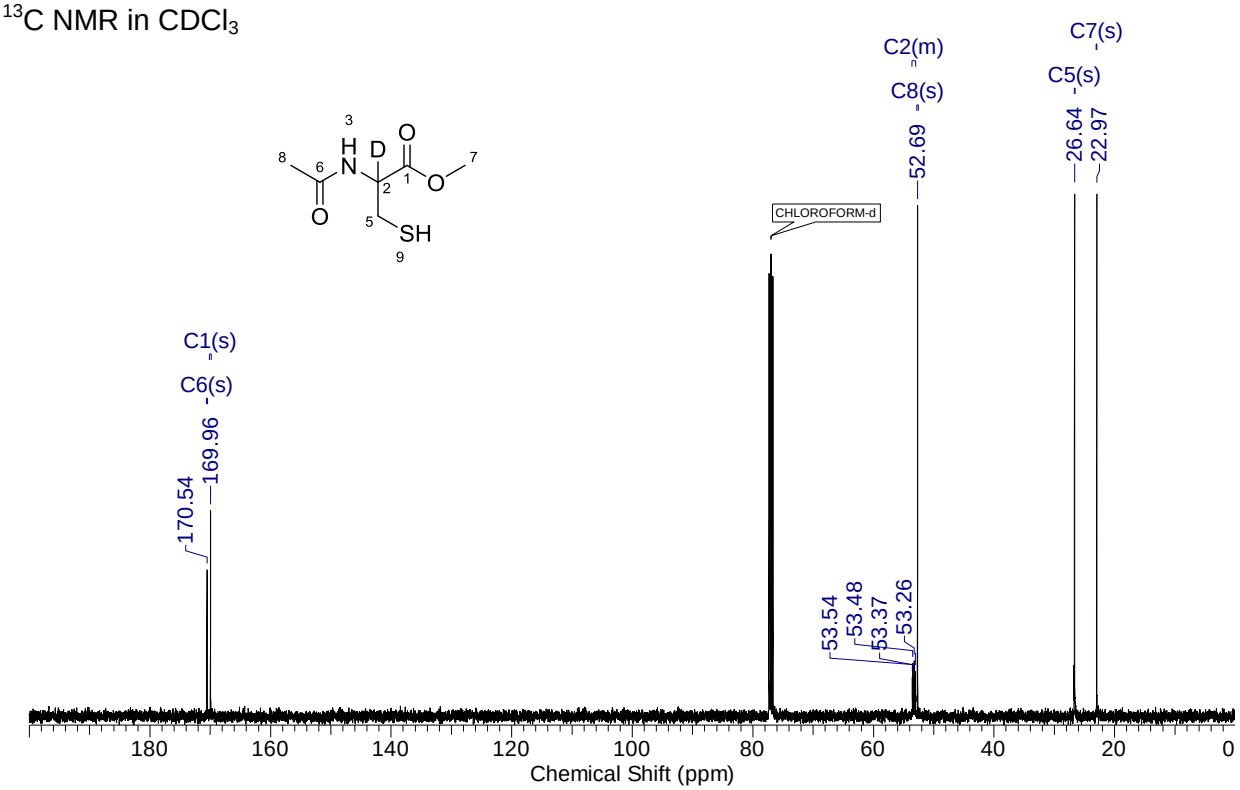


^{1}H NMR (400 MHz, DEUTERIUM OXIDE pH 2) δ ppm 2.03 (s, 3 H), 2.93 (dd, J=19.56, 14.18 Hz, 2 H), 3.74 (s, 3 H).

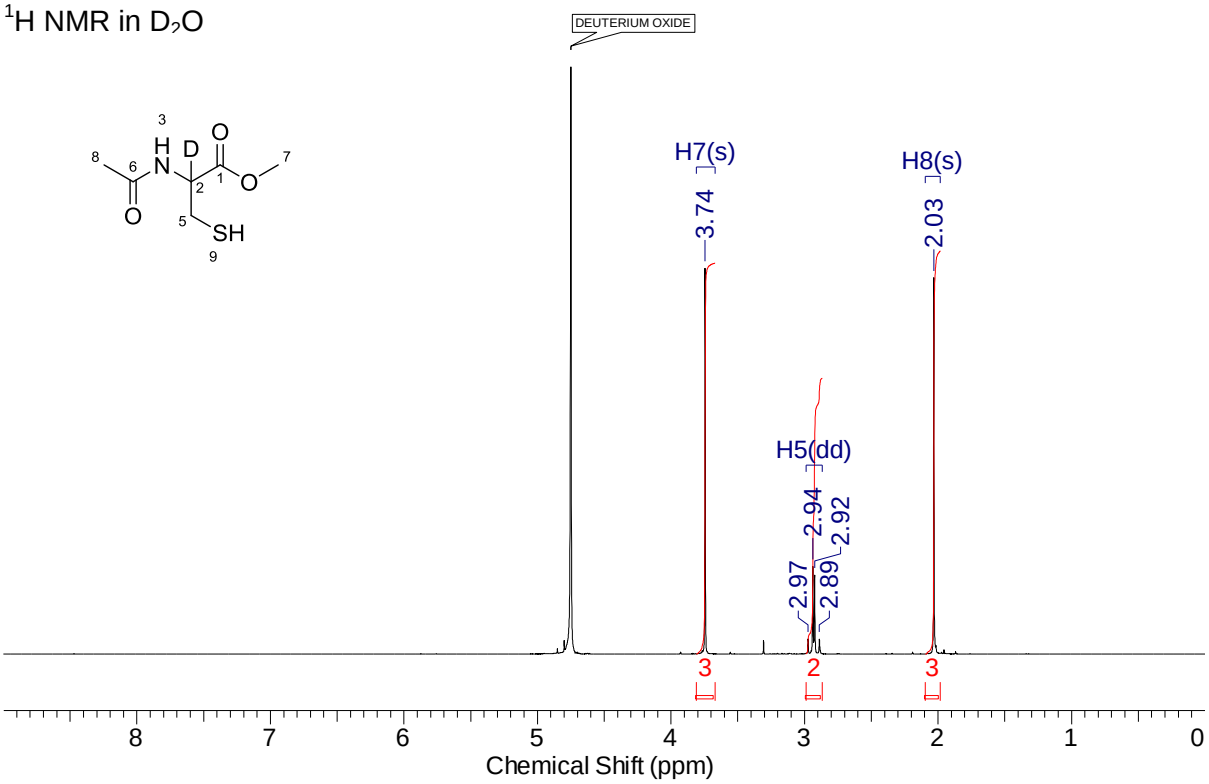
^{1}H NMR in $CDCl_3$



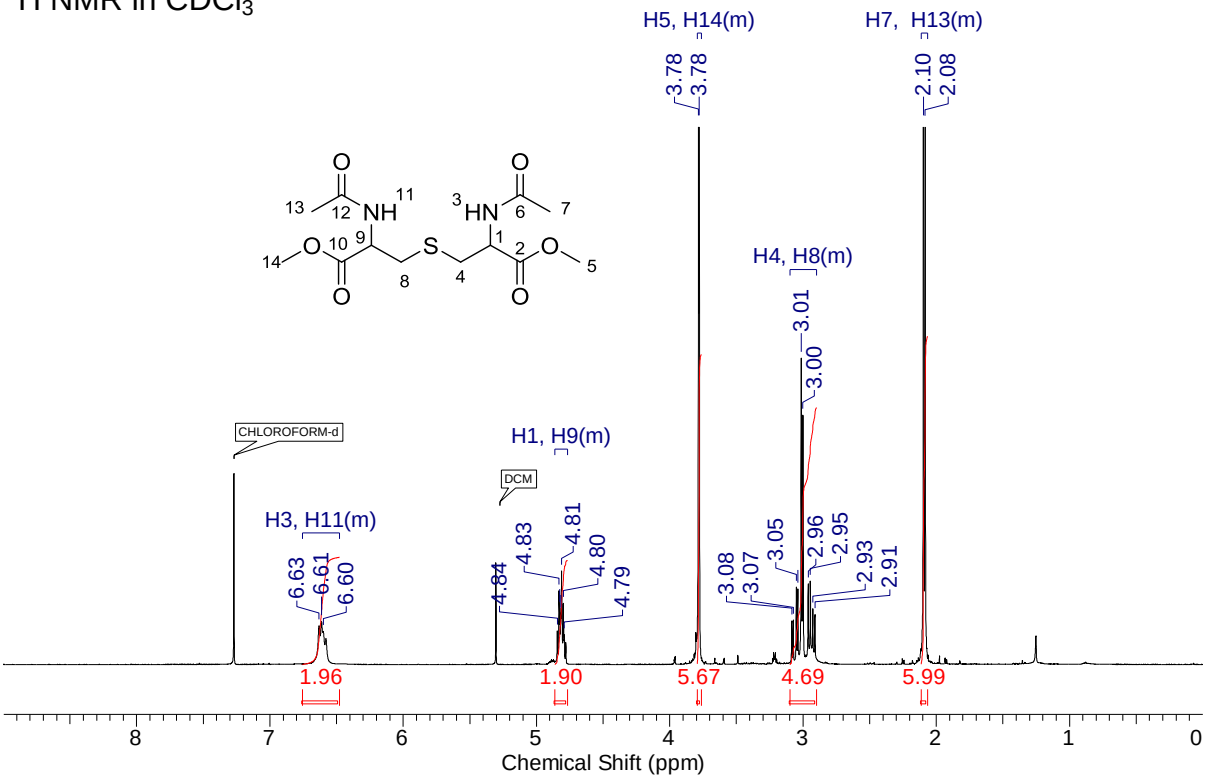
^{13}C NMR in $CDCl_3$



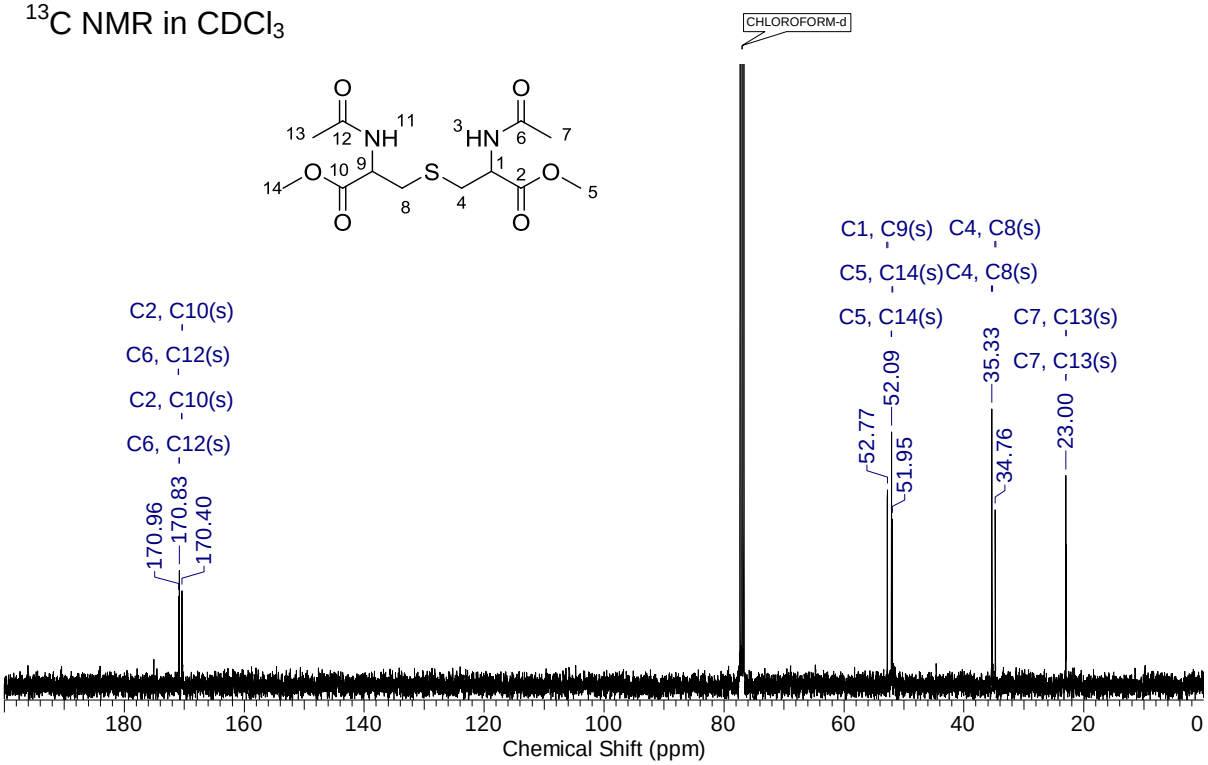
^{1}H NMR in D_2O


DEUTERIUM OXIDE
H7(s)
3.74
H8(s)
2.03
H5(dd)
2.94
2.92
2.97
2.89
3
2
3
8
7
6
5
4
3
2
1
0
Chemical Shift (ppm)

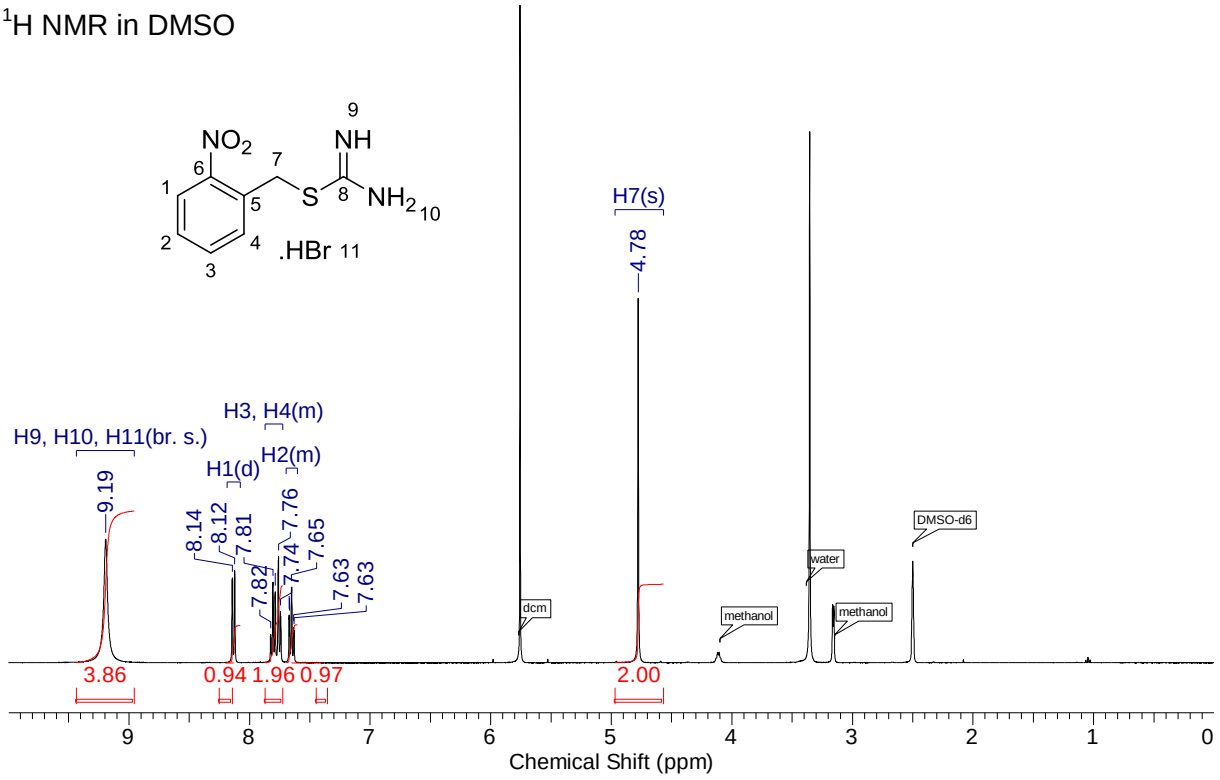
^{1}H NMR in $CDCl_3$



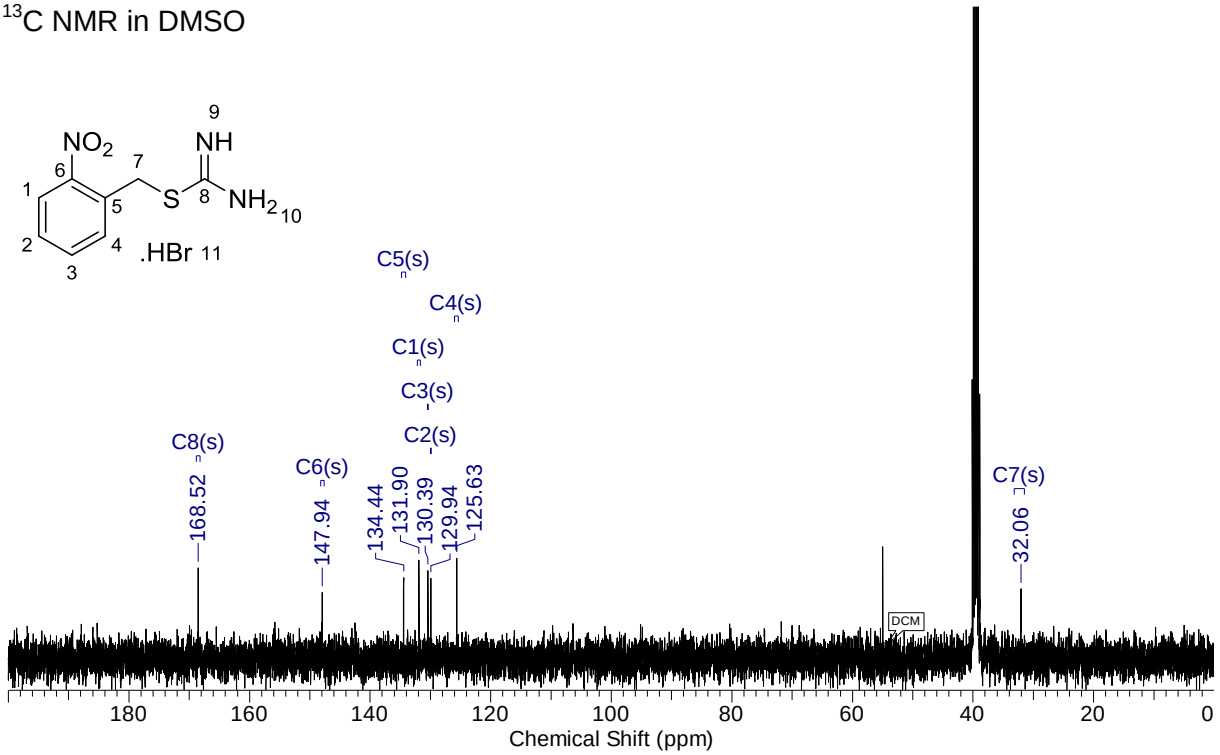
^{13}C NMR in $CDCl_3$



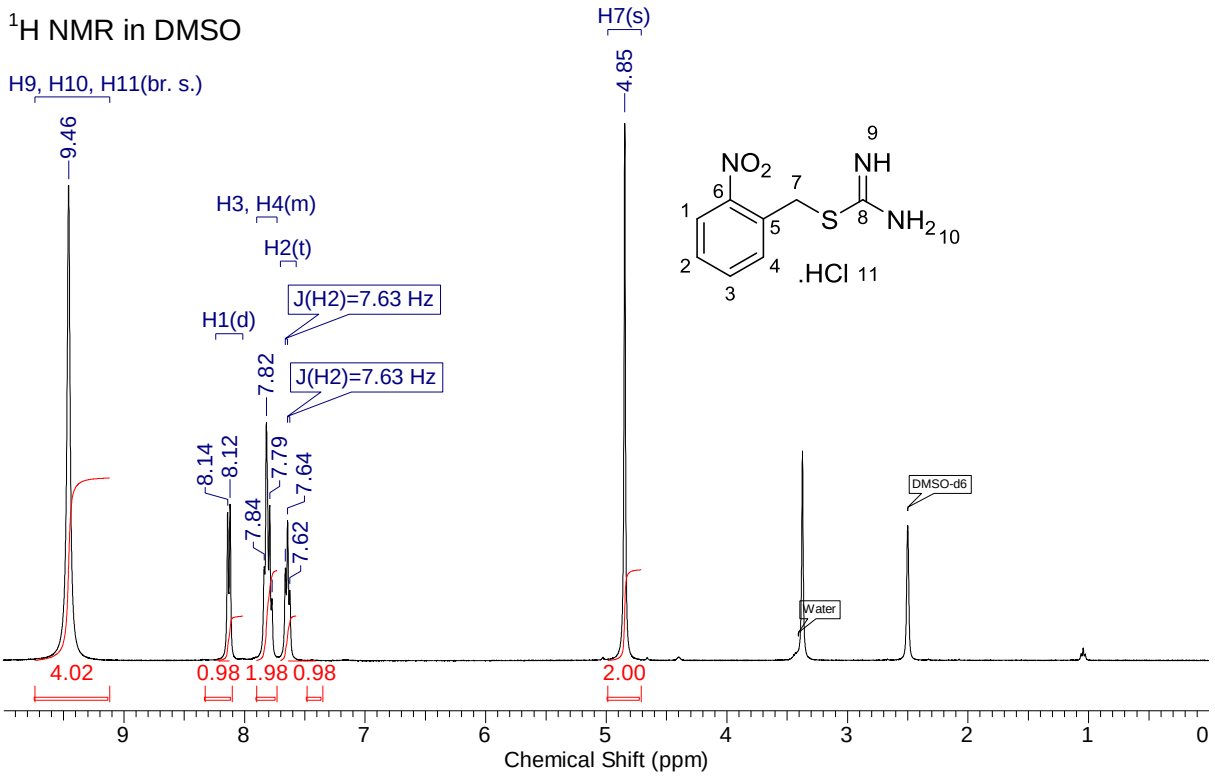
^{1}H NMR in DMSO



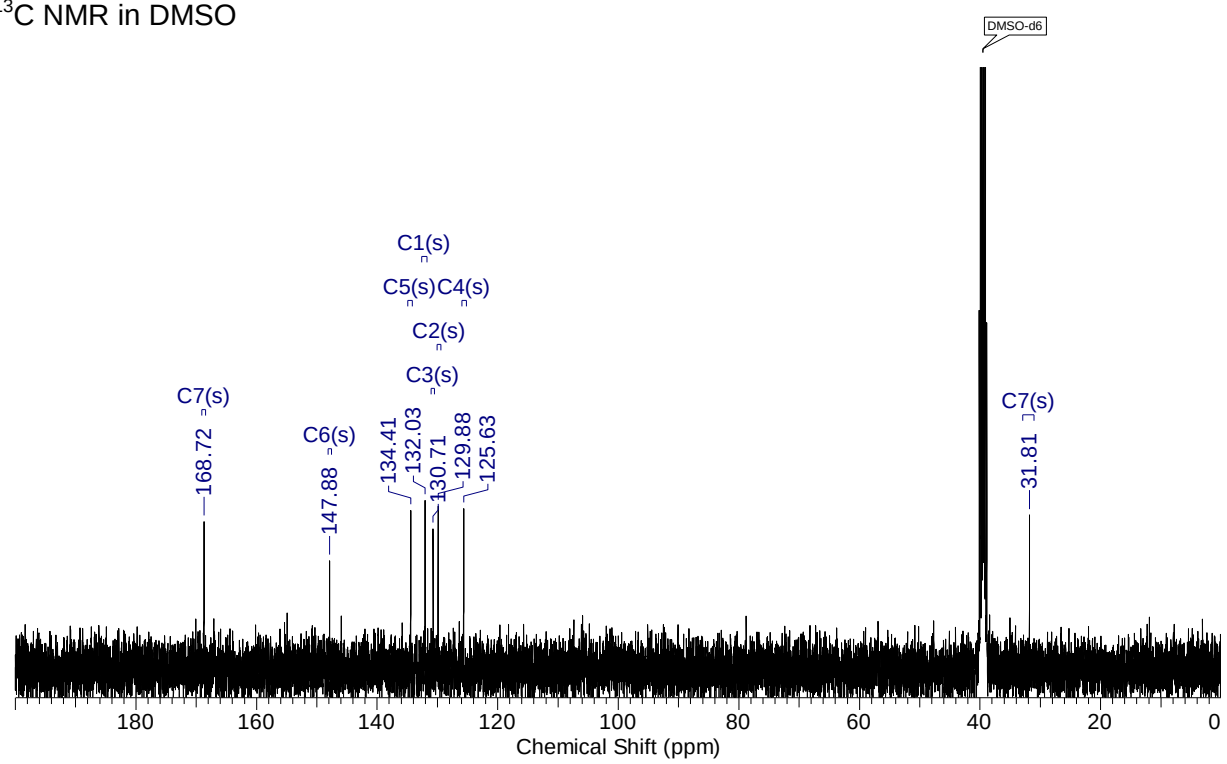
^{13}C NMR in DMSO



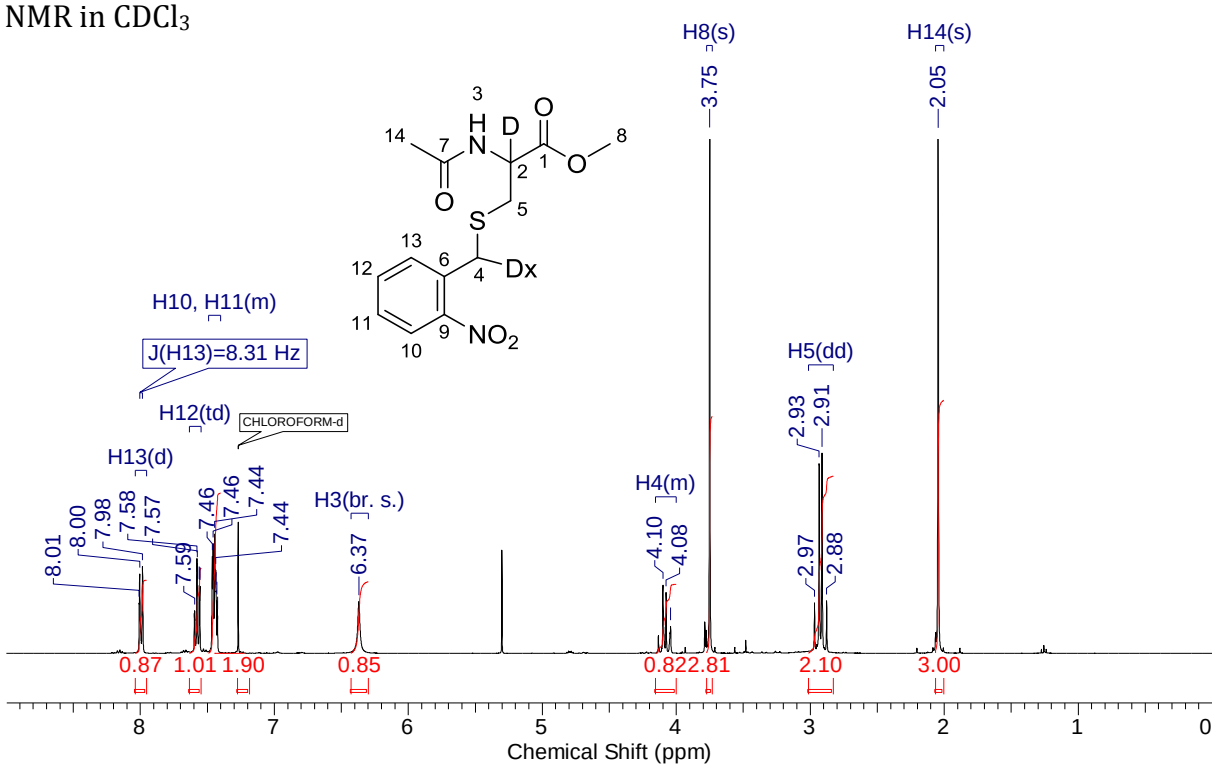
^{1}H NMR in DMSO



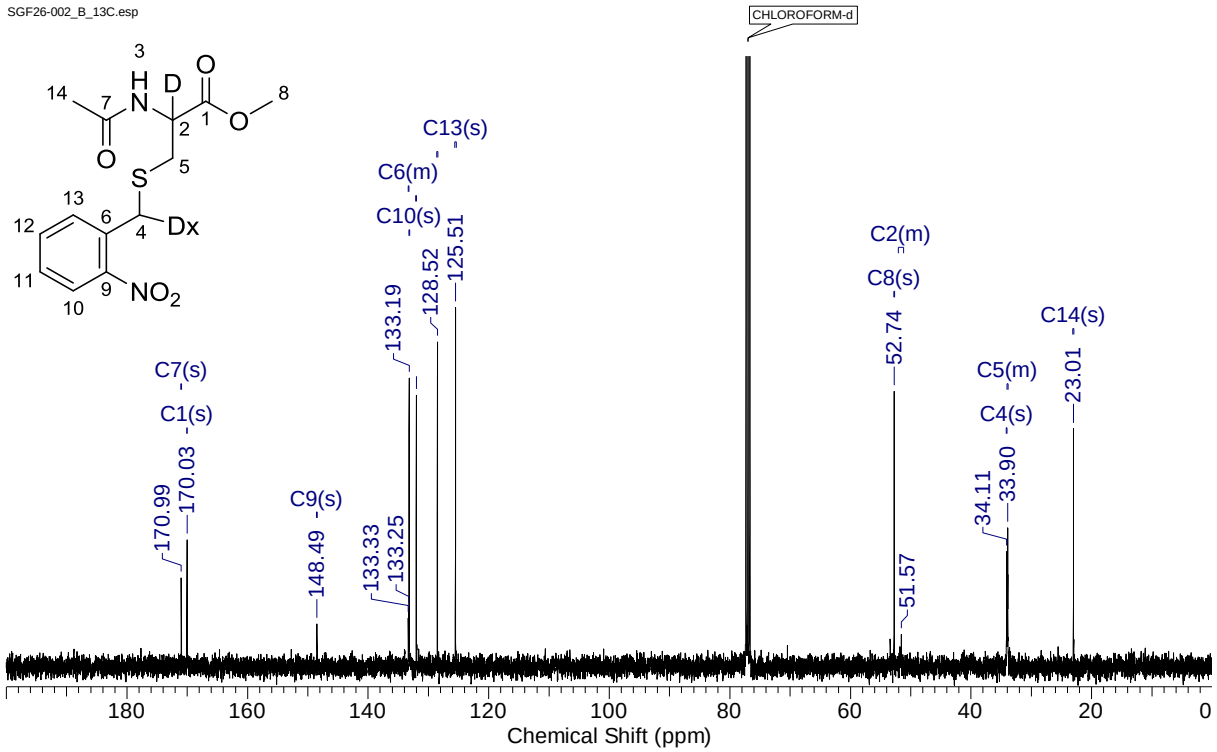
^{13}C NMR in DMSO



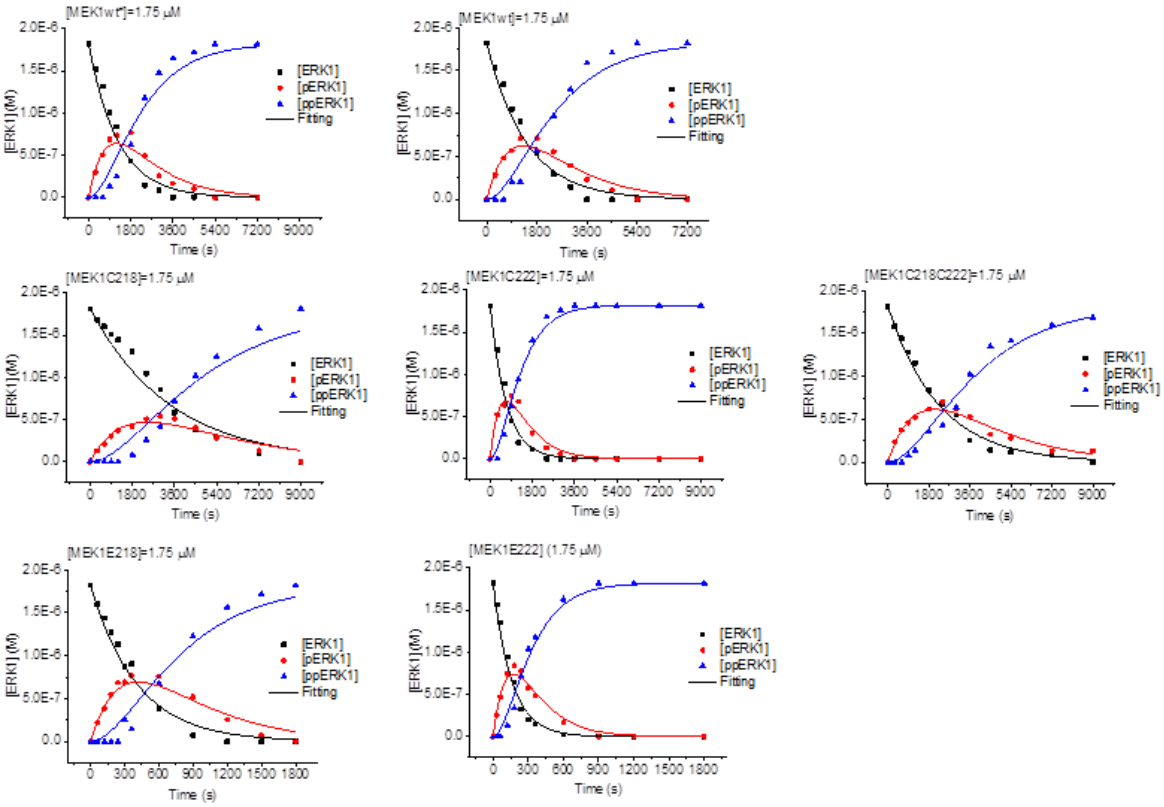
^{1}H NMR in $CDCl_3$

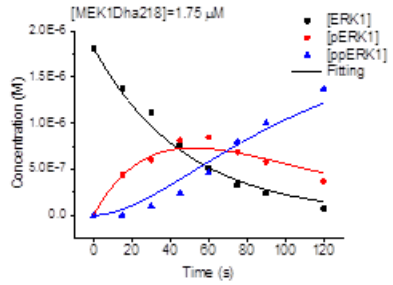


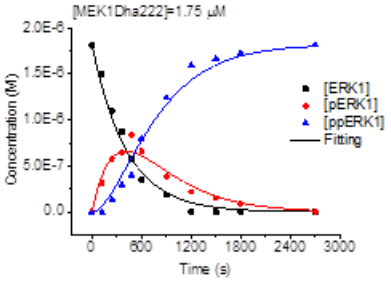
^{13}C NMR in $CDCl_3$

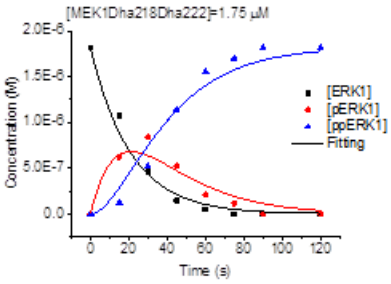


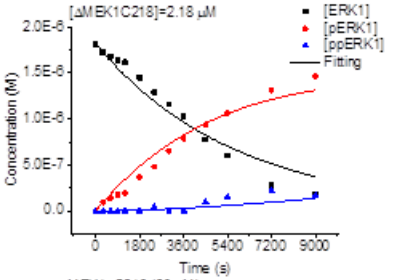
Appendix F: Progress curves of the phosphorylation of ERK1 by MEK1 constructs

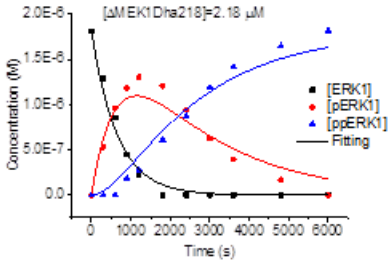


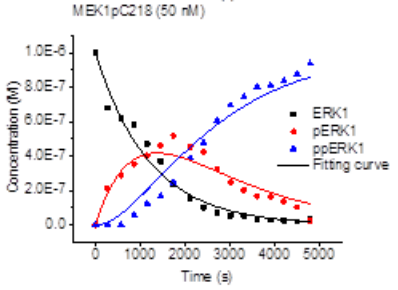

[MEK1Dha218]=1.75 µM
[ERK1]
[pERK1]
[ppERK1]
Fitting
Concentration (M)
Time (s)

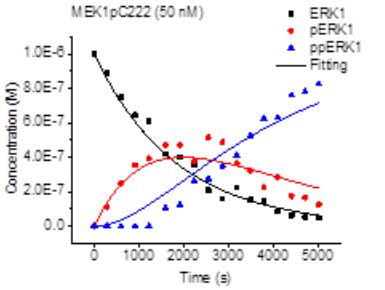

[MEK1Dha222]=1.75 µM
[ERK1]
[pERK1]
[ppERK1]
Fitting
Concentration (M)
Time (s)


[MEK1Dha218Dha222]=1.75 µM
[ERK1]
[pERK1]
[ppERK1]
Fitting
Concentration (M)
Time (s)


[ΔMEK1C218]=2.18 µM
[ERK1]
[pERK1]
[ppERK1]
Fitting
Concentration (M)
Time (s)


[ΔMEK1Dha218]=2.18 µM
[ERK1]
[pERK1]
[ppERK1]
Fitting
Concentration (M)
Time (s)


MEK1pC218 (50 nM)
ERK1
pERK1
ppERK1
Fitting curve
Concentration (M)
Time (s)


MEK1pC222 (50 nM)
ERK1
pERK1
ppERK1
Fitting
Concentration (M)
Time (s)

Appendix G: DFT calculations

Deprotonation TS:

N 2.056461 -1.559938 -0.262609 H 1.918129 -2.119001 -1.095545 C 0.971768 -1.569929 0.691708 H 1.3133 -1.997974 1.641833 C 0.498977 -0.150035 1.003993 C -0.137638 -2.48075 0.129947 H -0.257329 -0.123493 1.788532 H 1.357586 0.458337 1.287411 O 0.110457 -3.230344 -0.813155 N -1.326068 -2.397879 0.724558 C 3.28592 -1.08494 0.05528 O 3.51335 -0.561098 1.139448 C 4.343663 -1.257011 -1.006089 H 5.109485 -0.495375 -0.872489 H 4.806508 -2.239126 -0.883342 H 3.924928 -1.195221 -2.010599 C -2.435549 -3.198359 0.232316 H -2.204388 -4.262939 0.299552 H -3.305732 -2.977598 0.847659 H -2.655894 -2.945481 -0.806524 H -1.5541 -1.694022 1.452116 C -1.943433 0.991309 0.005211 C 0.331325 2.32514 -0.406946 C -1.991045 2.502512 0.301465 C -0.572526 2.958659 0.63921 H 0.124513 2.721106 -1.404689 H -2.359129 3.057943 -0.566229 H -2.663022 2.696178 1.140741 H -0.473801 4.044495 0.611544 H -0.270793 2.617211 1.632277 S -0.249515 0.579908 -0.504994 C 1.831167 2.394493 -0.116261 O 2.238575 2.830223 0.944063 N 2.626632 1.913849 -1.089427 H 3.627762 1.965368 -0.966089 H 2.271537 1.646194 -1.994912 C -2.900225 0.431471 -1.001511 O -2.614268 -0.490447 -1.757656 N -4.136622 0.980521 -0.966212 H -4.866219 0.558471 -1.520261 H -4.39981 1.670619 -0.281316 O -2.398037 -0.40431 2.197573 H -2.324572 0.090554 3.01795 H -2.164571 0.361804 1.065988

Intramolecular deprot. TS:

N -0.055137 2.27985 0.504822 H 0.522366 2.770401 1.176509 C 0.68268 1.218677 -0.149875 H 0.708662 0.152156 0.69577 C 0.113493 0.537649 -1.292686 C 2.147938 1.548638 -0.119185 H -0.801396 0.933681 -1.713239 H 0.807599 0.190485 -2.050006 O 2.559536 2.475775 0.584456 N 2.983984 0.747777 -0.801598 C -1.266401 2.800326 0.251454 O -2.053988 2.357238 -0.595104 C -1.662235 3.978584 1.108995 H -2.533493 3.694269 1.701788 H -0.869301 4.316077 1.774369 H -1.958082 4.796778 0.451669 C

4.421558 0.927802 -0.702443 H 4.904168 0.13753 -1.274669 H 4.726323 1.896024 -1.104755 H 4.748239 0.868313 0.337481 H 2.644462 -0.008302 -1.387008 C 0.250844 -1.226663 0.990092 C -2.126919 -0.717602 0.177487 C -0.761837 -1.117341 2.136877 C -2.146987 -1.390862 1.53947 H -2.024092 0.360827 0.289957 H -0.54291 -1.839227 2.925738 H -0.742361 -0.118446 2.582436 H -2.315031 -2.462638 1.417017 H -2.951388 -0.975528 2.150012 S -0.547431 -1.307914 -0.561727 C -3.290428 -1.034665 -0.752101 O -3.74237 -2.163925 -0.854916 N -3.737705 0.034521 -1.432008 H -4.50909 -0.075259 -2.073707 H -3.336562 0.954422 -1.280171 C 1.437392 -2.080899 1.067573 O 1.995888 -2.55418 0.063866 N 1.941311 -2.317953 2.301717 H 2.838369 -2.774995 2.362474 H 1.607422 -1.827937 3.115365 O 2.242621 -1.59181 -2.454117 H 2.160734 -2.115384 -1.633854 H 2.939002 -2.002638 -2.973495

E2-elimination TS:

N -2.169316 0.71011 0.039708 H -2.297952 0.902922 1.025502 C -1.697156 -0.611391 -0.319471 H -2.364069 -1.310429 -1.244816 C -0.317585 -0.58401 -0.925469 C -1.739982 -1.488929 0.904122 H -0.053368 -1.530388 -1.400509 H -0.181203 0.252819 -1.608239 O -1.726867 -1.02523 2.050914 N -1.736772 -2.805599 0.639899 C -2.577721 1.614247 -0.874643 O -2.474002 1.424707 -2.08485 C -3.175922 2.892215 -0.333403 H -4.220105 2.937829 -0.649 H -3.12516 2.966008 0.752011 H -2.656551 3.740017 -0.782508 C -1.714963 -3.790157 1.705391 H -1.582672 -4.777427 1.26752 H -0.883872 -3.586697 2.381705 H -2.641824 -3.776711 2.284747 H -2.011595 -3.062685 -0.309387 C 2.556865 0.081559 -0.505794 C 0.887766 1.132368 1.254912 C 3.231027 1.128856 0.411699 H 2.222386 0.548288 -1.434806 C 2.35834 1.340876 1.662805 H 0.234166 0.951257 2.10867 H 4.223836 0.796871 0.71413 H 3.334253 2.059814 -0.143617 H 2.611098 0.610539 2.432935 H 2.494148 2.33591 2.08718 S 1.031046 -0.449701 0.365158 C 0.413052 2.301996 0.395057 O 0.646133 2.362243 -0.801161 N -0.190067 3.277817 1.093561 H -0.437752 4.136445 0.623671 H -0.367058 3.206619 2.083404 C 3.321522 -1.185603 -0.878979 O 2.745339 -2.260802 -0.924846 N 4.614533 -1.010309 -1.19015 H 5.14318 -

1.798795 -1.533714 H 5.050353 -0.101808 -1.207649 O -2.717998 -2.282771 -2.014073 H -3.667928 -2.246648 -2.156125

Appendix H: Deuterium incorporation and optimisation

Deuterium incorporation estimation using LCMS on intact protein

Aliquot of 0.5 mg/mL aliquots of the desired protein were prepared in water. The aliquots were analysed using Xevo G2-S Q-Tof. The data was process using MaxLynx v4.1. Deconvolution was set up: output mass, ranges 10000:20000, resolution 0.04 Da/channel; damage model, uniform Gaussian width at half height 0.5 Da, minimum intensity ratios left 33% and right 33% and completion options, iterate to convergence. 2 independent runs were performed on each sample.

Percentage deuterium incorporation (%D) on intact protein was determined as follow:

$$M_{obs.}=M_{deut.}\times\%D+M_{non\text{-}deut.}\times(1-\%D)$$

with $M_{obs.}$= Mass observed, $M_{deut.}$= Mass of deuterated protein, %D= percentage of deuterated protein or percentage deuterium incorporation, $M_{non\text{-}deut.}$= mass of non-deuterated protein.

This equation was equivalent to:

$$\%D=(M_{obs.}-M_{non\text{-}deut.})/(M_{deut.}-M_{non\text{-}deut.})$$

As $M_{deut.}-M_{non\text{-}deut.}=1$, the percentage incorporation could be estimated as $\%D=M_{obs.}-M_{non\text{-}deut.}$.

Table S1. Determination of deuterium incorporation in H3C10 and H3C26 under different conditions

Protein	Conditions[a]	Calculated Mass (Da)[b]	%D
H3C10	1	15254.6407±0.0566	0
H3dC10	2	15254.9409±0.0283	30±6
H3dC10	3	15255.1812±0.0284	54±6
H3dC10	4	15255.5405±0.1416	90±15
H3dC10	5	15255.5405±0.0284	90±6
H3C26	1	15185.5605[c]	0
H3dC26	6	15186.1611[c]	60
H3dC26	5	15186.5205[c]	90

[a] See table S2; [b] The duplicated mass values for each sample were averaged and the error was calculated as the standard variation; [c] the mass was measured once.

Table S2. Conditions used for site-selective deuteration of cysteine residues on protein

Conditions	Protein	Na_3SPO_3	Buffer
1	lyophilised	Hydrate	50 mM Tris, 5 M GdnCl, pH8
2	lyophilised	Hydrate	50 mM Tris, 5 M GdnCl, pH 8 (70% deuterated)[b]
3	lyophilised	Hydrate	50 mM Tris, 5 M GdnCl, pH 8 deuterated[a]
4	lyophilised	Anhydrous[a]	50 mM Tris, 5 M GdnCl, pH 8 deuterated[a]
5	lyophilised in D_2O[a]	Anhydrous[a]	50 mM Tris, 5 M GdnCl, pH 8 deuterated[a]
6	lyophilised in D_2O[a]	Hydrate	50 mM Tris, 5 M GdnCl, pH 8 deuterated[a]

[a] See page 11; b. Buffer components directly added to D_2O

Examples of deconvoluted spectra for H3dC10

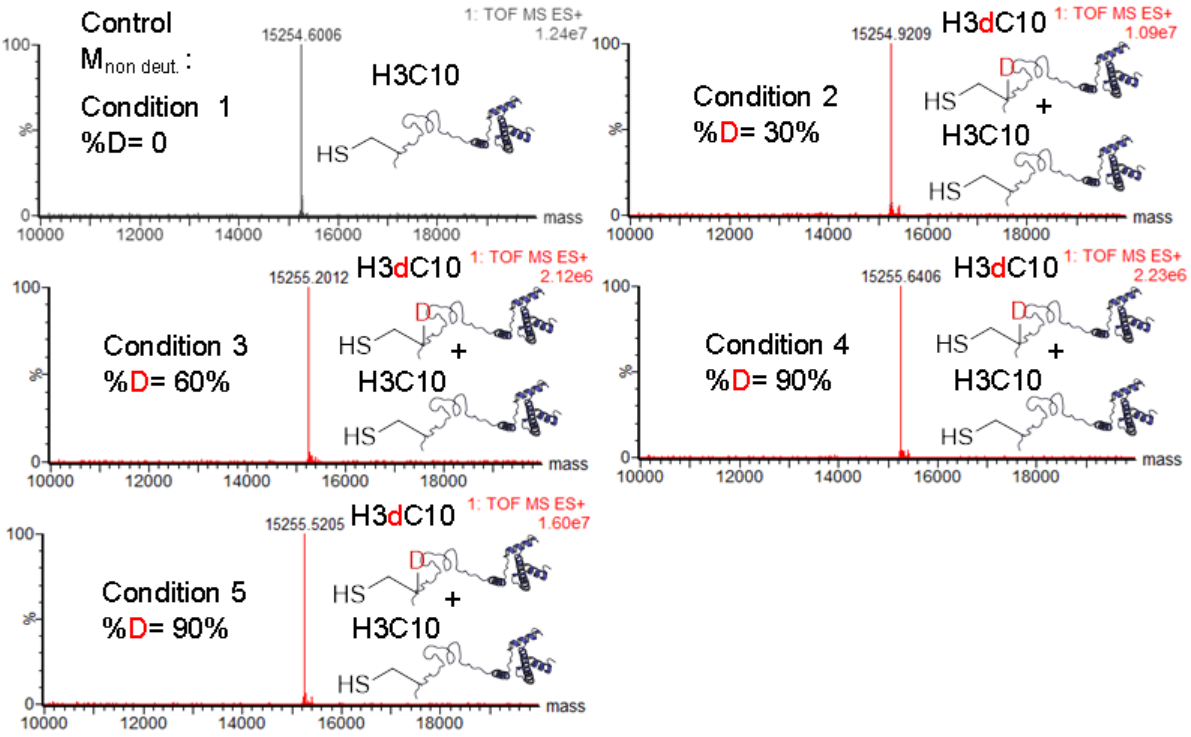


Figure S26

Examples of deconvoluted spectra for H3dC26

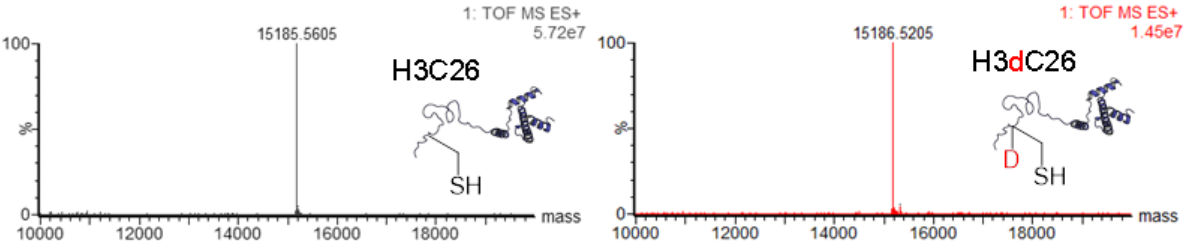


Figure S27

Control: isotopic pattern determination

The deconvolution was performed with a uniform Gaussian at half height of 0.04 Da to allow the observation of the isotopic pattern of the desired mass, which confirmed that the variation in mass was due to a deuterium addition. A typical spectra is showed below:

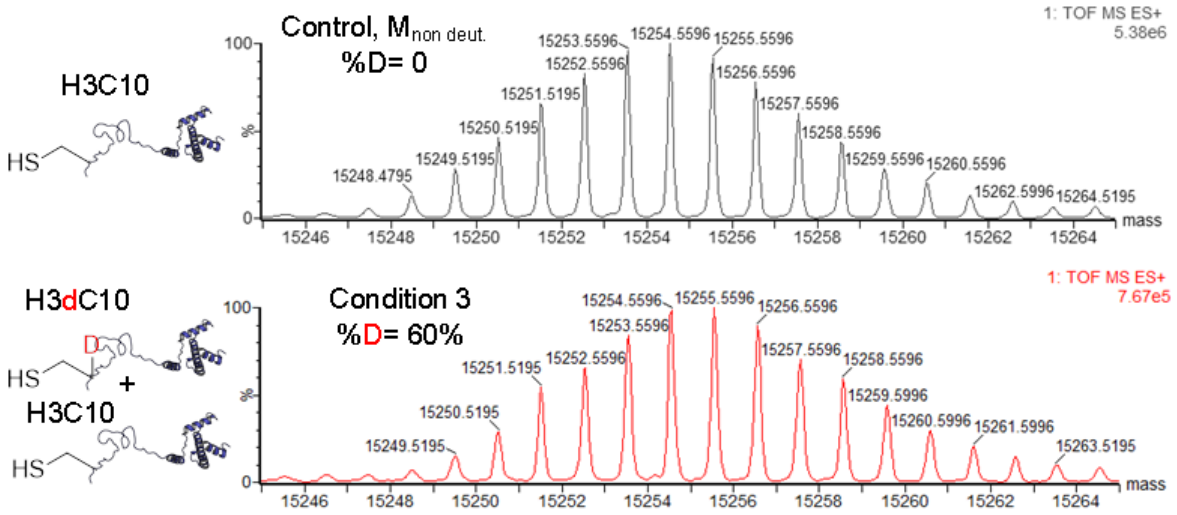


Figure S28

Validation of deuterium incorporation levels and conformation of location of deuterium by LCMS-MS of tryptic peptide

Typical protocol

An aliquot of H3 (1 mg/mL, 25 µL) was diluted with a 8 M aqueous solution of urea (75 µL). The sample was incubated at room temperature for 10 minutes with occasional vortexing. The protein was reduced with DTT (200mM solution in water) at 56 °C for 25 minutes and alkylated with iodoacetamide (400mM solution in water) at room temperature for 30 minutes in the dark. The sample was diluted with ammonium bicarbonate buffer and trypsin added (final enzyme:protein 1:50 (w/w) ratio). The sample was incubated overnight at 37 °C. The digested protein was diluted to a final concentration of 100 fmol/µL with MilliQ water. The sample was analysed by LCMS-MS.

Data processing as confirmation of location of deuterium

The data was processed using Mascot on own database containing H3 variants with variable modifications: custom-made deuterated carbamidomethyl on cysteine residue; peptide mass tolerance: ±20 ppm; fragment mass tolerance: ±0.5 Da; max missed cleavages: 3; Instrument type: default.

H3dC10

Protein sequence coverage: 68%; Matched peptides shown in **bold red**.

```
  1 MARTKQTARK CTGGKAPRKQ LATKAARKSA PATGGVKKPH RYRPGTVALR

 51 EIRRYQKSTE LLIRKLPFQR LVREIAQDFK TDLRFQSSAV MALQEASEAY

101 LVALFEDTNL AAIHAKRVTI MPKDIQLARR IRGERA
```

Start–End	Observed m/z	Mr(expt)	Mr(calc)	ppm	Score	Expect.	Peptide
6–18	358.6971	1430.7595	1430.7635	-2.80	52	5.6e-06	QTARKdC$_{CAM}$TGGKAPR

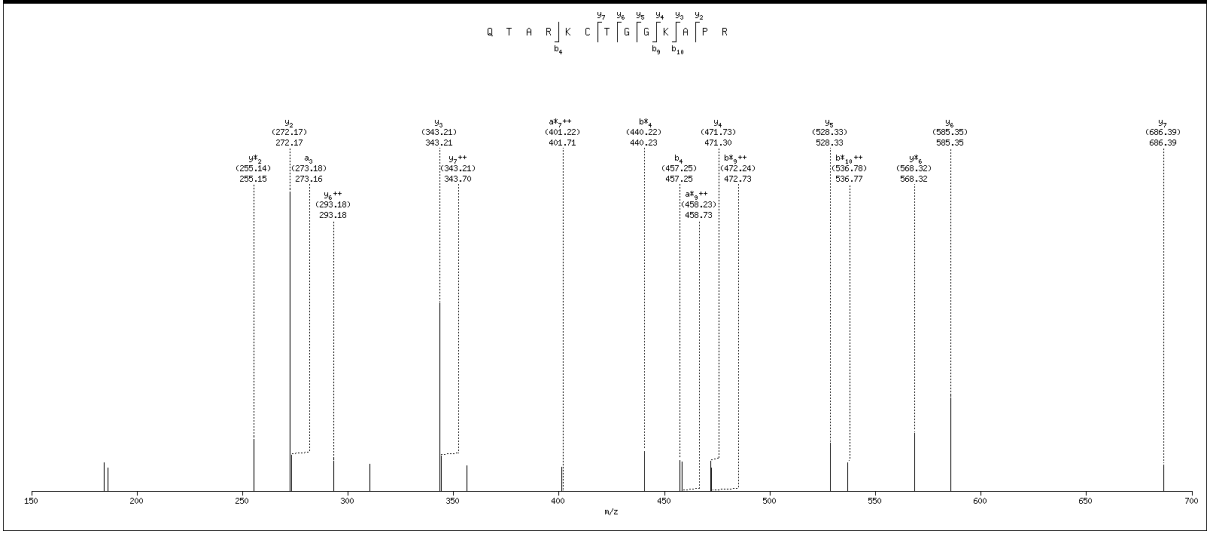


Figure S29

For H3dC26

Protein sequence coverage: 90%; Matched peptides shown in **bold red**.

```
  1  MARTKQTARK STGGKAPRKQ LATKAACKSA PATGGVKKPH RYRPGTVALR

 51  EIRRYQKSTE LLIRKLPFQR LVREIAQDFK TDLRFQSSAV MALQEASEAY

101  LVALFEDTNL AAIHAKRVTI MPKDIQLARR IRGERA
```

Start–End	Observed m/z	Mr(expt)	Mr(calc)	ppm	Score	Expect.	Peptide
25–37	609.8195	1217.6245	1217.6297	-4.26	57	2.2e-06	AAdC$_{CAM}$KSAPATGGVK

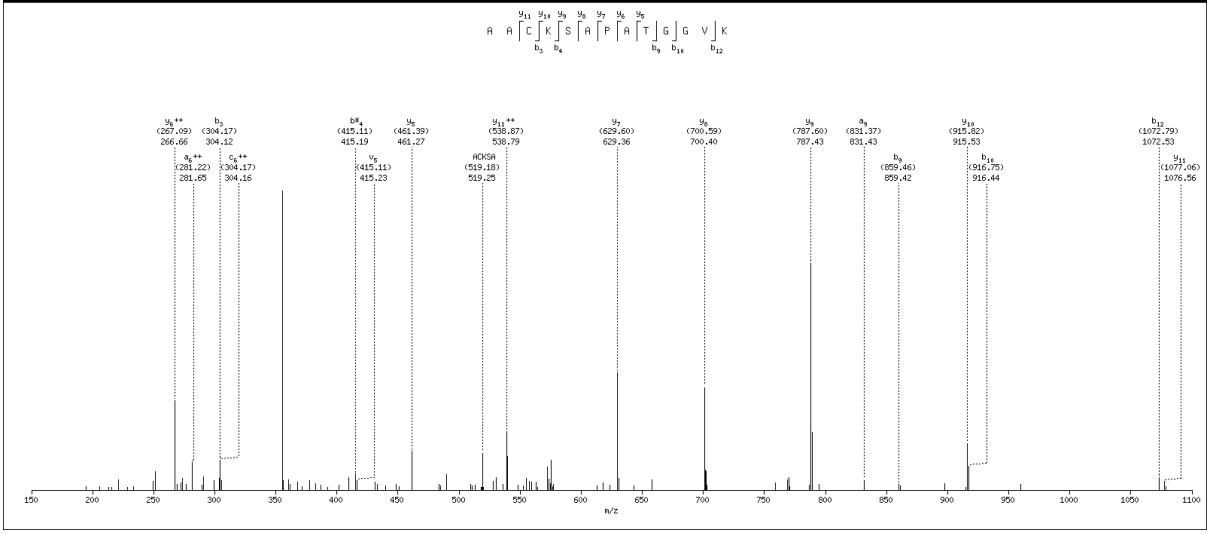


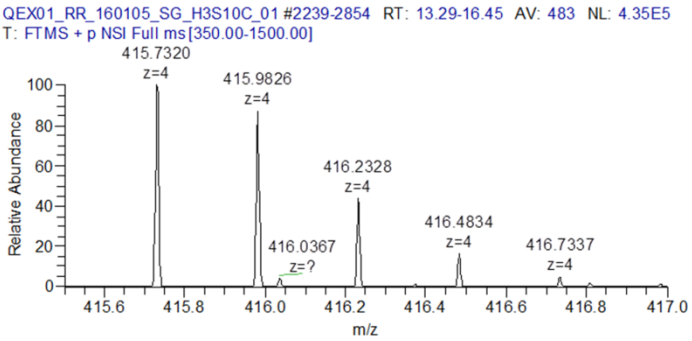
Figure S30

Deuterium incorporation estimation using LCMS on tryptic peptide

The data generated by LCMS-MS of tryptic digest were processed using Thermo Xcalibur Qual browser. MS1 was processed using chromatogram ranges: base peak, scan filter full ms, range: m/z value of peptide with automatic processing (mass tolerance 10 ppm and 4 decimals 4). The m/z values used were determined in the mascot analysis. The full desired peak was integrated to afford an average of the mass of the desired peptide. 3 different peptides were used as triplicates. For example, MS1 analysis of the peptide 4-18 TKQTARKC_{CAM}TGGKAPR, m/z 415.7320 (Figure S31) from the tryptic digests of H3dC10 batches prepared using conditions 1, 2, 3, 4 and 5 (see table S2) showed different mixture of TKQTARKC_{CAM}TGGKAPR (Figure S32) and TKQTARKdC_{CAM}TGGKAPR. The comparison between isotopic patterns of non-deuterated peptide and the sample peptide allowed the calculation of the percentage deuteration incorporation (Figure S33).

H3C10 peptide 4-18 TKQTARKC_{CAM}TGGKAPR $m/z_{calc.}$=415.7322, $m/z_{obs.}$=415.7320, z=4

Condition 1 MS1 spectra (Isotopic pattern)



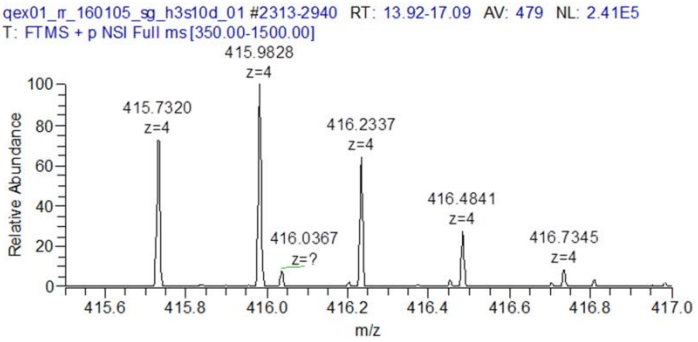
Relative intensities

m/z	Intensity	Relative (%)	Charge
415.7320	444257.5	100	4
415.9826	379098.8	85.33	4
416.2328	192332.2	43.29	4
416.4834	70806.5	15.94	4
416.7337	20483.4	4.61	4

Figure S31

H3dC10 peptide 4-18 TKQTARKdC_{CAM}TGGKAPR $m/z_{calc.}$=415.9838, z=4

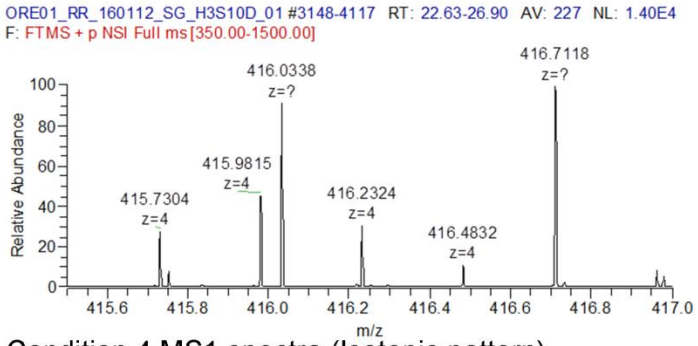
Condition 2 MS1 spectra (Isotopic pattern)



Relative intensities

m/z	Intensity	Relative (%)	Charge
415.7320	181181.2	74.71	4
415.9828	242499.2	100	4
416.2337	155601.2	64.17	4
416.4834	66534.4	27.44	4
416.7345	21275.1	4.61	4

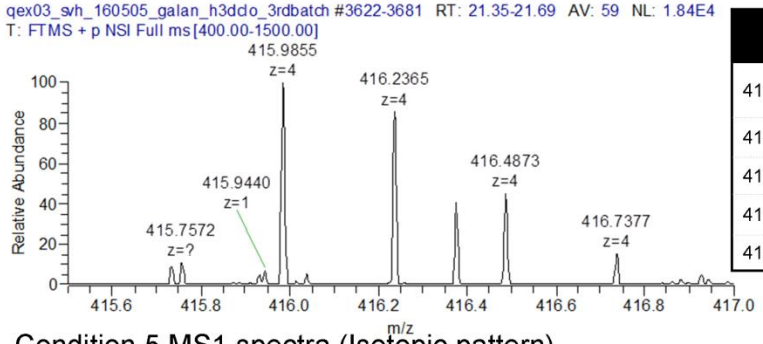
Condition 3 MS1 spectra (Isotopic pattern)



Relative intensities

m/z	Intensity	Relative (%)	Charge
415.7304	3970.6	28.03	4
415.9815	6540.2	46.18	4
416.2324	4326.6	30.55	4
416.4832	1538.5	10.86	4
416.7333	412.1	2.91	4

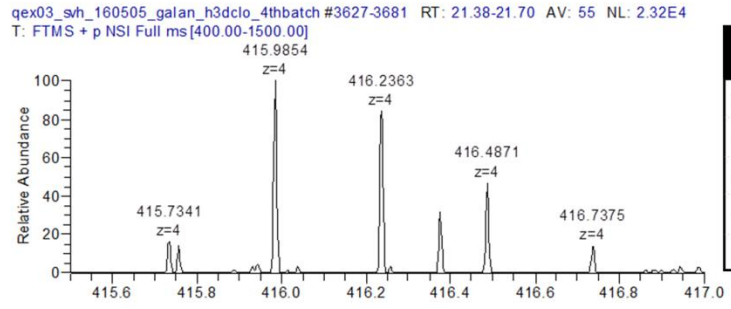
Condition 4 MS1 spectra (Isotopic pattern)



Relative intensities

m/z	Intensity	Relative (%)	Charge
415.7340	1628.2	8.61	4
415.9855	18908.1	100	4
416.2328	16045.7	84.86	4
416.4834	8440.8	44.64	4
416.7337	2930.9	15.50	4

Condition 5 MS1 spectra (Isotopic pattern)

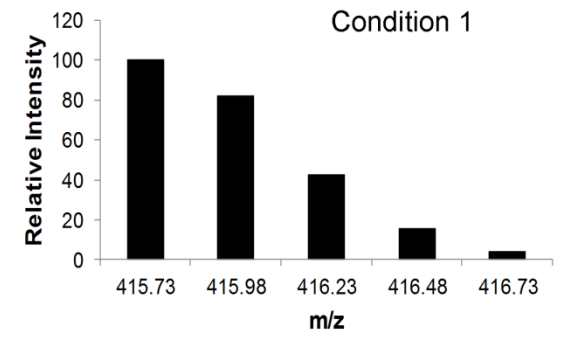


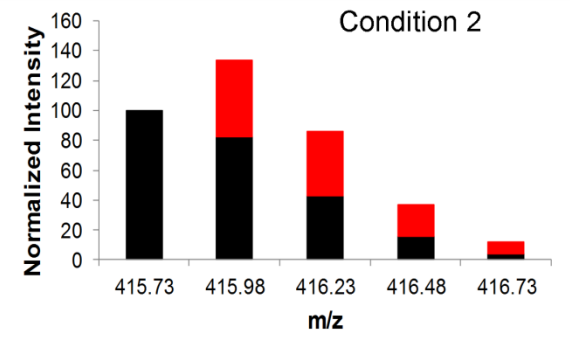
Relative intensities

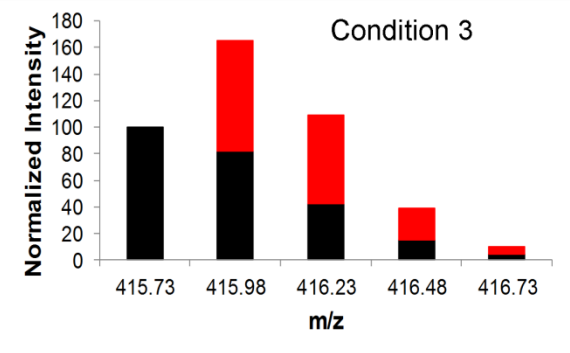
m/z	Intensity	Relative (%)	Charge
415.7341	3949.4	16.57	4
415.9854	23838.4	100	4
416.2363	19878.4	83.39	4
416.4871	11007.4	46.18	4
416.7375	3368.7	4.61	4

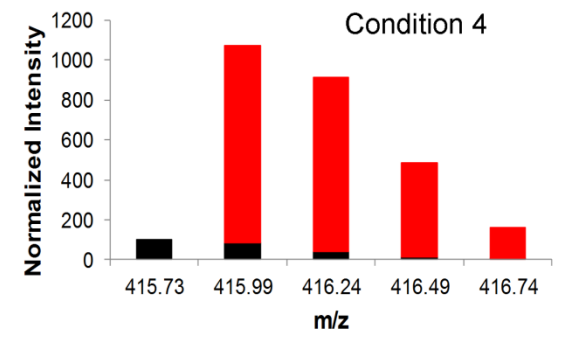
Figure S32

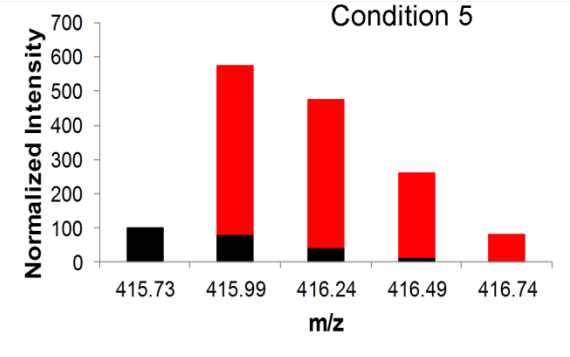
Representation of MS1 spectra showing isotopic pattern of mixtures deuterated/non deuterated tryptic peptide (4-18) depending of reaction conditions and %D estimation











Protein	Condition[a]	%D
H3C10	1	0
H3dC10	2	33±1[b]
H3dC10	3	41±5[b]
H3dC10	4	91[c]
H3dC10	5	83[c]

[a] See table S2; [b] 3 different peptides were used as triplicate, the results were the average values and the error, the standard deviation; [c] Only one peptide used.

Figure S33