



University of HUDDERSFIELD

University of Huddersfield Repository

Murphy, K.J.

An NMR and molecular modelling study on the solution conformation of Heparan Sulphate: new insights into the relationship between structure and function

Original Citation

Murphy, K.J. (2007) An NMR and molecular modelling study on the solution conformation of Heparan Sulphate: new insights into the relationship between structure and function. Doctoral thesis, University of Huddersfield.

This version is available at <http://eprints.hud.ac.uk/id/eprint/189/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>

An NMR and molecular modelling study on the solution conformation of Heparan Sulphate: new insights into the relationship between structure and function.

Kevin John Murphy

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

The University of Huddersfield.

February 2007

List of contents

INTRODUCTION

1. Introduction	Page 26
1.1 The conformation of N-glucosamine monosaccharides	Page 26
1.2 Cremer-Pople puckering parameters and NMR coupling constants	Page 29
1.3 The conformation of glucuronic acid and iduronate acid monosaccharides	Page 30
1.4 The conformation of Δ UA monosaccharides	Page 32
1.5 Protein core and HS chains- one functional unit	Page 34
2. Proteoglycans	Page 34
2.1 Syndecan HSPGs	Page 35
2.2 Syndecan ectodomains	Page 37
2.3 Syndecan transmembrane domains	Page 38
2.4 Syndecan cytoplasmic domains	Page 38
2.5 Syndecan expression	Page 40
2.6 Glypican HSPGs	Page 40
2.7 Glypican expression	Page 43
3. Extracellular matrix HSPGs	Page 43
3.1 Perlecan	Page 43
3.2 Agrin	Page 44
3.3 Collagen XVIII	Page 44
4. Limitations of core protein analyses	Page 45
5. HS chain structure is cell type specific	Page 45

6. Biosynthesis	Page 46
6.1 HS biosynthesis	Page 46
6.2 Chain initiation	Page 47
6.3 Chain elongation, HS vs CS/DS	Page 49
6.4 HS chain elongation	Page 50
6.5 Availability of PAPS	Page 51
6.6 N-deacetylation/N-sulphation	Page 51
6.7 Differences between heparin and HS biosynthesis	Page 52
6.8 NDST3 and 4	Page 53
6.9 Generation of GlcNH ₂ residues	Page 53
6.10 Epimerisation and 2-O- sulphation	Page 54
6.11 6-O- sulphation	Page 55
6.12 3-O- sulphation	Page 56
6.13 Sulfs	Page 57
6.14 Biosynthetic enzyme expression patterns	Page 58
6.15 Chain heterogeneity	Page 58
7. De-polymerisation techniques	Page 60
7.1 Enzymatic	Page 60
7.2 Chemical	Page 61
8. HS and heparin structural distinct molecules	Page 62
8.1 Domain structure of HS	Page 62
8.2 Alternating sequences and K5 lyase	Page 65
9. Ligand binding	Page 67

9.1 Protein binding sites for HS	Page 67
10. A footprinting approach	Page 69
10.1 Interferon- γ	Page 69
10.2 Interferon- γ structure	Page 70
10.3 Interferon- γ , HS binding sites	Page 72
10.4 The domain structure of Interferon- γ binding oligosaccharides	Page 72
10.5 HS binding prevents receptor binding possibly by steric hinderance	Page 72
10.6 Biological implications of HS- IFN- γ binding	Page 75
11. Other cytokines	Page 75
11.1 The domain structure of chemokine binding oligosaccharides	Page 76
11.2 The role of sulphation patterns within IFN- γ and cytokine binding oligosaccharides	Page 76
12. Endostatin	Page 77
12.1 The oligomerisation state of endostatin influences its biological activity	Page 77
12.2 HS binding sites within endostatin	Page 78
12.3 The domain structure of endostatin binding oligosaccharides	Page 80
12.4 Specific sulphation requirements for endostatin binding oligosaccharides	Page 80
13. Fibroblast growth factors (FGFs)	Page 81
13.1 The domain structure and minimal sulphation requirements of FGF binding oligosaccharides	Page 82
13.2 Different FGFs interact with different HS sulphation sequences?	Page 82

13.3 Biologically active, FGF binding, HS oligosaccharides	Page 84
13.4 Fibroblast growth factor receptors (FGFRs)	Page 84
13.5 FGF-FGFR-heparin complexes	Page 85
13.6a Different FGFs and FGFRs do not interact with different HS sulphation sequences or specific sulphation sequences? Part 1	Page 90
14. Antithrombin III and Factor Xa	Page 92
14.1 Synthetic Factor Xa inhibiting pentasaccharides	Page 94
14.2 Conformationally locked iduronate residues	Page 95
13.6b Different FGFs and FGFRs do not interact with different HS sulphation sequences or specific sulphation sequences? Part 2	Page 95
15. The balance of iduronate equilibria, a critical influence on oligosaccharide binding? a purification problem	Page 96
16. Iduronate and terminal Δ 4,5 uronic acid conformational analysis of heparin and synthetic oligosaccharides	Page 96
17. Theoretical coupling constants	Page 97
18. Heparan Sulphate, three-dimensional structure	Page 101
18.1 The torsional angles Φ and Ψ , and the exo-and endo-anomeric effect	Page 104
19. Molecular dynamic studies	Page 109
19.1 Dynamic modelling of N-glucosamine monosaccharides	Page 109
19.2 Dynamic modelling of iduronate monosaccharides	Page 110
19.3 Molecular modelling of terminal uronate monosaccharides	Page 110
20. Aims of this study	Page 111
21. Molecular dynamic partial charge assignment	Page 112

EXPERIMENTAL MATERIALS AND PROCEDURES

1. Experimental Materials	Page 115
2. Enzymatic depolymerisation of HS	Page 115
3. Gel filtration chromatography.	Page 116
3.1 Analytical gel filtration chromatography	Page 116
3.2 Preparative gel-filtration chromatography	Page 116
4. Strong Anion-exchange HPLC (SAX-HPLC).	Page 116
5. Quantification.	Page 117
6. Polyacrylamide gel electrophoresis (PAGE).	Page 118
7. Recovery of HS oligosaccharides dp6-5 and dp6-6 from polyacrylamide gels	Page 118
8. DEAE ion-exchange chromatography	Page 119
9. Disaccharide Analysis	Page 119.
10. Chemical de-2-O sulphation	Page 120
11. NMR spectroscopy	Page 120
12. Coupling constant measurement	Page 121
13. Molecular modelling	Page 122
13.1 Iduronate Coupling constant analysis	Page 122
13.2 Partial charge calculation methods- Method 1 (partial charge set 1)	Page 124
13.3 Calculation of potential energy surfaces across glycosidic linkages	Page 124
13.4 Whole molecule molecular dynamic simulations	Page 125
14. Additional computational modelling	Page 126
14.1 Partial charge calculation method - Method 2 (partial charge set 2)	Page 126

14.2 Partial charge calculation method - Method 3 (partial charge set 3) Page 127

RESULTS

1. Initial studies into the degradation of HS by the heparinase III enzyme Page 129
2. Bio-gel P10 gel filtration chromatography of HS digested to completion with heparinase III Page 133
3. SAX-HPLC chromatography of sized HS dp4-dp10 oligosaccharides Page 135
4. PAGE analysis of dp6 and dp8 SAX-HPLC peaks Page 138
5. Disaccharide analysis of selected dp6 and dp8 SAX-HPLC peaks. Page 140
6. Variation between SAX-HPLC profiles Page 142
7. Purification strategy. Page 144
8. PAGE purification of dp6-5 and dp6-6. Page 144
9. Preliminary NMR analysis of dp6-2 and dp6-3. Page 146
10. Full spectral assignment of dp6-2 and dp6-3. Page 148
11. Proton assignment of dp6-3. Page 161
12. Summary of chemical shift assignments for dp6-2 and dp6-3. Page 169
13. Preliminary NMR analysis of oligosaccharide HS 6-5. Page 174
14. Spectral assignment of dp6-5 and dp6-6. Page 176
15. Purification of two heparin derived hexasaccharides Page 181
16. NMR and spectral assignment of hep dp6-1 and dp6-2. Page 183
17. Summary of chemical shift assignments for hep dp6-1 and dp6-2. Page 184
18. De-2-O-sulphation. Page 188
19. Further evaluation of the de-2-O-sulphation procedure and de-sulphation of oligosaccharide, hep dp6-1. Page 194

20. De-2-O-sulphation of HS dp6-2 and dp6-3.	Page 198
21. Molecular dynamic simulations of iduronate residues.	Page 203
22. Molecular modelling.	Page 210
22.1 Static molecular modelling of HS dp6-3	Page 210
22.2 Calculation of potential energy surfaces.	Page 214
22.3 Geometry of the UA-GlcNS linkage (monosaccharide “a”-monosaccharide “b”)	Page 217
22.4 Geometry of the GlcNS-IdoUA(2S) linkage (monosaccharide “b”- monosaccharide “c”)	Page 222
22.5 Geometry of the IdoUA(2S)-GlcNS linkage (monosaccharide “c”- monosaccharide “d”)	Page 224
22.6 Geometry of the GlcNS- GlcUA linkage (monosaccharide “d”- monosaccharide “e”)	Page 227
22.7 Geometry of the GlcUA- GlcNAc (α or β) linkage (monosaccharide “e”- monosaccharide “f”)	Page 229
22.8 Summary for HS dp6-3	Page 231
23. NOE inter-proton distance restraints in the static molecular modelling of HS dp6-2, dp6-5 and dp6-6.	Page 234
23.1 Static molecular modelling of HS dp6-2.	Page 234
23.2 Static molecular modelling of HS dp6-5.	Page 240
23.3 Static molecular modelling of HS dp6-6.	Page 245
24. Molecular dynamic modelling	Page 249
24.1 HS dp6-2, -3, -5 and -6 molecular dynamics	Page 250

DISCUSSION

1. Initial studies into the degradation of HS by the heparinase III enzyme.	Page 259
2. SAX-HPLC chromatography of sized HS oligosaccharides	Page 259
3. SAX-HPLC and PAGE analysis	Page 260
4. PAGE and disaccharide analysis	Page 262
5. Variation between SAX-HPLC profiles	Page 262
6. PAGE purification of dp6-5 and dp6-6	Page 263
7. ¹³ C DEPT-135 and ¹³ C HSQC experiments.	Page 264
8. Evidence for the conservation of HS monosaccharide sequences across species barriers	Page 266
9. HS dp6-5, a novel monosaccharide sequence	Page 266
10. Molecular modelling GlcNAc β , gg to gt transitions	Page 271
11. Derivation of molecular modelling inter-proton distance restraints	Page 272
12. SAX-HPLC elution position, modelling insights and biological implications	Page 275
13. HS dp6-2	Page 276
14. HS dp6-3, differences to HS dp6-2 and possible biological implications	Page 283
15. A molecular re-orientation for HS dp6-5	Page 288
16. De-2-O-sulphation analysis	Page 294
16.1 Biological implications	Page 295
17. The influence of sulphation pattern on iduronate and terminal uronic acid equilibria	Page 299

17.1 Biological implications	Page 307
18. Structural features in addition to sulphation pattern and iduronate conformation may influence biological activity	Page 314
19. A Comparison of HS dp6 structural models	Page 314
20. Molecular dynamic studies	Page 317
21. Molecular dynamic evidence that internal 6-O-sulphation may cause an increased sampling rate of geometry 'b'	Page 324
22. PAGE evidence for a link between internal 6-O-sulphation and a chain kink	Page 325
23. Implication of the link between chain kinks and 6-O-sulphation for the biological activity of the Goodger decasaccharides	Page 327
24. Experimental evidence for kinked heparin oligosaccharides in free solution?	Page 328
25. Overall summary	Page 328
APPENDIX I	Pages 331-351
APPENDIX II	Pages 352-355
REFERENCES	Pages 356-372

FIGURES

Figure 1	C5-C6 rotamer conformations of N-acetyl- α -D-glucosamine	Page 28
Figure 2	Equilibrium conformations of methyl- α -L-iduronic acid.	Page 31
Figure 3	Equilibrium conformations of Δ 4,5 unsaturated uronic acid.	Page 33
Figure 4	Syndecan structure	Page 36
Figure 5	Glypican structure	Page 42
Figure 6	HS chain initiation	Page 48
Figure 7	A simplified scheme of heparan sulphate biosynthesis	Page 59
Figure 8	HS/heparin domain structure	Page 64
Figure 9	A refined model of the molecular structure of HS	Page 66
Figure 10	The crystal structure of a heparin hexasaccharide in complex with a FGF-2 monomer.	Page 68
Figure 11	The crystal structure of human interferon- γ	Page 71
Figure 12	The crystal structure of human interferon- γ in complex with two identical interferon- γ receptor molecules	Page 74
Figure 13	The crystal structure of human endostatin.	Page 79
Figure 14	The crystal structure of a 2:2:1 FGF1:FGFR2c:heparin dp10 complex.	Page 87

Figure 15	The crystal structure of a 2:2:2 FGF2:FGFR1c:heparin complex.	Page 89
Figure 16		Page 93
Figure 17	Two, three dimensional models of heparin dodcasaccharides.	Page 103
Figure 18	Φ and Ψ torsion angles defining the relative special orientation of each monosaccharide across a glycosidic linkage.	Page 106
Figure 19	Axial and equatorial orientations of C1 substituents and their preferred rotamer conformations due to the exo-anomeric effect.	Page 108
Figure 20	Analytical TSK3000PW gel-filtration profiles of HS after cleavage with heparinase III.	Page 131
Figure 21	Comparison of the analytical gel-filtration profiles of HS digested with 30mU of heparinase III for 64 or 160 hours.	Page 132
Figure 22	Preparative Bio-Gel P10 gel filtration profiles of HS digested to completion with heparinase III and the commercially available low molecular weight heparin product Tinzaparin sodium.	Page 134
Figure 23	SAX-HPLC chromatography of sized HS oligosaccharides.	Page 137
Figure 24	PAGE analyses of selected peaks from the HS dp6 and dp8 SAX-HPLC profile.	Page 139
Figure 25	Variation of the SAX-HPLC profile of dp6 size pools purified from different heparinase III digests of HS.	Page 143
Figure 26	Further purification of dp6-5 and dp6-6 by PAGE.	Page 145

- Figure 27** One-dimensional 500 mHz $^1\text{H-NMR}$ spectra of hexasaccharides 6-2 and 6-3. Page 147
- Figure 28** Assignment of the non-reducing terminal uronic acid, residue “a” of oligosaccharide dp6-2. Page 150
- Figure 29** Assignment of glucosamine, residue “b” of oligosaccharide dp6-2. Page 152
- Figure 30** Assignment of iduronic acid, residue “c” of oligosaccharide dp6-2. Page 154
- Figure 31** Assignment of glucosamine, residue “d” of oligosaccharide dp6-2. Page 156
- Figure 32** Assignment of iduronic acid, residue “e” of oligosaccharide dp6-2. Page 158
- Figure 33** Assignment of the reducing terminal glucosamine, residue “f” of oligosaccharide dp6-2. Page 160
- Figure 34** Identification of trans-glycosidic NOEs between IdoA(2S), residue “c” of dp6-3 and the preceding glucosamine residue “d”. Page 162
- Figure 35** Assignment of glucosamine, residue “d” of oligosaccharide dp6-2. Page 164
- Figure 36** Assignment of glucuronic acid, residue “e” of oligosaccharide dp6-3. Page 166
- Figure 37** Assignment of the reducing terminal glucosamine, residue “f” of oligosaccharide dp6-3. Page 168

- Figure 38** A summary of the ^1H chemical shift assignments for dp6-2 and dp6-3.
Page 171
- Figure 39** One-dimensional 500 MHz ^1H -NMR spectra of hexasaccharide 6-5.
Page 175
- Figure 40** A summary of the ^1H chemical shift assignments for dp6-5 and dp6-6.
Page 179
- Figure 41** SAX-HPLC chromatography of heparin derived hexasaccharides.
Page 182
- Figure 42** A summary of the ^1H chemical shift assignments for hep dp6-1 and dp6-2.
Page 186
- Figure 43** NMR analysis of heparin dp6-2 after successive rounds of chemical de-2-O-sulphation.
Page 190
- Figure 44** NMR analysis of heparin dp6-1 after chemical de-2-O-sulphation.
Page 195
- Figure 45** NMR analysis of HS dp6-2 and dp6-3 after chemical de-2-O-sulphation.
Page 200
- Figure 46** Cremer-Pople ring puckering parameters for the methyl glycoside of IdoUA during a 1000ps molecular dynamic simulation.
Page 206
- Figure 47** Cremer-Pople ring puckering parameters for the methyl glycoside of IdoUA(2S) during a 1000ps molecular dynamic simulation.
Page 208
- Figure 48** NOE build up curves for all trans-glycosidic NOEs observed within oligosaccharide HS dp6-3.
Page 212

- Figure 49** Potential energy surfaces calculated for each glycosidic linkage within HS dp6-3. Page 216
- Figure 50** Energy minimised structures for the UA-GlcNS glycosidic linkage at $\Phi = 180^\circ$ and $\Psi = 0^\circ$. Page 218
- Figure 51** Energy minimised structures for the UA-GlcNS glycosidic in a number of low energy geometries, as indicated by the potential energy surface calculations. Page 221
- Figure 52** Energy minimised structures for the GlcNS-IdoUA(2S) (1C_4 and 2S_0) glycosidic linkage at $\Phi = -40^\circ$ and $\Psi = -20^\circ$. Page 223
- Figure 53** Energy minimised structures for the IdoUA(2S)-GlcNS glycosidic linkage at $\Phi = 40^\circ$ and $\Psi = 20^\circ$ for IdoUA(2S) in the 1C_4 conformation and $\Phi = -40^\circ$ and $\Psi = -20^\circ$ for IdoUA(2S) in the 2S_0 conformation. Page 225
- Figure 54** An energy minimised structure for the IdoUA(2S) 2S_0 -GlcNS glycosidic linkage at $\Phi = 40^\circ$ and $\Psi = 20^\circ$ Page 226
- Figure 55** The energy minimised structure for the GlcNS 4C_1 -GlcUA 4C_1 glycosidic linkage at $\Phi = -40^\circ$ and $\Psi = -20^\circ$ Page 228
- Figure 56** Energy minimised structures for the GlcUA-GlcNAc (α or β) glycosidic linkage in a number of low energy geometries, as indicated by the potential energy surface calculations. Page 230
- Figure 57** Two final proposed model structure of oligosaccharide HS dp6-3. Page 233
- Figure 58** Potential energy surfaces calculated across the glycosidic linkages indicated, within HS dp6-2. Page 237

Figure 59	Two final proposed model structure of oligosaccharide HS dp6-2.	Page 239
Figure 60	Potential energy surfaces calculated across the glycosidic linkages indicated, within HS dp6-5.	Page 242
Figure 61	Two final proposed model structure of oligosaccharide HS dp6-5.	Page 244
Figure 62	Potential energy surfaces calculated across the glycosidic linkages indicated, within HS dp6-6.	Page 246
Figure 63	Two final proposed model structures of oligosaccharide HS dp6-6.	Page 248
Figure 64	Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation of HS dp6-2.	Page 254
Figure 65	Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation of HS dp6-3.	Page 256
Figure 66	Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation, HS dp6-5, model A.	Page 257
Figure 67	Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation, HS dp6-6, model A.	Page 258
Figure 68	An enlargement of a section from a 2D-NOESY spectrum (500ms mixing time) recorded for HS dp6-5.	Page 268
Figure 69	Enlargements of sections from 2D- COSY and TOCSY spectra recorded for HS dp6-5.	Page 270
Figure 70	HS dp6-2 molecular models.	Page 278

Figure 71	HS dp6-2 molecular models.	Page 280
Figure 72	A superimposition model of HS dp6-2.	Page 282
Figure 73	A superimposition model of HS dp6-3.	Page 286
Figure 74	HS dp6-2 and HS dp6-3 molecular models.	Page 287
Figure 75	HS dp6-5 molecular models.	Page 291
Figure 76	A superimposition model of HS dp6-5.	Page 293
Figure 77	The monosaccharide sequences of 4 biologically active and 2 biologically inactive decasaccharides.	Page 313
Figure 78	A comparison of of the molecular shape of all HS dp6 structural models.	Page 316
Figure 79	A second superimposition model of HS dp6-2.	Page 319
Figure 80	An extended dodecasaccharide superimposition model of HS dp6-2.	Page 321
Figure 81	Enlarged sections from the NOESY spectra (500ms mixing time) of HS dp6-2 and HS dp6-3.	Page 323
Figure 82	Two-dimensional 500 mHz ¹ H-NMR COSY spectrum, HS dp6-2.	Page 331
Figure 83	Two-dimensional 500 mHz ¹ H-NMR COSY spectrum, HS dp6-3.	Page 332
Figure 84	Two-dimensional 500 mHz ¹ H-NMR TOCSY spectrum, HS dp6-2.	Page 333
Figure 85	Two-dimensional 500 mHz ¹ H-NMR TOCSY spectrum, HS dp6-3.	Page 334

- Figure 86** Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-2.
Page 335
- Figure 87** Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-3.
Page 336
- Figure 88** Two-dimensional 500 MHz ^1H -NMR COSY spectrum, HS dp6-5.
Page 337
- Figure 89** Two-dimensional 500 MHz ^1H -NMR COSY spectrum, HS dp6-6.
Page 338
- Figure 90** Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, HS dp6-5.
Page 339
- Figure 91** Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, HS dp6-6.
Page 340
- Figure 92** Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-5.
Page 341
- Figure 93** Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-6.
Page 342
- Figure 94** Two-dimensional 500 MHz ^1H -NMR COSY spectrum, hep 6-1.
Page 343
- Figure 95** Two-dimensional 500 MHz ^1H -NMR COSY spectrum, hep 6-2.
Page 344
- Figure 96** Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, hep 6-1.
Page 345

Figure 97	Two-dimensional 500 MHz ¹ H-NMR TOCSY spectrum, hep 6-2.	Page 346
Figure 98	Two-dimensional 500 MHz ¹ H-NMR NOESY spectrum, hep 6-1.	Page 347
Figure 99	Two-dimensional 500 MHz ¹ H-NMR NOESY spectrum, hep 6-2.	Page 348
Figure 100	Two-dimensional 500 MHz ¹ H-NMR COSY spectrum, HS 6-3 after chemical de-2-O-sulphation	Page 349
Figure 101	Two-dimensional 500 MHz ¹ H-NMR COSY spectrum, hep 6-1 after chemical de-2-O-sulphation.	Page 350
Figure 102	Two-dimensional 500 MHz ¹ H-NMR COSY spectrum, hep 6-2 after undergoing two rounds of the chemical de-2-O-sulphation procedure	Page 351
Figure 103	Two-dimensional 500 MHz HSQC spectrum, HS dp6-2.	Page 352
Figure 104	Two-dimensional 500 MHz HSQC spectrum, HS dp6-3.	Page 353
Figure 105	Two-dimensional 500 MHz HSQC spectrum, hep dp6-1.	Page 354
Figure 106	Two-dimensional 500 MHz HSQC spectrum, hep dp6-2.	Page 355

TABLES

Table 1. A comparison of the conformational distribution of iduronate residues within a number of previously published oligosaccharide sulphation sequences.

Page 100

Table 2 Disaccharide analyses of selected peaks from the HS dp6 and dp8 SAX-HPLC profile.

Page 141

Table 3	¹ H and ¹³ C chemical shifts assignments (ppm) for hexasaccharides 6-2 and 6-3.	Page 173
Table 4	¹ H chemical shifts assignments (ppm) for hexasaccharides 6-5 and 6-6	Page 180
Table 5	¹ H chemical shifts assignments (ppm) for hep 6-1 and 6-2	Page 187
Table 6.	The ¹ H chemical shifts assignments (ppm) for heparin dp6-2 prior to and after undergoing de-2-O-sulphation.	Page 192
Table 7.	¹ H chemical shifts assignments (ppm) for heparin dp6-1 prior to and after undergoing de-2-O-sulphation.	Page 197
Table 8.	¹ H chemical shifts assignments (ppm) for HS dp6-3 prior to and after undergoing de-2-O-sulphation.	Page 202
Table 9.	The average value and standard deviation of the theoretical iduronate coupling constants as obtained by explicit water molecular dynamic simulations.	Page 209
Table 10	Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-3.	Page 232
Table 11	Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-2.	Page 238
Table 12	Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-5.	Page 243
Table 13	Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-6.	Page 247

Table 14 Conformational distributions of iduronate and terminal uronic acid monosaccharides within oligosaccharides characterised in this study and a number of previously published studies. Page 304-306

ABBREVIATIONS

Δ UA or UA	Δ 4,5 unsaturated uronic acid
2OSTs	Uronosyl-2-O-sulphotransferase enzymes.
3OSTs	Glucosaminyl 3-O-sulphotransferase enzymes.
6OSTs	Glucosaminyl 6-O-sulphotransferase enzymes.
aMan	Anhydromannose.
AT-III	Antithrombin III.
CHO	Chinese Hamster Ovary.
CS	Chondroitin Sulphate.
DS	Dermatan sulphate.
ER	Endoplasmic Reticulum.
EXT	Exostoses.
EXTL	Exotoses like.
FGFs	Fibroblast growth factors.
FGFRs	Fibroblast Growth Factor Receptors.
GAG	Glycosaminoglycan
Gal	Galactose.
GalNAc	N-acetyl galactosamine.
GalNAcT-I	N-acetyl galactosamine transferase I.

GalTI and GalTII	Galactosyltransferases I and II.
GlcATI	Glucuronosyltransferase I.
GlcNAc	N-acetyl glucosamine.
GlcNAc(6S)	6-O-sulphated N-acetylglucosamine.
GlcNAcT-I	N-acetyl glucosamine transferase I.
GlcNH ₂	Glucosamine.
GlcNS	N-sulphoglucosamine.
GlcNS(6S)	6-O-sulphated N-sulphoglucosamine.
GlcUA	Glucuronic acid.
Gly	Glycine.
GPI	Glycosyl-phosphatidylinositol.
HexUA	Hexuronic acid (either IdoA or GlcA).
HS	Heparan Sulphate.
HSPG	Heparan Sulphate Proteoglycan.
IdoUA	Iduronic acid.
IdoUA(2S)	2-O-sulphated iduronic acid.
IFN- γ	Inteferon- γ .
IL-8	Interleukin 8.
MIP1 α	Macrophage inflammatory protein 1 α .
NAc-domain	N-acetylated domain.
NA / NS domain	Alternating domain.
NDSTs	N-deacetylase / N-sulphotransferase enzymes.
PAPS	3'-phosphoadenosine, 5'-phosphosulphate.

PF-4	Platelet Factor 4.
PG	Proteoglycan.
PKC	Protein Kinase C.
Ser	Serine.
S-domain	Sulphated domain.
UDP	Uridine diphosphate.
Xyl	Xylose.
XT	Xylotransferase.

Abstract

A deeper insight into the structural biology of HS is key to understanding its near-universal functional role as a co-receptor for growth factors and morphogens. Due to the extreme difficulty in preparing homogeneous HS oligosaccharides for structural and functional studies, traditionally, oligosaccharides derived from the related molecule heparin are used as HS structural models. In this study a number of authentic HS derived hexasaccharides, in addition to heparin derived hexasaccharides, have been purified in sufficient quantity to permit a detailed NMR and molecular modelling based analysis of their three dimensional structure. The primary sequence of one HS derived oligosaccharide has never previously been published. Studies on all oligosaccharides and their chemically de-2-O-sulphated derivatives have revealed additional new insights into the structural influence of sulphate groups. Consistent with previous studies, at the monosaccharide level, sulphation was found to influence iduronate conformational behaviour. However, with the data presented, a number of gaps in the literature have now been filled, and it is now possible for the first time to predict the balance of iduronate conformational equilibria within any HS monosaccharide sequence. Sulphation was also found to influence the overall topology of the oligosaccharide chains themselves. In particular, for the first time NMR data is presented to show that local deviations may occur along the helical axis of each oligosaccharide when it is free in solution. Polyacrylamide gel electrophoresis data and molecular dynamic modelling data are presented to suggest that the rate at which local deviations occur may be influenced by the sulphation pattern contained within a particular oligosaccharide. The functional implications of these and other new structural insights are discussed, and in particular are related to a library of HS derived deca-saccharide structures previously tested for biological activity.

Acknowledgement

I would like to dedicate this thesis to my parents and to my wife Heather for all their help, encouragement and support throughout the course of this project. I would also like to dedicate this thesis to my Aunty Bridget who never got the chance to read it, not that she would have wanted to anyway.

I would like to thank John Pezzulo who set up the least squares fitting analysis spreadsheet and took the time to phone me all the way from Kissimmee, Florida. I would like to thank Dave Case from the Scripps Research Institute, California and guardian of the AMBER source code. Without him waving the AMBER academic fee a majority of this research work would not have been possible. I would also like to thank Dr Neil McLay who expertly recorded all the NMR spectra reported and Dr Lindsay Harding who carried out mass spectrometry analysis on a number of samples although this data is not presented. I would also like to thank Francois-Yves Dupradeau at the University of Amien, France. Since I contacted him nearly one year ago, I have been receiving one email from him every day. I don't think anyone could find a better person to collaborate with, and I have high hopes for our work together. I wish to acknowledge Sarah Goodger for all the hard work she put into her own thesis. Hopefully my hard work will help to get hers the publicity it deserves. Finally I would like to thank my supervisor Dave Pye for all his help and basically just letting me get on with it for the past three years.

1. Introduction

Heparan Sulphate (HS) is a member of a family of complex polysaccharides known as glycosaminoglycans (GAGs). They are made up of repeating C1, C4- linked disaccharide units, that are composed of a hexuronic acid, which is one of either β -D-glucuronic acid (GlcUA) or α -L-iduronic acid (IdoUA), and an N-acetyl (GlcNAc) or N-sulpho- α -D-glucosamine (GlcNS). In addition to N-sulphation, many of these constituent disaccharides carry one or several O-sulphate substituents. Commonly O-sulphation occurs at C2 of IdoUA to form the monosaccharide IdoUA(2S) and/or at C6 of GlcNS or GlcNAc to form the monosaccharides GlcNS(6S) and GlcNAc(6S).

HS binds to a variety of protein ligands (Esko and Selleck 2002) and has been shown to be involved in the regulation of a wide range of biological activities, including developmental processes (Lander and Selleck 2000; Perrimon and Bernfield 2000), angiogenesis (Robinson and Stringer 2001), blood coagulation (Marcum *et al.* 1986) and tumour metastasis (Liu *et al.* 2002).

A brief description of the conformational features of the common constituent HS monosaccharide units is given below.

1.1 The conformation of N-glucosamine monosaccharides

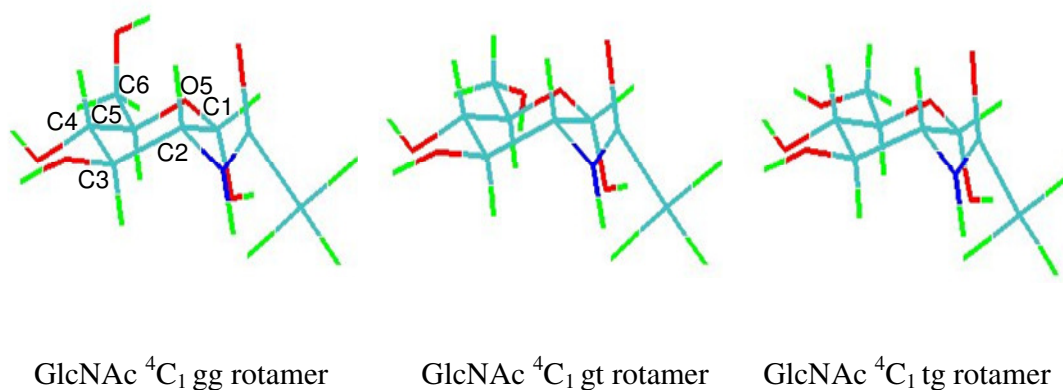
The GlcNAc, GlcNAc(6S), GlcNS and GlcNS(6S) monosaccharides (collectively termed N-glucosamines) of HS, and a related molecule heparin, exist exclusively in the 4C_1 chair conformation (Mulloy and Forster 2000). In common with other hexopyranose sugars within the 4C_1 conformation, there exists an equilibrium between rotamer conformations about the C5-C6 bond. These rotamers are characterised by the O6-C6-C5-O5 torsion

angle (w -angle). Three stable staggered rotamers are possible for the w -angle and are termed gauche-trans (gt), trans-gauche (tg) and gauche-gauche (gg). Gauche and trans referring to the orientation of the w -angle and the O6-C6-C5-C4 angle respectively (see Figure 1). Gluco-pyranosides display approximately equal populations of the gt and gg rotamers, with nearly a complete absence of the tg rotamer (Kirschner and Woods 2001). The tendency for the w -angle to adopt gauche conformations has been termed the gauche effect.

Figure 1 C5-C6 rotamer conformations of N-acetyl- α -D-glucosamine.

The pyranose ring of GlcNAc is stable in the 4C_1 conformation, as shown in each of the models below. In solution there exists an equilibrium between three different conformations about the C5-C6 bond. The three different conformations are termed gg, gt and tg (see text, for explanation of this nomenclature).

These conformations are characterised by a Cremer-Pople puckering parameter of $\theta = <10^\circ$ and NMR coupling constants of ${}^3J_{1,2} = 3.6$, ${}^3J_{2,3} = 10.2$, ${}^3J_{3,4} = 9.4$, ${}^3J_{4,5} = 9.9$, ${}^3J_{5,6a} = 4.0$, ${}^3J_{5,6b} = 2.0$ (van_Boeckel *et al.* 1987).



1.2 Cremer-Pople puckering parameters and NMR coupling constants

Cremer-Pople puckering parameters are merely a mathematical measurement of pyranose ring conformation (Cremer and Pople 1975; Cremer 1984). Their values are, however, influenced by the order in which the atoms within a ring system are selected for analysis (Forster and Mulloy 1993). Throughout this thesis work, atoms were selected for analysis in the order O5-C1-C2-C3-C4-C5. The result of this is that the Cremer-Pople parameters given in this study, are not directly comparable to a number of previous studies (Ragazzi *et al.* 1986; Ragazzi *et al.* 1993), where atoms were chosen in the order C1-C2-C3-C4-C5-O5. They are however comparable to the majority of previous work.

It is quite common to see NMR spectra in which the signal due to a particular proton appears as a collection of peaks (Carey 1992). The observed splitting of NMR signals is known as spin-spin splitting and the physical basis for this phenomenon is spin-spin coupling. Protons with different chemical shift values, which are connected by three bonds (vicinal protons), may be spin-spin coupled and give rise to split signals. Four bond couplings are weaker and not normally observable. The extent to which two vicinal nuclei are coupled is known as the coupling constant 3J . This is equal to the separation between adjacent lines of the signal of a particular proton. It has been shown that the value of a given coupling constant depends dramatically on the value of the dihedral angle between the two spin-spin coupled protons (Karplus 1959). This is known as the Karplus relationship. By measuring a particular coupling constant, it is possible to propose a dihedral angle between the two protons giving rise to this constant and *vice-versa*, using the Karplus equation

$$^3J = A + B \cos \theta + C \cos 2\theta$$

Where θ is a dihedral angle value and A, B and C are equation parameters.

Application of the Karplus equation depends critically on the values of the parameters A, B and C, which are obtained empirically from a set of experimental data (Becker 2000). For parameterisation of the equation as originally proposed by Karplus (Karplus 1959), values for A, B and C were determined by the analysis of ethane like molecules. Variations on the original Karplus equation, such as the Haasnoot equation attempt to take into account the effect of electronegative atoms such as oxygen on the relationship between dihedral angle and coupling constant values (Haasnoot *et al.* 1979).

1.3 The conformation of glucuronic acid and iduronate acid monosaccharides

In common with N-glucosamine monosaccharides, glucuronic acid monosaccharides are also stable in the 4C_1 chair conformation. However IdoUA and IdoUA(2S) residues (collectively termed iduronates) vary in conformation between the 4C_1 chair form, an alternate 1C_4 chair conformation, and a 2S_0 skew boat conformation (see Figure 2). When internally positioned within an oligosaccharide the 1C_4 and 2S_0 conformations are thought to pre-dominate (Ferro *et al.* 1990).

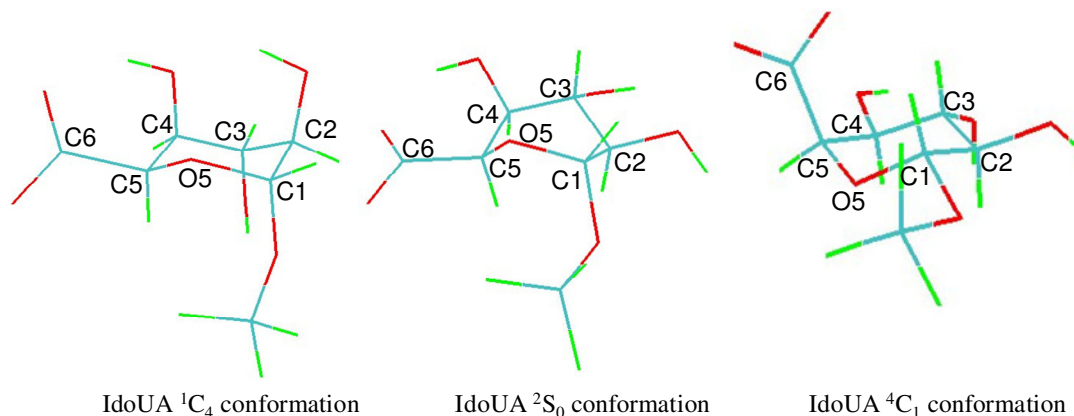
Figure 2 Equilibrium conformations of methyl- α -L-iduronic acid.

The pyranose ring of iduronic acid exists in equilibrium between three low energy conformations (shown below). The 4C_1 conformation is only thought to occur when iduronates form the terminal unit of an oligosaccharide. The 1C_4 and 2S_0 conformations are characterised by a Cremer-Pople puckering parameter of $\theta = >160^\circ$ (1C_4) and $\theta =$ approx 90° (2S_0). For the 2S_0 conformation a second parameter Φ_2 can be calculated. For 2S_0 , $\Phi_2 =$ approx 150° . The published NMR coupling constants for these conformations vary between studies (van_Boeckel *et al.* 1987; Forster and Mulloy 1993; Ojeda *et al.* 2002; Lucas *et al.* 2003). The following values are taken from the earliest cited work of van_Boeckel *et al.*

For 1C_4 , ${}^3J_{1,2} = 2.6$, ${}^3J_{2,3} = 3.5$, ${}^3J_{3,4} = 3.5$, ${}^3J_{4,5} = 1.1$

For 2S_0 , ${}^3J_{1,2} = 5.9$, ${}^3J_{2,3} = 10.3$, ${}^3J_{3,4} = 5.8$, ${}^3J_{4,5} = 4.3$

For 4C_1 , ${}^3J_{1,2} = 7.8$, ${}^3J_{2,3} = 9.7$, ${}^3J_{3,4} = 9.4$, ${}^3J_{4,5} = 4.6$



1.4 The conformation of Δ UA monosaccharides

Oligosaccharides frequently used for HS structural studies terminate at their non-reducing end in a Δ 4,5 unsaturated uronic acid residue (Δ UA or UA) (see enzymatic degradation of HS discussed later in this introduction). These monosaccharides exist in an equilibrium between the 1H_2 and 2H_1 half chair conformations (see Figure 3).

Figure 3 Equilibrium conformations of $\Delta 4,5$ unsaturated uronic acid.

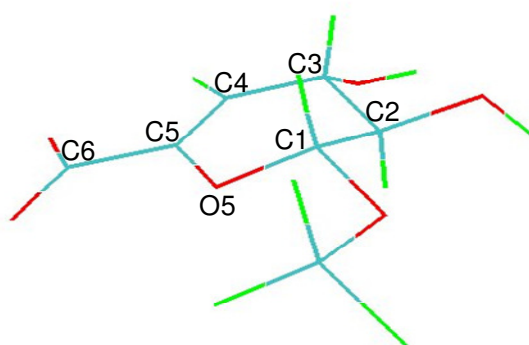
The pyranose ring of UA exists in an equilibrium between two low energy conformations shown below. These are characterised by Cremer-Pople puckering parameters of $\theta = \text{approx } 50^\circ$, $\Phi_2 = \text{approx } 90^\circ$ and $\theta = \text{approx } 130^\circ$, $\Phi_2 = \text{approx } 250^\circ$ for ${}^1\text{H}_2$.

NMR coupling constants are given as:

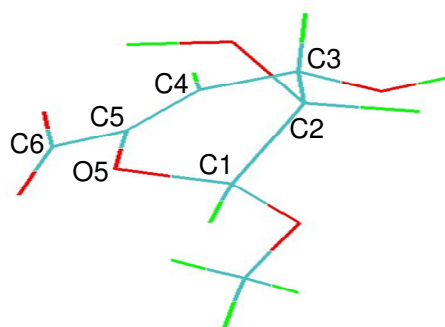
For ${}^2\text{H}_1$, ${}^3J_{1,2} = 8.27$, ${}^3J_{2,3} = 7.67$, ${}^3J_{3,4} = 2.78$

For ${}^1\text{H}_2$, ${}^3J_{1,2} = 2.87$, ${}^3J_{2,3} = 2.37$, ${}^3J_{3,4} = 4.67$

(Ragazzi *et al.* 1993)



UA ${}^2\text{H}_1$ conformation



UA ${}^1\text{H}_2$ conformation

1.5 Protein core and HS chains- one functional unit

HS chains commonly occur at the cell surface typically covalently linked to a protein core structure (Bernfield *et al.* 1992; David 1993). Proteins that are modified in this way are termed proteoglycans (PGs). Proteins that are specifically modified, through the attachment of HS chains are termed specifically Heparan Sulphate proteoglycans (HSPGs).

Both the protein and HS part of HSPGs are thought to play a functional role in their biological activity. For example, syndecan HSPGs via their HS chains, bind growth factor and extra-cellular matrix molecules extra-cellularly while also interacting with the intra-cellular cytoskeleton, via their protein core cytoplasmic domain (Bernfield *et al.* 1992). Covalently anchored to the protein core, usually several HS chains are presented for potential ligand binding interactions at the cell surface. By regulating the distribution of the core protein within the plasma membrane and the time at which core proteins are expressed, the cell may have a mechanism by which to limit the action of HS binding ligands to a place and time at the cell surface. The structure and function of the most common protein core units is discussed below.

2. Proteoglycans

PGs have traditionally been categorised on the basis of their cellular location, either at the cell surface, or within the extracellular matrix. The major cell membrane HSPGs are the transmembrane syndecans and the GPI anchored glypicans. Other minor forms of membrane HSPG include betaglycan (Andres *et al.* 1992), which is unusual in that it carries only one HS chain, and the V-3 isoform of CD44 present on keratinocytes and

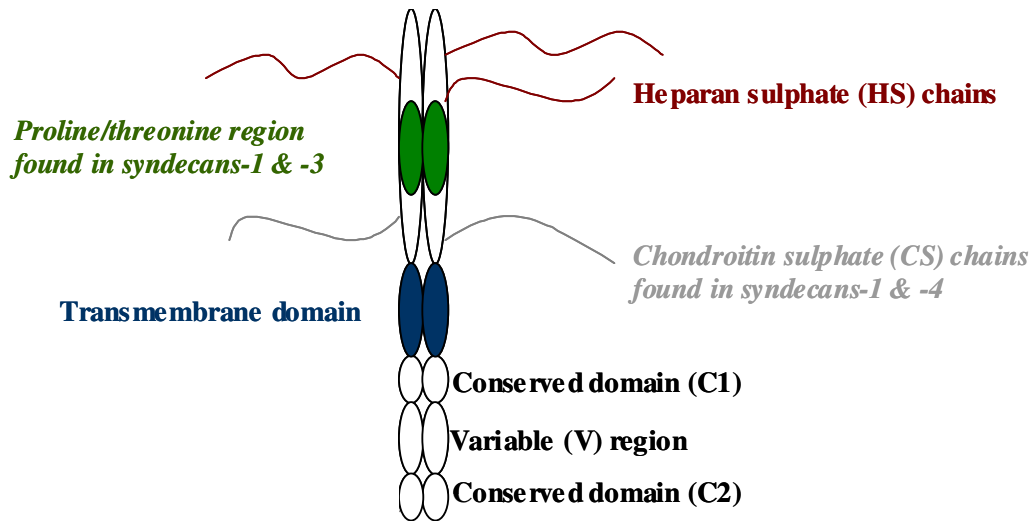
activated monocytes (Jackson *et al.* 1995). In the extracellular matrix, especially basement membranes the multi-domain perlecan, agrin and collagen XVIII core proteins are the main HS-bearing species.

2.1 Syndecan HSPGs

The syndecan family of HSPGs is comprised of four members. Syndecans 1 and 3 and syndecans 2 and 4 comprise separate subfamilies that have arisen by gene duplication and divergent evolution from a single ancestral gene (Carey 1997). The syndecan numbers 1-4 reflect the order in which the cDNAs for each family member were cloned. All are small type I membrane proteins, with an N-terminal signal peptide, an ectodomain, a single hydrophobic transmembrane domain and a short C-terminal cytoplasmic domain (Bernfield *et al.* 1992) (see Figure 4). The ectodomains show the least amount of amino acid sequence homology, no more than 10-20 % (David 1993). In contrast to the ectodomains the transmembrane and the short cytoplasmic tail domains of the four syndecans contain around 60-70% sequence homology (David 1993). The transmembrane domains contain an unusual alanine/glycine motif while the cytoplasmic domain is essentially composed of two regions of conserved sequence (C1 and C2) separated by a central variable region that is distinct for each family member (V) (see Figure 4).

Figure 4 Syndecan structure

The extracellular domains of the four syndecans contain several sites for HS attachment. The transmembrane domain is highly conserved between syndecan family members and contains an unusual glycine/alanine motif. The cytoplasmic domain contains two domains (C1 and C2) that are also highly conserved and flank a central variable (V) region. This Figure is reproduced with the permission of the author from the thesis of Goodger (Goodger 2003)



Proposed biological roles for syndecans are diverse. A change in the level of cell surface hypothalamic syndecan 3, in murine models, has been linked to an increase in food intake and obesity. Syndecans 1 and 4 have been linked with cellular invasion of the bacterium *Neisseria gonorrhoeae*; while a further role for syndecan 1 in modulating the Wnt signalling pathway has been elucidated. This pathway is involved in cell proliferation and differentiation during both development and disease. For a review of all these biological roles see (Bellin *et al.* 2003).

2.2 Syndecan ectodomains

Conserved within the amino acid sequences of the various ectodomains are glycosylation signal sequences that signal the attachment sites for modifying HS chains. All syndecans have a similar sequence motif for HS attachment towards the amino terminus (Carey 1997). This motif consists of a single serine-glycine amino acid sequence that is preceded or flanked by a sequence of acidic amino acids (Bernfield *et al.* 1992). In syndecans-1 and -4 sites for additional glycanation close to the cell membrane are occasionally modified by a galactosaminoglycan, chondroitin sulphate (CS) (Kokenyesi and Bernfield 1994; Shworak *et al.* 1994). A co-operative role between CS and HS chains has been shown in the binding of a number of growth factors (Deepa *et al.* 2004). The conservation of HS attachment sites within the protein structure of all members of the syndecan family demonstrates the vital role played by HS modification in syndecan function.

2.3 Syndecan transmembrane domains

The alanine/glycine motif found within the transmembrane domain has been shown to have a role in the dimerisation of syndecan-3 and probably plays a similar role in the dimerisation of other syndecans (Asundi and Carey 1995). It is also through the transmembrane domain that syndecans are able, in principle, to interact with other membrane proteins, this may affect their positioning in appropriate plasma membrane compartments (Carey 1997). The transmembrane domain has also been associated with the internalisation of extracellular ligands (Zimmerman and David 1999).

2.4 Syndecan cytoplasmic domains

One of the best-characterised roles for a syndecan cytoplasmic domain is in the recruitment of syndecan-4 into focal adhesions. Syndecan-4 is a common constituent of focal adhesions (Woods and Couchman 1994). Its recruitment into focal adhesions is stimulated by the treatment of fibroblasts with activators of protein kinase C (PKC) (Baciu and Goetinck 1995), and a PKC α specific signalling function has been identified (Oh *et al.* 1997). The mechanism by which PKC α signalling is activated is complex (for an extensive review see (Rapraeger 2001)) but key to the mechanism is the dimerisation of the V-region of the cytoplasmic domain (Oh *et al.* 1997). Dimerisation appears to be reduced by the phosphorylation of a conserved serine residue within the C1 cytoplasmic domain (Horowitz and Simons 1998). Phosphorylation of this conserved serine residue within C1, and tyrosine residues in syndecan 1, has been detected either *in-vitro* or *in-vivo* for all members of the syndecan family (Prasthofer *et al.* 1995; Reiland *et al.* 1996; Oh *et al.* 1997). While the inhibition of the dimerisation of the V-region of syndecan-4

may be a unique mechanism, it is conceivable that phosphorylation within C1, may represent a common mechanism for regulating V-region and cytoplasmic domain interactions.

At the opposite end of the cytoplasmic domain, another family of five structurally related proteins has been identified that recognise the specific tetrapeptide sequence EFYA, which is conserved in the C2 region of all syndecans (Rapraeger 2001). The five members so far identified are syntenin, synectin, syndesmos, CASK/Lin-2 and synbindin (Asundi and Carey 1995; Cohen *et al.* 1998; Ethell *et al.* 2000; Gao *et al.* 2000).

Overexpression of either synectin or syndesmos appears to result in an increased recruitment of syndecan-4 into focal adhesions (Baciu *et al.* 2000; Gao *et al.* 2000). CASK has been shown to bind syndecan-2 at synaptic junctions and is co-immunoprecipitated with syndecan-3 from brain extracts (Hsueh and Sheng 1999). It has also been shown to co-localise with syndecan-1 at the basolateral borders of epithelial cells (Cohen *et al.* 1998). Synbindin binds to the EFYA motif of syndecan-2 in dendritic spines (Ethell *et al.* 2000). Syntenin, the first member of this family to be identified, has been reported to be a component of cell adhesion sites and of microfilament bundles (Zimmermann *et al.* 2001).

The presence of a least one PDZ domain in all these proteins, and the experimental observations described above, suggests these are adaptor proteins and may connect the syndecans to the cytoskeleton (Rapraeger 2001). The significance of differing adaptor proteins binding to differing syndecans depending on their cellular location is unclear. While some insights have been made into the role phosphorylation plays in the functional relationship between the cytoplasmic C1 and V domains, mostly regarding syndecan-4,

questions remain as to the effect of phosphorylation within C1 on protein binding to C2 and *vice versa*.

2.5 Syndecan expression

Virtually all cells express at least one form of syndecan, most express multiple forms, and there are distinct patterns of syndecan expression that characterise individual cell types and tissues (Kim *et al.* 1994). For example, adult mouse kidney is found to contain mostly syndecan-4 mRNA while liver contains high levels of all syndecan mRNAs except syndecan-3. This reflects syndecan expression by the various cell types that compose these two types of tissue (Kim *et al.* 1994).

Syndecan expression is often associated with morphological transitions, or major changes in tissue organization. A burst of syndecan-3 expression corresponds with a period of oligodendrocyte differentiation and further formation of the central nervous system in new-born mice (Carey *et al.* 1997). Alterations in the levels of other syndecan family members, have also been shown to coincide with differentiation events in other cell types (reviewed by (Carey 1997)).

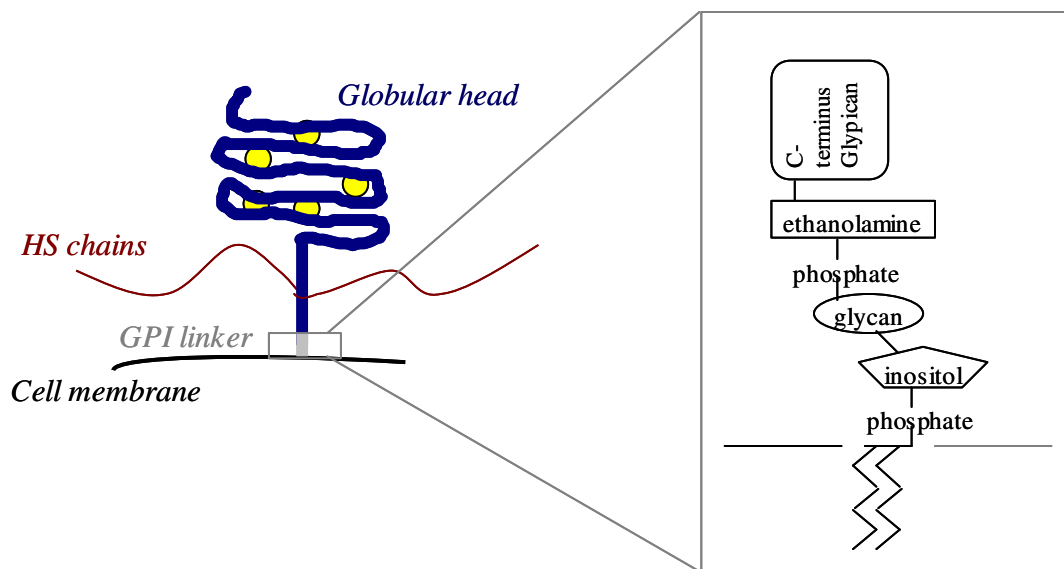
2.6 Glypican HSPGs

The glypican family of HSPGs are anchored to the cell surface via covalent linkage to glycosylphosphatidylinositol (GPI), and in contrast to the syndecans do not span the cell membrane (see Figure 5). Six glypicans have been identified in vertebrates, two in *Drosophila* and one in *C. elegans* (Filmus 2001).

All glypicans contain an N-terminal secretory signal peptide and a hydrophobic domain in the C-terminal region, required for the insertion of the GPI anchor. Comparing the amino acid sequences of the vertebrate glypicans to one another, the sequences vary from being 17% to 63% identical (De Cat and David 2001). The location of 14 cysteine residues is conserved between the glypicans, suggesting the existence of a highly similar three-dimensional structure. HS attachment is restricted to the last 50 amino acids in the C-terminus, placing HS chains close to the cell membrane (Saunders *et al.* 1997).

Figure 5 Glypican structure

The glypican core protein is covalently linked to the plasma membrane at its C-terminus via a GPI linker or anchor. Numerous disulphide bridges organize the core protein into a highly compact globular structure. HS attachment is restricted to the last 50 amino acids in the C-terminus. This figure is reproduced with the permission of the author from the thesis of Goodger (Goodger 2003).



2.7 Glypican expression

In terms of its cellular distribution, glypicans are predominantly expressed during development (De Cat and David 2001). Like the syndecans, glypican expression is also associated with morphological transitions or major changes in tissue organization. For example cerebroglycan, a member of the glypican family, is only expressed in the developing rat nervous system, specifically during neuronal differentiation (Stipp *et al.* 1994). Mutation studies on the *Drosophila* glypican or Dally have identified this as being required for correct development of the *drosophila* central nervous system (Nakato *et al.* 1995). While mutation of human glypican 3 leads to Simpson-Golabi-Behmel syndrome, characterised by pre- and post-natal overgrowth of multiple tissues and organs (Pilia *et al.* 1996).

3. Extracellular matrix HSPGs

3.1 Perlecan

The perlecan core protein consists of five-domains and has a mass of 470 kDa. It is located in most extracellular matrices and has been purified from a variety of human cell types (Knox *et al.* 2001). The precise role of perlecan *in-vivo* is unknown. Some of the biological functions of perlecan may be linked to the core proteins ability to interact with a variety of macromolecules, including β 1 integrins and laminin (Brown *et al.* 1997; Hopf *et al.* 1999). However the HS chains are believed to be the main contributors to its biological action, through ligand binding (Whitelock *et al.* 1999), and control of basement membrane permeability.

3.2 Agrin

The best characterised role for agrin is in the development of the neuromuscular junction during embryogenesis (Cole and Halfter 1996). Indeed, agrin's name is based on its involvement in the aggregation of acetylcholine receptors during synaptogenesis.

Recently, it has been shown that treatment of a cultured muscle cell line with chlorate reduces the spontaneous formation of acetylcholine receptor clusters (McDonnell and Grow 2004). Chlorate prevents the correct biosynthesis of HS and this observation suggests that the HS side chains of agrin may play a role in the formation of acetylcholine receptor clusters.

There are three potential HS attachment sites within the primary sequence of agrin but it is thought that only two of these actually carry HS chains when the protein is expressed (Groffen *et al.* 1998). Some of the biological roles of agrin require interactions through the protein core itself. While another biological role in the retention of anionic macromolecules, within the vasculature, has been suggested for HS at the glomerular or pulmonary alveolar basement membranes (Groffen *et al.* 1998).

3.3 Collagen XVIII

Collagen XVIII is the only currently known collagen that carries HS side chains (Halfter *et al.* 1998). Eight potential GAG attachment sites were identified in chick collagen XVIII and three of these were found to carry GAGs. These were located towards the N-terminus of the protein. In the naturally occurring protein it is believed that collagen XVIII is exclusively modified by HS attachment, although CS modifications were detected *in-vitro* (Dong *et al.* 2003).

Proteolytic cleavage within a C-terminal domain of collagen XVIII releases a 20 kDa anti-angiogenic protein, endostatin (O'Reilly *et al.* 1997). Endostatin has been shown to bind heparan sulphate along with a number of other components of the vascular basement membrane (Blackhall *et al.* 2003; Marneros and Olsen 2005). The exact details of the molecular interaction between endostatin and HS, is discussed in more detail under “section 11” of this introduction.

4. Limitations of core protein analyses

Studies into the expression of both cell membrane and extracellular matrix core proteins, yield useful information on the type of HSPGs present at the cell surface. These studies may also demonstrate how expression patterns alter. However, in the absence of a structural analysis of the HS chains linked to the protein core, limited information can be gained on their biological role. This is particularly the case with glypican core proteins.

Since glypicans have no transmembrane or cytoplasmic domains with which to make intracellular connections, any signalling functions must be indirect. Glypicans, via their HS chains, have the potential to capture any number of extracellular ligands, which may then signal through transmembrane receptors; variations in HS structure being able to influence the range of ligands with which a glypican or other HSPGs may interact.

5. HS chain structure is cell type specific

For syndecan-1 (Kato *et al.* 1994; Sanderson *et al.* 1994) and perlecan (Knox *et al.* 2002) it has been shown that the structure of the attached HS chains are specific for a particular cell type. Other studies have also detailed structural differences between HS isolated from

different cell types and tissues. However, within the same cell, syndecans-1 and -4 have been shown to carry similar HS chains with similar ligand binding properties (Zako *et al.* 2003). Likewise, syndecan-4 and glypican-1 proteoglycans extracted from the same cell type, have also been shown to possess HS chains with no major structural differences (Tumova *et al.* 2000). These extracted proteoglycans were also shown to have similar affinities for an extracellular ligand, the Hep II domain of fibronectin (Tumova *et al.* 2000).

Together these studies support the hypothesis that the synthesis of HS chains and core proteins of individual HSPGs, are independently regulated in any particular cell type and possibly tailored for the desired role(s) of that cell.

6. Biosynthesis

6.1 HS biosynthesis

Many different cell types produce HS chains with many different primary structures. Therefore there is room for a great deal of variability in the way HS chains are synthesised. However, essential to the formation of HS regardless of primary sequence, is a range of biosynthetic enzymes (Esko and Selleck 2002).

These enzymes consist of multiple glycosyltransferases, sulphotransferases and an epimerase. These enzymes also synthesise heparin, a GAG closely related in structure to HS, but much more heavily sulphated. Many of these enzymes have now been purified, molecularly cloned and their expression patterns studied. From early work on the fundamental stages of HS/heparin synthesis using a mouse mastocytoma cell free system

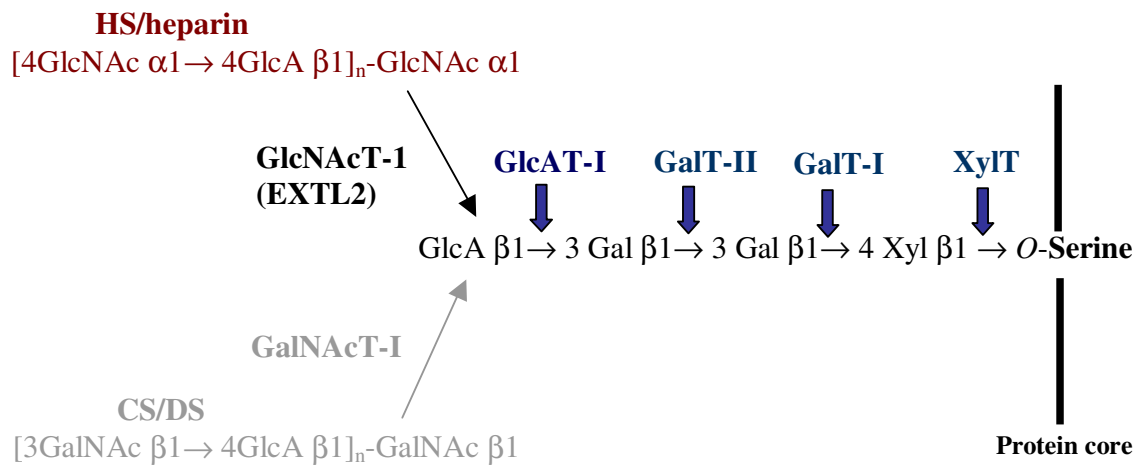
a lot is known about the order of enzyme reactions and enzyme specificity (Lindahl *et al.* 1998).

6.2 Chain initiation

HS synthesis initiates with the transfer of xylose from UDP-xylose by xylotransferase (XT) to specific serine residues. Attachment of two D-galactose (Gal) residues by galactosyltransferases I and II (GalTI and GalTII) and GlcUA by glucuronosyltransferase I (GlcATI) completes the formation of a core protein linkage tetrasaccharide, $\beta\text{GlcUA-1,3}\rightarrow\beta\text{Gal-1,3}\rightarrow\beta\text{Gal-1,4}\rightarrow\beta\text{Xyl}$ (see Figure 6). Xylose attachment to the core protein is thought to occur in the ER with further assembly of the linkage region and the remainder of the chain occurring in the golgi apparatus.

Figure 6 HS chain initiation

The pathways for HS and CS/DS biosynthesis diverge after the formation of a common core protein linkage tetrasaccharide (see text). Four separate enzymes XylT, GalT-I, GalT-II and GlcAT-I are responsible for the formation of this tetrasaccharide. The next enzyme to act either GlcNAcT-I or GalNAcT-I directs synthesis to HS/heparin or CS/DS (see text). This figure is reproduced with the permission of the author from the thesis of Goodger (Goodger 2003).



6.3 Chain elongation, HS vs CS/DS

The initial linkage tetrasaccharide is the same for the attachment of glucosaminoglycans (HS/heparin) and galactosaminoglycans (CS/DS) (Esko and Selleck 2002). The amino acid sequence Ser-Gly-X-Gly is the minimal sequence for xylosylation (Bourdon *et al.* 1985). After the formation of the linkage tetrasaccharide, the next step in chain formation is crucial to determining which of these two types of chain are synthesised. This is carried out by HexNAc transferases that were thought to differ from enzymes involved in chain polymerisation (Rohrman *et al.* 1985; Fritz *et al.* 1994), although it has also been reported that a GlcNAc transferase may initiate and also elongate HS chains (Kim *et al.* 2001). The attachment of GlcNAc results in the formation of HS/heparin, while the attachment of GalNAc results in the formation of CS/DS.

Studies into the protein sequence surrounding the attachment sites for HS and CS in syndecans-1 and -4, suggest this is key to determining the type of GAG chain synthesised (Kokenyesi and Bernfield 1994; Shworak *et al.* 1994). Synthetic addition of HS to both serine residues in a serine glycine repeat (SGSG) is favoured when it is flanked by acidic residues towards the N-terminal and hydrophobic residues towards the C-terminal. Analysis of three CS attachment sites in syndecan-1 showed that all three shared a similar sequence, a single sequence of serine-glycine flanked on both sides by acidic residues (Kokenyesi and Bernfield 1994). The HexNAc transferase enzymes may bind to the tetrasaccharide linkage region and a domain may then interact with the core protein. The local conformation of the core protein or peptide domains adjacent to the attachment site will determine, which HexNAc transferase enzyme then adds the first amino sugar.

Just as the conserved sequence of a HS attachment site between protein cores does not guarantee that both proteins will be glycosylated within this sequence, neither does the presence of an attachment motif within an individual protein. Syndecan-1 has five potential GAG attachment sites, but contains different numbers of chains in different cells (Bernfield *et al.* 1992).

6.4 HS chain elongation

After attachment of the first GlcNAc residue elongation of the tetrasaccharide linker continues by the stepwise addition of GlcUA and GlcNAc residues. GlcUA and GlcNAc residues are transferred from their respective UDP-sugar nucleotides and attached to the non-reducing end of the growing chain. This is carried out by one or more related enzymes, whose genes are members of the exostoses (EXT) gene family of tumour suppressors.

The EXT family consists of six members in total, EXT 1-3, and three related genes termed EXT like, EXTL1-3. A GlcNAc transferase activity has been detected for the gene products of all three EXTL genes (Kim *et al.* 2001). However, their role *in-vivo* is unknown, as mutants lacking these enzymes have not been produced. However, murine mutants have been produced that lack either EXT1 or EXT2 gene products (Lin *et al.* 2000; Stickens *et al.* 2005). Mice with either EXT1 (Lin *et al.* 2000) or EXT2 (Stickens *et al.* 2005) knock out mutations lack HS and die during embryogenesis as a result of developmental abnormalities.

Mutations at the EXT 1-3 gene loci in humans leads to the development of hereditary multiple exostoses, which are characterised by multiple cartilaginous tumours, suggesting

an important role for HS in bone growth (McCormick *et al.* 2000). Sequence comparisons between the various EXTs suggest they contain two catalytic domains, each of the domains being independently responsible for GlcNAc or GlcUA transferase activity (Esko and Selleck 2002). It is not yet understood how chain elongation is coordinated between this group of enzymes. The cell free mastocytoma model was used to show that HS chain polymerisation can proceed in the absence of any subsequent chain modification reactions (Lind *et al.* 1998).

Chain modification

6.5 Availability of PAPS

As the chain polymerises, it undergoes a series of modification reactions carried out by four classes of sulphotransferases and an epimerase. The availability of the sulphate donor PAPS is crucial to the activity of the sulphotransferases (Silbert 1967; Silbert 1967). Therefore, both the enzyme systems responsible for PAPS formation, and the PAPS transmembrane transport system that makes PAPS available in the golgi lumen for HS biosynthesis (Abeijon *et al.* 1997), will influence the final structure of the HS/heparin chain.

6.6 N-deacetylation/N-sulphation

Until comparatively recently, the first polymer modification was thought to be the N-deacetylation/N-sulphation of GlcNAc residues into GlcNS (see Figure 7 for a general overview of HS biosynthesis). This was thought to be prerequisite for all subsequent modification reactions and is carried out by one or more members of a family of four

GlcNAc N-deacetylase/N-sulphotransferase enzymes (NDSTs). Recent NDST genetic knockout studies have called the prerequisite requirement for N-deacetylation/N-sulphation into doubt (see 6-O-sulphation section below).

In early studies, it was shown that NDST enzymes could recognize and act on any N-acetylated residue in the forming polymer (Lindahl *et al.* 1973; Hook *et al.* 1975). Therefore the modification of GlcNAc residues should occur randomly throughout the chain. However, HS N-sulphated residues are mainly grouped together (see HS-domain structure, discussed later) and separated by regions of N-acetylation where GlcNAc remains unmodified.

6.7 Differences between heparin and HS biosynthesis

Heparin is much more highly modified by N-sulphation than HS (~85% GlcNS in heparin versus ~40-50% in HS (Gallagher and Walker 1985)). *In-vivo* heparin is stored in the secretory granules of mast cells. NDST2 was originally cloned from mast cells (Eriksson *et al.* 1994) and was found to have a much higher N-deacetylase/N-sulphotransferase activity than NDST1, which had earlier been cloned from rat liver (Hashimoto *et al.* 1992). A mutation in NDST2 resulted in a selective decrease in the sulphation of mast cell heparin (Forsberg *et al.* 1999). Overexpression of NDST2 in cell lines increased the level of GlcNS from 40% to 80%, making HS the cells produced more heparin like (Cheung *et al.* 1996). All these observations suggest that NDST2 is primarily responsible for the biosynthesis of heparin and that NDST1 is responsible for the biosynthesis of HS. NDST knock out mice have also been shown to have a defect in heparin synthesis linked to the failure of mast cell differentiation (Forsberg *et al.* 1999). However, a CHO cell

mutant defective for NDST1 but not NDST2 can still produce HS containing 25% GlcNS (Aikawa and Esko 1999). Suggesting that rather than being solely responsible for the biosynthesis of heparin, NDST2 also plays a role in the formation of HS. This is further supported by the finding that mRNAs for both enzymes, are abundantly expressed by all tissues and cells that make HS and heparin, all be it that cells producing heparin have a predominance of NDST2 (Aikawa *et al.* 2001). Thus it would seem that both the levels of enzyme expression and the specific properties of the NDST isoform apparently determine the extent of N-sulphation. It is not yet known what it is about NDST2 that enables it to catalyse the formation of the much more highly sulphated heparin polymer.

6.8 NDST3 and 4

Little is known about the significance of the expression pattern of NDST3 and 4 or their activities (Esko and Selleck 2002). NDST3 has a greater deacetylation activity than sulphotransferase activity and NDST4 has a greater sulphotransferase activity than deacetylation activity. Expression of both enzymes is thought to be restricted to adult brain and foetal tissues (Aikawa *et al.* 2001).

6.9 Generation of GlcNH₂ residues

Due to the N-deacetylase and N-sulphotransferase activity being carried out by the same enzyme, N-sulphation is normally tightly coupled to N-desulphation. GlcNH₂ residues resulting from an apparent uncoupling of the two activities have been found in heparin and HS (Toida *et al.* 1997; Westling and Lindahl 2002). The activity of NDST3 is a prime candidate for the generation of such residues.

Murine embryonic stem cells in which the NDST1 and NDST 2 genes are knocked out, still produce HS chains with a number of free amine groups (Holmborn *et al.* 2004). In these cells NDST3 mRNA can be detected at low levels, further supporting this as the enzyme responsible for the generation of GlcNH₂ residues. However, the low levels at which it is detected, and the restricted tissue expression pattern of this protein, calls this into question. Possibly, there may exist an as yet un-discovered de-acetylase enzyme.

It has been shown that a GlcNH₂ unit forms part of the binding site in HS for herpes simplex virus glycoprotein D (Shukla *et al.* 1999) and is implicated in the binding of L-selectin (Norgard_Sumnicht and Varki 1995), suggesting the occurrence of GlcNH₂ may be under tight regulatory control.

6.10 Epimerisation and 2-O- sulphation

Epimerisation is catalysed by one enzyme, the GlcUA C5 epimerase. This enzyme epimerises certain GlcUA units to IdoUA. Substrate recognition requires that the GlcN residue linked to the non-reducing side of a potential GlcUA target be N-sulphated (Jacobsson *et al.* 1984). Studied in isolation the GlcUA C5-inversion is freely reversible. Prolonged incubation of either O-desulphated heparin (IdoUA, ~90% of total HexUA) or N-sulpho-heparosan, (GlcUA, 100% of total HexUA) with soluble epimerase yielded similar products with 30-35% IdoUA, representing equilibrium (Hagner Mcwhirter *et al.* 2000). *In-vivo*, this equilibrium is clearly not established. Heparin chains may contain ~90% IdoUA (Esko and Selleck 2002). It is the action of another enzyme, uronosyl-2-O-sulphotransferase (2OST), and the resulting 2-O-sulphation of IdoUA that prevents its re-conversion to GlcUA (Valla *et al.* 2001).

This enzyme acts on both IdoUA and GlcUA units but sulphates IdoUA at a faster rate, under most conditions, when substrates containing both units are provided (Rong *et al.* 2001). In heparin 2-O-sulphation nearly always modifies IdoUA residues. In HS the extent of 2-O-sulphation ranges from 50-90% of the available IdoUA residues and as a result, IdoUA residues without 2-O- modification are present in the mature HS chain. It may be that the speed at which the HS chain is synthesised, and moved away from the golgi, that prevents the establishment of an equilibrium, and conversion of IdoUA units that escaped 2-O- modification back to GlcUA. For heparin, bio-synthesis can occur in less than 1 min (Hook *et al.* 1975).

6.11 6-O- sulphation

Three glucosaminyl 6-O-sulphotransferases (6OSTs) have been identified. In a mutant cell line completely lacking 2-O-sulphotransferase activity, HS chains were found to contain a higher level of 6-O-sulphation than HS from normal cells (Merry *et al.* 2001). This suggests that 2-O-sulphation may play a role in regulating the action of the 6-O-sulphotransferase enzymes and occur prior to the action of the 6OST enzymes. However, the fact that 6-O-sulphation occurred at all in the absence of 2-O-sulphation suggests that a strict order of 2-O- then 6-O- sulphation may not exist. In support of this, the different 6-OST isoforms were found to have preferences for the sulphation of GlcNS linked to either IdoUA or GlcUA but the activity of no isoform was found to be reliant on the presence of 2-O-sulphation (Habuchi *et al.* 2000). More recently in a mutant cell line completely lacking 6-O-sulphotransferase activity, HS chains were found to contain a higher level of 2-O-sulphation than HS from normal cells (Kamimura *et al.* 2006).

In murine embryonic stem cells in which the NDST1 and NDST2 genes were knocked out, 6-O-sulphation was never the less detected (Holmborn *et al.* 2004). This result places 6-O-sulphation at an early stage of HS bio-synthesis.

While 6-O-sulphation is certainly not essential in order for N-sulphation to occur, it may influence N-sulphation levels and positioning (see HS domain structure, discussed later). Perhaps not surprisingly, given that N-sulphation is not required in order for 6-O-sulphation to occur, GlcNAc units that are 6-O-sulphated are frequently found in the mature HS chain.

6.12 3-O- sulphation

At least five glucosaminyl 3-O-sulphotransferases (3OSTs) exist. The action of four of these, 3OST1, 3OST2, 3OST3A and 3OST3B have been characterised. The 2-O-sulphation of the uronic acid, either IdoUA or GlcUA, adjacent to the target GlcNS(6S) prevents the action of 3OST1 (Zhang *et al.* 2001) while it is required for the substrate recognition and action of the other 3OST isoforms (Liu *et al.* 1999; Liu *et al.* 1999). The 3OST1 enzyme requires the target GlcNS residue to be 6-O-sulphated. However the action of other members of the family does not require GlcNS to be 6-O-sulphated (Esko and Selleck 2002).

The 3OST enzymes are of key importance in the biosynthesis of at least two characterised ligand binding sites, antithrombin III (Lindahl *et al.* 1984) and herpes simplex virus glycoprotein D (Shukla *et al.* 1999).

6.13 Sulfs

Although not strictly bio-synthetic enzymes the extracellular sulphatases (Sulfs) play an important role in modifying HS structure. The Sulfs are a family of 6-O-endosulphatases. The first member of this family to be discovered was termed QSulf1; “Q” signifying Quail, the origin of the cDNA library from which the gene for this enzyme was identified (Dhoot *et al.* 2001). Subsequent homologs of this enzyme have been found in a range of species from *Caenorhabditis elegans* (*C. elegans*), to mice and humans. The original nomenclature for the quail derived enzyme has stuck and MSulf1 and HSulf1 refer to murine or human derived protein isoforms respectively.

A related enzyme Sulf2 has been identified in vertebrates and more recently by a re-screening of the quail cDNA library (Ai *et al.* 2006). To date the QSulf1 and QSulf2 isoforms are the most well characterised.

Both enzymes exhibit the same substrate specificity towards disaccharides of the sequence -IdoUA(2S)-GlcNS(6S)-, removing the 6-O-sulphate group from the GlcNS monosaccharide (Ai *et al.* 2006). This modification could have profound consequences for the binding of a number of protein ligands (see FGF-2 discussed later). In contrast to the true HS bio-synthetic enzymes, QSulf1 does not remain localised to the golgi apparatus and is transported to the cell surface together with HSPGs. It has been shown that its sulphatase activity exists both at the cell surface and while temporarily localised to the golgi (Ai *et al.* 2003). The discovery of this family of enzymes adds an additional layer of complexity to the already complex picture of HS biosynthesis.

6.14 Biosynthetic enzyme expression patterns

There is a differential tissue expression pattern of the various 6OSTs and 3OSTs (Shworak *et al.* 1999; Habuchi *et al.* 2000). Brain tissue contains large amounts of 6OST2 and 3OST1, 2 and 4 while liver contains mostly 6OST1 and 3OST3A and 3B. Variations in the level of expression of 6OST may also influence the level of N-sulphation and in turn 2-O-sulphation. Changes in the level of Sulf expression may also influence HS primary structure. It is conceivable that individual cell types express different levels of different isoforms of the HS biosynthetic enzymes, allowing the creation of unique patterns of N-, 2-O-, 6-O- and 3-O-sulphation.

6.15 Chain heterogeneity

In the absence of a 'code' that specifies the synthesis of particular saccharide sequences, a great deal of research has been undertaken in order to characterise the action of the enzymes involved in HS biosynthesis (as detailed above). The modifications that occur are non-random and to an extent interdependent. To add further complexity to the system they are also quantitatively incomplete.

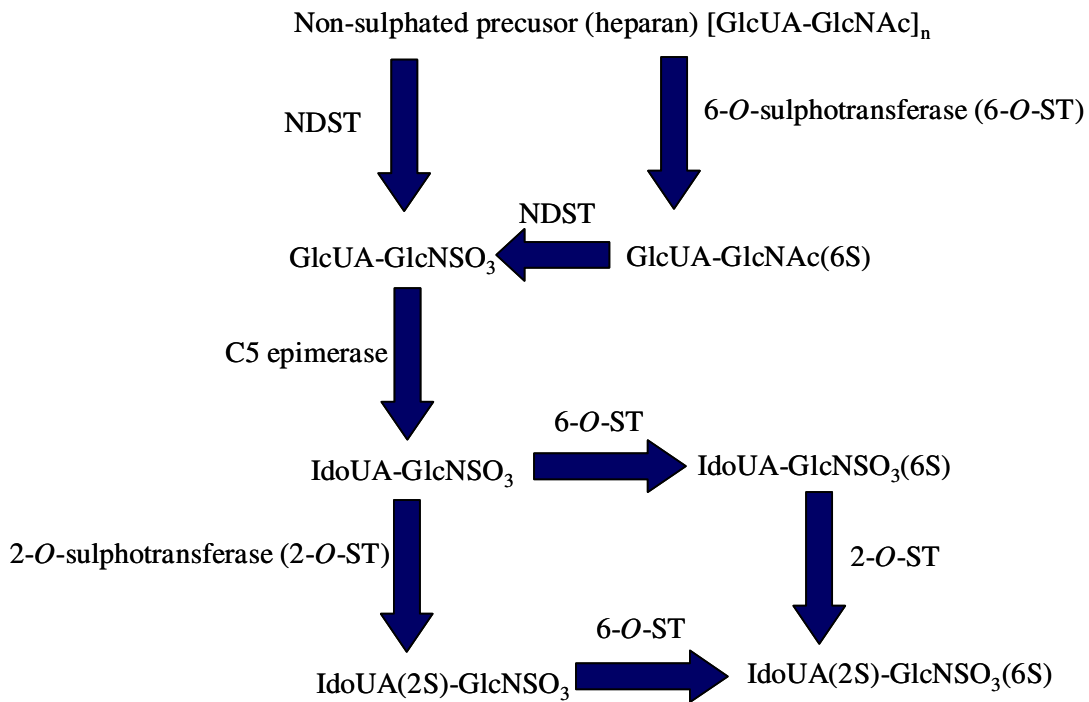
In mature HS, possible substrates for the O-sulphotransferase enzymes can be found that have escaped sulphation (bearing in mind that some 6-O-sulphates may have been removed by the action of the Sulf enzymes). The end result is a very high level of structural complexity within any individual HS chain.

A simplified scheme for HS biosynthesis is given in Figure 7.

Figure 7 A simplified scheme of heparan sulphate biosynthesis

In mature HS 40-50% of the GlcNAc residues present in the heparan precursor are modified to GlcNSO₃. All other potential modifications to heparan, as shown below, are also quantitatively incomplete and result in heparan sulphate chains with a highly complex primary structure.

Not shown are the rare modifications carried out by the 3OST enzymes, the formation of GlcNH₂ residues and 6-O-sulphatase modifications carried out by Sulf enzymes.



7. De-polymerisation techniques

Either chemical or enzymatic de-polymerisation techniques or a combination of the two underlie the vast majority of analyses carried out on the structure of HS and heparin.

7.1 Enzymatic

The enzymes traditionally used to digest heparin or HS for structural studies are naturally produced by the soil bacterium *Flavobacterium heparinum*. This bacterium is capable of utilizing either heparin or HS as its sole carbon and nitrogen source. In order to do this it produces a range of enzymes such as lyases, glucuronidases, sulphoesterases and sulphamidases (Gallihier *et al.* 1981). It is the lyases that have mainly been used in HS/heparin structural studies. The bacterium produces three lyases, heparinases I, II and III and each has distinct substrate specificities as detailed below (Linhardt *et al.* 1990; Desai *et al.* 1993).

Heparinase I	GlcNSO ₃ (±6S) – IdoUA(2S)
Heparinase II	GlcNSO ₃ /Ac(±6S) – IdoUA(±2S)
or	GlcNSO ₃ /Ac(±6S) – GlcUA
Heparinase III	GlcNSO ₃ /Ac(±6S) – GlcUA/IdoUA (with a preference for GlcUA)

The lyases cleave the HS/heparin chain by a β -elimination mechanism. This action generates an unsaturated double bond between C4 and C5 of the non reducing terminal uronate residue (Linker and Hovingh 1972; Linhardt *et al.* 1988). This C4-C5 unsaturated uronate is termed Δ UA or UA. It is a sensitive UV chromophore (max absorption at 232

nm) and allows the rate of an enzyme digest to be followed as well as providing a convenient method for detecting the fragments produced by enzyme digestion.

Due to their substrate specificities there are a limited number of cleavage sites for each enzyme within heparin/HS chains. An individual enzyme can be used to either cleave the chain at all the possible cleavage sites, or its action can be controlled so that only a partial number of cleavage sites will be acted upon. When used in combination, the enzymes will digest a HS/heparin chain into its component disaccharide units. The various component disaccharides can then be identified by their position of elution on HPLC or capillary electrophoresis. In this way the complete disaccharide composition of the chain can be established.

Unfortunately, by generating an unsaturated uronate residue, asymmetry at C5 of the uronic acid is lost. This asymmetry distinguishes IdoUA from GlcUA and so it is not possible to identify the uronic acid component of the disaccharides. It is only possible to define the proportion of mono-, di- and tri- sulphated disaccharides and not proportions of IdoUA and GlcUA.

7.2 Chemical

Nitrous acid can be used to chemically depolymerise heparin/HS and avoids the loss of uronate asymmetry. Nitrous acid can be used at pH 1.5 or at a higher pH of 4. Under both conditions nitrous acid effects deaminative cleavage of the chain (Shively and Conrad 1976). At both 'high' (4) and 'low' (1.5) pH, deaminative cleavage occurs between GlcNSO₃-HexA units, all be it at a slower rate at the higher pH, but deaminative cleavage

between GlcNH₂-HexUA units only occurs at high pH. The deamination reaction, and therefore chain cleavage, is regardless of O-sulphation carried by either unit.

At low pH deaminative cleavage results in the release of inorganic SO₄, and the conversion of GlcNSO₃ to anhydromannose (aMan). The preference for cleavage of the chain at either GlcNSO₃-HexUA or GlcNH₂-HexUA, depending on pH, can be exploited in HS structural studies (Edge and Spiro 1985). Low pH nitrous acid treatment is also an excellent method to distinguish N-sulphated polysaccharides such as HS and heparin from non N-sulphated polysaccharides such as CS and DS; CS and DS polysaccharides being un-susceptible to nitrous acid cleavage.

8. HS and heparin structural distinct molecules

When HS chains become highly modified, the boundary between what is classed as HS and heparin becomes unclear. It is only through the use of agents that can cleave either HS or heparin at specific linkages, and the subsequent analysis of the fragments generated, that major differences in the structures of the two molecules becomes clear. It is the domain structure of HS that distinguishes it from heparin.

8.1 Domain structure of HS

Heparin is almost completely depolymerised by nitrous acid at low pH to di- or tetrasaccharides (Cifonelli and King 1977). Regions of the HS chain under the same conditions remain intact. A 1985 study analysed the size of the nitrous acid cleavage products of HS from a number of different sources (Gallagher and Walker 1985). It was shown that in the HS analysed most N-sulphated units were present in alternating or

contiguous sequences, and that there was therefore a strong tendency for N-acetylated and N-sulphated disaccharides to be separated, generating domains (see Figure 8). Due to the substrate specificities of the biosynthetic enzymes, epimerised and 2-O-sulphated disaccharides are predominantly located within the N-sulphated domains. These have subsequently been termed S-domains and the regions of consecutive N-acetylation termed N-Ac domains.

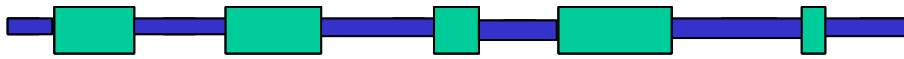
Within the S-domains 50-60% of the hexosamine residues were identified as IdoUA and most of these were 2-O-sulphated. This domain structure is typical of all HS chains analysed subsequently. The disaccharides that form the heparin chain are much more extensively modified and as a result heparin lacks the distinctive domain structure of HS (see Figure 8).

Figure 8 HS/heparin domain structure

The domain structure of HS is made up of blocks of highly modified sequence (S-domains) separated by blocks of unmodified sequence (N-Ac domains).

Heparin is essentially one continuous S-domain interrupted by the odd region of low sulphation.

Heparan sulphate



Heparin



8.2 Alternating sequences and K5 lyase

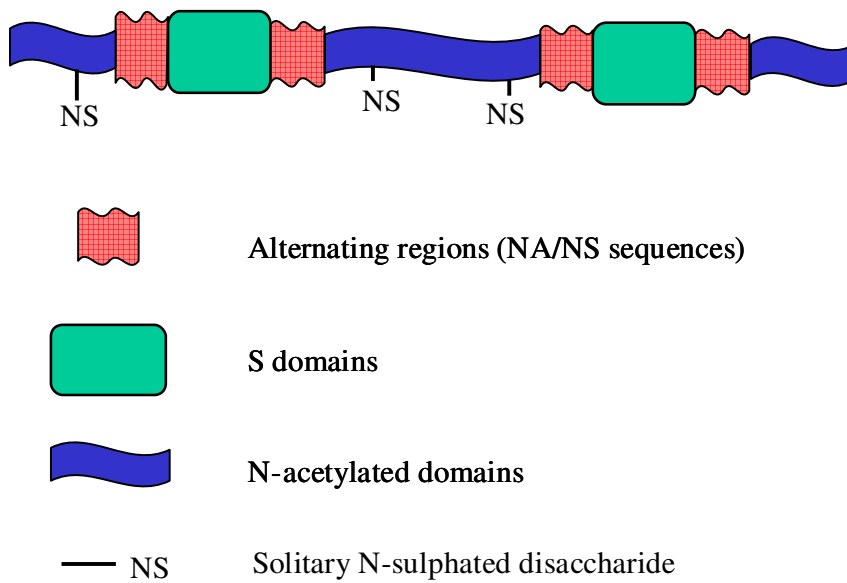
Typically blocks of N-acetylation and N-sulphation were also detected in HS purified from a number of bovine organs by Maccarana *et al.* (Maccarana *et al.* 1996). This study also demonstrated that a great deal of variability exists within blocks of alternating N-acetylated and N-sulphated disaccharides, termed NA/NS or alternating domains previously identified by Gallagher and Walker (1985) (Gallagher and Walker 1985). These sequences frequently contained none sulphated IdoUA residues and more than half of the total number of 6-O-sulphate groups within the chain. These alternating domains were presumed to occur at the boundary between the S-domains and N-acetylated domains. This presumption was proved correct by the structural analysis of oligosaccharides generated by the use of the K5 lyase enzyme (Murphy *et al.* 2004).

This enzyme cleaves the HS chains where contiguous un-interrupted regions of eight or more N-acetylated disaccharides occur. In the resulting oligosaccharides, alternating sequences were only found in conjunction with S-domain type sequences. This led to a refinement of the model for the domain structure of HS that was presented in Figure 8 (see Figure 9).

In this refined model S-domains are flanked on either side by alternating domains. From the initial K5 lyase study it was not possible to establish whether this is actually the case. It may be that alternating domains only occur to one side or are more extended to one side of the S-domains.

Figure 9 A refined model of the molecular structure of HS.

The S-domains and the N-acetylated sequences are arranged in an alternating and equidistant manner, with solitary N-sulphated disaccharides scattered within the NAc-domains. Flanking the S-domains and forming the border between the S-domains and NAc-domains are the NA/NS domains. Only two S-domains are shown.



9. Ligand binding

All ligands that have been found to bind heparin also bind HS. The S-domains within HS resemble heparin in structure, although less 6-O-sulphated. The majority of protein interactions analysed have been found to occur via the S-domains (Gallagher and Lyon 2000; Coombe and Kett 2005).

The number of protein ligands that have been shown to bind HS/heparin is considerable. Including all members of individual protein families that interact with HS/heparin, the number of protein ligands is probably around 30 to 40. The discussion below focuses on specific proteins. For these proteins the HS/heparin structural requirements for the interaction have been particularly well characterised, or their interaction with HS illustrates the role HS domain structure may play in ligand binding.

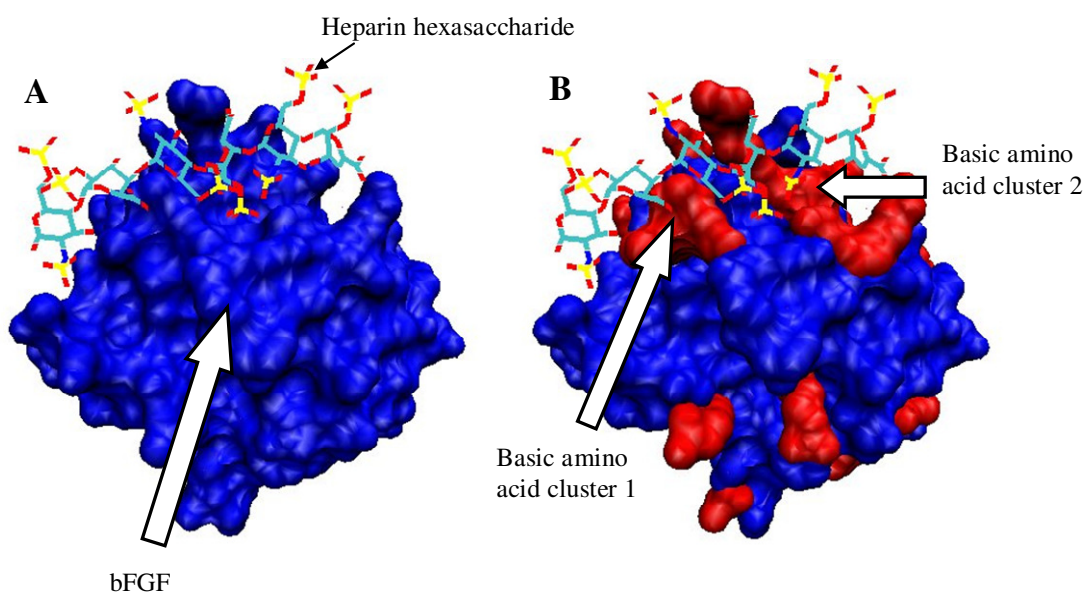
9.1 Protein binding sites for HS

The binding sites for HS/heparin within a protein ligand are usually characterised by extensive regions of positive charge on the surface of the protein. These result from the close positioning of basic amino acid side chains in the secondary or tertiary structure of a protein or protein complex. As an illustration of this, two such regions of positive charge can be seen on the surface of a FGF-2 monomer in complex with a heparin hexasaccharide (see Figure 10).

Figure 10 The crystal structure of a heparin hexasaccharide in complex with a FGF-2 monomer.

The coordinates for this crystal structure were taken from the protein data base, code 1BFC (Faham *et al.* 1996).

For both models A and B the molecular surface of the FGF-2 protein was calculated using the surf algorithm as implemented in the VMD molecular graphics program (Humphrey *et al.* 1996). In model A the entire surface of the protein is coloured blue, in model B all basic amino acid residues at the protein surface are marked out in red. In model B a number of electrostatic contacts can be seen between two clusters of basic amino acids at the protein surface and the hexasaccharide structure.



It can be clearly seen in Figure 10 that the heparin binding site consists of two regions of basic amino acids, separated by a region of neutral amino acids over which the heparin oligosaccharide spans. This type of binding site makes it difficult to definitively identify a consensus amino acid sequence within proteins for HS/heparin binding. Amino acid sequences separated by great distances within the primary structure of a protein, may fold to form a HS/heparin binding site. Never the less the sequences XBBXBX and XBBBXXBX have been suggested as consensus sequences to which HS/heparin binding may occur (Cardin and Weintraub 1989). “B” representing an amino acid with a basic charge (usually arginine or lysine) and “X” an uncharged or hydrophobic amino acid. In most cases these sequences will not form the whole, but part of a HS binding site.

10. A footprinting approach

A ‘footprinting’ approach has been adopted in the study of HS interactions with a number of proteins (Lortat_Jacob *et al.* 1995; Stringer and Gallagher 1997; Spillmann *et al.* 1998; Stringer *et al.* 2002). Footprinting involves enzymatic depolymerisation of the HS chain in the presence of a high concentration of ligand. Regions of the HS chain that are then protected from enzymatic cleavage by ligand binding are structurally examined. The first protein for which a footprinting approach was adopted was interferon- γ (Lortat_Jacob *et al.* 1995).

10.1 Interferon- γ

Interferon- γ (IFN- γ) is a cytokine secreted by many cells involved in an immune response (Schroeder *et al.* 2004). The first stage of biological activity involves its binding to a

specific cell membrane receptor, IFN- γ R (Sadir *et al.* 1998). In contrast to many other HS binding ligands, where HS binding promotes biological activity. The binding of IFN- γ to HS actually inhibits its biological activity (Sadir *et al.* 1998).

10.2 Interferon- γ structure

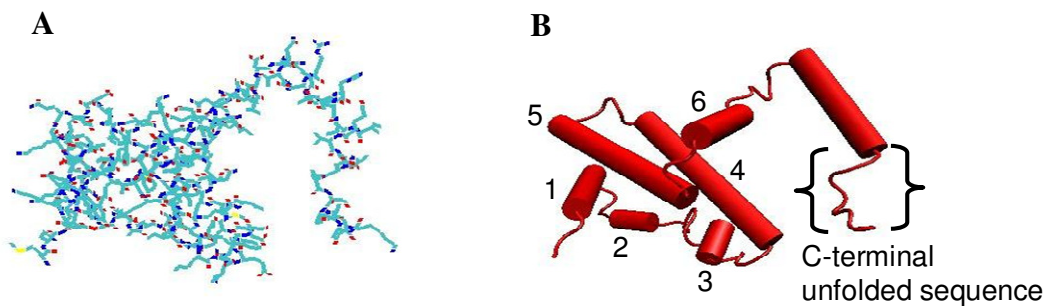
The IFN- γ monomer is comprised of a core of six α -helices and an extended, unfolded region towards the C-terminal of the protein (see Figure 11). It is thought to be biologically active in a dimeric form, the dimer forming by anti-parallel interlocking of the two monomers (Thiel *et al.* 2000) (see Figure 11).

Figure 11 The crystal structure of human interferon- γ

The coordinates for this crystal structure and the crystal structure shown subsequently in Figure 12 were taken from the protein data base, code 1FG9 (Thiel *et al.* 2000). Models were then generated using the VMD molecular graphics program (Humphrey *et al.* 1996).

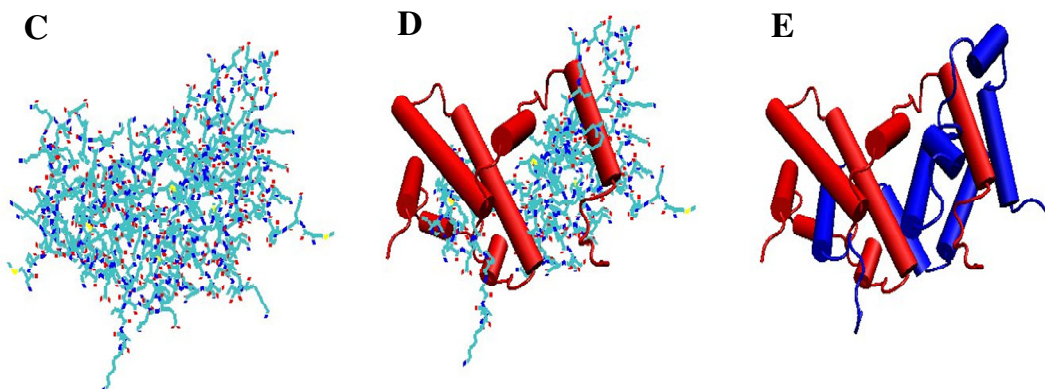
Monomeric interferon- γ

In model A a single protein monomer is shown by a line representation. In the cartoon representation shown in model B the six α -helices that comprise the core of the structure are numbered.



Dimeric interferon- γ

The biologically active IFN- γ dimer is formed through the anti-parallel interlocking of the two monomers as shown in models C-E



The models shown in Figure 11 are all truncated at their C-termini by 17 amino acids; full length IFN- γ being 143 amino acids in length, the models being 126 amino acids in length.

10.3 Interferon- γ , HS binding sites

Affinity for HS resides solely within the deleted sequence of 17 amino acids (Vanhaverbeke *et al.* 2004). In this sequence two clusters of basic amino acids occur, ¹²⁵KTGKRKR¹³¹ and ¹³⁷RGRR¹⁴⁰, termed D1 and D2 respectively. HS interacts with both of these sequences, with the D1 sequence contributing most to the interaction (Lortat-Jacob and Grimaud 1991).

10.4 The domain structure of Interferon- γ binding oligosaccharides

Using the footprinting approach, the binding of dimeric IFN- γ was found to protect a HS fragment 21-24 disaccharide units in length. This consisted of two S-domains (each 3-4 units in length) separated by a region rich in N-acetylated residues (15-16 units in length). Each S-domains is thought to interact with a separate HS binding site, one at each C-terminal of either protein monomer, with the N-acetylated region acting to span the distance between the two binding sites (Lortat-Jacob *et al.* 1995).

10.5 HS binding prevents receptor binding possibly by steric hinderance

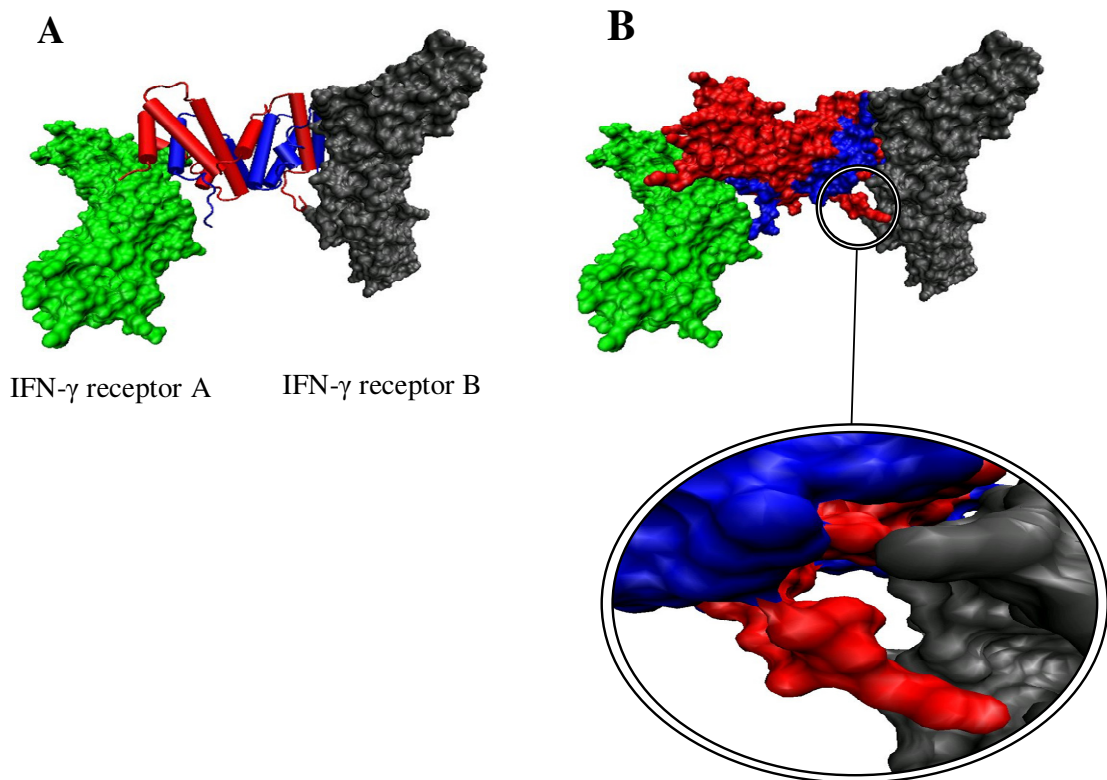
If the models shown in Figure 11 are expanded to include the co-ordinates of two of the three IFN- γ receptor molecules also included in the 1FG9 protein data base file (pdb), it

becomes clear why HS binding inhibits the biological activity of IFN- γ (Sadir *et al.* 1998)
(see Figure 12).

Figure 12 The crystal structure of human interferon- γ in complex with two identical interferon- γ receptor molecules.

For both models the molecular surface of the two IFN- γ receptors, and for model B the IFN- γ dimer, was calculated using the surf algorithm as implemented in the VMD molecular graphics program (Humphrey *et al.* 1996).

In model A the IFN- γ dimer is shown in a cartoon representation. The enlargement shows the close proximity of the red C-terminal domain of one IFN- γ monomer and the grey receptor molecule. In this close proximity a number of electrostatic interactions may occur between the two molecules (see text for exact residues involved). Similar interactions may occur between the C-terminal domain of the blue IFN- γ monomer and the receptor shown in green. In model B the view of this interaction is obscured by the front edge of the receptor molecule and is not enlarged.



It should be noted that the authors of the 1FQ9 pdb file, note that the electron density beyond residue 120 of each IFN- γ was weak (Thiel *et al.* 2000). The enlarged C-terminal interactions between lysine 125 and a number of glutamic acids (173 and 175) within the receptor protein, may represent one of many partially populated conformational states. Despite this, it can be seen how HS binding to an extended C-terminal domain may compete with, and block receptor binding for the same region of IFN- γ .

10.6 Biological implications of HS- IFN- γ binding

The functional significance of the HS- IFN- γ complex is unclear. It has been shown that the presence of the D1 amino acid sequence, in the absence of HS, increases the rate at which IFN- γ -IFN- γ R complexes form (Sadir *et al.* 1998). Interaction between the D1 sequences and receptor molecules may be the first step in complex formation. Binding of HS protects the D1 sequence from proteolytic cleavage (Lortat-Jacob and Grimaud 1991), possibly saving it for a later interaction with a receptor molecule. After protection, how IFN- γ or specifically the D1 sequences are then freed from HS binding in order to interact with receptor molecules remains to be determined.

11. Other cytokines

The most well known function of the cytokines is their regulation of leukocyte migration, but they also play a role in controlling other biological functions such as development, angiogenesis and neuronal patterning to name a few (Lortat_Jacob *et al.* 2002). Under conditions required for structural studies, like IFN- γ , other cytokines are dimeric or tetrameric in structure. *In-vivo* it is unclear whether it is the multimeric or monomeric

form of the protein that binds cell surface receptors as the individual monomeric units may dissociate faster than receptor binding occurs. However, a role has been suggested in the stabilisation of the multimeric state for HS at the cell surface; multimerisation appears to be strongly enhanced by the presence of HS (Lortat_Jacob *et al.* 2002).

11.1 The domain structure of chemokine binding oligosaccharides

Similar fragments of variable length and sulphation to the fragments with high affinity for IFN- γ , have been isolated using a footprinting approach from cytokines such as platelet factor 4 (PF-4) (Stringer and Gallagher 1997), interleukin 8 (IL-8) (Spillmann *et al.* 1998) and macrophage inflammatory protein 1 α (MIP1 α) (Stringer *et al.* 2002). It has been suggested that binding to this type of extended HS sequence may be a common feature of multimeric cytokines (Gallagher and Lyon 2000).

11.2 The role of sulphation patterns within IFN- γ and cytokine binding oligosaccharides

Little is known about the specific sulphation features of these saccharides, only a general overview of their organisation into various sulphation domains. Within these oligosaccharides the function of the N-Ac domain may be to space the S-domains correctly for binding to multiple binding sites at the protein surface to occur. Little is known about how variations in the sulphation patterns within the NA/NS domains or the S-domains themselves effect protein binding.

Endostatin is another protein that binds to a similar type of HS sequence. However in this case many of the saccharide structural features that influence endostatin binding have been examined.

12. Endostatin

As mentioned earlier under the general introduction to collagen XVIII, endostatin is derived from the C-terminal of collagen XVIII by proteolytic cleavage (O'Reilly *et al.* 1997). Collagen XVIII/endostatin is a common component of almost all epithelial and endothelial basement membranes.

12.1 The oligomerisation state of endostatin influences its biological activity

At the C-terminal of collagen XVIII lies a 38 kDa NC1 domain ("NC" for None Collagenous). This domain contains the 20kDa endostatin domain but also a trimerisation domain. Without cleavage of the so called 'hinge region' in between the trimerisation domain and the start of the endostatin domain, the endostatin domain is trimerised (Sasaki *et al.* 1998). Cleavage within the hinge region by elastase and/or cathepsin L releases monomeric endostatin which has been shown to have an anti-angiogenic effect *in-vitro* and *in-vivo* (O'Reilly *et al.* 1997). Opposing this, trimeric endostatin is a potent motogen (Kuo *et al.* 2001) and this biological activity could actually be described as pro-angiogenic.

The loss of endostatin, due to genetic mutation in mice and humans, does not lead to a major increase or decrease in angiogenic activity (Marneros and Olsen 2005). Suggesting

endostatin is one of many pro-angiogenic and anti-angiogenic factors that work together to regulate this process.

12.2 HS binding sites within endostatin

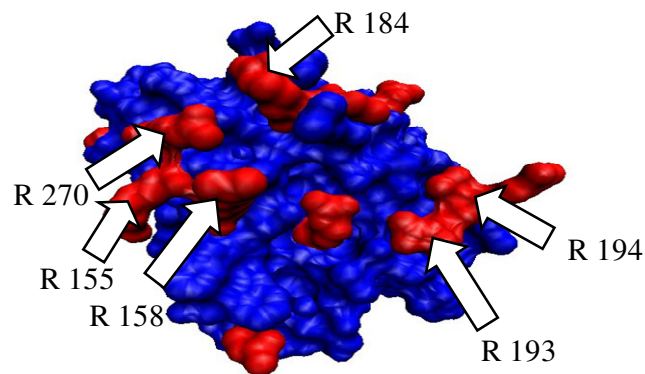
The HS component of a glypican proteoglycan has been identified as an endothelial cell surface binding site for endostatin (Karumanchi *et al.* 2001). Six arginine residues, R155, R158, R193, R194, R184 and R270, have been identified as being crucial for HS/heparin binding to endostatin; loss of either of these groups resulting in a loss of heparin binding (Sasaki *et al.* 1999).

Within the tertiary structure of endostatin these arginine residues form two distinct clusters. R155, R158, R184 and R270 forming one cluster and R193 and R194 forming the other (see Figure 13). These two clusters are in marked contrast to the single cluster observed at the protein surface of FGF-2 (see Figure 10).

Figure 13 The crystal structure of human endostatin.

Coordinates for this structure were taken from the protein database file code 1KOE (Hohenester *et al.* 1998).

The molecular surface of the molecule was calculated as in Figure 10. Basic amino acid residues are coloured red. The six specific arginine residues crucial for HS/heparin binding are labelled.



12.3 The domain structure of endostatin binding oligosaccharides

As can be seen in Figure 13, in order to interact with both arginine clusters a single HS/heparin oligosaccharide must span across a considerable area of the protein surface. This is reflected in a minimal length, of dp12, for heparin oligosaccharides found to bind efficiently to endostatin (Sasaki *et al.* 1999). Subsequently it has been shown that binding oligosaccharides need not be composed of contiguously N-sulphated disaccharides. Both Kreugar *et al.* (Kreuger *et al.* 2002), and Blackhall *et al.* (Blackhall *et al.* 2003), have shown that HS derived oligosaccharides, composed of two short S-domain type sequences separated by NA/NS and N-acetylated type sequences, bind with a similar affinity to heparin derived oligosaccharides.

12.4 Specific sulphation requirements for endostatin binding oligosaccharides

Several apparently conflicting reports have been published as to the importance of HS sulphation on endostatin binding. Blackhall *et al.* have shown, through a filter binding study, that HS completely devoid of 2-O-sulphation displays an identical affinity for endostatin when compared to HS containing a normal level of 2-O-sulphation (Blackhall *et al.* 2003). Contradictory to this Ricard-Blum *et al.* have shown, through a BIAcore study, that de-2-O-sulphated heparin displays a marked decrease in affinity for endostatin when compared to intact heparin (Ricard-Blum *et al.* 2004). Endostatin also fails to bind a melanoma cell line displaying a 90% reduction in normal 2-O-sulphation levels (Karumanchi *et al.* 2001). To date however, perhaps the most comprehensive study on the role of HS sulphation on endostatin binding has been carried out by Clamp *et al.* (Clamp *et al.* 2006).

In their studies a panel of Chinese hamster ovary cells (CHO cells) with specific defined defects in their HS biosynthetic processes were utilised. Oligomeric endostatin was found to induce morphological changes in the same CHO cells producing normal cell surface HS. By comparing the morphological features of mutant and normal CHO cells in response to added oligomeric endostatin, the effect of specific HS structural defects could be investigated. In this study it was found that 2-O-sulphation was not required in order for oligomeric endostatin to elicit a morphological response from mutant CHO cells. A 50% reduction in N-sulphation was also not detrimental to an endostatin morphological response. Within this reduced N-sulphated cell line an increase in the proportion of O-sulphate groups found within NA/NS type sequences was detected. Previous studies had already implicated these type of sequences in endostatin binding (Kreuger *et al.* 2002; Blackhall *et al.* 2003). Endostatin is the first protein for which it has been proposed that NA/NS domains rather than S-domains form the principal binding site within HS. Clamp *et al* have also shown that binding of the oligomeric form of endostatin to cell surface HS may be essential for its biological activity. Competition between monomeric endostatin and oligomeric endostatin for the same HS binding site may regulate their biological activities.

13. Fibroblast growth factors (FGFs)

The fibroblast growth factor family of proteins, in humans, encompasses 23 members. These control the proliferation, migration, differentiation and survival of a wide variety of cell types (Ornitz and Itoh 2001). Most fibroblast growth factors (FGFs) have a high

affinity for HSPGs and require HS to activate one of five types of cell surface receptor (Ornitz and Itoh 2001; Sleeman *et al.* 2001).

13.1 The domain structure and minimal sulphation requirements of FGF binding oligosaccharides

Oligosaccharides composed solely of S-domain type sulphation sequences have been found to bind many members of the FGF family. Specifically HS binding sequences for FGF-1 and 2 have been particularly well characterised (Ishihara 1994; Faham *et al.* 1996; Kreuger *et al.* 2001).

A minimal HS binding sequence for FGF-2 has been deduced from x-ray crystallography studies on a heparin-FGF2 complex (Faham *et al.* 1996). The structure of this complex was shown earlier in Figure 10. The minimal HS binding sequence consists of a pentasaccharide sequence GlcNS-IdoUA(2S)-GlcNS-IdoUA-GlcNS. Interestingly this sequence contains only a single O-sulphate group (see next section 12.2).

FGF-1 interacts with a similar minimal sequence although it also appears to recognise 6-O- sulphate groups (Ishihara 1994). An O-sulphated tri-saccharide motif consisting of IdoUA(2S)-GlcNS(6S)-IdoUA(2S) has been shown to be present in all saccharides of six to eight disaccharides in length that bind with high affinity to FGF-1 (Kreugar *et al.*, 1999).

13.2 Different FGFs interact with different HS sulphation sequences?

The rationale behind many HS-FGF binding studies has been to demonstrate that different members of the FGF family bind to different sulphation sequences within the

HS polymer. Support for this rationale has recently come from a comprehensive screen of the HS binding structures for FGFs-2, 4, 7, 10 and 18 (Ashikari-Hada *et al.* 2004). FGF-8 was also screened, however this was shown not to bind any of the oligosaccharides analysed. It has been suggested that this was due to the choice of a none heparin binding isoform of this protein (Kreuger *et al.* 2005). The binding affinity of a number of variably O-sulphated, entirely N-sulphated octasaccharides, for these FGFs was assessed (Ashikari-Hada *et al.* 2004). The FGFs could be divided into 4 groups, as detailed below.

Group 1	FGF-2	requires 2-O-sulphation but not 6-O-sulphation
Group 2	FGF-10	requires 6-O-sulphation but not 2-O-sulphation
Group 3	FGF-18	requires 2-O-sulphation or 6-O-sulphation
Group 4	FGF-4 and FGF-7	require both 2-O- and 6-O-sulphation

The group 4 classification of FGF-7 is not supported by the previous study of Ostrovsky *et al.*, which places FGF-7 in group 3 (Ostrovsky *et al.* 2002).

The specific number of each O-sulphate group required varied between the FGFs. FGF-2 bound to an octasaccharide carrying a lone 2-O-sulphate group with appreciable affinity, as predicted from x-ray crystallography studies (Faham *et al.* 1996). FGF-10 required the presence of three 6-O- sulphate groups. FGF-18 also required three 6-O-sulphate groups or alternatively two or three 2-O-sulphate groups. Results for FGFs -4 and -7 were not reported.

The different requirements for O-sulphate groups can be interpreted as a difference in the HS sulphation patterns that binds to each FGF.

13.3 Biologically active, FGF binding, HS oligosaccharides

With many FGFs, a saccharide sequence longer than the minimal FGF binding sequence is required in order for them to elicit a biological response. With FGF-2 it has also been shown that additional O-sulphate groups are also required.

The HS sequence of a biologically active FGF-2-HS deca-saccharide complex must comprise a minimum core of three IdoA(2S)-GlcNS repeats substituted with one or at most two 6-O-sulphate groups (Guimond *et al.* 1993; Ishihara *et al.* 1995; Pye *et al.* 2000). The 6-O-sulphate groups are not required for FGF-2 binding to HS and it has been proposed that they may be involved in interactions with a receptor molecule (Gallagher and Lyon 2000).

13.4 Fibroblast growth factor receptors (FGFRs)

The FGFRs are members of the receptor tyrosine kinase family of proteins. When a ligand binds to the extracellular domain of these receptors, receptor dimerisation occurs. Each monomer then phosphorylates a number of conserved tyrosines within the cytoplasmic domain of its dimeric partner. This then leads to a series of downstream intracellular signalling events. FGFR dimerisation is therefore a key feature in initiating a biological response (Ornitz and Itoh 2001). Alternative splicing of four FGFR genes (FGFR1-4) results in the production of over 48 different isoforms of the FGFR receptors that vary in their ligand binding properties and kinase domains (Coutts and Gallagher 1995; Duchesne *et al.* 2006). All share a common extracellular region composed of three immunoglobulin (Ig) like domains (D1-D3).

Interactions with FGFs occur via FGFR domains D2 and D3. Each receptor can be activated by several FGFs. In many cases the FGFs themselves can also activate more than one receptor, this is not the case with FGF-7 however which can only activate FGFR2b. A gene for a fifth FGFR protein, FGFR5, has also been identified (Sleeman *et al.* 2001). In contrast to FGFRs 1-4 it lacks a cytoplasmic tyrosine kinase domain and one isoform, FGFR5 γ , only contains the extracellular domains D1 and D2. Due to its later discovery, little is known about the ligand binding activity of this protein. A comprehensive review of FGF receptors and which FGFs are able to activate individual splice variants here is not warranted, for such a review for FGFRs 1-4 see (Ornitz *et al.* 1996).

In the presence of either FGF-1 or -2 various ectodomain splice variants (FGFRs 1-4) have been shown to have an affinity for HS derived oligosaccharides. In some cases the affinity being enhanced by the presence of one or other of the FGFs (Jastrebova *et al.* 2006). This and the previous work detailed (Guimond *et al.* 1993; Ishihara *et al.* 1995; Pye *et al.* 2000), supports the hypothesis that biologically active oligosaccharides must interact with both receptor molecules, in addition to the FGF protein itself. This has been shown to be the case in 2 crystal structures of FGF-FGFR-heparin complexes (Pellegrini *et al.* 2000; Schlessinger *et al.* 2000).

13.5 FGF-FGFR-heparin complexes

Two crystal structures have been published that show how HS may interact with both receptor and FGF at the cell surface to stabilise a ternary complex (see Figures 14 and 15) (Pellegrini *et al.* 2000; Schlessinger *et al.* 2000).

Figure 14 The crystal structure of a 2:2:1 FGF1:FGFR2c:heparin dp10 complex.

The co-ordinates for this structure were taken from the protein data bank file 1E00 (Pellegrini *et al.* 2000). Molecular surfaces for the FGFs (green) and FGF receptors (grey and purple respectively) were calculated as in Figure 10. The heparin oligosaccharide is shown by a line representation. The whole complex is shown in model A.

Within the complex and due to their close proximity, a number of electrostatic interactions may exist between various negatively charged heparin carboxyl groups and sulphate groups and positively charged basic amino acid residues exposed at the surface of both FGF molecules and the single FGFR molecule coloured grey. The close proximity of these groups can be seen more clearly in enlargements B and C. The side chains of all basic amino acids are coloured red in enlargement C.

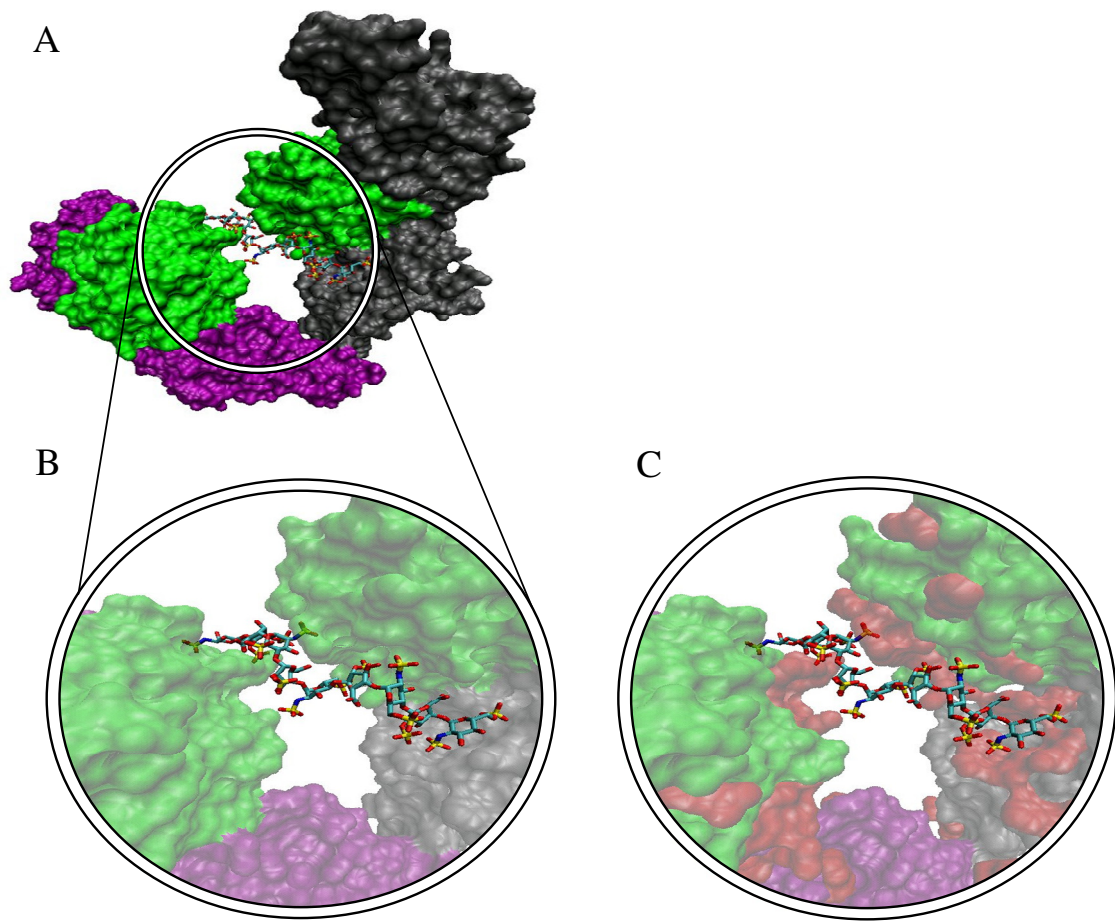


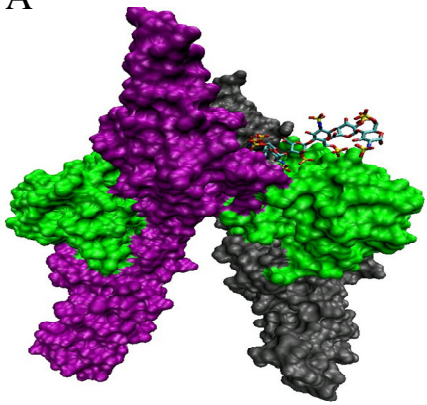
Figure 15 **The crystal structure of a 2:2:2 FGF2:FGFR1c:heparin complex.**

The co-ordinates for this structure were taken from the protein data bank file 1FQ9 (Schlessinger *et al.* 2000).

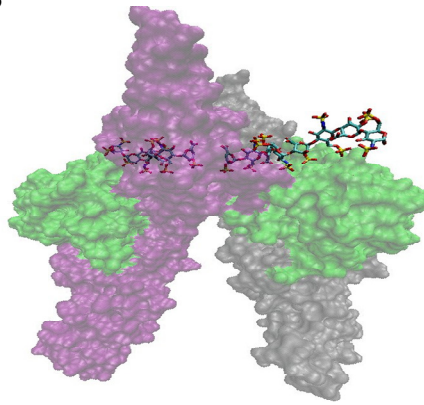
Molecular surfaces for the FGFs (green) and FGF receptors (grey and purple respectively) were calculated as in Figure 10. The heparin oligosaccharide is shown by a line representation.

From model A it is not immediately obvious that a two oligosaccharides run along the top edge of the complex. This can be more clearly seen in the transparent representation of model B. For model C, model B was rotated to a position where the two heparin oligosaccharides could be more clearly viewed from above. In enlargements D and E the non-reducing end of each oligosaccharide can clearly be seen positioned in the centre of the complex. In enlargement E basic residues are again marked out in red, due to their close proximity again a number of electrostatic interactions are possible between these and various heparin carboxyl and sulphate groups. Two 6-O-sulphate groups which have been shown to be essential for biological activity (see text) are marked with arrows.

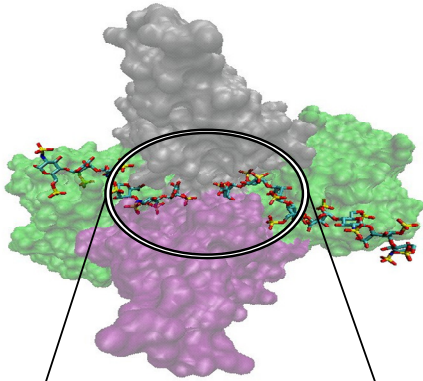
A



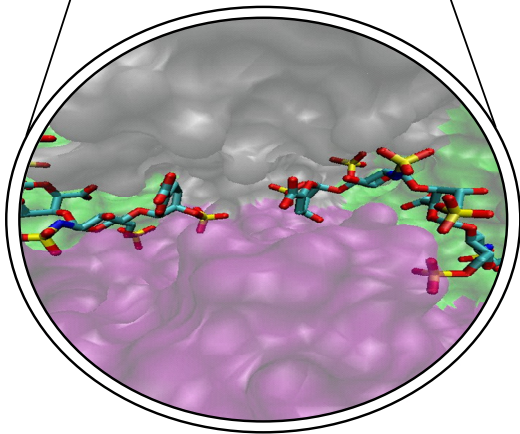
B



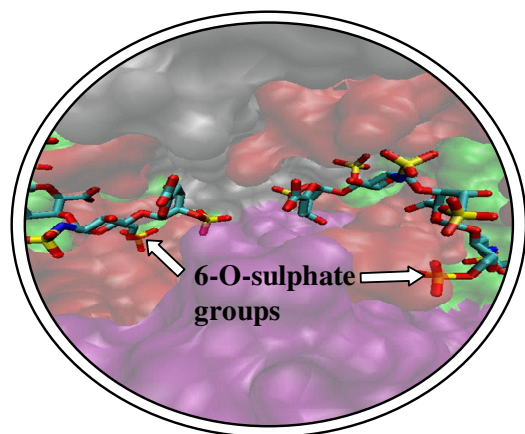
C



D



E



The models shown in Figures 14 and 15 are for different FGFs (FGF-1 and -2) and different receptor ectodomains (FGFR2c and FGFR1c). The Pelligrini model (Figure 14) puts HS/heparin oligosaccharides at the heart of complex formation and biological activity. It can be seen how without stabilisation by the centrally positioned oligosaccharide FGFR dimerisation may not occur and a biological response not initiated. The Schlessinger model (Figure 15) gives heparin/HS less of a key role. An extensive area of contact exists between both FGFRs and it is doubtful whether the presence of an oligosaccharide in the complex stabilises this interaction still further. Differences between these models may be due to experimental procedure. For example Schlessinger *et al* diffused heparin oligosaccharides into pre-formed crystals of the FGF and FGFR. While the crystals analysed by Pelligrini *et al.* were formed with oligosaccharides *in-situ*. Differences may also simply be the result of using different FGFs and splice variant FGFR ectodomains.

It has recently been suggested that the binding of a single FGF-1 ligand alters the local conformation of oligosaccharide chain. The altered conformation then has an increased affinity for the binding of a second FGF-1 ligand and in this way facilitates the formation of a 2:2:1 Pelligrini type complex (Robinson *et al.* 2005).

13.6a Different FGFs and FGFRs do not interact with different HS sulphation sequences or specific sulphation sequences? Part 1

The long held assumption that different FGFs interact with different sulphation sequences has recently been challenged (Kreuger *et al.* 2005; Kreuger *et al.* 2006). A number of sequenced HS derived oligosaccharides were analysed for their binding affinity to various

FGFs (FGFs -1, -2 , -4, -7 and -8b). A considerable degree of overlap was detected in many of their binding characteristics. It was also found that rather than interacting with specific sequences the affinity of the FGFs for the oligosaccharides was mainly due to the charge density of the oligosaccharide. Similar charge density rather than sequence effects have also recently been shown for receptor binding interactions (Jastrebova *et al.* 2006). The first conclusion that FGFs share binding sites is not too controversial. It can be imagined how competition between the various growth factors for a limited number of HS binding sites might provide a type of regulatory mechanism. This is similar to the proposal for competition to exist between the endostatin monomer and trimer in the regulation of endostatin biological activity. The second conclusion that binding is related to charge density rather than sequence cannot fit with studies where specific sequences of sulphation have been shown to either possess or not possess biological activity (Pye *et al.* 1998; Guimond and Turnbull 1999). A good example of a study of this type being the thesis work of Goodger (Goodger 2003).

From this work, a deca-saccharide carrying 7 sulphate groups,

UA-GlcNAc-IdoUA-GlcNS-IdoUA(2S)-GlcNS(6S)-IdoUA(2S)-GlcNS-HexUA-GlcNAc(6S)

was unable to elicit a biological response from FGF-2 in combination with the FGFR1 IIIc receptor. A deca-saccharide carrying 6 sulphate groups was however,

UA-GlcNAc-IdoUA-GlcNS-IdoUA(2S)-GlcNS(6S)-IdoUA-GlcNS-HexUA-GlcNAc(6S).

This result directly contradicts the new charge density theory.

A possible explanation for this apparent contradiction may lie in the conformational flexibility of iduronate residues. This issue will be re-examined in the light of a discussion of another HS binding protein antithrombin III.

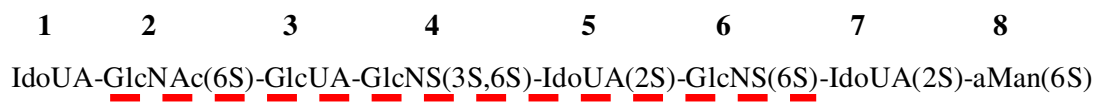
14. Antithrombin III and Factor Xa

Historically the first and most well characterised ligand-binding interaction is that of antithrombin III (AT-III). When bound to heparin, the rate at which AT-III inactivates thrombin and other proteases involved in blood clotting, most notably factor Xa, dramatically increases (Bjork and Lindahl 1982). However, it is the HS present at the cell surface of vascular endothelial cells that is one of the factors responsible for maintaining the non-thrombogenic properties of blood vessel walls, rather than heparin (Marcum *et al.* 1986).

A sequence of five monosaccharides within a heparin derived AT-III binding octasaccharide (underlined in the octasaccharide sequence shown in Figure 16) was found to be the minimal sequence required for AT-III binding (Lindahl *et al.* 1984). Any loss of sulphation or alteration of any of these five residues resulted in a dramatic drop in the affinity of the octasaccharide for AT-III, although N-acetylation or sulphation was subsequently found to be tolerated at position 2 (Conrad 1998).

Figure 16

The sequence of an octasaccharide with high affinity for ATIII (Lindahl *et al.* 1984). The minimal antithrombin-binding pentasaccharide (units 2-6) is underlined in red.



Inactivation of factor Xa by an activated antithrombin–HS complex requires only the specific pentasaccharide able to bind antithrombin (Danielsson *et al.* 1986). However, inactivation of thrombin requires the pentasaccharide to be part of a much longer sequence that is then thought to bind both thrombin and antithrombin bringing them into close proximity and facilitating their interaction (Danielsson *et al.* 1986).

An extended highly sulphated ‘biologically active’ sequence encompassing both a thrombin binding site and the antithrombin binding pentasaccharide would be relatively common within a heparin chain. However, within HS, such a long consecutive sequence of sulphation would be rare and it is more likely that the thrombin and antithrombin binding sites occur within two separate S-domains, linked by a region of alternating or N-acetylated sequence.

14.1 Synthetic Factor Xa inhibiting pentasaccharides

The potential application of a Factor Xa inhibiting pentasaccharide as a clinical anticoagulant has prompted much research into its synthesis using synthetic chemistry techniques. This research has culminated in the marketing of the anticoagulant product Arixtra® by Sanofi-Aventis. Arixtra® contains a synthetic pentasaccharide almost identical in structure to the minimal antithrombin III binding sequence. A number of derivatives of this pentasaccharide have also been produced and these have shed new light on the functional role of a number of specific sulphate groups within this sequence (Sisu *et al.* 2003).

14.2 Conformationally locked iduronate residues

The minimal AT-III binding pentasaccharide contains a single iduronate residue at position 5 (see Figure 16). This can be synthetically locked into the 1C_4 or 2S_0 conformations (Das *et al.* 2001). Only oligosaccharides containing the locked 2S_0 conformation were able to bind and activate AT-III. If the iduronate is retained locked in this conformation but the N- and 6-O- sulphate groups removed from the N-glucosamine residue to its reducing side, AT-III binding and biological activity is maintained (Sisu *et al.* 2003).

The conclusion here is that the removed sulphate groups are not required to directly interact with AT-III. The role of these sulphate groups, may be to influence the conformational equilibrium of the adjacent iduronate more towards the 2S_0 conformation necessary for binding to AT-III to occur.

13.6b Different FGFs and FGFRs do not interact with different HS sulphation sequences or specific sulphation sequences? Part 2

A similar iduronate influencing role could be envisioned for strategically positioned sulphate groups within biologically active HS/heparin oligosaccharide sequences. It may be that the apparent correlation between an increase in charge density, (or more specifically an increase in the number of sulphate groups) is related to an increase in the chance that one or several of the additional sulphate groups is able to influence the conformational equilibria of a conformationally critical iduronate residue.

15. The balance of iduronate equilibria, a critical influence on oligosaccharide binding? A purification problem

Typically iduronate conformational details are not available for HS oligosaccharides screened in binding or biological assays. These details can only be gained through the use of NMR techniques, which require a minimum of 0.5mg of material for analysis. Typical functional analysis studies can work with <10µg of material.

Although HS itself has been available in bulk quantities for a number of years (Griffin *et al.* 1995), few authentic HS oligosaccharides larger than dp4 have been well characterised by NMR techniques that are needed, not only for iduronate, but for more general conformational studies (Hileman *et al.* 1997).

This is due to the extreme difficulty in obtaining a homogeneous sample in the quantities required for NMR analysis. The number of potential sulphation sequences and consequently sample heterogeneity, increases dramatically as the chain length increases (Hileman *et al.* 1997).

16. Iduronate and terminal Δ4,5 uronic acid conformational analysis of heparin and synthetic oligosaccharides

Due to the purification problem, iduronate conformation has only been analysed within synthetic or heparin derived oligosaccharides (reviewed by (Ferro *et al.* 1990)) Due to heparins high level of sulphation it is much easier to purify homogeneous oligosaccharides in high quantities for NMR analysis and other structural studies.

The conformational equilibrium has been shown to be influenced by the sulphation state of neighbouring N-glucosamine residues on both sides of the iduronate as well as 2-O-

sulphation of the iduronate residue itself (Ferro *et al.* 1990). For example iduronate in the tetrasaccharide sequence GlcUA-GlcNS(6S,3S)-IdoUA(2S)-GlcNS(6S) has a conformational equilibrium of 32%:68% ${}^1C_4: {}^2S_0$. While in the tetrasaccharide sequence GlcUA-GlcNS(6S)-IdoUA(2S)-GlcNS(6S) the conformational balance is altered to 56%:44% ${}^1C_4: {}^2S_0$.

Like iduronates the balance of terminal Δ 4,5 uronic acid conformational equilibrium is influenced by the sulphation state of adjacent glucosamines and by 2-O- sulphation. For example, the sequence UA-GlcNS(6S) exists as 46%:54% ${}^2H_1: {}^1H_2$, while the sequence UA(2S)-GlcNS(6S) exists as 8%:92% ${}^2H_1: {}^1H_2$ (Ragazzi *et al.* 1993).

A survey of the literature reveals that for both iduronate and terminal uronic acid residues, there are significant gaps in the sequences within which their conformational equilibrium has been studied. For example, no reference could be found to the balance of the iduronate conformational equilibrium in the sequence -GlcNS-IdoUA(2S)-GlcNAc-.

17. Theoretical coupling constants

The 1990 review by Ferro *et al.* (Ferro *et al.* 1990) used a least squares fitting technique, to fit experimental coupling constant data to theoretical coupling constant values, and derive the percentage of iduronate in each of its three low energy conformations. This technique has subsequently been used in a number of similar studies. However, all of these studies use different theoretical values to fit to their experimental data (Forster and Mulloy 1993; Ojeda *et al.* 2002; Lucas *et al.* 2003). Still more theoretical values were published prior to 1990 (van_Boeckel *et al.* 1987), although these were not used in a least-squares fitting analysis.

It is not clear to what extent the results of a least-squares analysis are influenced by the choice of theoretical coupling constants. Nor was it clear which should be the correct choice of theoretical values for a least squares analysis of the data presented in this thesis. In order to examine these two issues, as part of this thesis work a three conformational least squares fit was performed to previously published data. The conformational fit was calculated as detailed below.

Using the solver add-in of Microsoft Excel the sum of the squares of the differences between each of the experimentally observed coupling constants and a weighted average of the theoretical coupling constants was minimised. The weighted average of the theoretical constants was represented by the formula

$$a {}^3J_{x,y} {}^1C_4 + b {}^3J_{x,y} {}^4C_1 + c {}^3J_{x,y} {}^2S_0$$

where ${}^3J_{x,y}$ = the theoretical value for each coupling constant ${}^3J_{1,2}$, ${}^3J_{2,3}$, ${}^3J_{3,4}$ and ${}^3J_{4,5}$.

The sum of the squares of the differences was minimised by varying parameters a, b and c subject to the constraints that their sum should equal 1 and each could take a value ≥ 0 but ≤ 1 . The percentages of each conformation reported are parameters a, b or c expressed as a percentage of their sum total 1. The spread sheet to carry out this analysis was set out by Professor John Pezzulo, Georgetown University, Washington DC.

Table 1 shows the distribution of iduronate conformations best-fitting the measured experimental coupling constants for a number of previously published oligosaccharide structures.

Table 1. A comparison of the conformational distribution of iduronate residues within a number of previously published oligosaccharide sulphation sequences.

A three conformation least-squares fit was carried out for experimentally determined coupling constants taken from (Ferro *et al.* 1990) (oligosaccharides 1, 2, 3, 5 and 6) and (Lucas *et al.* 2003) (oligosaccharide 4). The fit was to theoretical values as published by (van_Boeckel *et al.* 1987; Ferro *et al.* 1990; Forster and Mulloy 1993; Ojeda *et al.* 2002; Lucas *et al.* 2003).

The iduronate residue to which the coupling constants refer are marked in red.

Note, the sequences of oligosaccharides 1 and 3 were published incorrectly by Ferro *et al.* by reference to the original work cited by them of van Boeckel *et al.*(van_Boeckel *et al.* 1987)

Oligosaccharide	Coupling constants				Contribution of conformer to equilibrium (%)			Reference for theoretical values
	$^3J_{1,2}$	$^3J_{2,3}$	$^3J_{3,4}$	$^3J_{4,5}$	1C_4	2S_0	4C_1	
1. GlcNS-IdoUA-GlcNS	3.3	5.4	3.6	2.4	73 63 62 76 71	27 29 38 24 29	0 8 0 0 0	van Boeckel <i>et al.</i> (1987) Ferro <i>et al.</i> (1990) Forster and Mulloy (1993) Ojeda <i>et al.</i> (2002) Lucas <i>et al.</i> (2003)
2. GlcNAc-IdoUA-GlcNAc	3.9	6.6	3.6	3.2	56 46 45 58 57	44 47 55 42 43	0 7 0 0 0	van Boeckel <i>et al.</i> (1987) Ferro <i>et al.</i> (1990) Forster and Mulloy (1993) Ojeda <i>et al.</i> (2002) Lucas <i>et al.</i> (2003)
3. GlcNAc(6S)-IdoUA(2S)-GlcNS	1.4	3.2	2.8	2.1	100 95 94 100 99	0 1 6 0 0	0 4 0 0 1	van Boeckel <i>et al.</i> (1987) Ferro <i>et al.</i> (1990) Forster and Mulloy (1993) Ojeda <i>et al.</i> (2002) Lucas <i>et al.</i> (2003)
4. GlcNS-IdoUA(2S)-GlcNS-IdoUA(2S)-GlcNS	3.6	4.2	3.6	3.5	78 74 64 80 78	22 2 24 20 0	0 24 8 0 21	van Boeckel <i>et al.</i> (1987) Ferro <i>et al.</i> (1990) Forster and Mulloy (1993) Ojeda <i>et al.</i> (2002) Lucas <i>et al.</i> (2003)
5. GlcNS(6S)-IdoUA(2S)-GlcNS	2.8	5.4	3.1	2.5	76 65 65 79 74	24 35 35 21 26	0 0 0 0 0	van Boeckel <i>et al.</i> (1987) Ferro <i>et al.</i> (1990) Forster and Mulloy (1993) Ojeda <i>et al.</i> (2002) Lucas <i>et al.</i> (2003)
6. GlcNS(6S,3S)-IdoUA(2S)-GlcNS	3.9	7.3	3.9	3.3	48 40 38 49 49	52 53 62 51 51	0 7 0 0 0	van Boeckel <i>et al.</i> (1987) Ferro <i>et al.</i> (1990) Forster and Mulloy (1993) Ojeda <i>et al.</i> (2002) Lucas <i>et al.</i> (2003)

As can be seen in Table 1, the choice of theoretical coupling constants can have an effect on the results of the least-squares fitting procedure. This is most dramatically seen for oligosaccharide 4.

The study of Forster and Mulloy (Forster and Mulloy 1993) cited in Table1, examined the molecular dynamics of iduronates using a computational model. The coupling constants then published and used to produce the conformational percentages given in Table 1, were an average over the structures observed during the course of a 200ps molecular dynamic simulation. This approach eliminates the high level of bias introduced by choosing to fit to an arbitrary single set of coupling constants. These results are, however, by their nature dependent on the charge model and molecular dynamic force field used. It was felt that an improvement on the accuracy of these values might be offered by repeating this work using the modern carbohydrate specific Glycam04 molecular dynamics force field (Case *et al.* 2005).

18. Heparan Sulphate, three-dimensional structure

Many NMR studies of polymeric heparin or heparin derived, rather than HS derived oligosaccharides have been published, again due to the ease of heparin oligosaccharide purification (Loganathan *et al.* 1990; Yamada *et al.* 1992; Mulloy *et al.* 1993; Yamada *et al.* 1993; Mulloy *et al.* 1994; Larnkjaer *et al.* 1995; Pervin *et al.* 1995; Mikhailov *et al.* 1996; Mikhailov *et al.* 1996; Tsuda *et al.* 1996; Yates *et al.* 1996; Cros *et al.* 1997; Mikhailov *et al.* 1997; Yamada *et al.* 1998; Hricovini *et al.* 1999; Yates *et al.* 2000; Chuang *et al.* 2001; Guerrini *et al.* 2001; Hricovini *et al.* 2001; Chuang *et al.* 2002;

Guerrini *et al.* 2002; Hricovini *et al.* 2002; Hakansson and Caffrey 2003; Vanhaverbeke *et al.* 2004; Angulo *et al.* 2005).

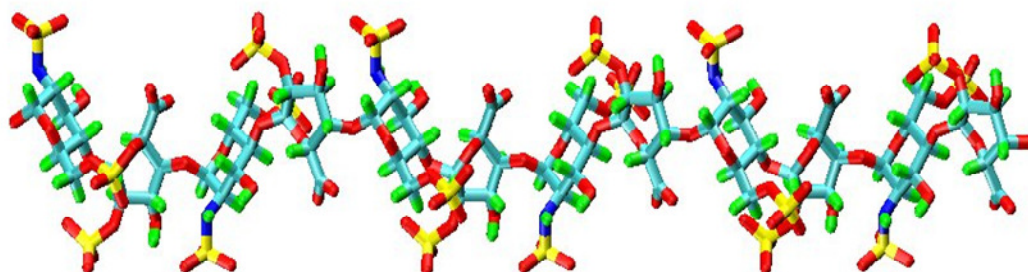
Most NMR studies are limited to proton assignment however two 3D-models have been produced for a heparin dodecasaccharide using a combination of NMR and molecular modelling techniques (see Figure 17) (Mulloy *et al.* 1993).

Figure 17 Two, three dimensional models of heparin dodcasaccharides.

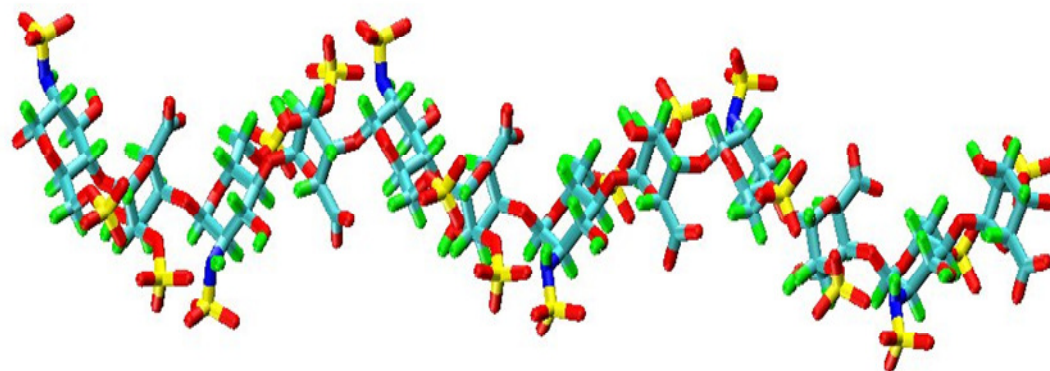
The coordinates for both structures were taken from the protein data base, code 1HPN (Mulloy *et al.* 1993). The models are composed of 6 repeats of a GlcNS(6S)-IdoUA(2S) disaccharide unit.

In model A all IdoUA(2S) monosaccharides are in the 2S_0 conformation. In model B all IdoUA(2S) monosaccharides are in the 1C_4 conformation.

A



B



NMR studies on chemically de-sulphated heparin derivatives suggest that removal of specific types of sulphate group does not unduly affect the overall helical conformation of the oligosaccharides presented in Figure 17 (Mulloy *et al.* 1994). On this basis the helical conformations of these models are assumed to be adopted by all HS S-domain type sequences (Mulloy and Forster 2000).

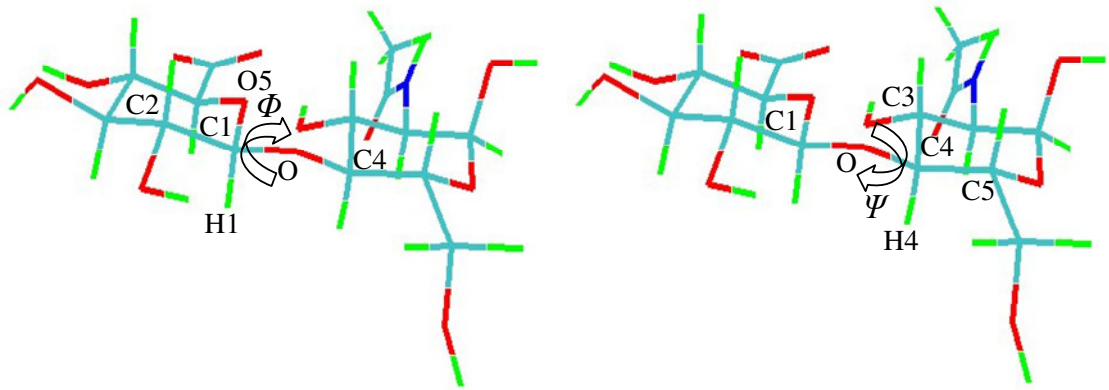
Chemical de-sulphation techniques though, lack the subtlety to reproduce many of the sulphation sequences found within authentic HS derived sequences (Merry *et al.* 1999; Vives *et al.* 1999; Kreuger *et al.* 2001). It may be that subtle sulphational differences produce subtle conformational differences within authentic HS oligosaccharides.

18.1 The torsional angles Φ and Ψ , and the exo-and endo-anomeric effect

Defining the relative orientation of each monomeric unit in the Figure 17 models are the Φ and Ψ torsion angles across each glycosidic linkage. The authors of these models defined these torsional angles as $\Phi = \text{H1-C1-O-C4}$ and $\Psi = \text{C1-O-C4-H4}$ (see Figure 18). Definition of these torsional angles is an area of inconsistency in HS/heparin oligosaccharide modelling. Some authors define these angles as $\Phi = \text{O5-C1-O-C4}$ and $\Psi = \text{C1-O-C4-C5}$ (Verli and Guimaraes 2004; Becker *et al.* 2005) (see Figure 18). While others use a torsion angle of $\Psi = \text{C1-O-C4-C3}$ (Imberty and Perez 2000) or $\Phi = \text{C2-C1-O-C4}$ and $\Psi = \text{C1-O-C4-C3}$ (Raman *et al.* 2003) (see Figure 18). Occasionally how these angles were defined is not declared (Mikhailov *et al.* 1996; Mikhailov *et al.* 1997). While each definition is equally valid, the lack of consistency makes a comparison of data between studies a difficult if not impossible task, particularly when definitions are not

declared. The definition adopted by Mulloy *et al.* slightly predominates in the literature and was adopted throughout this study.

Figure 18 Φ and Ψ torsion angles defining the relative special orientation of each monosaccharide across a glycosidic linkage.





The exo-anomeric effect is reported to cause a considerable barrier to free rotation about the Φ and Ψ torsion angles (Collins and Ferrier 1998). The exo-anomeric effect causes axial or equatorial groups attached to C1 to prefer differing rotamer states. These rotamer states are presented as a Newman projection along the C1-glycosidic oxygen bond in Figure 19. The exo-anomeric effect is the result of the glycosidic oxygen preferring to adopt a conformation in which its lone pair orbitals overlap with the non-bonded electrons of the endocyclic O5-C1 bond (Bubb 2003).

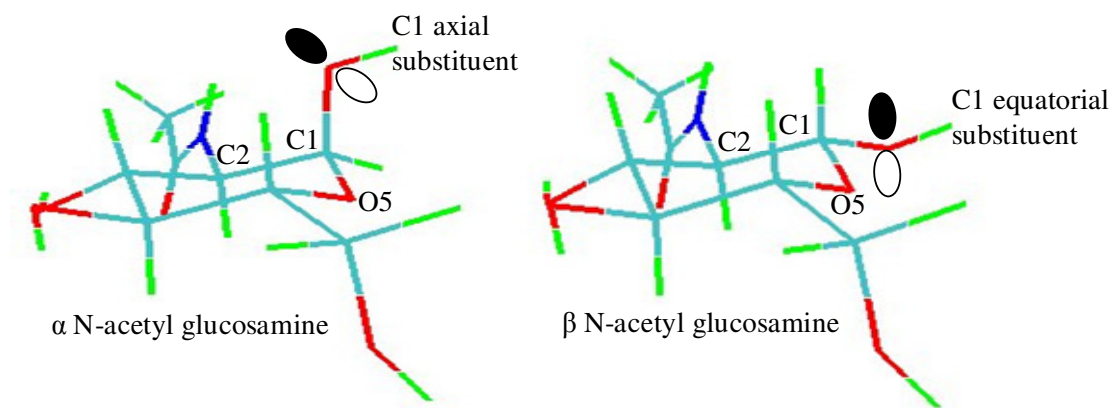
The intra-ring equivalent of the exo-anomeric effect is the endo-anomeric effect. For pyranoses, the usually energetically preferred equatorial orientation of substituents in six membered ring systems, is opposed by interactions between lone pair electrons of the endo-cyclic oxygen (O5) and electronegative substituents attached to C1; these substituents favouring an axial orientation (see Figure 19) (Bubb 2003).

According to this effect, all C1 substituents within HS/heparin have a preference for an axial orientation. This may be the cause of the formation of α C1-C4 glycosidic linkages between all HS/heparin structural units as opposed to β C1-C4. (note, through a quirk of carbohydrate naming nomenclature C1-C4 bonds between glucuronic acid and a N-glucosamine are classed as β C1-C4. The C1 substituent of β -glucuronic acid is actually axially orientated, as with all other internal HS/heparin structural units).

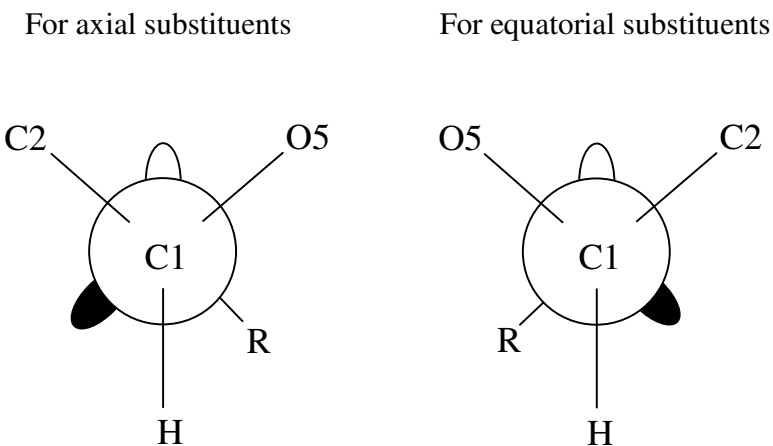
For axially orientated systems the endo- and exo- anomeric effects oppose each other (Collins and Ferrier 1998). This makes the exo- anomeric effect less of a barrier to rotation about the Φ and Ψ torsion angles of HS/heparin. By how much this barrier is lowered is unclear.

Figure 19 Axial and equatorial orientations of C1 substituents and their preferred rotamer conformations due to the exo-anomeric effect.

Lone pair orbitals on the oxygen substituent are represented as  and .



Preferred rotamer conformations due to the endo-anomeric effect:



The R group from the above two structures is a single hydrogen atom. However, the R group in an oligosaccharide, will be the C4 of a linking monosaccharide.

19. Molecular dynamic studies

NMR derived structures represent an average of many structures that occur in solution. This is particularly the case with oligosaccharide structures, which are known to be particularly flexible in solution (Imberty and Perez 2000). An idea can be gained as to how much a modelled structure may sample alternate conformations, through the use of molecular dynamic modelling techniques.

A review of the insights gained and the problems with HS/heparin molecular dynamic modelling is discussed below.

19.1 Dynamic modelling of N-glucosamine monosaccharides

The gauche effect is notoriously hard to replicate in molecular dynamic modelling simulations. It has been shown that for methyl α -D-glucopyranoside (essentially GlcNAc but with a hydroxyl group in place of the NAc substituent) a simulation time of 50ns is required to produce conformational ensembles that display the correct, experimentally determined, rotamer populations (38:5:57, gt:tg:gg) (Kirschner and Woods 2001). Explicit solvent molecules must also be included in the calculation. A 50ns explicit solvent simulation, for any modelled system, is only realistically possible through the use of expensive multi-processor hardware.

In the modelling of N-glucosamines, within HS/heparin oligosaccharides rotamer populations are usually assigned as 'gg' where static models are produced (Mulloy *et al.* 1993) (see Figure 17), or initiated in the 'gg' conformation when molecular dynamic simulations are carried out (Mikhailov *et al.* 1997; Ojeda *et al.* 2002; Lucas *et al.* 2003; Verli and Guimaraes 2004). In all these studies the rotamer conformation remained in the

'gg' conformation throughout the length of the simulation. The modelled preference for a 'gg' conformation is mainly based on the experimentally observed rotamer equilibrium for methyl α -D-glucopyranoside.

19.2 Dynamic modelling of iduronate monosaccharides

To date, no molecular dynamic simulation has been able to reproduce experimentally determined iduronate conformer populations. In one study, a single transition from the 2S_0 to the 4C_1 conformation was reported (Forster and Mulloy 1993), while in another study torsional angle restraints were employed to actively prevent 'ring flipping' (Verli and Guimaraes 2004). The presumption here is that the applied restraints do not artificially influence the dynamics of the modelled system. However, in many studies iduronates remain in their starting conformation throughout a molecular dynamic simulation, without the application of additional torsion angle restraints (Ojeda *et al.* 2002; Angulo *et al.* 2003; Lucas *et al.* 2003). A clear advance in the field would be the demonstration that models may sample alternate iduronate ring conformations, at a rate consistent with experimental data. The influence of sulphation pattern on the rate of sampling could then be assessed.

19.3 Molecular modelling of terminal uronate monosaccharides

Few models have been generated that terminate in this unit. However, Mikhailov *et al.* (Mikhailov *et al.* 1996) carried out a number of molecular dynamic simulations of a tetrasaccharide terminating in UA(2S). The starting conformation for this residue in all simulations was 1H_2 . Only in one simulation was a conformational change observed to

$^2\text{H}_1$. In other simulations the conformation remained un-altered or an intermediate conformation between the two half chair conformations became stabilised. The fact that an intermediate conformation existed for almost the entire duration of one simulation, suggests there was a parameterisation problem with this residue in this model system.

20. Aims of this study

Over many years of study, a great deal of NMR structural information has been gathered for heparin or heparin derived oligosaccharides (see section 16, for a comprehensive reference list from 1990-2005). Only one previous HS-NMR study has characterised authentic HS derived, di-, tetra- and a single hexasaccharide. This study was limited to just proton assignment (Hileman *et al.* 1997). The primary aim of the work presented in this thesis was to address this imbalance.

A preliminary SAX-HPLC analysis of authentic HS hexasaccharides derived from a bulk source of commercial porcine HS was carried out. This revealed that the considerable peak overlap, observed on HPLC analysis by Hileman *et al.*, did not occur with these samples. It should, therefore, have been possible to purify more than the single hexasaccharide isolated by Hileman *et al.* and possibly longer oligosaccharides for NMR analysis. This would represent a first in the field of HS conformational analysis.

Once a number of authentic HS oligosaccharides had been purified and sequenced by NMR, it became clear that an analysis of their iduronate and terminal UA conformations could fill a number of gaps in the literature. For a comparison a number of heparin derived hexasaccharides were also produced. Chemical de-2-O-sulphation of these

heparin and HS oligosaccharides were carried out in order provide a greater insight into the role 2-O-sulphate groups play in influencing conformational equilibria.

It was also sought to produce an improved set of coupling constants for iduronate conformational analysis, using computational molecular dynamic modelling, in a way similar to the coupling constant set previously produced by Forster and Mulloy (Forster and Mulloy 1993).

The second aim of the work presented, was to use the NMR structural data gathered to provide new information on the relationship between HS structure and function. Molecular models were produced for all HS oligosaccharides on which NMR structural data had been gathered. It was found that static models could only provided limited new structure/functional information. In particular static molecular modelling could not explain the basis on which a number of deca-saccharide structures either possessed or did not possess biological activity (Goodger 2003). Molecular dynamic modelling was then used to try and account for the significant degree of conformational flexibility all oligosaccharides possess in solution (Imberty and Perez 2000). By analysing the data from dynamic studies in was hoped to provide a link between oligosaccharide sulphation pattern and biological activity that was missing from the thesis work of Goodger (Goodger 2003).

21. Molecular dynamic partial charge assignment

Partial charge assignments are a problem area for the field of molecular dynamic modelling. They are necessary in order for through space interactions between atoms to

be considered in dynamic simulations. In reality, however, partial charges are a fictitious concept and as such cannot be measured experimentally.

Many different methods and variations of methods exist with which to assign partial charges (reviewed in (Leach 2001)). Usually, molecular modelling force-field torsional parameters are calculated using a single set of partial charges calculated by a given method. Changing partial charge values from the values at which torsion parameters were derived, should strictly require a re-evaluation of the torsional parameters to fit with these new values, this is seldom however carried out. For use with a particular force field, it is thought good practice to calculate partial charges for new systems using a method similar to, if not identical to, the original method used in the force fields development, although this is not always carried out, see (Verli and Guimaraes 2004). Verli and Guimaraes applied Lowdin type partial charges to a force field originally parameterised using a different partial charge calculation method.

Partial charge values for the HS oligosaccharide models of this study are particularly problematic. While a majority of the torsional parameters needed for the models are contained in the glycam04 parameter set, which is available on line at http://www.glycam.com/gl_params.html and distributed free with the AMBER 8 molecular modelling software suite (Case *et al.* 2004), parameters needed to cover sulphate, carboxyl and the C4-C5 double bond of UA are not. Parameters for these are available in the literature or from other force fields, but have been developed using alternate partial charge calculation methodologies. Ideally, what is needed is a new force field, or a further development of the glycam parameter set to cover these missing parameters.

Using existing parameters an evaluation of three different partial charge sets was made in this thesis work, on HS dp6-2 and -3 structural models. Partial charge calculation method one is the most similar in principal to that employed by the glycam04 developers (Professor. Karl. Kirschner, personal communication). Partial charge methods two and three are significantly more rigorous and represent an early stage in the development of a new force field with which to model HS/heparin oligosaccharides (under development in collaboration with Dr. Francois Dupradeau, University of Amien, Amien, France).

EXPERIMENTAL MATERIALS AND PROCEDURES

1. Experimental Materials

Bio-Gel P-10 (fine grade) was purchased from Bio-Rad (Hemel Hempsted, Herts, UK). PD-10 prepacked, disposable Sephadex G-25 columns were purchased from Amersham Pharmacia Biotech. ProPac PA-1 semi-preparative strong anion-exchange HPLC columns were from Dionex (Camberley, Surrey, UK). The TSK3000PW gel filtration column was purchased from phenomenex (Macclesfield, UK). Heparinase I (*Flavobacterium heparinum*; EC 4.2.2.7), heparinase II (*F. heparinum*; no EC number assigned) and heparinase III (*F. heparinum*; EC 4.2.2.8) were purchased from Grampian Enzymes (Orkney, UK). The disaccharide standard Δ UA(2S)-GlcNS(6S) was also purchased from Grampian Enzymes. Porcine intestinal mucosal Heparan Sulphate was purchased from Celsus laboratories (Cincinnati, OH, USA). Tinzaparin Sodium (Innohep®) was purchased from Leo Pharmaceuticals (Princes Risborough, UK). D₂O was purchased from GOSS scientific (Essex, UK). Acrylamide and HPLC grade sodium chloride were purchased from VWR international (Poole, UK). Trizma-base, Glycine, Azure A, DEAE-Sephacel and sodium hydroxide were purchased from Sigma (Dorset, UK) as was the Shigemi NMR microtube.

EXPERIMENTAL PROCEDURES

2. Enzymatic depolymerisation of HS.

Porcine intestinal mucosal HS was depolymerised using the enzyme Heparinase III. To 200mg of HS 2ml of 0.1mM sodium acetate, 0.1mM calcium acetate, pH 7.0 and 10mU of heparinase III was added. For preparative digests a further three additions of 10 mU of

heparinase were made at 16hr intervals. In total 50 identical digests were carried out using 10g of HS starting material. All digestions were carried out at 25°C.

3. Gel filtration chromatography.

3.1 Analytical gel filtration chromatography

A 10µl aliquot containing 200µg of HS was applied to a TSK3000PW gel filtration column. The column was pre-equilibrated with 0.15M NaCl, 20mM Na₂HPO₄, at pH 7.2 (phosphate buffered saline (PBS)) and the sample eluted from the column using the same PBS at a flow rate of 0.4ml/min. The eluant was monitored in-line for absorbance at 232nm.

3.2 Preparative gel-filtration chromatography

Each 200mg digest was applied to a Bio-Gel P10 column (3 x 120cm) eluted with 0.25 M NH₄HCO₃ at a flow rate of 9ml h⁻¹. Likewise 3 vials of tinzaparin sodium were applied and eluted from the column. Fractions of 3ml were collected, their absorbance at 232nm measured and relevant fractions pooled. After initial evaporation to dryness the concentration of NH₄⁺ ions was further reduced by repeatedly re-dissolving the samples in 10ml H₂O and evaporating to dryness using rotary evaporation.

4. Strong Anion-exchange HPLC (SAX-HPLC).

HS tetrasaccharide, hexasaccharide and heparin hexasaccharide samples were applied to a ProPac PA-1 column (9 x 250mm) pre-equilibrated with distilled water adjusted to pH 3.5 with HCl. Elution was effected with an NaCl gradient of 0-1M NaCl over 80 mins at

a flow rate of 4ml min^{-1} . The NaCl gradient for the HS octasaccharides was 0-1.2M NaCl over 90 mins and the gradient for the HS deca-saccharides was 0-1.2M NaCl over 180 mins. All gradients were identical to those used by Merry *et al.* (Merry *et al.* 1999). The eluate was monitored in-line for absorbance at 232nm and 4ml fractions were collected. Relevant fractions were pooled and added to identical fractions from previous separation runs. These were concentrated by rotary evaporation, desalted on a PD10 column eluted with water, concentrated and then lyophilised.

5. Quantification.

Initial quantification was carried out using the following method.

A known quantity of the disaccharide standard $\Delta\text{UA}(2\text{S})\text{-GlcNS}(6\text{S})$ (molecular weight 664 Da) was applied to a TSK3000PW gel filtration column. The column was pre-equilibrated with PBS and the sample eluted from the column using the same PBS at a flow rate of 0.4ml min^{-1} . The eluate was monitored in-line for absorbance at 232nm. The area under the $\Delta\text{UA}(2\text{S})\text{-GlcNS}(6\text{S})$ peak was calculated to obtain a nmoles/ml per area unit.

The purified samples were re-dissolved in $100\mu\text{l}$ of water and $5\mu\text{l}$ then applied and eluted from the TSK column. The amount of material applied to the column was calculated by dividing the area under the sample peak by the previously calculated nmoles/ml per area unit. From this the concentration and total quantity of purified material could be calculated.

Later in the study enough material had been purified to allow direct quantification of the sample, after lyophilisation to dryness, by weight on a standard laboratory balance.

6. Polyacrylamide gel electrophoresis (PAGE).

To each well 10µg of HS hexasaccharide or octasaccharide in 20µl of a 20 % glycerol, 1% phenol red solution was loaded. Samples were initially ran through a 1.5 cm stacking gel (5% acrylamide/ 2% cross-linker). Then through a 16cm x 12cm x 0.75cm resolving gel (30% acrylamide / 5% cross-linker) at a constant voltage of 200V until the phenol red marker reached the bottom of the gel. The discontinuous buffer system of Laemmli was used (Laemmli 1970) with 0.375M Tris/HCl in the resolving gel and 0.125M Tris/HCl in the stacking gel at pHs of 8.5 and 6.5 respectively. Tank buffer consisted of 25mM Tris/ 0.192M glycine pH 8.3. Oligosaccharide bands were visualised by staining with 0.25% aqueous Azure A for 5 mins under constant agitation. Gels were de-stained by repeated washing in water for up to 30mins.

7. Recovery of HS oligosaccharides dp6-5 and dp6-6 from polyacrylamide gels

This was done using a variation of the method published by Vives *et al.* (Vives *et al.* 2001). For preparative purposes bands corresponding to HS dp6-5 and HS dp6-6 were cut out from the gel and pooled into a single tube. To this 5mls of 2M NaCl was added and the sample left overnight. The 5mls was then applied to a Bio-Gel P2 column (1.5cm x 5 cm) equilibrated with distilled water. The sample was then washed through the column until the blue Azure A colour was 1cm from the bottom. The large sample volume collected was then concentrated by rotary evaporation and further desalted by passage over PD-10 desalting columns eluted with water, and lyophilised to dryness. Removal of the Azure A organic dye in this way meant that PD10 columns were not contaminated and could be easily re-used.

8. DEAE ion-exchange chromatography

DEAE-Sephacel columns (1.5cm x 3.5 cm) were equilibrated with PBS. Dry lyophilised samples were re-dissolved in 1ml H₂O and applied to the column. Acrylamide contaminants were then eluted from the column using 10-20ml of PBS. Oligosaccharides were then eluted from the column using 1M NaCl, 20mM Na₂HPO₄, at pH 7.2. Oligosaccharide samples were concentrated by rotary evaporation, desalted by passage over PD-10 desalting columns eluted with water, and lyophilised to dryness before further analysis.

9. Disaccharide Analysis.

A volume corresponding to 100µg of sample was lyophilised to dryness. The sample was then re-dissolved in 20µl of 0.1mM sodium acetate, 0.1 mM calcium acetate, pH 7.0 and 2 milliunits of heparinases I, II and III were added. All digests were incubated at 37 °C for 16 hrs and the total digest volume was 50µl. After digestion into their component disaccharides, all samples were made up to a volume of 1ml by the addition of water-adjusted to a pH of 3.5 with HCl, and applied to a Pro-Pac PA-1 SAX-HPLC column (9 x 250mm) pre-equilibrated in water adjusted to pH 3.5 with HCl. Elution was effected by an increasing NaCl gradient of 0-1M over 45 mins at a flow rate of 4ml min⁻¹. Peak identification was made by comparison to the elution position of known disaccharide standards.

10. Chemical de-2-O sulphation

The total quantity of sample available was lyophilised down to dryness. A 1M solution of sodium hydroxide was prepared and the sample re-dissolved at a ratio of 1:4 sodium hydroxide to water, to a final volume of 1ml and a final concentration of sodium hydroxide of 0.2M. The sample was then lyophilised to dryness and re-dissolved in 900 μ l of water. The sodium hydroxide present was neutralised by the addition of 100 μ l of glacial acetic acid and the sample de-salted by passage over a PD-10 de-salting column before lyophilisation to dryness.

This procedure constituted one round of de-2-O- sulphation, samples were then subjected to NMR spectroscopy analysis (see below) before additional rounds of de-sulphation were undertaken on some samples.

11. NMR spectroscopy.

All NMR spectra were recorded by Dr. Neil M^cLay (Dept. of Chemical and Biological Sciences, University of Huddersfield).

For NMR measurements, the total quantity of material available was dissolved in 250 μ l 0.15M NaCl, 20mM Na₂HPO₄ (PBS), ²H₂O at pH 7.2 and then repeatedly exchanged in 1ml ²H₂O with intermediate lyophilization. For analysis samples were contained in a 5-mm Shigemi tube at a sample volume of 250 μ l.

The ¹H-NMR spectra were acquired on a 500 MHz. Bruker NMR spectrometer. Reported chemical shifts are in ppm relative to tetramethylsilane. The TOCSY and NOESY spectra were recorded in phase sensitive mode using states-TPPI. A total of 256 experiments of 2048 data points were recorded with 128 scans over a spectral width of 2741 KHz in both

dimensions. The FIDs were zero filled in F1 and multiplied by phase shifted sine functions in both dimensions to give data sets of 1024x1024 points.

The DQF-COSYs were recorded in magnitude mode using the same size data set as above but with the number of scans reduced to 64. The FIDs were zero filled in F1 and multiplied by sine functions in both dimensions prior to transformation.

HSQC spectra were recorded in phase-sensitive acquisition mode using Echo-anti-echo with sensitivity improvement. A total of 256 experiments of 1024 data points were recorded, using 16 scans and a spectral width of 2480Hz (^1H) in F2 and 20833Hz (^{13}C) in F1. The FIDs were multiplied by phase shifted sine functions in both dimensions and linear prediction applied to the F1 dimension to give final data sets of 1024 x 1024 data points.

12. Coupling constant measurement

In most cases it was possible to measure multiplet coupling constants directly from the recorded transformed 1D-NMR spectra. In cases where it was not possible, due to peak overlap, using the MestRe-C NMR analysis program the following procedure was adopted.

The first derivative of the spectrum was calculated and a Hilbert transformation applied to re-generate the imaginary part of the spectrum. Phase correction was then applied to make all analysed peaks positive. Line fitting (de-convolution) was then used to optimise peak numbers, positioning and shape composing an individual multiplet. By default, MestReC uses the Levenberg-Marquardt and Downhill Simplex fitting methods in combination. Optimisation was deemed to be successful when the sum of the areas of the

optimised peaks reproduced the peak shape of the analysed multiplet to a high degree of accuracy. Coupling constants were then measured between optimised de-convoluted peaks.

13. Molecular modelling

Molecular modelling was carried out using the AMBER 8 molecular modelling suite (Case *et al.* 2004). Generously provided free of charge by Professor David Case (The Scripps Research Institute, California, USA). For modelling purposes all carboxyl, sulphate and sulphamate groups were modelled in their de-protonated form and all charge models calculated accordingly.

13.1 Iduronate Coupling constant analysis

Initial coordinates for the methyl glycosides of IdoUA and IdoUA(2S) in the 1C_4 conformation were based on geometries taken from the protein data bank file 1HPN (Mulloy *et al.* 1993). These structures were then submitted to the antechamber module of AMBER 8 and AM1-BCC charges (Jakallan *et al.* 2002) and GAFF atom types assigned (Wang *et al.* 2004). The 2S_0 and 4C_1 geometries were taken from a 5ns generalized Born implicit solvent molecular dynamic (MD) simulation of the 1C_4 models at a temperature of 400K. At this temperature a number of transitions occur between all three conformations (data not shown). AM1-BCC charges were then re-assigned for randomly chosen 2S_0 and 4C_1 structures.

For explicit water simulations, each conformation was placed in a truncated octahedral box of explicit TIP3P water molecules. An 8Å buffer of water was generated around the

monosaccharide to the edge of the water box in each dimension. The charge on the monosaccharides was then neutralised through the addition of sodium ions using the addions method implemented in the xleap module. Standard molecular modelling techniques were then used to equilibrate the system prior to a 1ns 300K molecular dynamics (MD) run, as follows.

During equilibration the initial position of water molecules and sodium ion(s) were minimised, while the conformation of the sugar was restrained ($50 \text{ Kcal mol}^{-1} \text{ \AA}^{-1}$) (1500 steps steepest descent minimisation, 1500 steps conjugate gradient minimisation). The restraints were then removed and the whole system minimised (1000 steps steepest descent minimisation, 1500 steps conjugate gradient minimisation) before weak restraints were re-applied to the sugar ($10 \text{ Kcal mol}^{-1} \text{ \AA}^{-1}$) and the system allowed to heat up to 300K at constant volume for 20ps. Restraints were again removed and the density of the water was then allowed to relax during 100ps of MD using a constant pressure periodic boundary with an average pressure of 1 atm. During this final equilibration step and the subsequent MD run the temperature was maintained at 300K. In all equilibration steps and MD runs the Particle Mesh Ewald electrostatic treatment was implemented using the sander default settings and a cut off distance of 6\AA . The SHAKE algorithm was also applied to all bonds involving hydrogen and a 2fs MD time step was used. The system coordinates were output to a trajectory file every pico-second, resulting in a final trajectory file for each MD run consisting of 1000 frames/structures.

The program mdXvu (see <http://sourceforge.net/projects/mdxvu>) was used to analyse the explicit water MD trajectories and extract theoretical coupling constants for the sugar in each frame according to the Haasnoot *et al.* equation (Haasnoot *et al.* 1979). It was also

used to monitor values of the Cremer-Pople ring puckering parameters θ and ϕ_2 (pyranose ring atoms were selected in the order O5-C1-C2-C3-C4-C5).

13.2 Partial charge calculation methods- Method 1 (partial charge set 1)

Initial coordinates for all monosaccharides were based on geometries taken from the protein data bank file 1HPN (Mulloy *et al.* 1993). All monosaccharides were modelled with a hydroxyl group capping carbons 1 and 4 of the pyranose ring, where a glycosidic linkage normally occurs. These structures were then submitted to the antechamber module of AMBER 8 and AM1-BCC charges assigned (Jakallan *et al.* 2002). The hydroxyl group from carbon position 1 (except for units needed to form a non-reducing end terminus), and the hydrogen of the hydroxyl group attached to carbon position 4 were then deleted from each monosaccharide structure. Their charges were then spread equally among the remaining atoms, adjusting each charge slightly.

13.3 Calculation of potential energy surfaces across glycosidic linkages

Calculations were carried out using the Glycam04 force field for carbohydrates, supplied with the AMBER software suite. Additional parameters for O- and N-sulphates groups were added (Huige and Altona 1995) and parameters for the C4 to C5 double bond of non-reducing end uronates were taken from the General Amber Force Field (GAFF). All 1-4 and greater, electrostatic interactions and van der Waals interactions were treated equally (SCEE=SCNB=1) (Kirschner and Woods 2001).

Conformational energy maps were calculated for each glycosidic linkage as a function of the Φ and Ψ torsion angles as defined in the introduction (section 17.1). A dielectric

constant of 4 was used (Chuang *et al.* 2000). Φ and Ψ were varied from -160° to $+180^\circ$ and disaccharide structures were energy minimised every 20° with an Φ and Ψ torsion restraint of $10000 \text{ Kcal mol}^{-1} \text{ \AA}^{-1}$ until the RMS derivative for the whole structure was less than 0.01.

13.4 Whole molecule molecular dynamic simulations

Molecular dynamic simulations were carried out using the same modified version of the glycam04 parameter set used in the potential energy surface calculations (see section 13.3 above).

The hexasaccharides were neutralised, through the addition of sodium ions, using the additions method implemented in the xleap module, and placed in a truncated octahedral box of explicit TIP5P water molecules. The use of TIP5P water necessitates the addition of parameters to describe the oxygen lone pairs present in this water model. These were taken from the general amber force field (GAFF). An 8\AA buffer of water was generated around the hexasaccharide to the edge of the water box in each dimension.

Each system was then equilibrated prior to a 1ns production molecular dynamics (MD) simulation. The equilibration and production molecular dynamic protocol were essentially the same as described under section 13.1. The only difference being that minimisation of the whole system without restraints (1000 steps steepest descent minimisation, 1500 steps conjugate gradient minimisation) was repeated twice. This was due to the fact that the TIP5P water box generated by xleap/AMBER is not pre-equilibrated, unlike the TIP3P water box.

14. Additional computational modelling

14.1 Partial charge calculation method - Method 2 (partial charge set 2)

Partial charge calculation methods 2 and 3 were carried out in collaboration with Dr. Francois-Yves Dupradeau (University of Amien, France).

Initial coordinates for all monosaccharides were based on geometries taken from the protein data bank file 1HPN (Mulloy *et al.* 1993). All monosaccharide structures were capped with an O-Methyl group at carbon position 1 and a hydroxyl group at carbon position 4. For all monosaccharide structures, N-acetyl, sulphamate and sulphate groups were then replaced by hydroxyl groups. Quantum mechanical geometry optimisation of these structures was then carried out at the HF/6-31G* level of theory, using the Gaussian 98 software package. No dihedral angle restraints were applied during the optimisation. This results in structures containing a maximum number of intra-molecular hydrogen bonds. The geometry of N-methyl amide, N-methyl sulphamate and methyl sulphate were optimised independently.

For each optimised monosaccharide conformation a molecular electrostatic potential (MEP) computation was carried out at the HF/6-31G* level, using the CHELPG algorithm as implemented in the Gaussian package. Using the RESP program (part of the AMBER 8 molecular modelling suite) partial charges were then calculated for all atoms that reproduce the calculated MEP.

For this calculation four molecular orientations based on the rigid body re-orientation algorithm (RBRA) implemented in the R.E.D. program were used (atom sets used in the RBRA approach were C1, C3, C5 and C5, C3, C1 and C2, C4, O5, and O5, C4, C2). Two molecular orientations based on the RBRA approach were also used for N-methyl amide,

(atom sets were; N-methyl carbon, N-methyl nitrogen, carbonyl carbon and carbonyl carbon, N-methyl nitrogen, N-methyl carbon), N-methyl sulphamate (atom sets were; N-methyl carbon, N-methyl nitrogen, sulphamate sulphur, and sulphamate sulphur, N-methyl nitrogen, N-methyl carbon) and methyl sulphate (atom sets were; methyl carbon, ester oxygen, sulphate sulphur, and sulphate sulphur, ester oxygen, methyl carbon).

As specified by the Glycam04 force field developers, the charge value of the hydrogen atoms linked to any sp³ carbon atoms were set to zero during the fitting procedure using intra-molecular charge constraints.

The charge values for atoms involved in potential glycosidic linkages were calculated using an inter-molecular charge constraint where the charge values of the C1-methyl and HO4-O4 hydroxyl groups of the two considered monosaccharides were required to have a net charge of zero. Likewise, the N-acetyl, sulphamate, and sulphate groups were linked to the correct position of their monosaccharides using an inter-molecular charge constraint set to a net charge of zero. These were between the target monosaccharide hydroxyl group and the N-methyl group of N-methyl amide or the methyl group of N-methyl sulphamate or alternatively methyl sulphate, respectively.

The whole procedure was automatically carried out using a beta version of the R.E.D. IV program, currently under development in the laboratory of Dr. Dupradeau (University of Amien, Amien, France).

14.2 Partial charge calculation method - Method 3 (partial charge set 3)

Partial charge calculation method 3 was similar to method 2. The only difference being that during the initial quantum mechanical geometry optimisation of the structures, care

was taken to avoid the formation of any intra-molecular hydrogen bonds. This was done by constraining a number of dihedral angles during the initial quantum mechanical geometry optimisation.

For Δ UA, a dihedral constraint of -60° was used for the HO3-O3-C3-H3 dihedral angle in the ${}^2\text{H}_1$ conformation. For the α -methyl glycoside of glucose (gg rotamer) a dihedral constraint of 180° was used for the HO4-O4-C4-H4 dihedral angle. All other dihedral angles involving hydroxyl groups, in all structures, were initially set to 180° and allowed to move freely during optimisation.

RESULTS

1. Initial studies into the degradation of HS by the heparinase III enzyme.

Trial heparinase III digests were carried out in order to ascertain the minimum amount of enzyme needed to digest 200mg of celsus HS to completion. Progress of digestion was monitored through changes in the gel-filtration profile of the sample. These profiles are shown in Figure 20.

After 4 additions of 10mU of heparinase III and incubation for an 88 hour period (total) digestion can be said to be complete. The addition of a further 10mU of enzyme and incubation for a further 24 hour period did not alter the gel filtration profile significantly (compare Figure 20, panels D and E). Incubation time as opposed to amount of enzyme added was also investigated. See Figure 21. Incubation with 30 mU of enzyme for a time period additional to 64 hrs was not sufficient to bring about a significant amount of further digestion. This result also verified the approach of adding the enzyme in 10 mU aliquots to the sample. The addition of a further 10mU of enzyme at 64 hrs did result in further digestion (compare Figure 20, Panel D with Figure 21, Panel A).

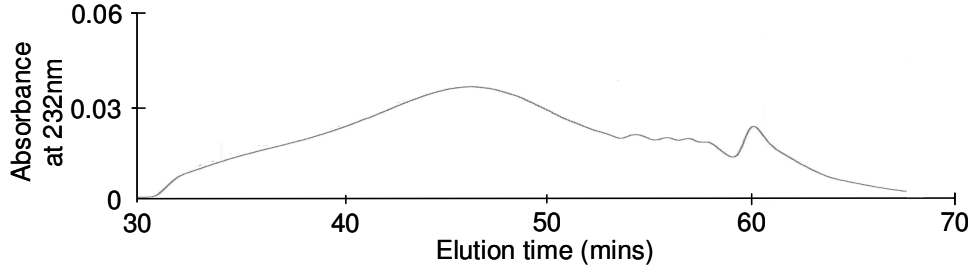
Figure 20 Analytical TSK3000PW gel-filtration profiles of HS after cleavage with heparinase III.

Samples of 10 μ g from a 200mg total digest of Celsus HS by the heparinase III enzyme were chromatographed on a TSK3000PW analytical gel-filtration column (30 x 0.75 cm).

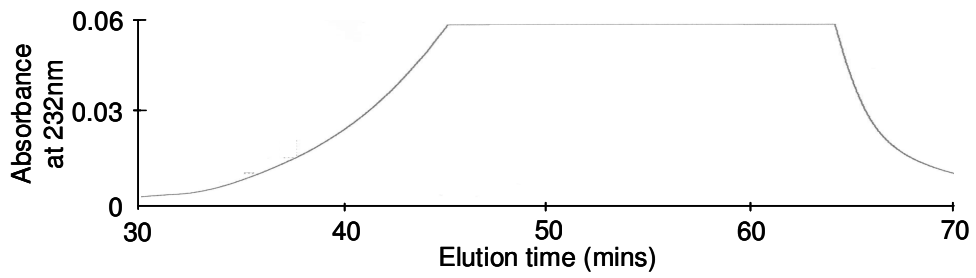
Panel A, 16 hrs after the addition of 10mU of heparinase III; Panel B, 24 hrs after the addition of a further 10mU of heparinase III; Panel C, 24 hrs after the addition of a further of 10mU of heparinase III; Panel D, 24 hrs after the addition of a further of 10 mU heparinase III; Panel E, 24 hrs after the addition of a further of 10mU of heparinase III.

Figure 20

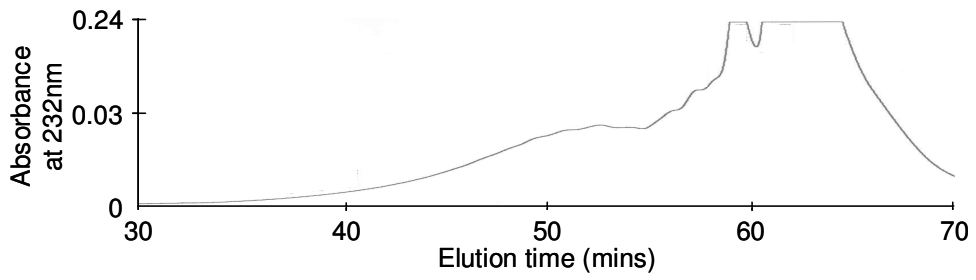
Panel A



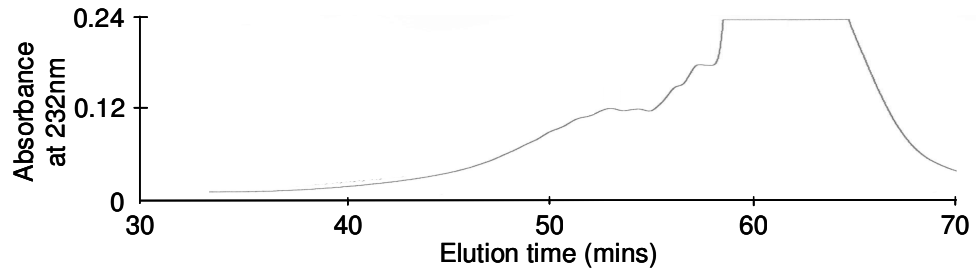
Panel B



Panel C



Panel D



Panel E

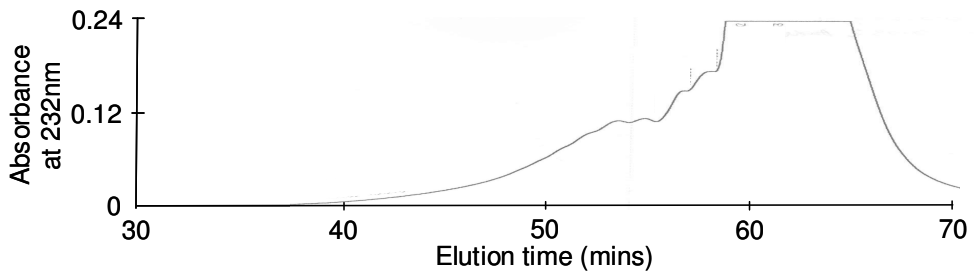


Figure 21

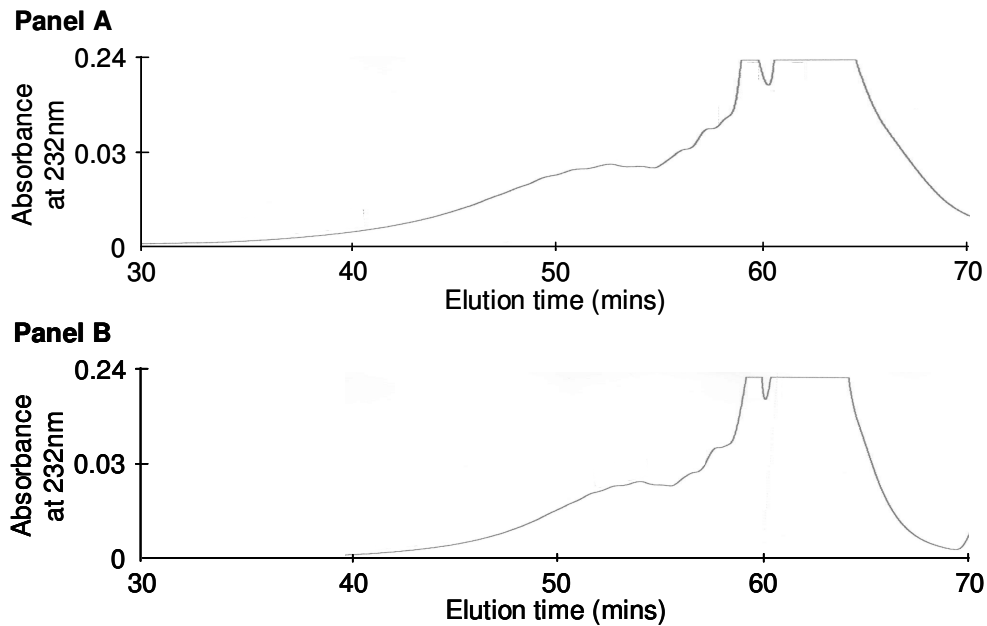


Figure 21 Comparison of the analytical gel-filtration profiles of HS digested with 30mU of heparinase III for 64 or 160 hours.

Celsus HS (200mg total) was digested with 30mU of heparinase III for 64 or 160 hrs. Samples of 10 μ g were then chromatographed on a TSK3000PW analytical gel-filtration column (0.75cm x 30cm). Panel A shows the gel filtration profile after 64 hrs of incubation and panel B after 160 hrs incubation.

2. Bio-gel P10 gel filtration chromatography of HS digested to completion with heparinase III.

A typical Bio-Gel P10 filtration profile is shown in Figure 22 (panel A) for HS digested to completion with heparinase III. A typical Bio-Gel P10 filtration profile for the commercial low molecular weight heparin product Tinzaparin Sodium (Innohep®) is also shown (Figure 22, panel B).

A series of oligosaccharides ranging in size from dp2 up to dp18 were separable for the HS sample. In comparison to the heparin profile the oligosaccharide peaks for HS are considerably broader. The elution position for HS and heparin derived oligosaccharides of the same length are also different. V_0 and V_t remained un-altered for both profiles corresponding approximately to fractions 50 and 165 respectively.

Fractions corresponding to separate peaks were either pooled and stored (dp2 and >dp10), or taken on for further analysis (dp4-dp10).

Figure 22

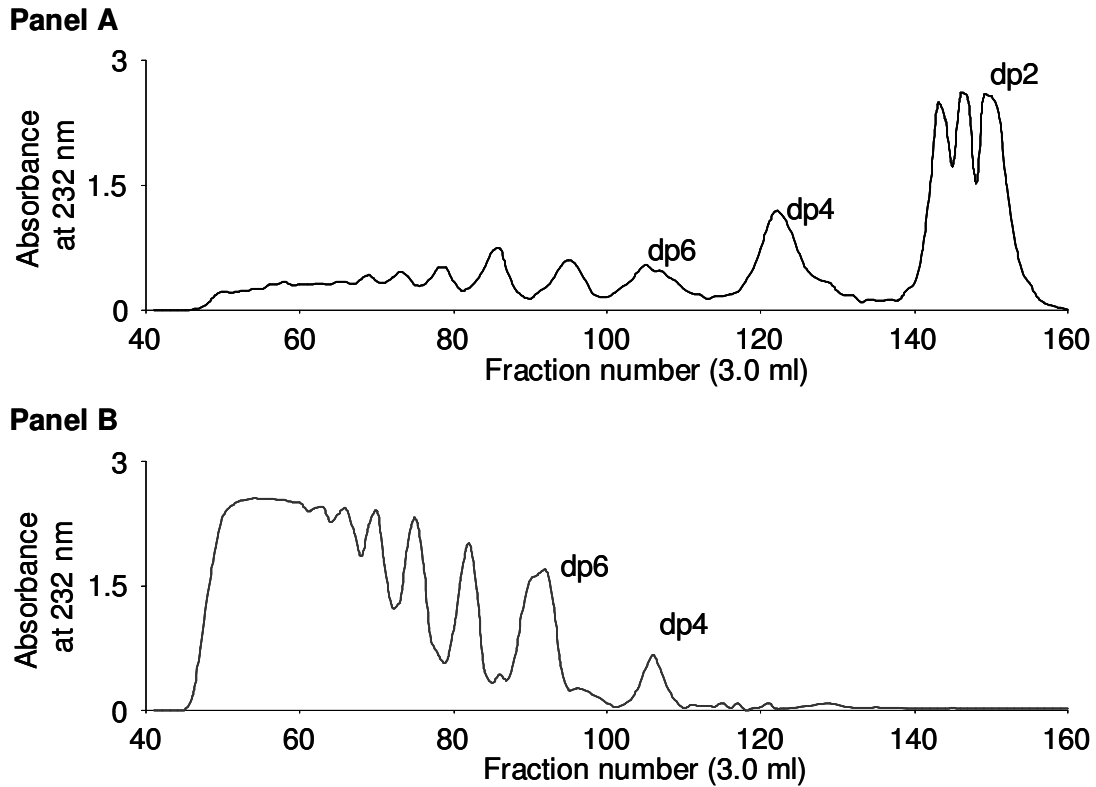


Figure 22 Preparative Bio-Gel P10 gel filtration profiles of HS digested to completion with heparinase III and the commercially available low molecular weight heparin product Tinzaparin sodium.

HS (200 mg) digested to completion by the addition of 40mU of heparinase III over an 88 h period and Tinzaparin sodium (1 vial, approx 300 mg) were chromatographed on a Bio-Gel P10 gel-filtration column (3cm x 120cm). Panel A, heparinase III digested HS; Panel B, Tinzaparin sodium.

3. SAX-HPLC chromatography of sized HS dp4-dp10 oligosaccharides.

In order to examine the number of charged species of oligosaccharide within the HS dp4-dp10 size pools, a sample of each was applied to a ProPac PA-1 SAX-HPLC column. Elution gradients are as described under experimental procedures. These results are shown in Figure 23.

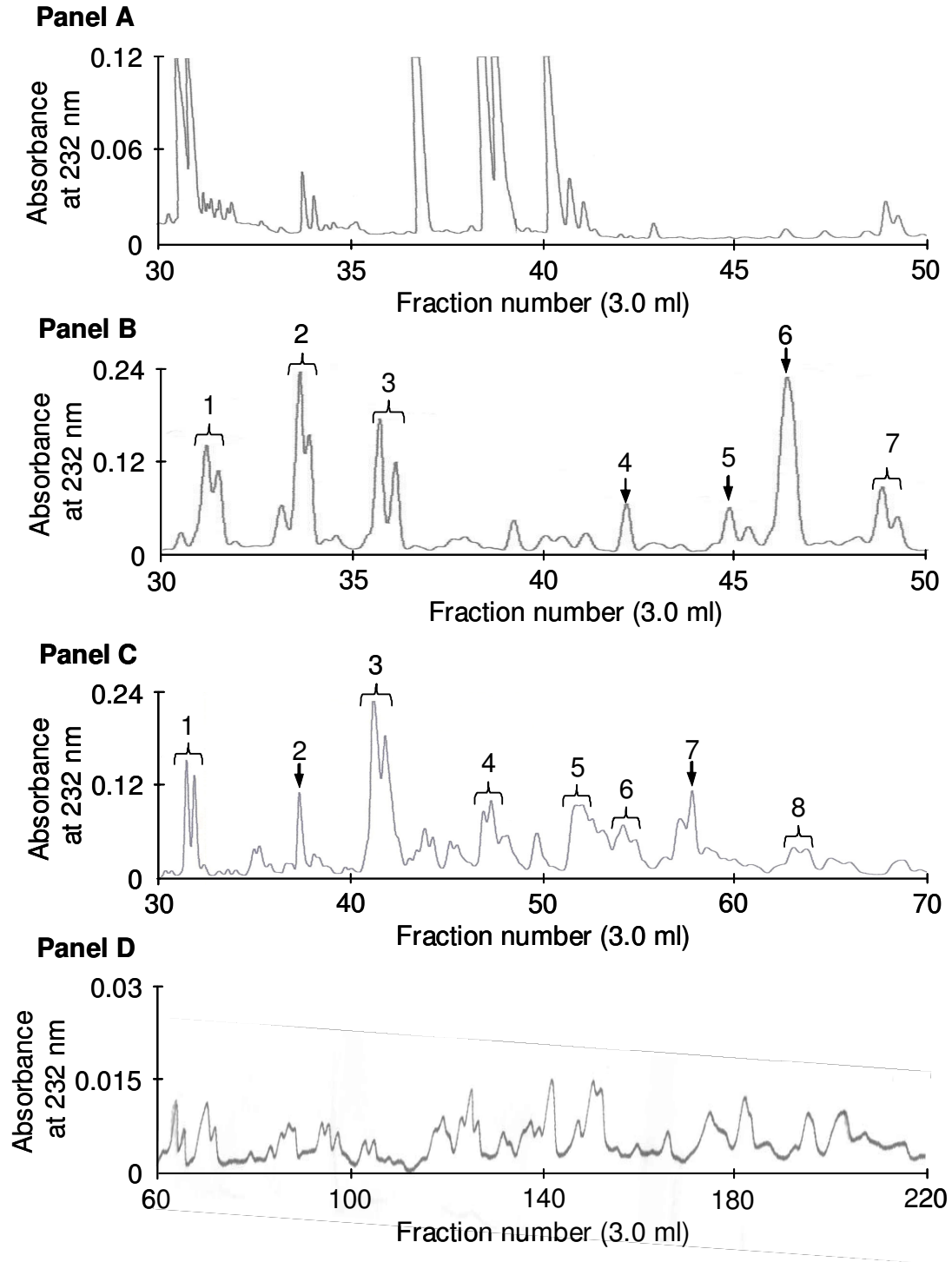
Due to the large amount of previous work undertaken on the dp4 oligosaccharides (Hileman *et al.* 1997), their predicted lack of biological activity (see introduction) and the high degree of heterogeneity encountered in the dp10 pool, further analytical efforts were directed towards the dp6 and dp8 peaks numbered in Figure 4, panels B and C.

Figure 23 SAX-HPLC chromatography of sized HS oligosaccharides.

The dp4-dp10 oligosaccharide pools generated by heparinase III digestion of HS and purified by Bio-Gel P10 chromatography (see Figure 22) were individually applied to a ProPac PA-1 SAX-HPLC column (250mm x 9cm) (see experimental procedures for gradients used). Each peak or groups of peaks collected and taken on further analysis by PAGE are numbered (panels B and C).

Panel A, dp4 pool; Panel B, dp6 pool; Panel C, dp8 pool; Panel D, dp10 pool.

Figure 23



4. PAGE analysis of dp6 and dp8 SAX-HPLC peaks.

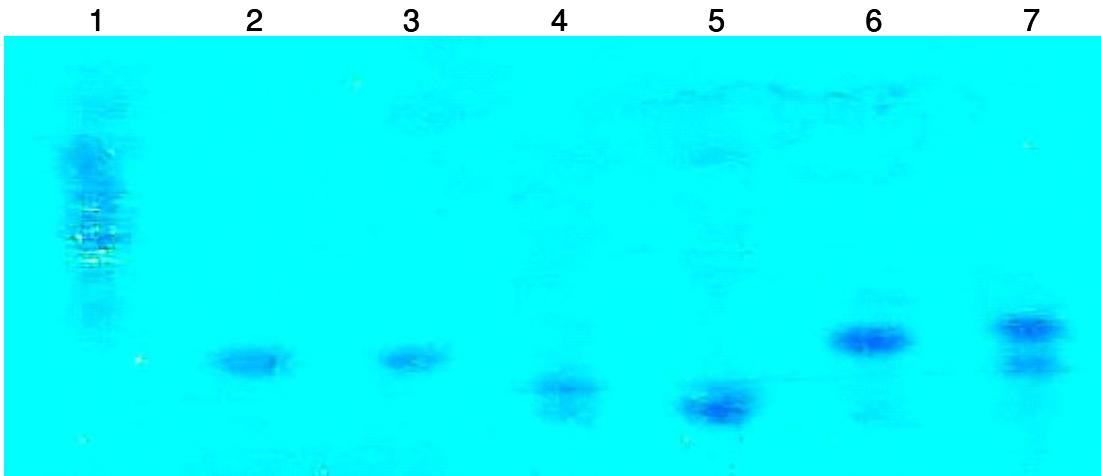
Dp6 peaks 1 to 7 and dp8 peaks 1 to 8 were analysed by PAGE. These results are shown in Figure 24.

In the cases of dp6-2, dp6-3, dp8-1 and dp8-3 the SAX-HPLC profile indicated two charged species to be present, these pools were made up of two closely eluting peaks (see Figure 23, panels B and C). On PAGE analysis only one charged species was detected. In other cases, most clearly seen in Figure 24 for dp8-7, a single peak from the SAX-HPLC profile was separated into more than one charged species on PAGE analysis. For dp8-7, one major and one minor band were observed. Similarly, more than one band could be seen for dp6-4 and dp8-2 on initial de-staining of the gel.

Based on these results, peaks containing predominantly a single charged oligosaccharide species were analysed for their disaccharide composition.

Figure 24

Panel A



Panel B

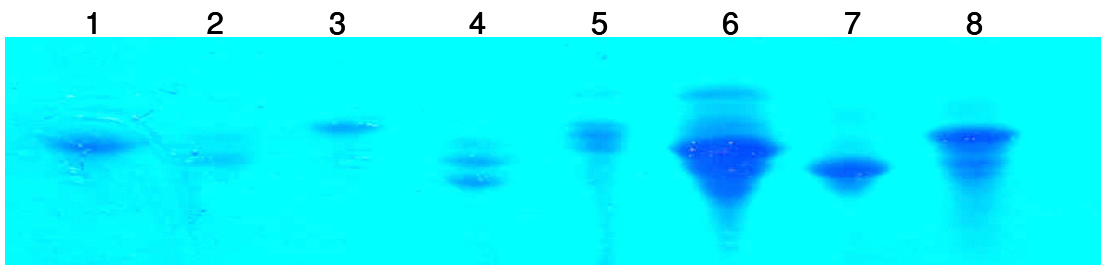


Figure 24 PAGE analyses of selected peaks from the HS dp6 and dp8 SAX-HPLC profile.

Individual peaks or groups of peaks from the dp6 and dp8 SAX-HPLC profile (see Figure 23) were pooled, desalted and quantified (see experimental methods).

Approximately 20 μ g of material was applied to each lane.

Panel A shows the dp6 analysis, panel B the dp8 analysis. Lane numbers correspond to the peaks labelled in Figure 23, panels B and C.

5. Disaccharide analysis of selected dp6 and dp8 SAX-HPLC peaks.

The disaccharide analysis of selected dp6 and dp8 SAX-HPLC peaks shown to contain predominantly a single charged oligosaccharide species by PAGE, are shown in Table 2.

These results clearly indicate that although a single species is indicated on the PAGE analysis of the dp8 pools, a significant number of minor species must also be present, warranting their further purification. In contrast all dp6 pools analysed were shown to contain mainly a single charged species of oligosaccharide (see discussion).

Table 2 Disaccharide analyses of selected peaks from the HS dp6 and dp8 SAX-HPLC profile.

Individual peaks or groups of peaks from the dp6 and dp8 SAX-HPLC profile (see Figure 23), that were indicated to contain predominantly one charged species of oligosaccharide by PAGE analysis (see Figure 24), were analysed for their disaccharide composition. Oligosaccharides were broken down into their component disaccharides by combined heparinase digestion and analysed by SAX-HPLC (see experimental procedures).

	Oligosaccharide size and pool number						
	dp6- 2	dp6- 3	dp6- 5	dp6- 6	dp8- 1	dp8- 3	dp8- 7
Disaccharide	Disaccharide as a % of total disaccharides						
UA-GlcNAc	37	34	0	37	48	40	8
UA-GlcNS	34	33	38	36	14	26	39
UA-GlcNAc(6S)	0	0	62	0	7	6	24
UA(2S)-GlcNAc	0	0	0	0	0	0	0
UA-GlcNS(6S)	0	0	0	0	11	3	7
UA(2S)-GlcNS	28	34	0	0	7	9	7
UA(2S)-GlcNS(6S)	0	0	0	26	13	16	16

6. Variation between SAX-HPLC profiles.

The peaks dp6-2, dp6-3, dp6-5 and dp6-6 here after referred to as oligosaccharides dp6-2, dp6-3, dp6-5 and dp6-6 were deemed suitable initial targets for purification in the quantities required for NMR analysis.

However, on initiating large scale purification of these oligosaccharides, significant variation was observed in the quantities of dp6-5 and dp6-6, present in dp6 size pools from different digests. These results are shown in Figure 25. Frequently accurate pooling of dp6-5 and dp6-6 without a high level of cross contamination was un-achievable (see Figure 25, panel B).

Figure 25

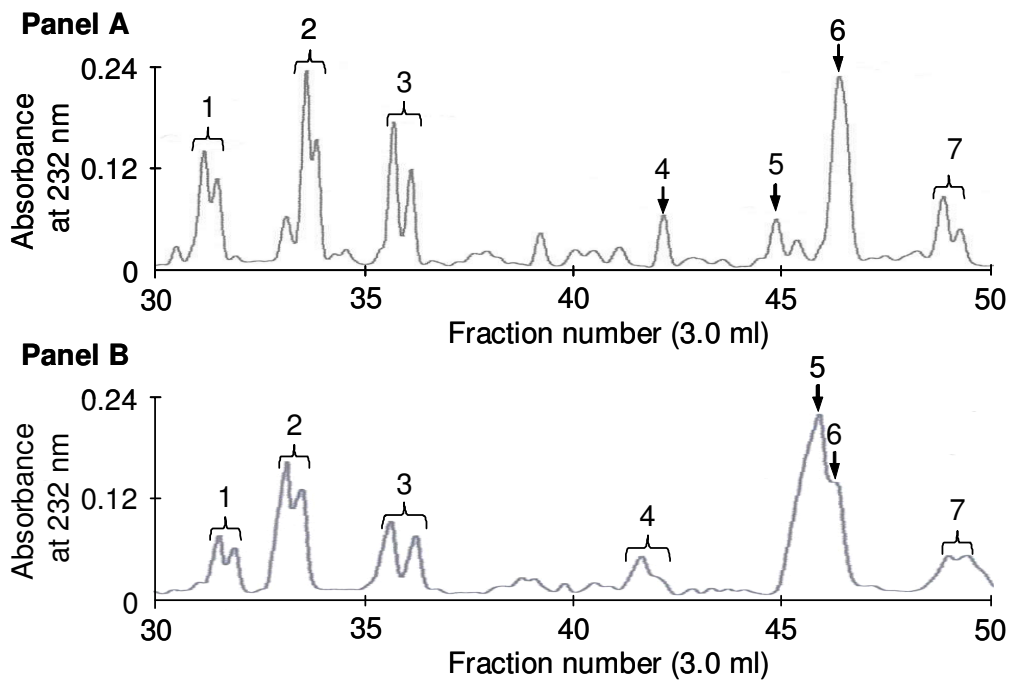


Figure 25 Variation of the SAX-HPLC profile of dp6 size pools purified from different heparinase III digests of HS.

Dp6 oligosaccharides generated by heparinase III digestion of HS and purified by Bio-Gel P10 chromatography (see Figure 22) were applied to a ProPac PA-1 SAX-HPLC column (250mm x 9cm). Panel A and panel B show the SAX-HPLC profile of dp6 size pools purified from two different 200mg HS digests.

7. Purification strategy.

The strategy adopted was to target the purification of oligosaccharides dp6-2 and dp6-3. When dp6-5 and dp6-6 were clearly resolved into two peaks, which was in a minority of cases, these peaks were pooled separately otherwise a single peak was pooled for later further purification using PAGE.

In total to produce 2mg each of dp6-2 and dp6-3, **fifty** identical 200mg HS digests were carried out using a total of 10 grams of HS and 40 mU of enzyme per digest (2 U total).

8. PAGE purification of dp6-5 and dp6-6.

Where a single peak had been pooled containing both oligosaccharides dp6-5 and dp6-6 these samples were concentrated, desalted and subjected to PAGE (see experimental procedures). An example of the banding pattern observed is shown in Figure 26.

Bands corresponding to oligosaccharides dp6-5 and dp6-6 were cut out from the gel, extracted from the polyacrylamide and the azure A stain removed (see experimental procedures). The PAGE separated oligosaccharides were then added to the oligosaccharides separated by SAX-HPLC alone.

In total approximately 1mg of dp6-5 was produced and 0.5mg of dp6-6.

Figure 26

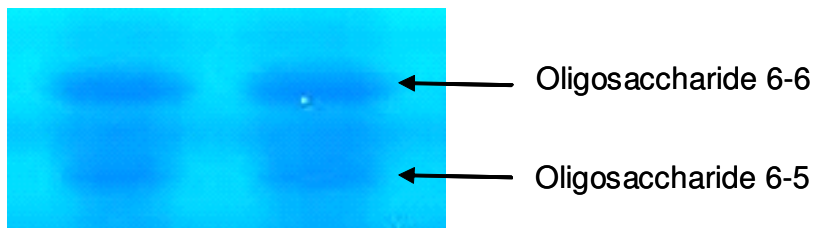


Figure 26 Further purification of dp6-5 and dp6-6 by PAGE.

Single broad peaks containing both oligosaccharides dp6-5 and dp6-6 (see Figure 25, panel B) were pooled, desalted and quantified (see experimental procedures). Approximately 20 μ g of material was applied to each lane. Bands corresponding to oligosaccharides dp6-6 and dp6-5 are labelled.

9. Preliminary NMR analysis of dp6-2 and dp6-3.

Oligosaccharides dp6-2 and dp6-3 were analysed by ^1H 1-D NMR spectroscopy. These results are shown in Figure 27. By reference to literature chemical shift values, reported mainly for heparin derived oligosaccharides (Chuang *et al.* 2001), a number of chemical shifts could be putatively assigned. These are labelled in Figure 27.

Figure 27

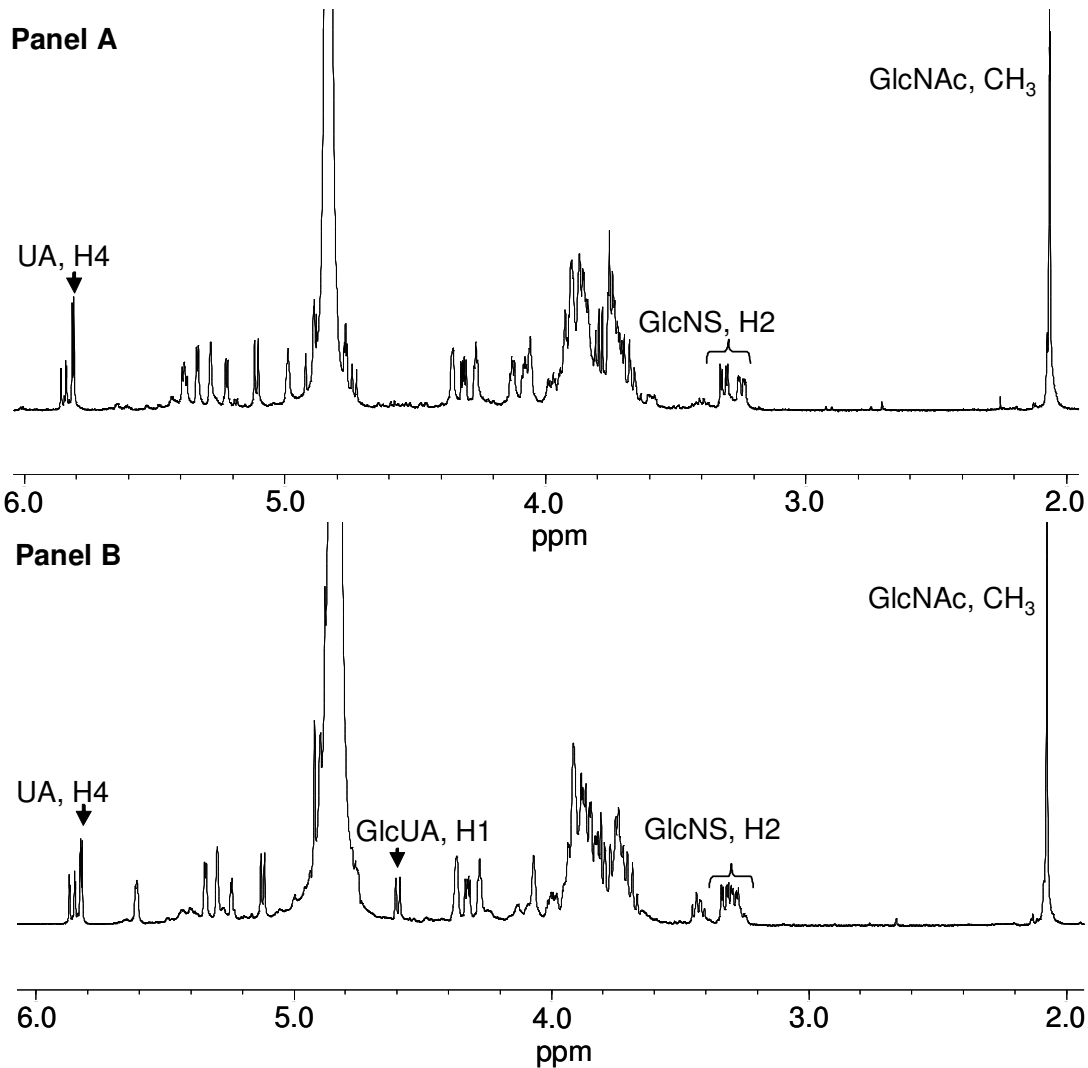


Figure 27 One-dimensional 500 MHz ¹H-NMR spectra of hexasaccharides 6-2 and 6-3.

Spectra were recorded in PBS at 25°C.

Panel A, dp6-2; Panel B, dp6-3.

Chemical shifts that could be putatively assigned by reference to reported literature values are labelled.

10. Full spectral assignment of dp6-2 and dp6-3.

Two dimensional COSY, TOCSY and NOESY NMR spectra were recorded for both oligosaccharides, these are shown in appendix I. The chemical shifts at around 2ppm, which correspond to the methyl protons of any N-acetyl groups, are not shown for all spectra in appendix I. No through-bond cross-peaks were detected in the COSY and TOCSY spectra for these three protons. Notably, also no through-space connectivity's were detected in all NOESY spectra. ^{13}C DEPT-135 and ^{13}C HSQC spectra were also recorded for these oligosaccharides and are shown in appendix II (see discussion).

COSY spectra (dp6-2, Figure 82, page 331, dp6-3, Figure 83, page 332),

TOCSY spectra (dp6-2, Figure 84, page 333, dp6-3, Figure 85, page 334)

NOESY spectra (dp6-2, Figure 86, page 335, dp6-3, Figure 87, page 336)

DEPT-135/HSQC spectra (dp6-2, Figure 103, page 352, dp6-3, Figure 104, page 353)

As a demonstration of the assignment process a detailed breakdown of the proton assignment of dp6-2 (Figures 28-33) and dp6-3 (Figures 34-37) are given.

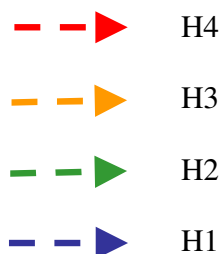
Assignment of proton chemical shifts using 2D-NMR experiments is generally regarded as standard practice, and a detailed breakdown of the process is not given for all subsequently characterised oligosaccharides.

For assignment purposes the non-reducing terminal uronic acid residue in each case was lettered residue "a" with each subsequent residue lettered accordingly, ending at the reducing terminus in residue number "f".

Figure 28 Assignment of the non-reducing terminal uronic acid, residue “a” of oligosaccharide dp6-2.

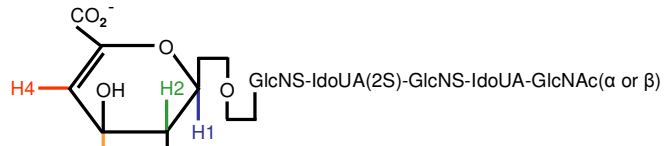
Using the COSY spectrum (Panel A) starting at the chemical shift for H4 assigned to 5.814ppm, H3 was assigned to 4.314ppm, H2 to 3.801ppm and H1 to 5.11ppm.

The coloured arrows (Panel A) correspond to cross-peaks involving a particular proton or the track of this proton across the spectrum. Cross-peaks forming the UA ring system are sequentially linked by the arrows.

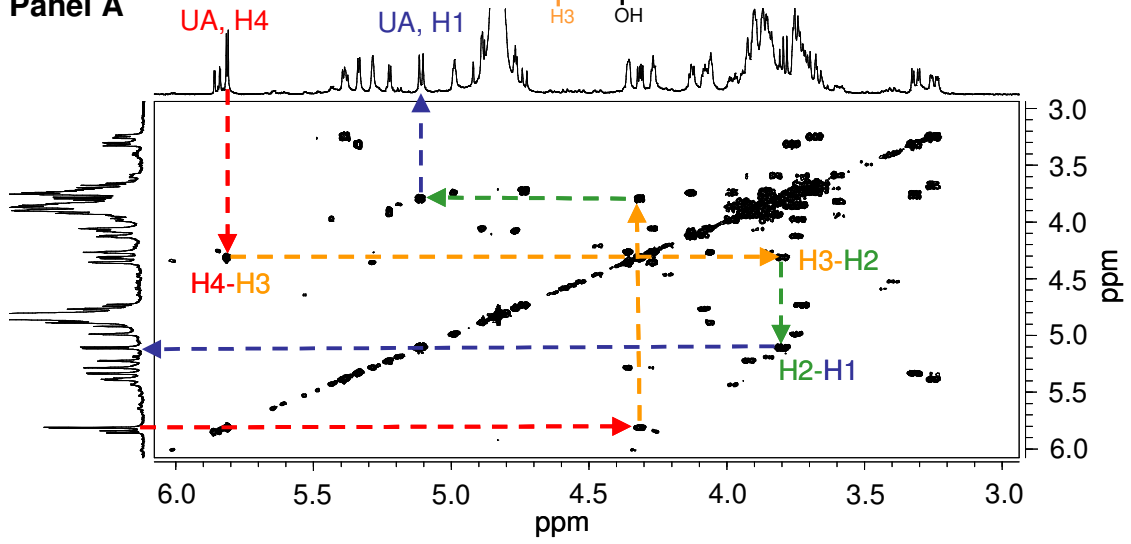


A single trans-glycosidic cross-peak was then identified in the NOESY spectrum (Panel B) from H1 of UA to an unassigned proton of the linking glucosamine residue at 3.866ppm. This cross-peak is circled in Panel B

Figure 28



Panel A



Panel B

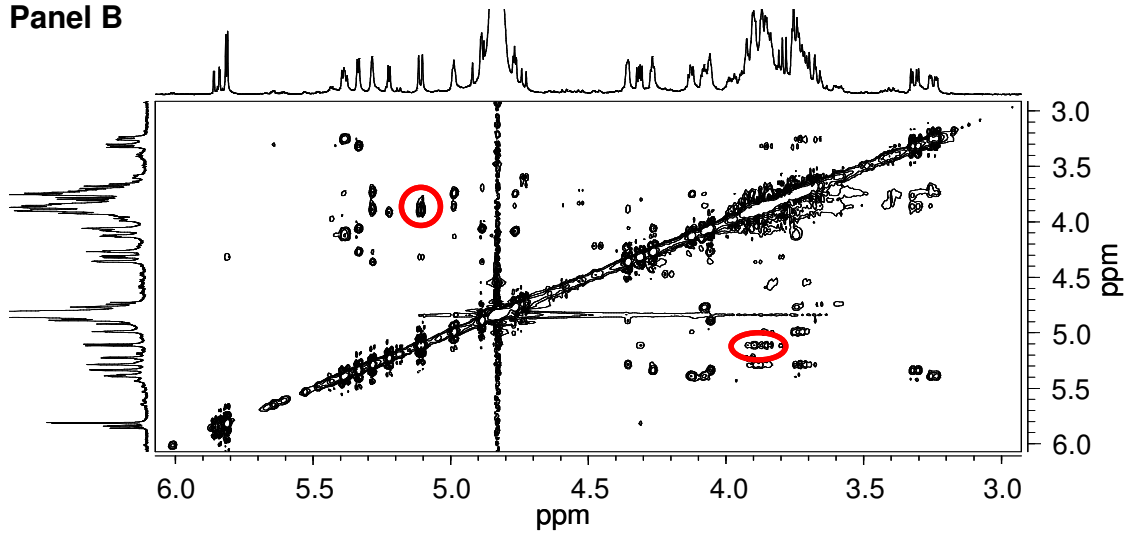
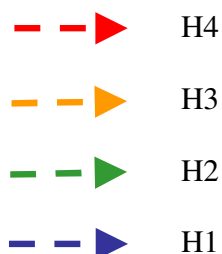


Figure 29 Assignment of glucosamine, residue "b" of oligosaccharide dp6-2.

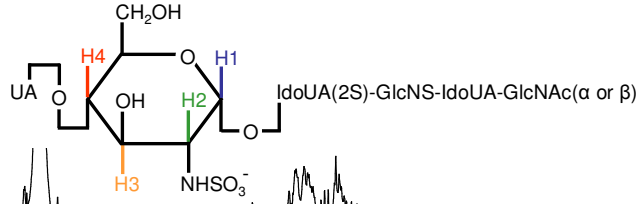
Using the TOCSY spectrum (Panel A) the anomeric proton H1 for the ring system containing the 3.866ppm unassigned glucosamine proton (circled in Figure 28, Panel B) was identified. This anomeric proton was assigned to 5.336ppm. Using the COSY spectrum (Panel B) H2 was then assigned to 3.314ppm, H3 to 3.761ppm and H4 to 3.866ppm, identifying this as the proton involved in the trans-glycosidic NOE from residue "a". H5 was later assigned to 3.837ppm and both H6 protons to a range between 3.837ppm and 3.925ppm, by analysis of a subsequently recorded HSQC spectrum (see discussion). Assignment of H2 to 3.314ppm identified this glucosamine as being N-sulphated.

The coloured arrows (Panel A and B) again correspond to cross-peaks involving a particular proton or the track of this proton across the spectrum. Cross-peaks for protons H1 to H4 forming part of the GlcNS ring system are sequentially linked by the arrows.

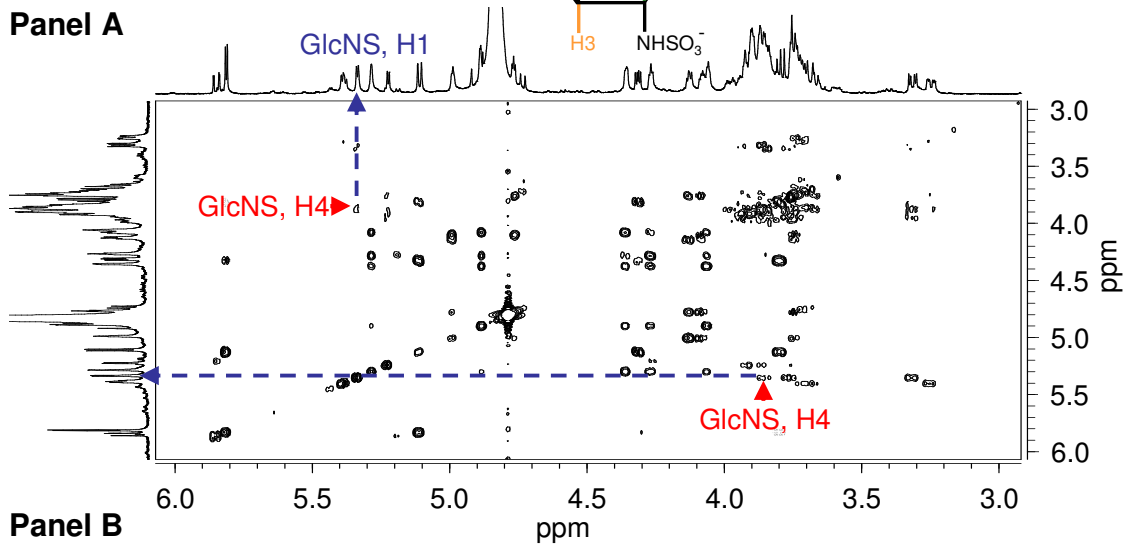


Two trans-glycosidic cross-peaks were identified in the NOESY spectrum (Panel C) from H1 of GlcNS to unassigned protons in the preceding IdoUA residue at 4.059ppm and 4.266ppm. These cross-peaks are circled.

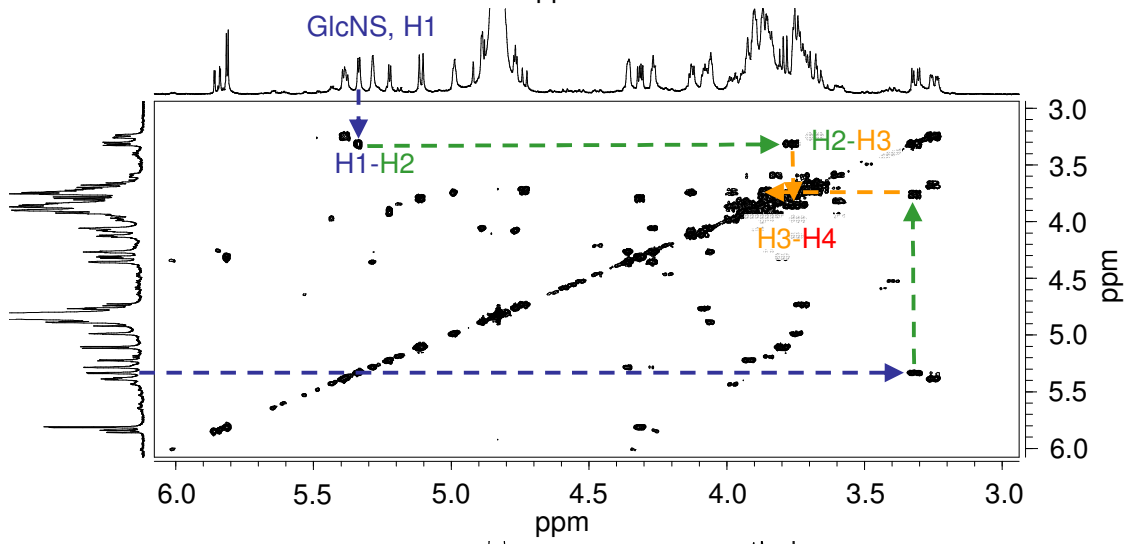
Figure 29



Panel A



Panel B



Panel C

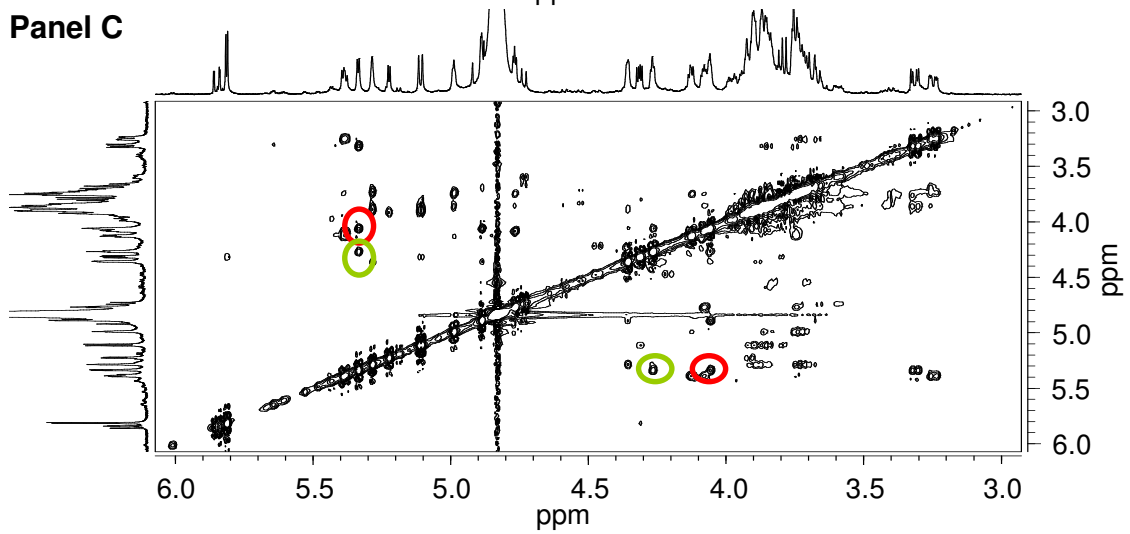


Figure 30 Assignment of iduronic acid, residue “c” of oligosaccharide dp6-2.

Using the TOCSY spectrum (Panel A) the anomeric proton H1 for the ring system containing the 4.059ppm and 4.266ppm unassigned iduronic acid protons (circled in Figure 29, Panel C) was identified. This anomeric proton was assigned to 5.285ppm. Using the COSY spectrum (Panel B) H2 was then assigned to 4.356ppm, H3 to 4.266ppm, H4 to 4.059ppm and H5 to 4.885ppm, identifying H3 and H4 respectively as the protons involved in the trans-glycosidic NOE from residue “b”. For clarity only assignments originating from H1 on the horizontal axis are tracked (Panel B). Assignment of H2 to 4.356ppm identified this iduronic acid as being 2-O-sulphated. The correspondence of the coloured arrows (Panel A and B) are with cross-peaks involving a particular proton or the track of this proton across the spectrum, as in Figures 28 and 29, with the addition of a H5 proton label/track.

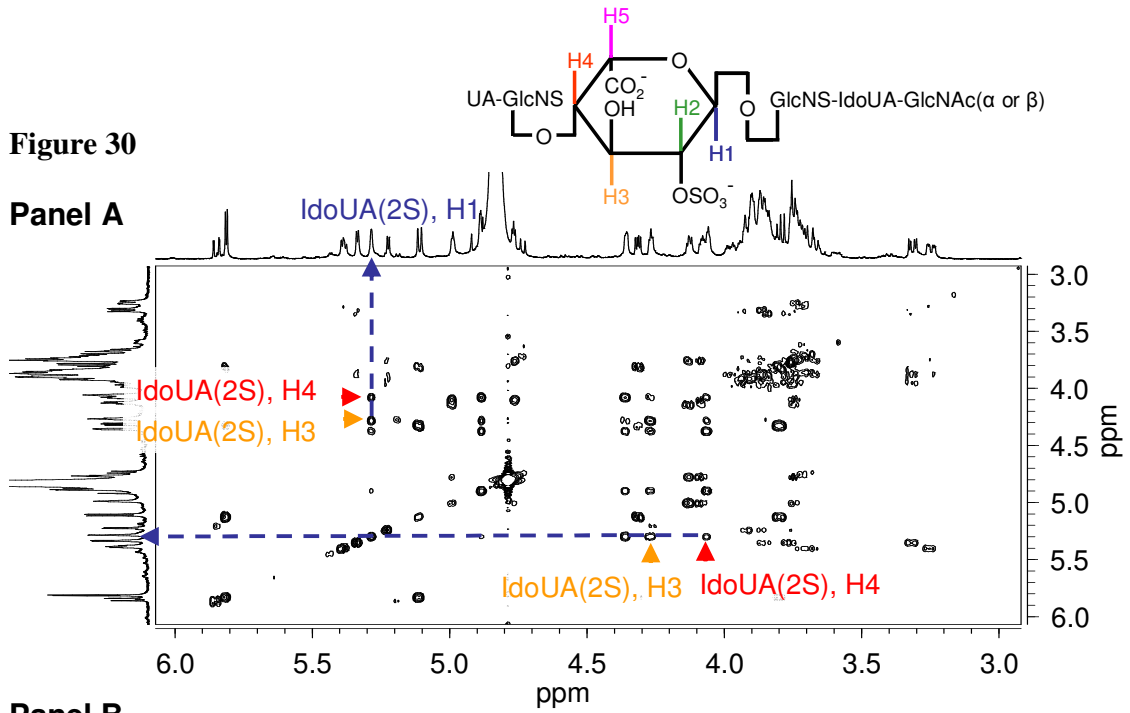
— —▶ H5

Cross-peaks forming the IdoUA(2S) ring system are sequentially linked.

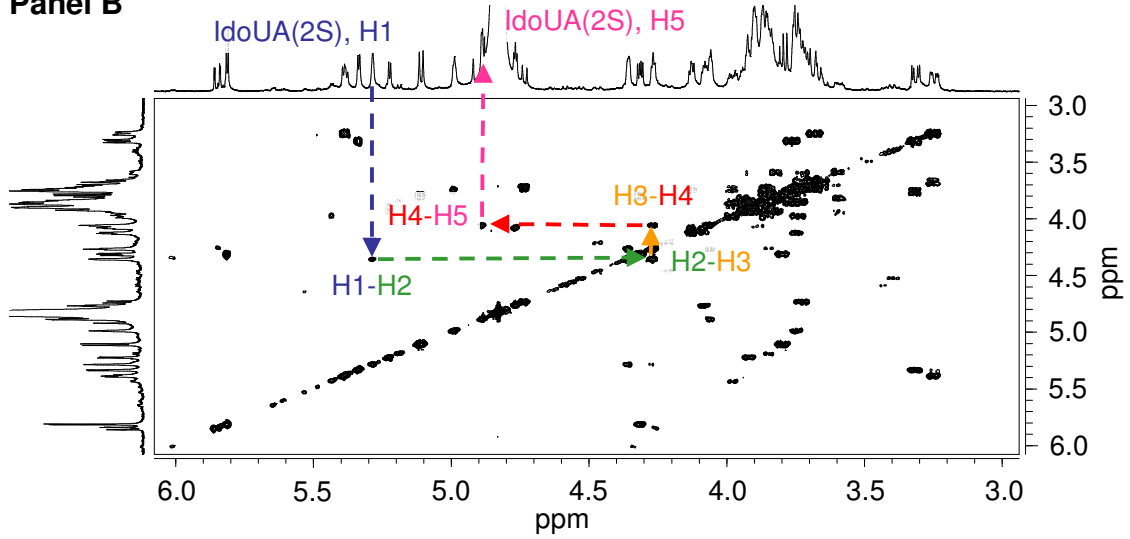
Two trans-glycosidic cross-peaks were identified in the NOESY spectrum (Panel C) from H1 of IdoUA(2S) to two unassigned protons in the preceding glucosamine residue at 3.72ppm and 3.889ppm. These cross-peaks are circled.

Figure 30

Panel A



Panel B



Panel C

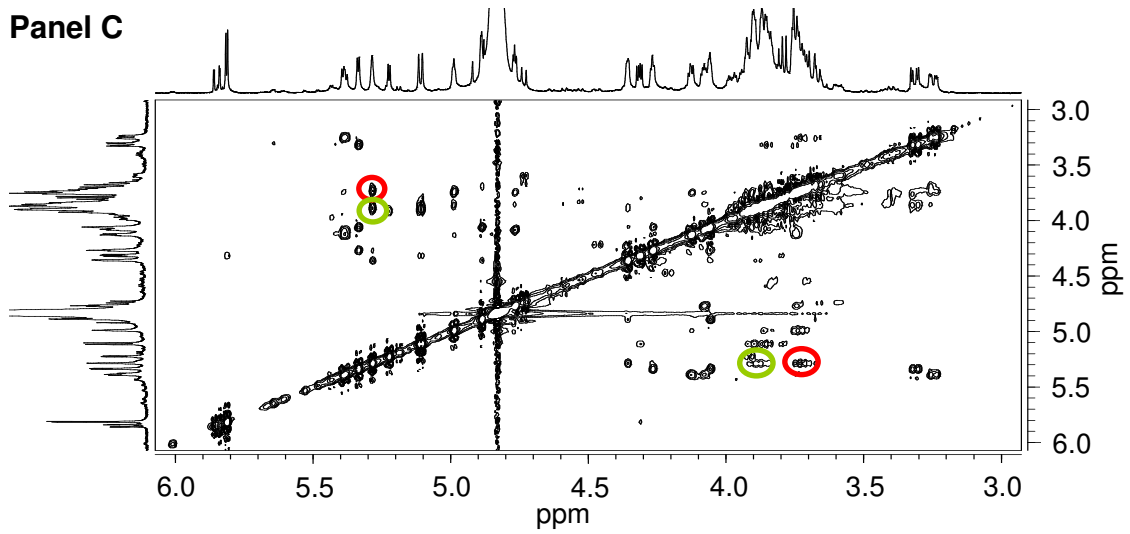


Figure 31 Assignment of glucosamine, residue “d” of oligosaccharide dp6-2.

Using the TOCSY spectrum (Panel A) the anomeric proton H1 for the ring system containing the 3.72ppm and 3.889ppm unassigned glucosamine protons (circled in Figure 30, Panel C) was identified. This anomeric proton was assigned to 5.391ppm. Using the COSY spectrum (Panel B) H2 was then assigned to 3.247ppm, H3 to 3.743ppm and H4 to 3.72ppm, identifying H4 as one of the two protons involved in trans-glycosidic NOEs from residue “c”. H5 was later assigned to 3.854ppm and both H6 protons to a range of between 3.889-3.905ppm (by HSQC analysis, see discussion). This assignment of H6 identified one of the H6 protons as involved in a trans-glycosidic NOE with residue “c”. Assignment of H2 to 3.247ppm identified this glucosamine as being N-sulphated. The correspondence of the coloured arrows (Panel A and B) are with cross-peaks involving a particular proton or the track of this proton across the spectrum as in Figures 28-30, with the addition of a H6 label.

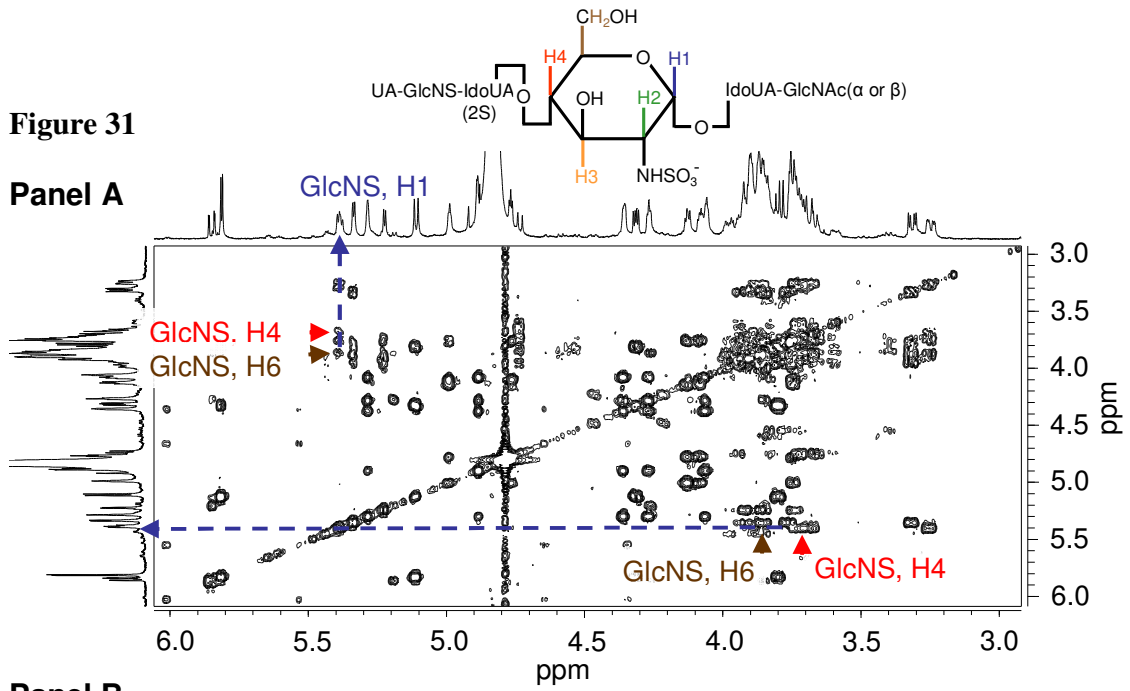
— — ► H6

Cross-peaks for protons H1 to H4 forming part of the GlcNS (residue 4) ring system are sequentially linked.

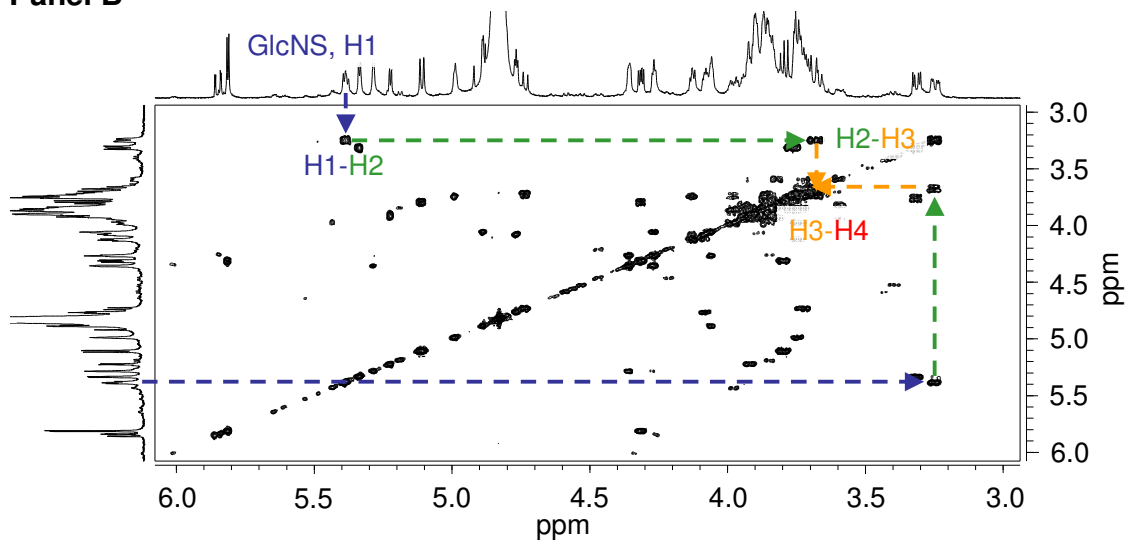
Two trans-glycosidic cross-peaks were identified in the NOESY spectrum (Panel C) from H1 of GlcNS to unassigned protons in the preceding IdoUA residue at 4.079ppm and 4.128ppm. These cross-peaks are very close together in the spectrum and circled.

Figure 31

Panel A



Panel B



Panel C

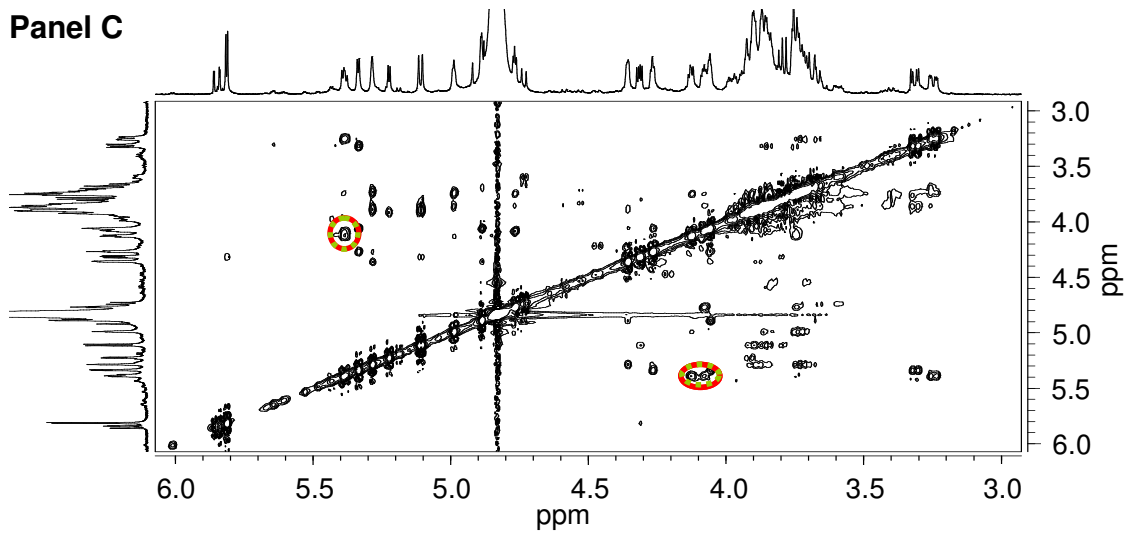


Figure 32 Assignment of iduronic acid, residue “e” of oligosaccharide dp6-2.

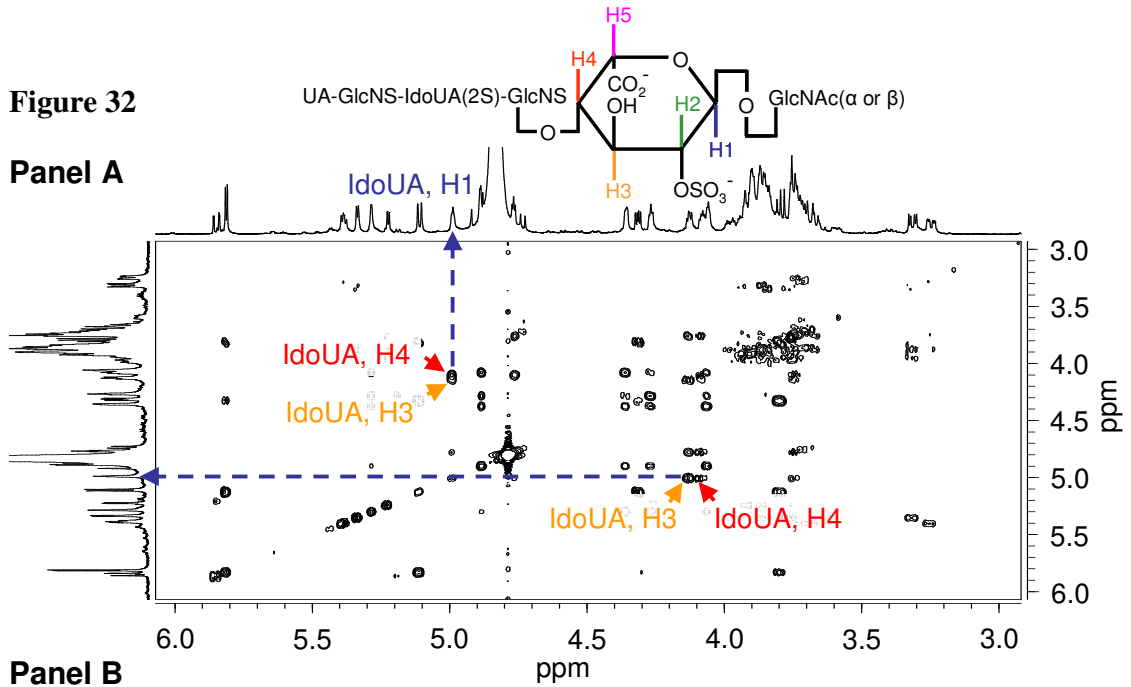
Using the TOCSY spectrum (Panel A) the anomeric proton H1 for the ring system containing the 4.079ppm and 4.128ppm un-assigned iduronic acid protons (circled in Figure 31, Panel C) was identified. This anomeric proton was assigned to 4.989ppm. Using the COSY spectrum (Panel B) H2 was then assigned to 3.743ppm, H3 to 4.128ppm, H4 to 4.079ppm and H5 to 4.767ppm, identifying H3 and H4 respectively as the protons involved in the trans-glycosidic NOE from residue “c”. For clarity only assignments originating from H1 on the horizontal axis are tracked (Panel B).

See Figures 28-31 for a key to the coloured arrows annotation scheme used in panels A and B. Cross-peaks forming the IdoUA ring system are sequentially linked by the arrows.

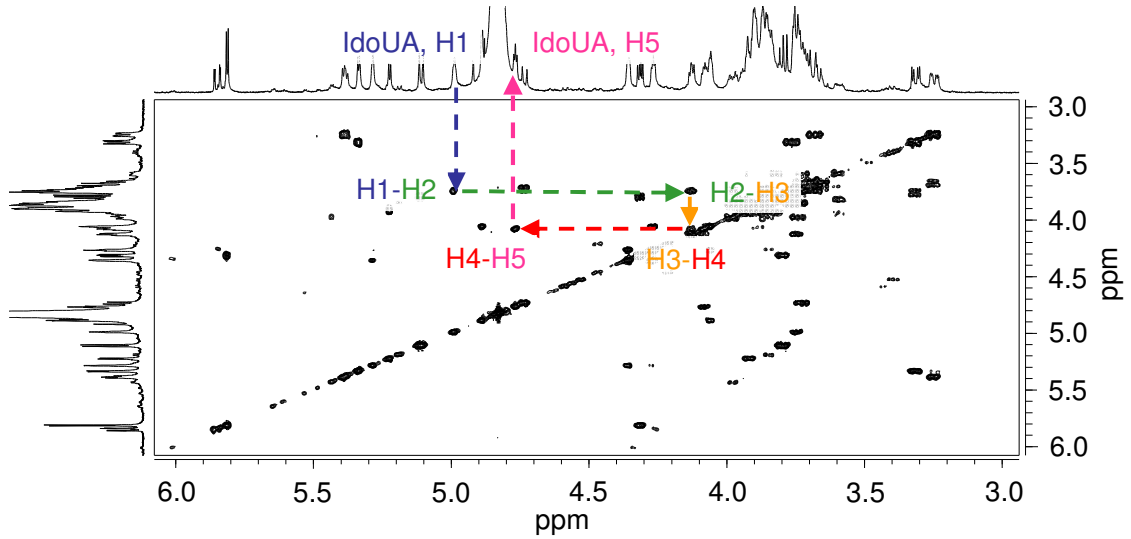
A single trans-glycosidic cross-peak was identified in the NOESY spectrum (Panel C) from H1 of IdoUA(2S) to an unassigned proton in the preceding glucosamine residue at 3.86ppm. This is circled.

Figure 32

Panel A



Panel B



Panel C

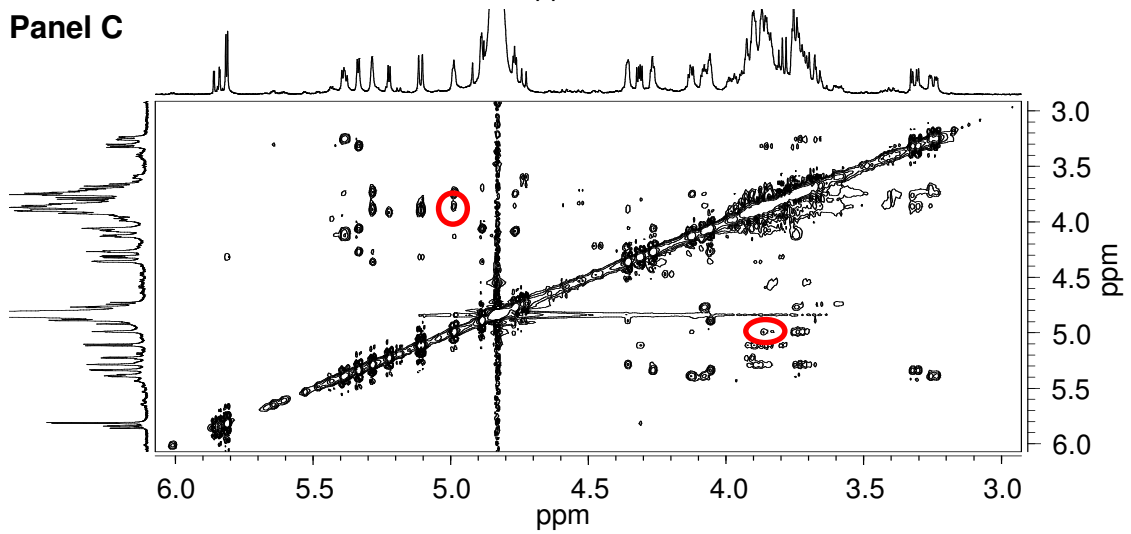


Figure 33 Assignment of the reducing terminal glucosamine, residue “f” of oligosaccharide dp6-2.

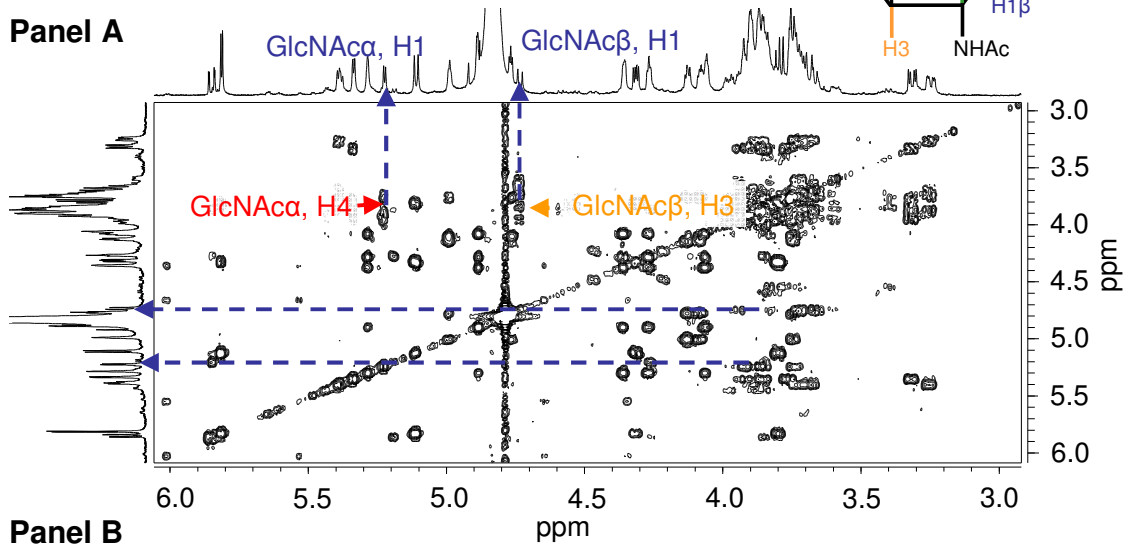
Using the TOCSY spectrum (Panel A) two ring systems were identified containing the 3.86ppm unassigned glucosamine proton (circled in Figure 32, Panel C). These ring systems correspond to the presence of the terminal GlcNAc residue in both the α and β anomeric configurations. The anomeric proton for each configuration was assigned to 5.224ppm (α) and 4.734ppm (β). Using the COSY spectrum (Panel B) H2 was then assigned to 3.905ppm (α), 3.725 (β) and H3 to 3.837ppm (α), 3.86ppm (β). Assignment of H2 to these values confirmed the glucosamine to be N-acetylated. H4 were later assigned to 3.86ppm (α), 3.931 (β) as were H5 to 3.979ppm (α) and 3.561ppm (β) and H6s to 3.725-3.754ppm (α) and 3.808-3.837 (β) (by HSQC analysis, see discussion). This identified H3 (β) and H4 (α) as involved in the trans-glycosidic NOE with residue “e”.

See Figures 28-31 for a key to the coloured arrows annotation scheme used in panels A and B. Cross-peaks for protons H1 to H3 forming part of both ring systems respectively are sequentially linked by the arrows.

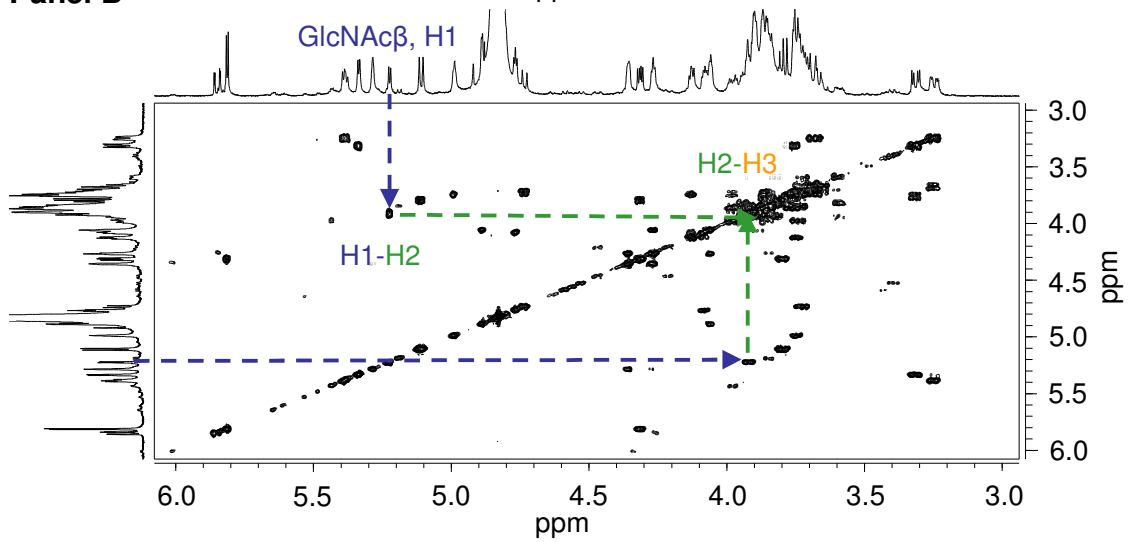
Figure 33



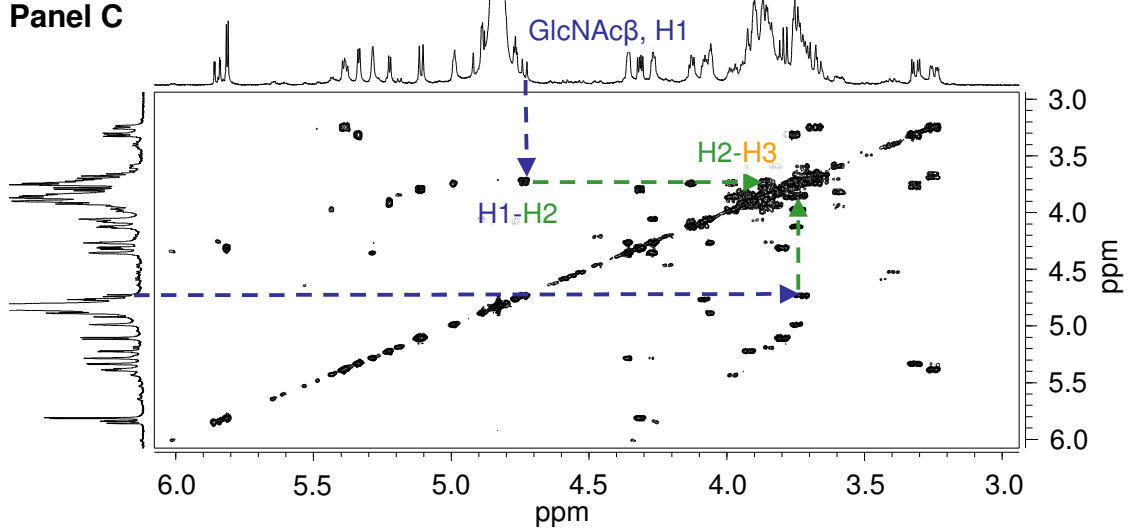
Panel A



Panel B



Panel C



These 2-D-NMR analyses established the primary sequence of dp6-2 to be

UA-GlcNS-IdoUA(2S)-GlcNS-IdoUA-GlcNAc(α or β)

a. b. c. d. e. f.

Letters signify residue designation and relative position within the oligosaccharide to the non-reducing end (lettered 'a').

11. Proton assignment of dp6-3.

A detailed proton assignment of the first 3 residues of dp6-3 established these residues to be identical in nature and sequence to those of dp6-2. Chemical shift assignments are summarised at the end of this section in Table 3.

Variation between the two oligosaccharides occurred in the nature of the uronic acid residue at position "e" and in the chemical shift values of residue "d". A detailed breakdown of the proton assignment of dp6-3, residues d-f is given in Figures 34-37.

Figure 34

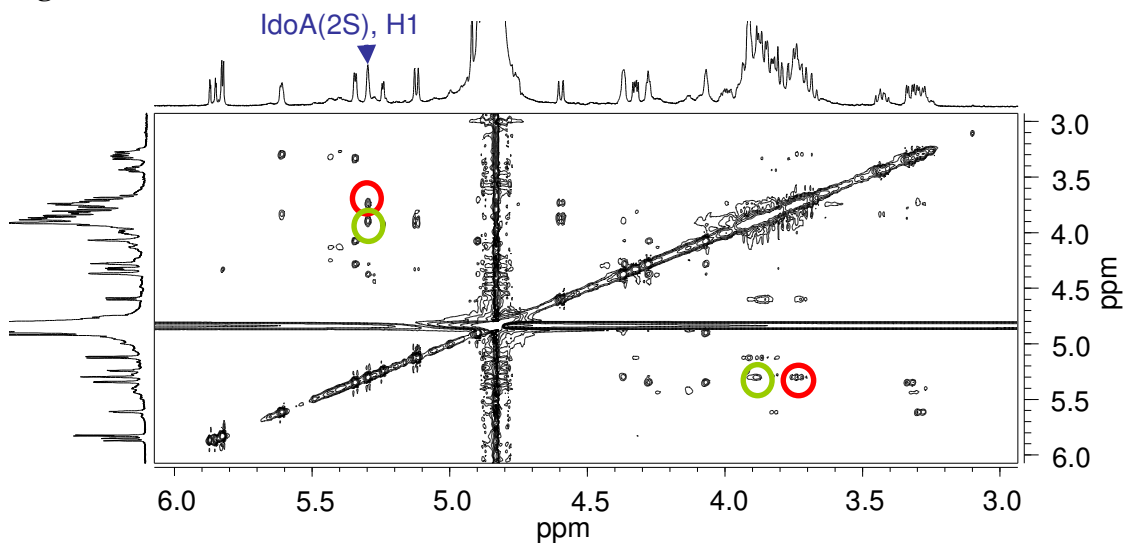


Figure 34 Identification of trans-glycosidic NOEs between IdoA(2S), residue “c” of dp6-3 and the preceding glucosamine residue “d”.

Two trans-glycosidic cross-peaks were identified in this NOESY spectrum, from H1 of IdoA(2S) to two unassigned protons in the preceding glucosamine residue at 3.741ppm and 3.891ppm. These cross-peaks are circled.

Figure 35 Assignment of glucosamine, residue “d” of oligosaccharide dp6-2.

Using the TOCSY spectrum (Panel A) the anomeric proton H1 for the ring system containing the 3.741ppm and 3.891ppm unassigned glucosamine protons (circled in Figure 34) was identified. This anomeric proton was assigned to 5.613ppm (contrast this with an assignment of 5.391ppm for the equivalent proton of dp6-2). Using the COSY spectrum (Panel B) H2 was then assigned to 3.288ppm, H3 to 3.683ppm and H4 to 3.741ppm, identifying H4 as one of the two protons involved in trans-glycosidic NOEs from residue “c”. H5 was later assigned to 3.823ppm and both H6 protons to a range of between 3.891-3.966ppm (by HSQC analysis, see discussion). This assignment of H6 identified one of the H6 protons as forming a trans-glycosidic NOE with residue “c”. Assignment of H2 to 3.288ppm identified this glucosamine as being N-sulphated.

See Figures 28-31 for a key to the coloured arrows annotation scheme used in panels A and B. Cross-peaks for protons H1 to H4 forming part of the GlcNS ring system are sequentially linked by the arrows.

A single trans-glycosidic cross-peaks were identified in the NOESY spectrum (Panel C) from H1 of GlcNS to an unassigned protons in the proceeding GlcUA residue at 3.828ppm. This cross-peak is circled.

Figure 35

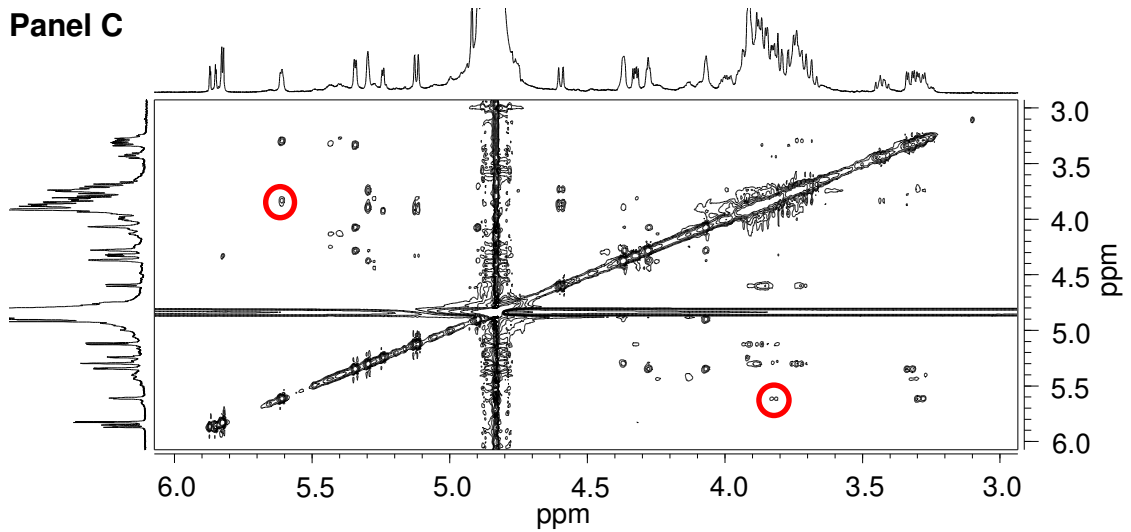
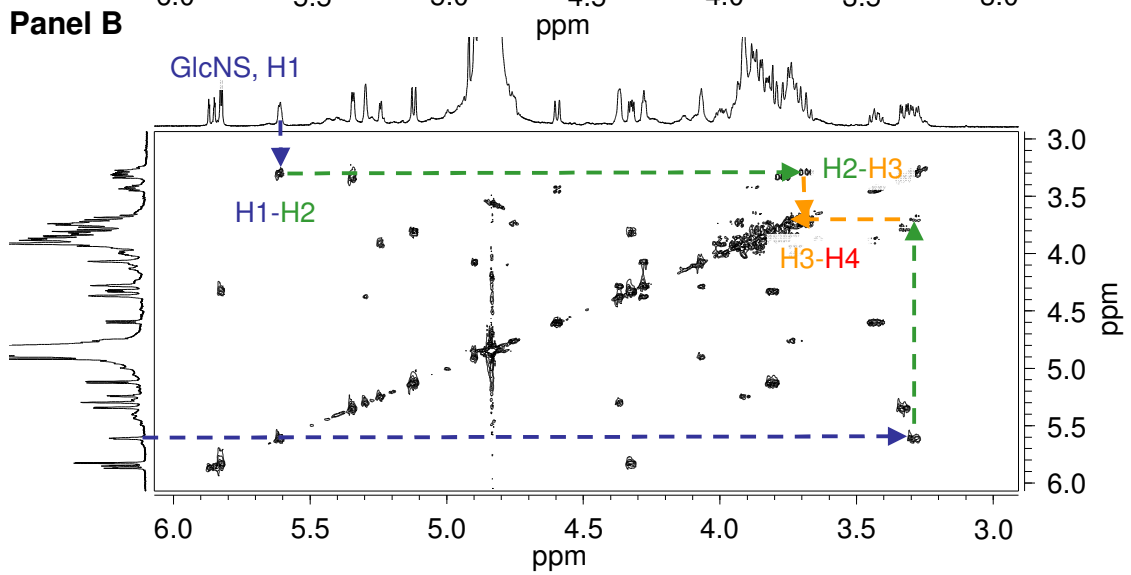
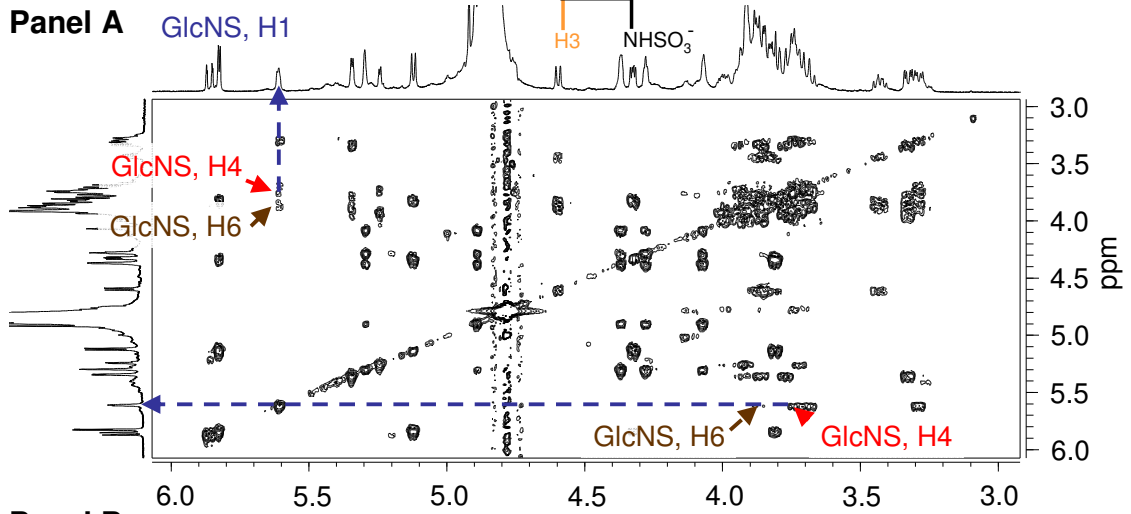
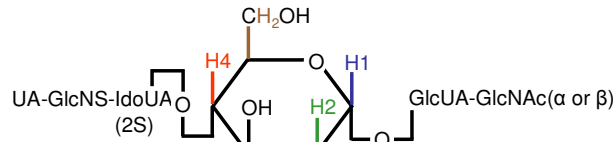


Figure 36 Assignment of glucuronic acid, residue “e” of oligosaccharide dp6-3.

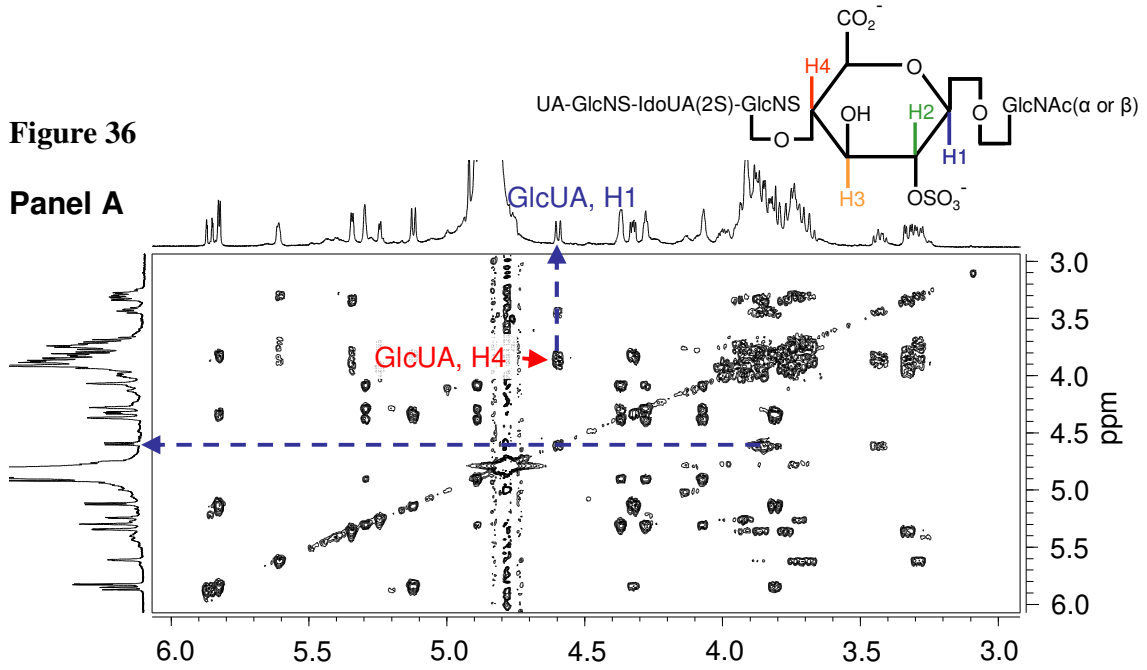
Using the TOCSY spectrum (Panel A) the anomeric proton H1 for the ring system containing the 3.828ppm unassigned uronic acid proton (circled in Figure 35, Panel C) was identified. This anomeric proton was assigned to 4.596ppm identifying this residue as GlcUA not IdoUA as in dp6-2. Using the COSY spectrum (Panel B) H2 was then assigned to 3.428ppm, H3 to 3.895ppm and H4 to 3.828ppm, identifying H4 as the proton involved in the trans-glycosidic NOE from residue “d”. H5 was subsequently assigned to 3.813ppm (by HSQC analysis).

See Figures 28-31 for a key to the coloured arrows annotation scheme used in panels A and B. Cross-peaks for protons H1 to H4 forming part of the GlcUA ring system are sequentially linked by the arrows.

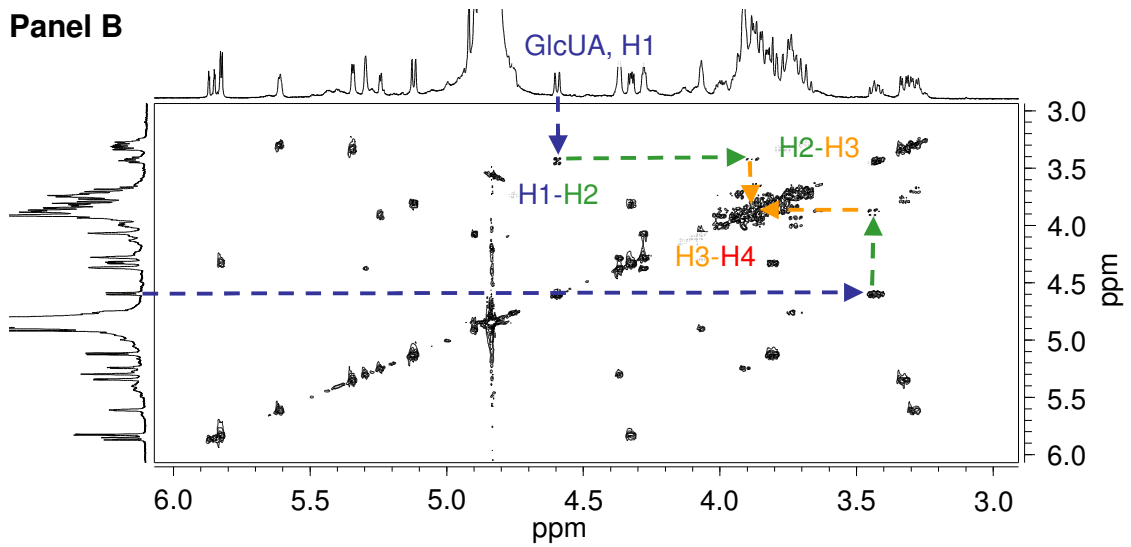
A single trans-glycosidic cross-peak was identified in the NOESY spectrum (Panel C) from H1 of GlcUA to an unassigned proton in the preceding glucosamine residue at 3.726ppm. This is circled.

Figure 36

Panel A



Panel B



Panel C

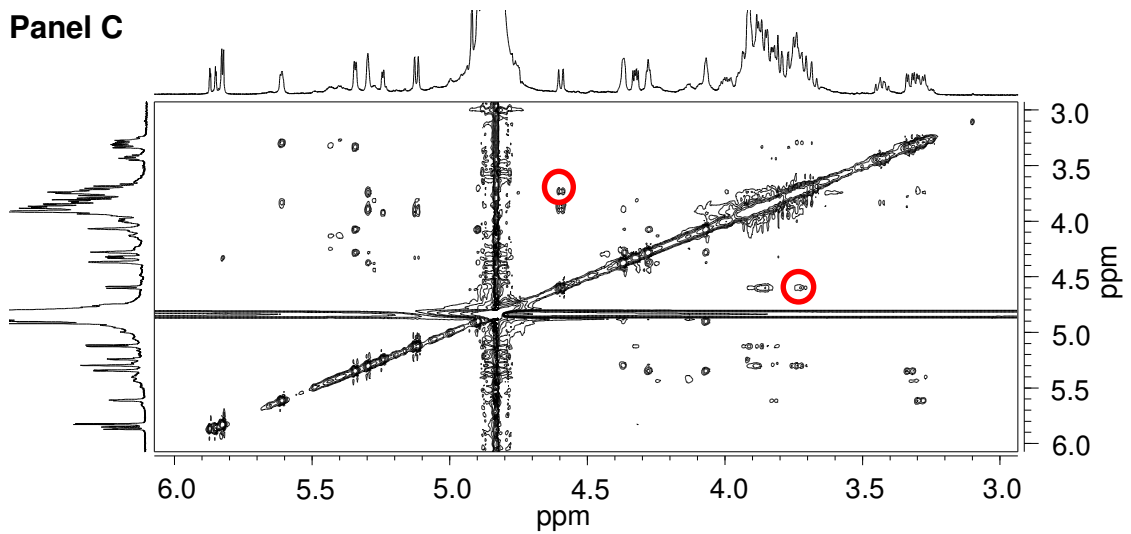


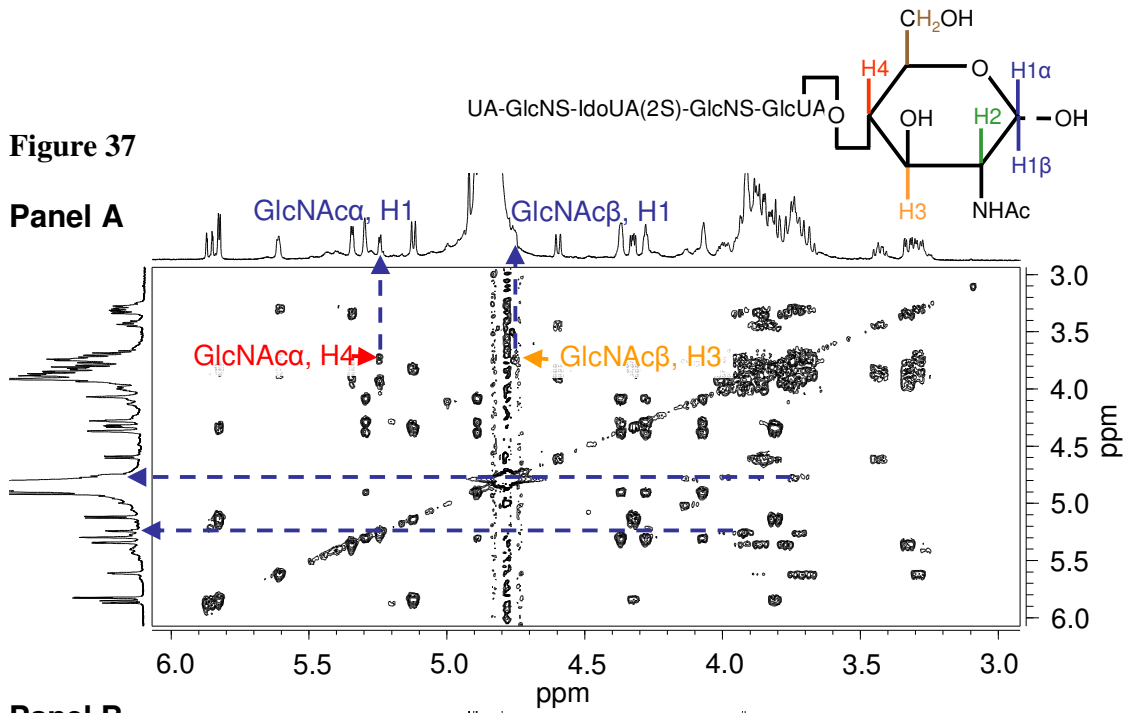
Figure 37 Assignment of the reducing terminal glucosamine, residue “f” of oligosaccharide dp6-3.

Using the TOCSY spectrum (Panel A) two ring systems were identified containing the 3.726ppm unassigned glucosamine proton (circled in Figure 36, Panel C). These ring systems, as with oligosaccharide dp6-2, correspond to the presence of the terminal GlcNAc residue in both the α and β anomeric configurations. The anomeric proton for each configuration was assigned to 5.242ppm (α) and 4.757ppm (β). Using the COSY spectrum (Panel B) H2 was then assigned to 3.905ppm (α), 3.735 (β) and H3 to 3.934ppm (α), 3.721ppm (β). Assignment of H2 to these values confirmed the glucosamine to be N-acetylated. H4 were later assigned to 3.726ppm (α), 4.0 (β) as were H5 to 3.996ppm (α) and 3.644ppm (β) and H6s to 3.842-3.939ppm (α) and 3.847-3.89 (β) (by HSQC analysis, see discussion). This identified H3 (β) and H4 (α) as involved in the trans-glycosidic NOE with residue “e”.

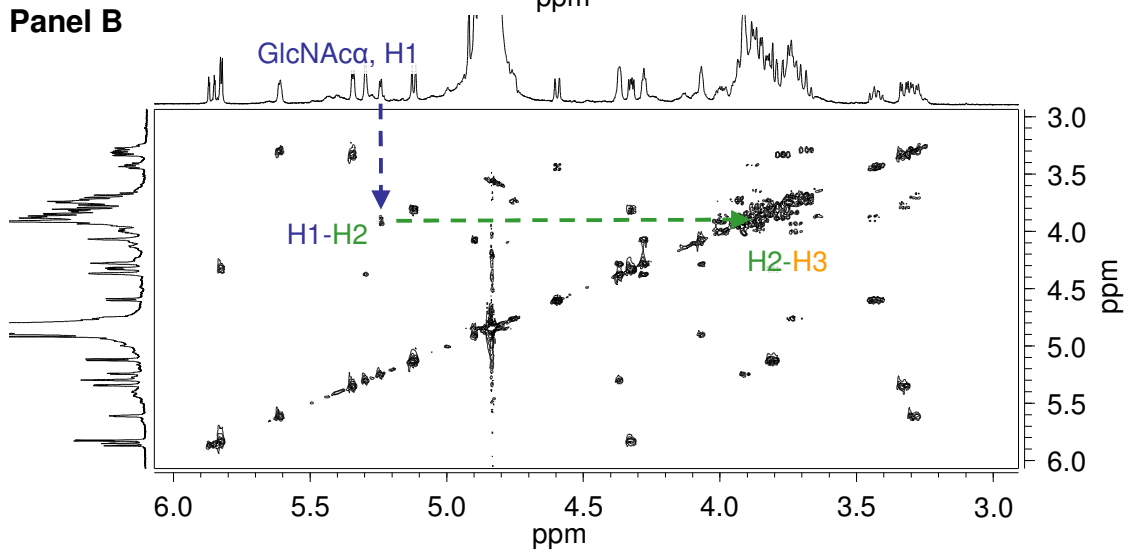
See Figures 28-31 for a key to the coloured arrows annotation scheme used in panels A and B. Cross-peaks for protons H1 to H3 forming part of both ring systems respectively are sequentially linked by the arrows.

Figure 37

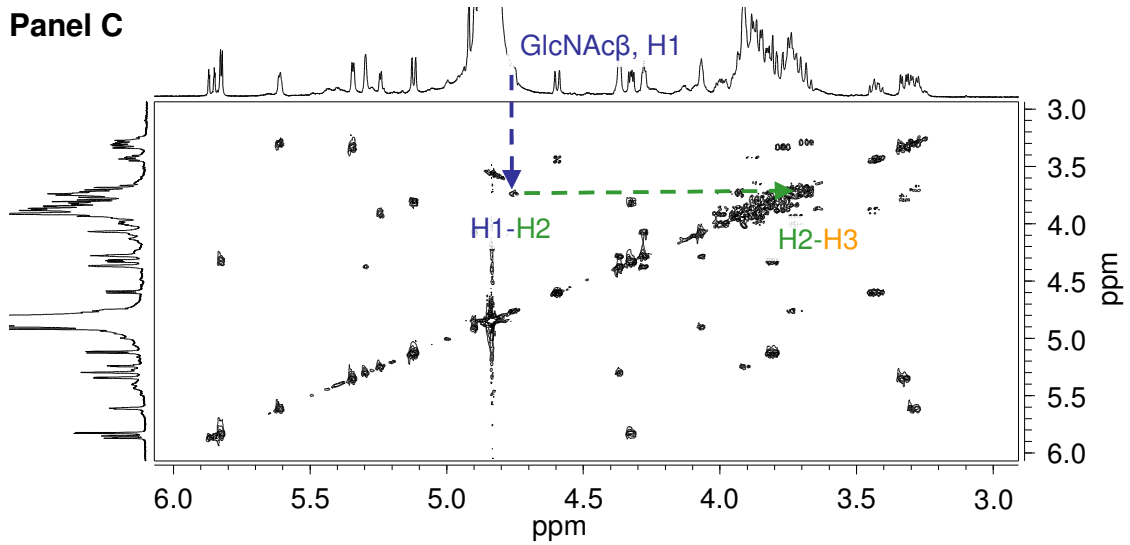
Panel A



Panel B



Panel C



These 2-D-NMR analyses established the primary sequence of dp6-3 to be

UA-GlcNS-IdoUA(2S)-GlcNS-GlcUA-GlcNAc(α or β)

a. b. c. d. e. f.

Letters again signifying residue designation and relative position within the oligosaccharide to the non-reducing end (lettered 'a').

For both dp6-2 and dp6-3, the NMR monosaccharide sequencing results are in perfect agreement with the disaccharide analysis results for each oligosaccharide (see Table 2).

12. Summary of chemical shift assignments for dp6-2 and dp6-3.

The chemical shift assignments for dp6-2 and dp6-3 are summarised in Figure 38 and in the preceding Table (Table 3). The coupling constants for each proton were measured as described under experimental procedures and are shown in brackets along with chemical shift values in Table 3.

Figure 38 A summary of the ^1H chemical shift assignments for dp6-2 and dp6-3.

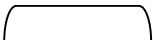
For assignment of the non-anomeric chemical shifts grouped  see Table 3.

Figure 38

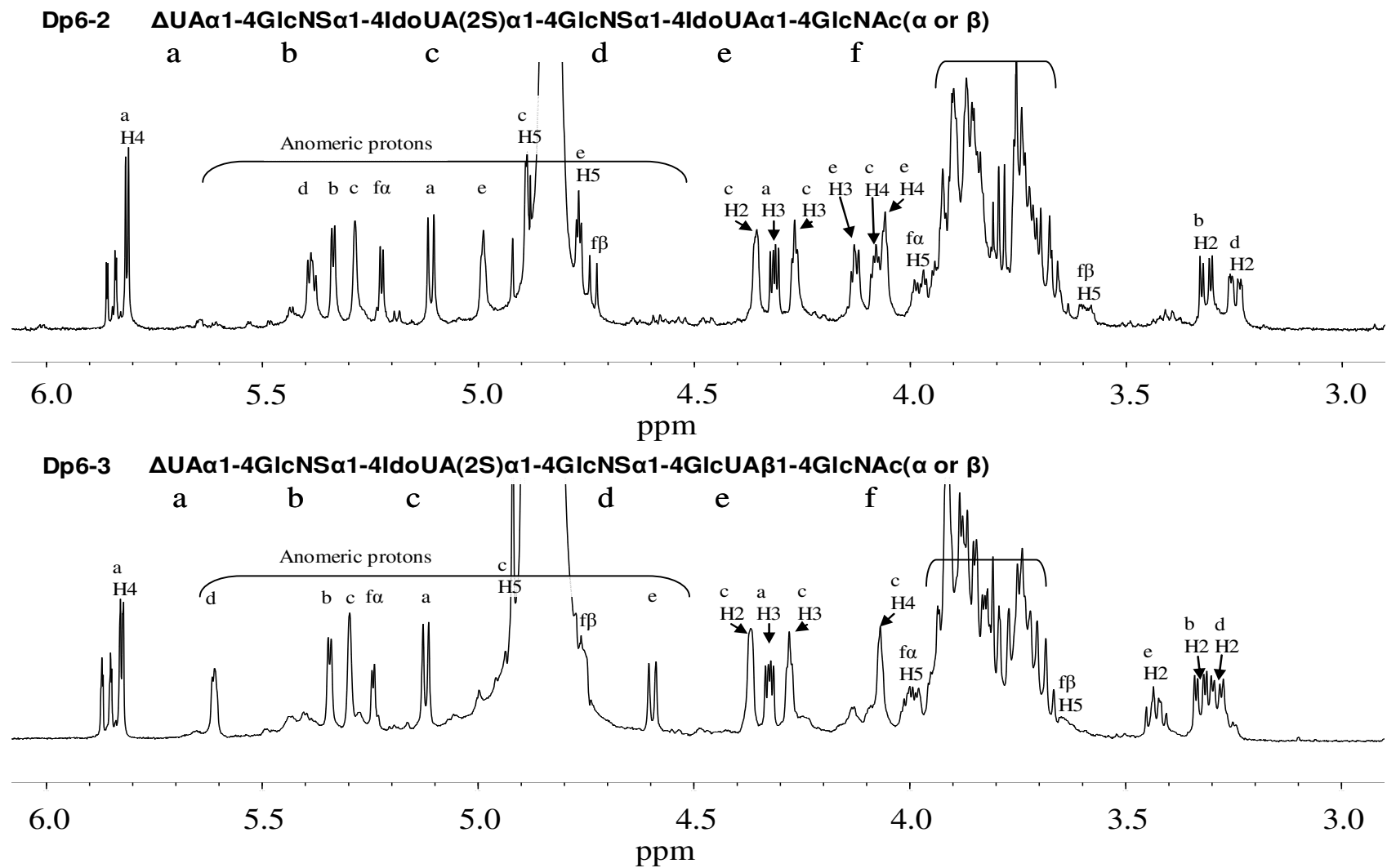


Table 3 ^1H and ^{13}C chemical shifts assignments (ppm) for hexasaccharides 6-2 and 6-3.

Residue	Oligosaccharide				Oligosaccharide			
	6-2				6-3			
1. ΔUA	H1	5.11 (6.9)	C1	100.947	H1	5.122 (6.6)	C1	101.108
	H2	3.801 (6.3)	C2	70.694	H2	3.813 (6.0)	C2	70.693
	H3	4.314 (3.1)	C3	67.296	H3	4.325 (3.4)	C3	67.296
	H4	5.814	C4	107.903	H4	5.825	C4	107.741
			C5	n.d.			C5	n.d.
			C6	n.d.			C6	n.d.
2. GlcNS	H1	5.336 (3.4)	C1	97.226	H1	5.343 (3.5)	C1	97.063
	H2	3.314 (10.7)	C2	57.913	H2	3.327 (10.7)	C2	57.913
	H3	3.761	C3	n.d.	H3	3.77	C3	n.d.
	H4	3.866	C4	n.d.	H4	3.885	C4	n.d.
	H5	3.837	C5	n.d.	H5	3.85	C5	n.d.
	H6	3.837-3.925	C6	n.d.	H6	3.797-3.966	C6	n.d.
	H6'	3.837-3.925			H6'	3.797-3.966		
3. IdoUA(2S)	H1	5.285 (1.6)	C1	99.167	H1	5.297 (1.6)	C1	99.167
	H2	4.356 (4.1)	C2	74.415	H2	4.368 (4.1)	C2	74.576
	H3	4.266 (3.15)	C3	67.62	H3	4.279 (3.15)	C3	67.458
	H4	4.059 (2.2)	C4	75.547	H4	4.068 (1.5)	C4	75.547
	H5	4.885	C5	68.105	H5	4.896	C5	68.267
			C6	n.d.			C6	n.d.
4. GlcNS	H1	5.391 (3.4)	C1	95.446	H1	5.613 (3.7)	C1	97.549
	H2	3.247 (10.1)	C2	58.075	H2	3.288 (10.1)	C2	58.236
	H3	3.743	C3	n.d.	H3	3.683	C3	n.d.
	H4	3.72	C4	n.d.	H4	3.741	C4	n.d.
	H5	3.854	C5	n.d.	H5	3.823	C5	n.d.
	H6	3.889-3.905	C6	n.d.	H6	3.891-3.966	C6	n.d.
	H6'	3.889-3.905			H6'	3.891-3.966		
5. HexUA	H1	4.989 (2.5)	C1	101.755	H1	4.596 (7.8)	C1	102.24
	H2	3.743 (5.0)	C2	69.238	H2	3.428	C2	72.797
	H3	4.128 (3.3)	C3	68.429	H3	3.895	C3	n.d.
	H4	4.079 (2.8)	C4	74.9	H4	3.828	C4	n.d.
	H5	4.767	C5	69.238	H5	3.813	C5	n.d.
			C6	n.d.			C6	n.d.
6. GlcNAc (α)	H1	5.224 (3.4)	C1	90.593	H1	5.242 (3.1)	C1	90.431
	H2	3.905	C2	n.d.	H2	3.905	C2	n.d.
	H3	3.837	C3	n.d.	H3	3.934	C3	n.d.
	H4	3.86	C4	n.d.	H4	3.726	C4	n.d.
	H5	3.979	C5	n.d.	H5	3.996	C5	n.d.
	H6	3.725-3.754	C6	n.d.	H6	3.842-3.939	C6	n.d.
	H6'	3.725-3.754			H6'	3.842-3.939		
	NAc	2.064	NAc	21.998	NAc	2.075	NAc	21.998
6. GlcNAc (β)	H1	4.734 (8.1)	C1	94.961	H1	4.757	C1	94.799
	H2	3.725	C2	n.d.	H2	3.735	C2	n.d.
	H3	3.86	C3	n.d.	H3	3.721	C3	n.d.
	H4	3.931	C4	n.d.	H4	4	C4	n.d.
	H5	3.591	C5	n.d.	H5	3.644	C5	n.d.
	H6	3.808-3.837	C6	n.d.	H6	3.847-3.89	C6	n.d.
	H6'	3.808-3.837			H6'	3.847-3.89		
	NAc	2.064	NAc	21.998	NAc	2.075	NAc	21.998

Coupling constants (in Hz) to proton directly below are given in parentheses
n. d. Not determined because of overlapping signals

13. Preliminary NMR analysis of oligosaccharide HS 6-5.

Oligosaccharide 6-5 was analysed by ^1H 1-D NMR spectroscopy. This result is shown in Figure 39, Panel A.

Unexpectedly, a large number of aliphatic protons were detected (chemical shift values ranging between 1.0 and 2.5ppm) in the original sample (see Figure 39, panel A). With the exception of the chemical shift detected at 2.08ppm, which could be assigned to the methyl protons of the two N-acetyl groups present (see Table 2). These aliphatic chemical shifts were assumed to have arisen from contaminating polyacrylamide, which had carried through into the NMR analysis. This was removed by the use of an additional DEAE gel chromatography purification step (see experimental procedures). The results of the further purification, on the 1-D NMR spectrum of HS dp6-5 are shown in Figure 39, panels B and C. Similar results were obtained with oligosaccharide 6-6 (data not shown).

Figure 39

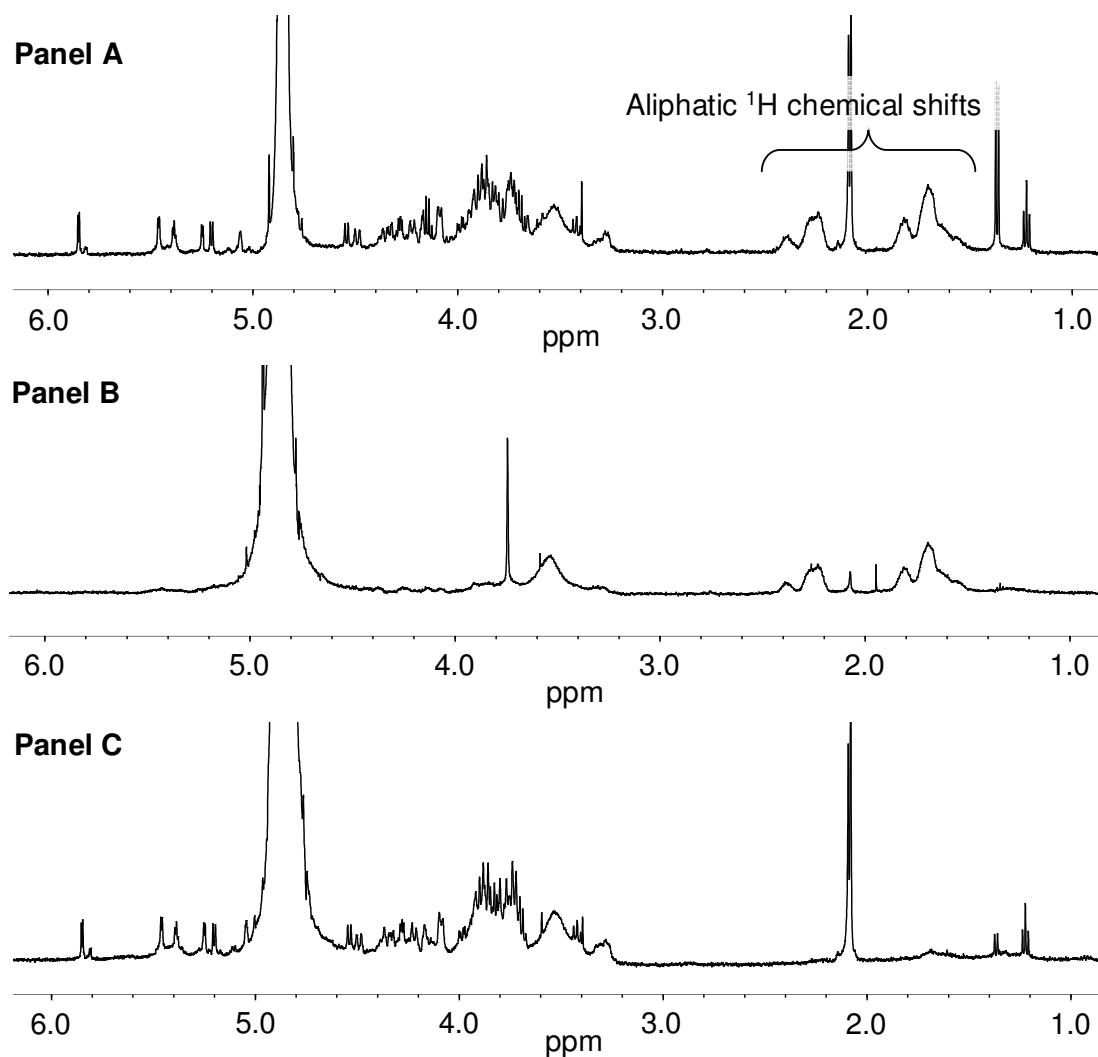


Figure 39 One-dimensional 500 MHz ¹H-NMR spectra of hexasaccharide 6-5.

Spectra were recorded in 0.15M NaCl, 20mM Na₂HPO₄, ²H₂O at pH 7.2 and 25°C. Panel A, dp6-5 prior to DEAE chromatography; Panel B, material that did not bind to the DEAE column; Panel C, material bound to the DEAE column and eluted using 1M NaCl. See experimental procedures for exact experimental detail of the DEAE chromatographic procedure.

14. Spectral assignment of dp6-5 and dp6-6.

Two dimensional COSY, TOCSY and NOESY NMR spectra were recorded for oligosaccharides dp6-5 and dp6-6. These are presented in appendix I

COSY spectra (dp6-2, Figure 88, page 337, dp6-3, Figure 89, page 338),

TOCSY spectra (dp6-2, Figure 90, page 339, dp6-3, Figure 91, page 340)

NOESY spectra (dp6-2, Figure 92, page 341, dp6-3, Figure 93, page 342)

The additional loss of material as a result of the PAGE and DEAE purification steps meant it was not possible to acquire ^{13}C spectra for these saccharides (see discussion).

Proton chemical shifts were assigned as has been outlined in detail for oligosaccharides dp6-2 and dp6-3 (see Figures 28-33 and 34-37) these results are summarised in Figure 40 and the proceeding Table (Table 4). These assignments established the primary sequence of dp6-5 to be:



a. b. c. d. e. f.

Letters signify residue designation and position within the monosaccharide sequence.

See discussion on the placement of GlcUA and IdoUA at residues “c” and “e” respectively within this sequence. The primary sequence of dp6-6 was established to be:



a. b. c . d. e. f.

Due to the low amount of material present for dp6-6 and the close proximity of the anomeric protons of residues “a” and “e” to the water peak, an NOE could not be directly established across the glycosidic linkages “a-b” and “e-f”. However the identity of these residues could be clearly ascertained and their position in the sequence then deduced.

Figure 40 A summary of the ^1H chemical shift assignments for dp6-5 and dp6-6.

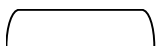
A low level of cross contamination of dp6-5 with dp6-6 and *vice-versa* was detected. This is most clearly seen around the chemical shift for residue “a” H4 when comparing both spectra. Contaminating peaks are not labelled. For assignment of the non-anomeric chemical shifts grouped  see Table 4. With both saccharides only the α anomeric configuration of residue “f” could be detected.

Figure 40

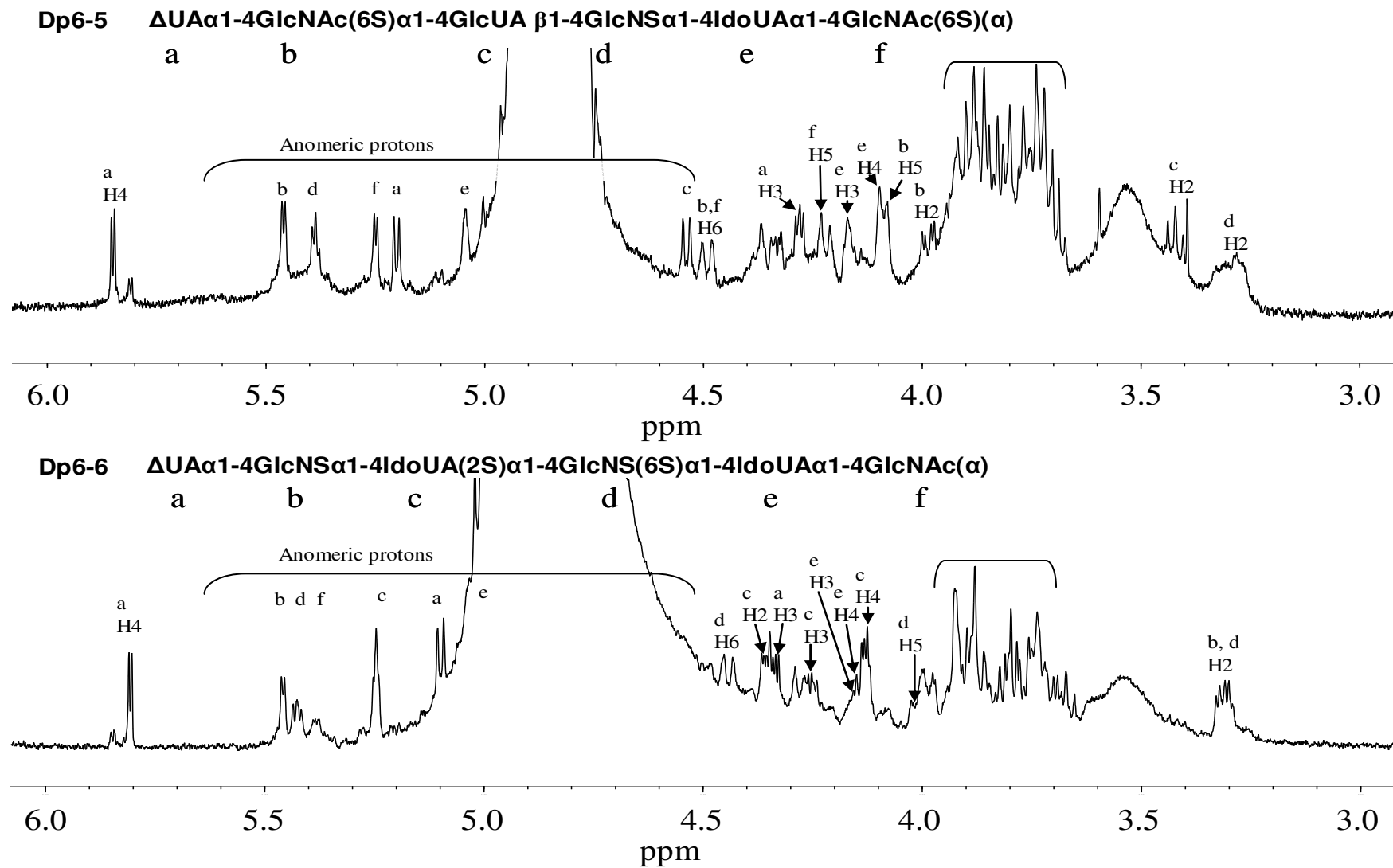


Table 4 ^1H chemical shifts assignments (ppm) for hexasaccharides 6-5 and 6-6

Oligosaccharide 6-5			Oligosaccharide 6-6		
Residue		Chemical Shift (ppm)	Residue		Chemical Shift (ppm)
1. ΔUA	H1	5.202 (6.0)	1. ΔUA	H1	5.088(6.9)
	H2	3.858(5.0)		H2	3.791(6.0)
	H3	4.279(3.4)		H3	4.336(3.5)
	H4	5.849		H4	5.806
2. GlcNAc(6S)	H1	5.459(3.8)	2. GlcNS	H1	5.448(3.5)
	H2	3.986(9.8)		H2	3.326(10.4)
	H3	3.872		H3	3.749
	H4	3.754		H4	n. d.
	H5	4.219		H5	n. d.
	H6	4.492(10.7)		H6	n. d.
H6'	n. d.	H6'	n. d.		
3. GlcA	H1	4.538(7.9)	3. IdoUA(2S)	H1	5.251(2.8)
	H2	3.42		H2	4.36(5.6)
	H3	3.722		H3	4.257(2.2)
	H4	3.819		H4	4.123
	H5	n. d.		H5	4.773
4. GlcNS	H1	5.388(3.8)	4. GlcNS(6S)	H1	5.417(3.5)
	H2	3.267		H2	3.295(10.4)
	H3	3.738		H3	3.67
	H4	3.894		H4	3.81
	H5	n. d.		H5	4.012
	H6	n. d.		H6	4.452(10.1)
H6'	n. d.	H6'	n. d.		
5. IdoUA	H1	5.044(2.5)	5. IdoUA	H1	5.009
	H2	3.814(5.3)		H2	3.728
	H3	4.17(4.1)		H3	4.157
	H4	4.097(3.4)		H4	4.152
	H5	4.785		H5	4.767
6. GlcNAc(6S) (α)	H1	5.248	6. GlcNAc (α)	H1	5.364(3.4)
	H2	3.961		H2	3.889
	H3	3.906		H3	n. d.
	H4	4.082		H4	n. d.
	H5	4.098		H5	n. d.
	H6	4.492		H6	n. d.
	H6'	n. d.		H6'	n. d.
NAc	2.08	NAc	2.08		

Where determined coupling constants (in Hz) to proton directly below are given in brackets
n. d. Not determined

15. Purification of two heparin derived hexasaccharides.

The dp6 size pool from a Bio-Gel P10 gel filtration profile of Tinzaparin sodium (see Figure 22) was applied to a SAX-HPLC column, and the oligosaccharides eluted using a linear salt gradient (see experimental procedures). A typical SAX-HPLC elution profile is shown in Figure 41. Consistent with a number of previous studies two dominant peaks were detected in this size pool (numbered 1 and 2 in Figure 41) (Larnkjaer *et al.* 1995; Pervin *et al.* 1995; Chuang *et al.* 2001). These peaks have been shown to be due to the presence of a single oligosaccharide species here after referred to as hep dp6-1 and hep dp6-2.

Figure 41

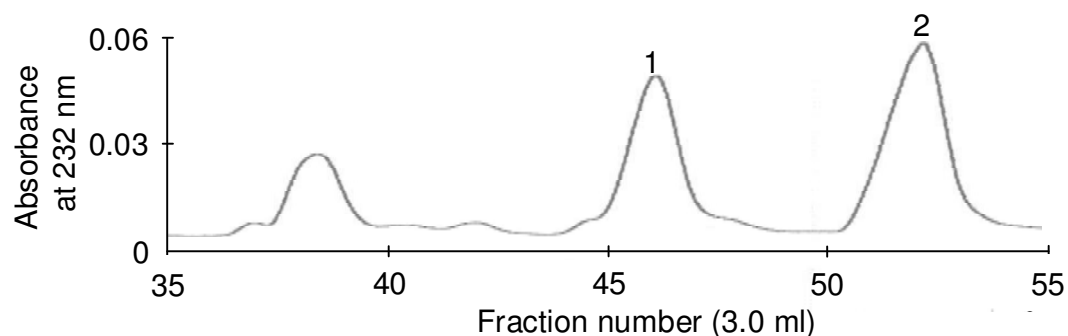


Figure 41 SAX-HPLC chromatography of heparin derived hexasaccharides.

The dp6 oligosaccharides of Tinzaparin sodium (generated by partial heparinase I digestion of heparin) were purified by Bio-Gel P10 chromatography (see Figure 22) then applied to a ProPac PA-1 SAX-HPLC column (250mm x 9cm) and eluted using a linear salt gradient (see experimental procedures for gradient used). Column eluate was monitored for absorbance at 232nm. Peaks numbered 1 and 2 were collected and taken on for further analysis by NMR spectroscopy.

16. NMR and spectral assignment of hep dp6-1 and dp6-2.

Two dimensional COSY, TOCSY and NOESY NMR spectra were recorded for oligosaccharides hep dp6-1 and hep dp6-2. These are presented in appendix I. ^{13}C DEPT-135 and ^{13}C HSQC spectra were also recorded for these oligosaccharides and are shown in appendix II (see discussion).

COSY spectra (hep dp6-1, Figure 94, page 343, hep dp6-2, Figure 95, page 344)

TOCSY spectra (hep dp6-1, Figure 96, page 345, hep dp6-2, Figure 97, page 346)

NOESY spectra (hep dp6-1, Figure 98, page 347, hep dp6-2, Figure 99, page 348)

DEPT-135/HSQC spectra (hep dp6-1, Figure 105, page 354, hep dp6-2, Figure 106, page 355)

These 2-D-NMR analyses enabled the primary sequence of hep dp6-1 to be confirmed as:

UA(2S)-GlcNS(6S)-IdoUA(2S)-GlcNS(6S)-GlcUA-GlcNS(6S)(α)

a. b. c. d. e. f.

As established in the literature (Larnkjaer *et al.* 1995; Pervin *et al.* 1995; Chuang *et al.* 2001) Letters again signifying residue designation and position within the monosaccharide sequence.

Hep dp6-2 was established to be:

UA(2S)-GlcNS(6S)-IdoUA(2S)-GlcNS(6S)-IdoUA(2S)- GlcNS(6S)(α)

a. b. c. d. e. f.

17. Summary of chemical shift assignments for hep dp6-1 and dp6-2.

The chemical shift assignments for hep dp6-1 and dp6-2 are summarised in Figure 42 and in the preceding Table (Table 5). The coupling constants for each proton were measured, as described under experimental procedures, and are shown in brackets along with chemical shift values in Table 5.

Figure 42 A summary of the ^1H chemical shift assignments for hep dp6-1 and dp6-2.

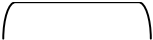
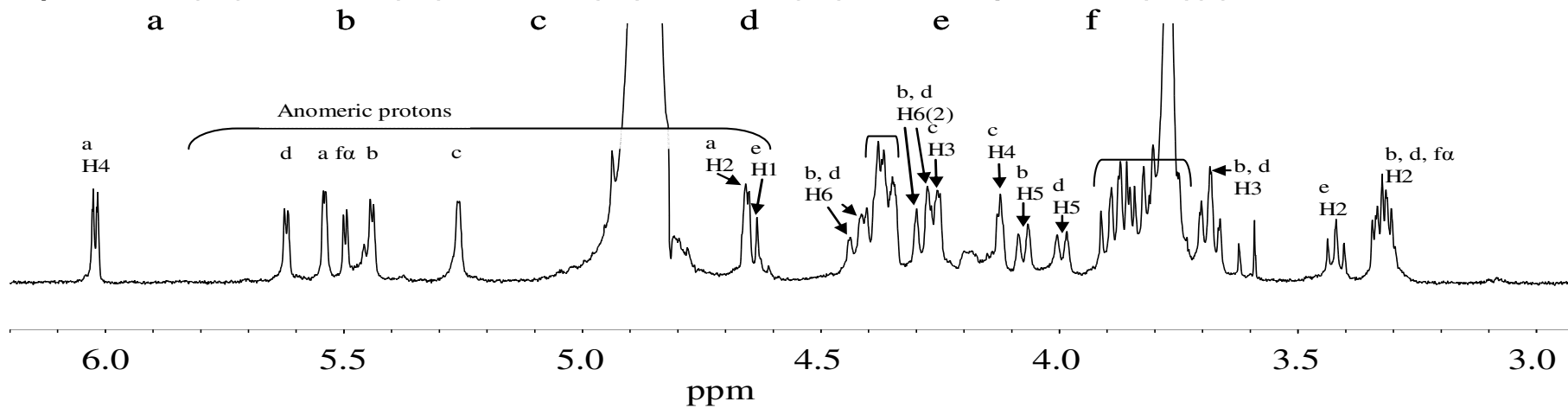
For assignment of the non-anomeric chemical shifts grouped  see Table 5. With both saccharides only the α anomeric configuration of residue "f" could be detected.

Figure 42

Hep 6-1 Δ UA(2S) α 1-4GlcNS(6S) α 1-4IdoUA(2S) α 1-4GlcNS(6S) α 1-4GlcUA β 1-4GlcNS(6S)(α)



Hep 6-2 Δ UA(2S) α 1-4GlcNS(6S) α 1-4IdoUA(2S) α 1-4GlcNS(6S) α 1-4IdoUA(2S) α 1-4GlcNS(6S)(α)

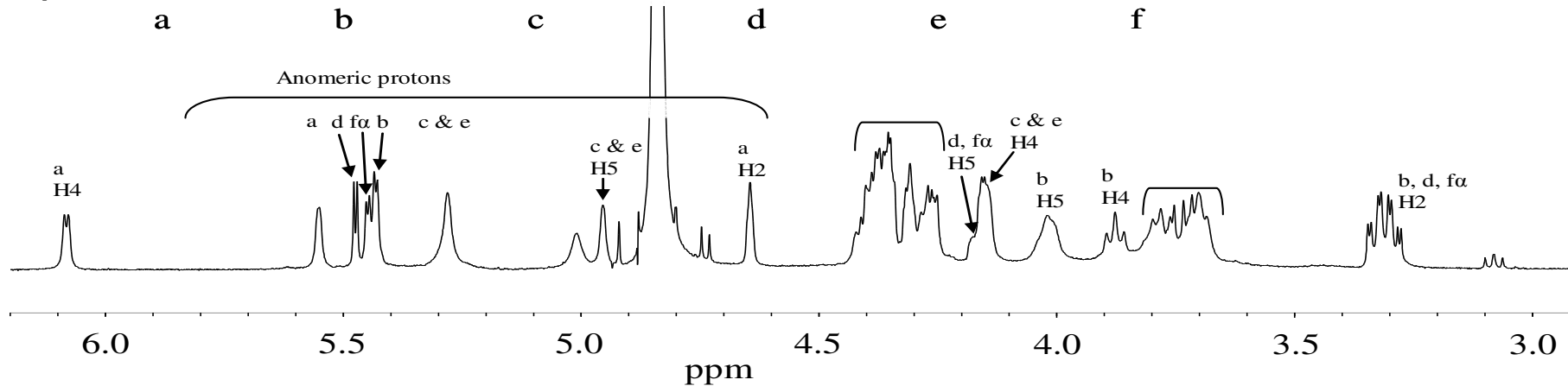


Table 5 ^1H chemical shifts assignments (ppm) for hep 6-1 and 6-2

Residue	Oligosaccharide				Oligosaccharide			
	Hep 6-1				Hep 6-2			
1. ΔUA(2S)	H1	5.542(2.2)	C1	97.387	H1	5.555(1.9)	C1	96.902
	H2	4.652	C2	74.415	H2	4.649	C2	74.414
	H3	4.341(5.4)	C3	62.928	H3	4.357(6.1)	C3	62.766
	H4	6.02	C4	105.962	H4	6.081	C4	106.77
			C5	n. d.			C5	n. d.
			C6	n. d.			C6	n. d.
2. GlcNS(6S)	H1	5.433(3.4)	C1	97.064	H1	5.429(3.7)	C1	97.282
	H2	3.332	C2	57.428-58.237	H2	3.336(11.6)	C2	58.075
	H3	3.682	C3	69.723	H3	3.698	C3	69.237
	H4	3.864	C4	n. d.	H4	3.882	C4	77.812
	H5	4.07	C5	68.914	H5	4.033	C5	69.076
	H6	4.291	C6	n. d.	H6	n. d.	C6	66.325
	H6'	or 4.401			H6'	n. d.		
3. IdoUA(2S)	H1	5.254(2.5)	C1	99.167	H1	5.28	C1	99.167
	H2	4.367(5.6)	C2	75.547	H2	4.357	C2	75.062
	H3	4.245(4.1)	C3	68.914	H3	4.307	C3	68.267
	H4	4.123(3.2)	C4	76.194	H4	4.139	C4	76.356
	H5	4.847	C5	68.914	H5	4.951	C5	68.914
			C6	n. d.			C6	n. d.
4. GlcNS(6S)	H1	5.613(3.8)	C1	97.387	H1	5.45(3.7)	C1	96.74
	H2	3.315	C2	57.428-58.237	H2	3.313(11.6)	C2	57.913
	H3	3.686	C3	69.561	H3	3.726	C3	69.237
	H4	3.798	C4	n. d.	H4	3.792	C4	n. d.
	H5	3.984	C5	69.076	H5	4.341	C5	n. d.
	H6	4.262	C6	n. d.	H6	4.347-4.419	C6	n. d.
	H6'	or 4.424			H6'	4.347-4.419		n. d.
5. IdoUA(2S) or GlcUA	H1	4.632(9.8)	C1	102.079	H1	5.28	C1	99.167
	H2	3.322(8.2)	C2	72.797	H2	4.357	C2	75.547
	H3	3.882	C3	n. d.	H3	4.265	C3	68.59
	H4	3.828	C4	n. d.	H4	4.156	C4	75.87
	H5	n. d.	C5	n. d.	H5	4.951	C5	68.914
			C6	n. d.			C6	n. d.
6. GlcNS(6S) (α)	H1	5.483(3.7)	C1	91.078	H1	5.472(3.1)	C1	90.916
	H2	3.309	C2	n. d.	H2	3.291(11.6)	C2	57.589
	H3	3.759	C3	69.4	H3	3.754	C3	77.03
	H4	4.192	C4	n. d.	H4	4.161	C4	n. d.
	H5	4.378	C5	n. d.	H5	4.269	C5	n. d.
	H6	4.248-4.305	C6	n. d.	H6	4.318-4.408	C6	n. d.
	H6'	4.248-4.305		n. d.	H6'	4.318-4.408		n. d.

Coupling constants (in Hz) to proton directly below are given in brackets
n. d. Not determined

18. De-2-O-sulphation.

After establishing the primary sequence of HS dp6-2, dp6-3 and hep dp6-1 and hep dp6-2., these oligosaccharides were subjected to chemical de-2-O-sulphation. As mentioned in the aims section of the introduction, this was done in order provide a greater insight into the role 2-O-sulphate groups play in influencing monosaccharide conformational equilibria.

The efficiency and selectivity of this procedure was evaluated using heparin oligosaccharide dp6-2. After initially undergoing one round of de-sulphation (see experimental procedures) samples were then analysed by 1-D NMR spectroscopy before being re-subjected to additional rounds of de-sulphation. The 1-D NMR spectrum recorded after each round of de-sulphation is shown in Figure 43. After the second round of de-sulphation a 2-D COSY spectrum was recorded (see Appendix I, Figure 102, page 351) and the chemical shifts assigned (see Table 6).

Figure 43 NMR analysis of heparin dp6-2 after successive rounds of chemical de-2-O-sulphation.

Panel A, heparin dp6-2 prior to the de-sulphation procedure; Panel B, after one round of de-sulphation; Panel C, after two rounds of de-sulphation; Panel D, after three rounds of de-sulphation; Panel E, after four rounds of de-sulphation.

The peak area of the resonance at 6.02 is given as a percentage of its area plus that of the resonance at 5.83 ppm in each panel (see text).

The most significant changes in the chemical shifts after de-2-O-sulphation involved protons within the UA and IdoUA residues (see Table 6) and these are marked with arrows.

Figure 43

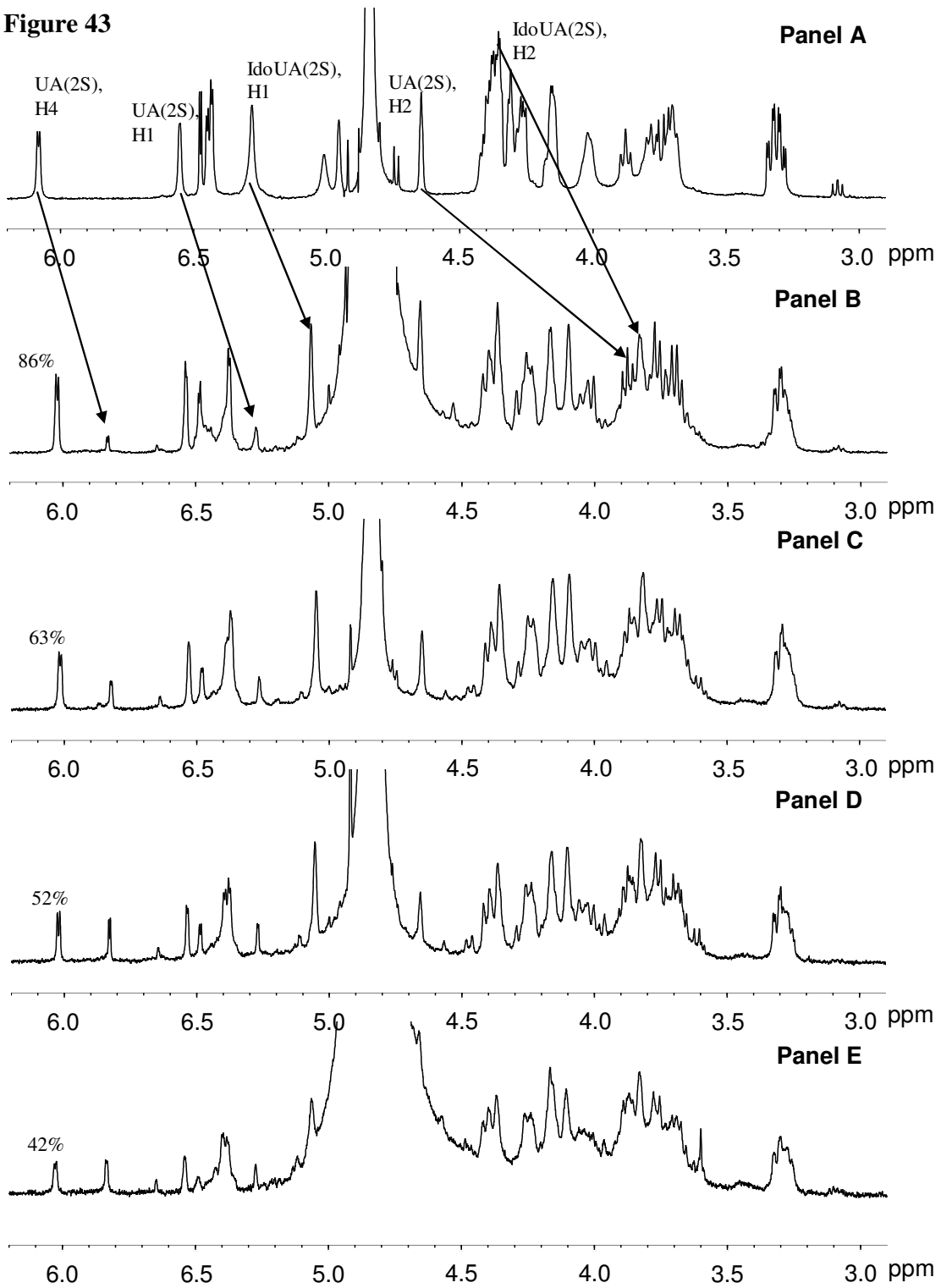


Table 6. The ^1H chemical shifts assignments (ppm) for heparin dp6-2 prior to and after undergoing de-2-O-sulphation.

The difference column refers to the difference between the protons original chemical shift value and the chemical shift value after de-2-O-sulphation.

Residue		Original (ppm)		After de-2-O-sulphation (ppm)	Difference (ppm)
1. ΔUA(2S)	H1	5.555(1.9)	H1	5.26	-0.295
	H2	4.649	H2	3.853	-0.796
	H3	4.357(6.1)	H3	4.407	+0.05
	H4	6.081	H4	5.83(3.4)	-0.251
2. GlcNS(6S)	H1	5.429(3.7)	H1	5.373(3.4)	-0.056
	H2	3.336(11.6)	H2	3.322	-0.014
	H3	3.698	H3	3.701	+0.003
	H4	3.882	H4	3.882	0
	H5	4.033	H5	4.012	-0.021
	H6	n. d.	H6	n. d.	n. d.
H6'	n. d.	H6'	n. d.	n. d.	
3. IdoUA(2S) or IdoUA	H1	5.28	H1	5.046	-0.234
	H2	4.357	H2	3.803	-0.554
	H3	4.307	H3	4.17	-0.137
	H4	4.139	H4	4.096	-0.043
	H5	4.951	H5	4.848	-0.103
4. GlcNS(6S)	H1	5.45(3.7)	H1	5.385(3.7)	-0.065
	H2	3.313(11.6)	H2	3.26	-0.053
	H3	3.726	H3	3.69	-0.036
	H4	3.792	H4	3.746	-0.046
	H5	4.341	H5	4.181	-0.16
	H6	4.347-4.419	H6	4.26-4.36	-0.087--0.059
H6'	4.347-4.419	H6'	4.26-4.36		
5. IdoUA(2S) or IdoUA	H1	5.28	H1	5.046	-0.234
	H2	4.357	H2	3.803	-0.554
	H3	4.265	H3	4.17	-0.137
	H4	4.156	H4	4.096	-0.043
	H5	4.951	H5	4.848	-0.103
6. GlcNS(6S) (α)	H1	5.472(3.1)	H1	5.481(3.7)	+0.009
	H2	3.291(11.6)	H2	3.3	+0.009
	H3	3.754	H3	3.752	-0.002
	H4	4.161	H4	4.187	+0.026
	H5	4.269	H5	4.243	-0.026
	H6	4.318-4.408	H6	4.351-4.492	+0.033-+0.084
H6'	4.318-4.408	H6'	4.351-4.492	+0.033-+0.084	

Coupling constants (in Hz) to proton directly below are given in brackets
n. d. Not determined

The 2-O-sulphate group of UA(2S) (residue “a”) was found to be resistant to removal using this de-sulphation procedure; removal of the 2-O-sulphate group resulting in a shift in the H4 resonance from 6.02ppm to 5.83ppm, among other resonance shifts (see Figure 43). After each round of de-sulphation a resonance at 6.02ppm could be clearly detected (see Figure 43, Panels B-E). However, each successive round of de-sulphation resulted in an increase in the strength of the signal at 5.83ppm relative to the strength of the signal at 6.02ppm (labelled in Figure 43). This corresponds to an increase in the number of H4 protons from a non-2-O-sulphated UA residue and a fall in the number of H4 protons for a sulphated IdoUA. This infers partial rather than total resistance of UA(2S) to de-2-O-sulphation.

In contrast, the 2-O-sulphate groups of non-terminal IdoUA(2S) residues were highly susceptible to removal. All 2-O-sulphate groups could be said to have been removed by one round of the procedure (compare IdoUA(2S) H1 resonance in Figure 43, Panels A and B). Slight changes occurred in the chemical shifts of all glucosamine protons. These changes were small in comparison to changes in the UA and IdoUA(2S) chemical shifts and indicates that the de-sulphation procedure targeted selective removal of 2-O-sulphate groups. As concluded in previous studies (Jaseja *et al.* 1989; Piani *et al.* 1993), N- and 6-O-sulphate groups were not removed by this chemical procedure.

19. Further evaluation of the de-2-O-sulphation procedure and de-sulphation of oligosaccharide, hep dp6-1.

All hep dp6-2 material was used in the initial evaluation of the de-2-O-sulphation procedure described above. Further evaluation of the procedure was carried out using heparin oligosaccharide dp6-1. This oligosaccharide sample was subjected to a variation of the chemical de-2-O-sulphation procedure used in the initial evaluation with heparin dp6-2 (see section 20 above).

After one round of de-sulphation as carried out for heparin dp6-2, the concentration of NaOH used to form an alkaline solution of the oligosaccharide for the next round of lyophilisation was increased by 4 times, from 0.2M to 0.8M (see experimental procedures). After lyophilisation to dryness the sample was analysed by 1-D NMR spectroscopy. The 1-D NMR spectra prior to de-2-O-sulphation and after two rounds of de-sulphation at 0.2M and 0.8M NaOH are shown in Figure 44. A 2-D COSY spectrum was recorded (see Appendix I, Figure 101) and the proton chemical shifts assigned (Table 7).

Figure 44

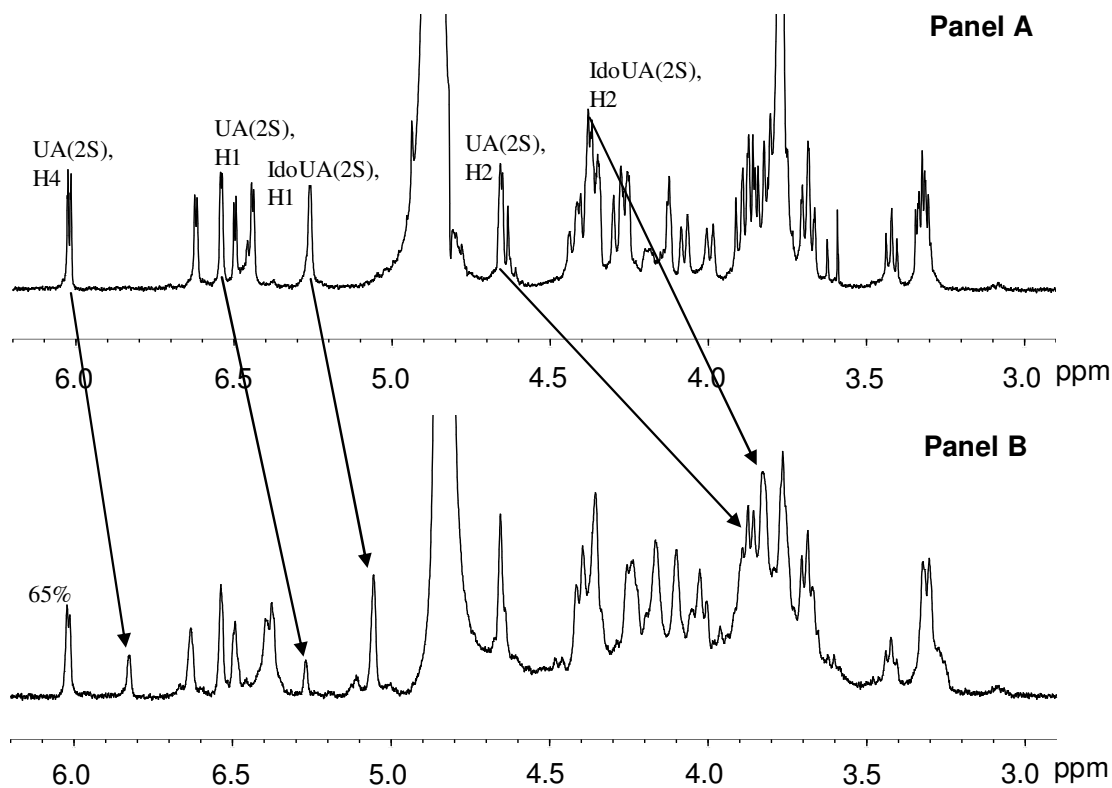


Figure 44 NMR analysis of heparin dp6-1 after chemical de-2-O-sulphation.

Panel A, heparin dp6-1 prior to the de-sulphation procedure; Panel B, after one round of de-sulphation using 0.2M NaOH and a subsequent round using 0.8M NaOH. The peak area of the resonance at 6.02 is given as a percentage of the its area plus that of the resonance at 5.83 ppm in each panel (see text). The most significant changes in the chemical shifts after de-2-O-sulphation involved protons within the UA and IdoUA residues (see Table 7) and these are marked with arrows.

Table 7. ^1H chemical shifts assignments (ppm) for heparin dp6-1 prior to and after undergoing de-2-O-sulphation. The difference column refers to the difference between the protons original chemical shift and the chemical shift value after de-2-O-sulphation.

Residue		Original (ppm)		After de-2-O-sulphation (ppm)	Difference (ppm)
1. ΔUA(2S)	H1	5.542(2.2)	H1	5.277(1.8)	-0.265
	H2	4.652	H2	3.86	-0.792
	H3	4.341(5.4)	H3	4.41(2.8)	+0.069
	H4	6.02	H4	5.832	-0.188
2. GlcNS(6S)	H1	5.433(3.4)	H1	5.37(3.4)	-0.063
	H2	3.332	H2	3.315	0.017
	H3	3.682	H3	3.689	+0.007
	H4	3.864	H4	3.759	-0.105
	H5	4.07	H5	4.042	-0.028
	H6	4.291	H6	4.243	-0.048
	H6'	or 4.401	H6'	or 4.364	-0.037
3. IdoUA(2S) or IdoUA	H1	5.254(2.5)	H1	5.055(1.6)	-0.199
	H2	4.367(5.6)	H2	3.83(5.3)	-0.537
	H3	4.245(4.1)	H3	4.168(3.4)	-0.77
	H4	4.123(3.2)	H4	4.102(2.8)	-0.021
	H5	4.847	H5	4.859	+0.012
4. GlcNS(6S)	H1	5.613(3.8)	H1	5.626(4.4)	+0.013
	H2	3.315	H2	3.321	+0.006
	H3	3.686	H3	3.684	+0.002
	H4	3.798	H4	3.759	-0.039
	H5	3.984	H5	4.042	+0.058
	H6	4.262	H6	4.243	-0.019
	H6'	or 4.424	H6'	or 4.364	-0.06
5. GlcUA(2S)	H1	4.632(9.8)	H1	4.652(7.2)	+0.02
	H2	3.322(8.2)	H2	3.426(9.1)	+0.104
	H3	3.882	H3	3.9	+0.018
	H4	3.828	H4	n. d.	n. d.
	H5	n. d.	H5	n. d.	n. d.
6. GlcNS(6S) (α)	H1	5.483(3.7)	H1	5.489(4.1)	+0.006
	H2	3.309	H2	3.321	-0.012
	H3	3.759	H3	3.764	+0.005
	H4	4.192	H4	4.188	-0.004
	H5	4.378	H5	4.203	-0.175
	H6	4.248-4.305	H6	4.354-4.42	+0.106-+0.115
	H6'	4.248-4.305	H6'	4.354-4.42	+0.106-+0.115

Coupling constants (in Hz) to proton directly below are given in brackets
n. d. Not determined

As monitored through the relative intensity of the H4 signals at 6.02 and 5.83 (see Figure 44, above) with hep dp6-1 a second round of de-sulphation using 0.8M NaOH did not result in the removal of an additional number of sulphate groups. This was over an equivalent two rounds of de-sulphation using 0.2M NaOH. Compare percentages of 37% of the 2-O-sulphate groups removed after two rounds of de-sulphation using 0.2M NaOH (see Figure 43, panel C) and 35% removed using one round of 0.2M NaOH and one round of 0.8M NaOH (see Figure 44, panel B).

This suggests there is a limit to the number of UA(2S) 2-O-sulphate groups that can be removed in one pass of this well established procedure, and that this limit can not simply be overcome by increasing the concentration of NaOH used (see discussion).

20. De-2-O-sulphation of HS dp6-2 and dp6-3.

De-2-O-sulphation of HS dp6-2 and dp6-3 were carried out using two rounds of the procedure at 0.2M NaOH. The 1D-NMR spectra before and after de-2-O-sulphation are shown in Figure 45. A full assignment of HS dp6-2 was not carried out as it was clear from the comparison to 1-D spectrum, prior to de-sulphation, that severe overlap occurred between resonances for the two non-sulphated iduronic acid residues that were now present in this sequence (see Figure 45). A COSY spectrum was recorded only for dp6-3 (see Appendix I, Figure 100) and chemical shifts only assigned for this oligosaccharide (Table 8). Changes in the chemical shifts of HS dp6-2 on de-2-O-sulphation as shown in Figure 45 are inferred from a comparison to the 1D-spectra of the sample prior to de-2-O-sulphation.

Figure 45 NMR analysis of HS dp6-2 and dp6-3 after chemical de-2-O-sulphation.

Panel A, HS dp6-2 prior to the de-sulphation procedure; Panel B, HS dp6-2 after one round of de-sulphation using 0.2M NaOH; Panel C, HS dp6-3 prior to the de-sulphation procedure; Panel D, HS dp6-3 after one round of de-sulphation using 0.2M NaOH. The most significant changes in the chemical shifts after de-2-O-sulphation involved protons within the IdoUA residue (see Table 8 for dp6-3 changes) and these are marked with arrows.

Figure 45

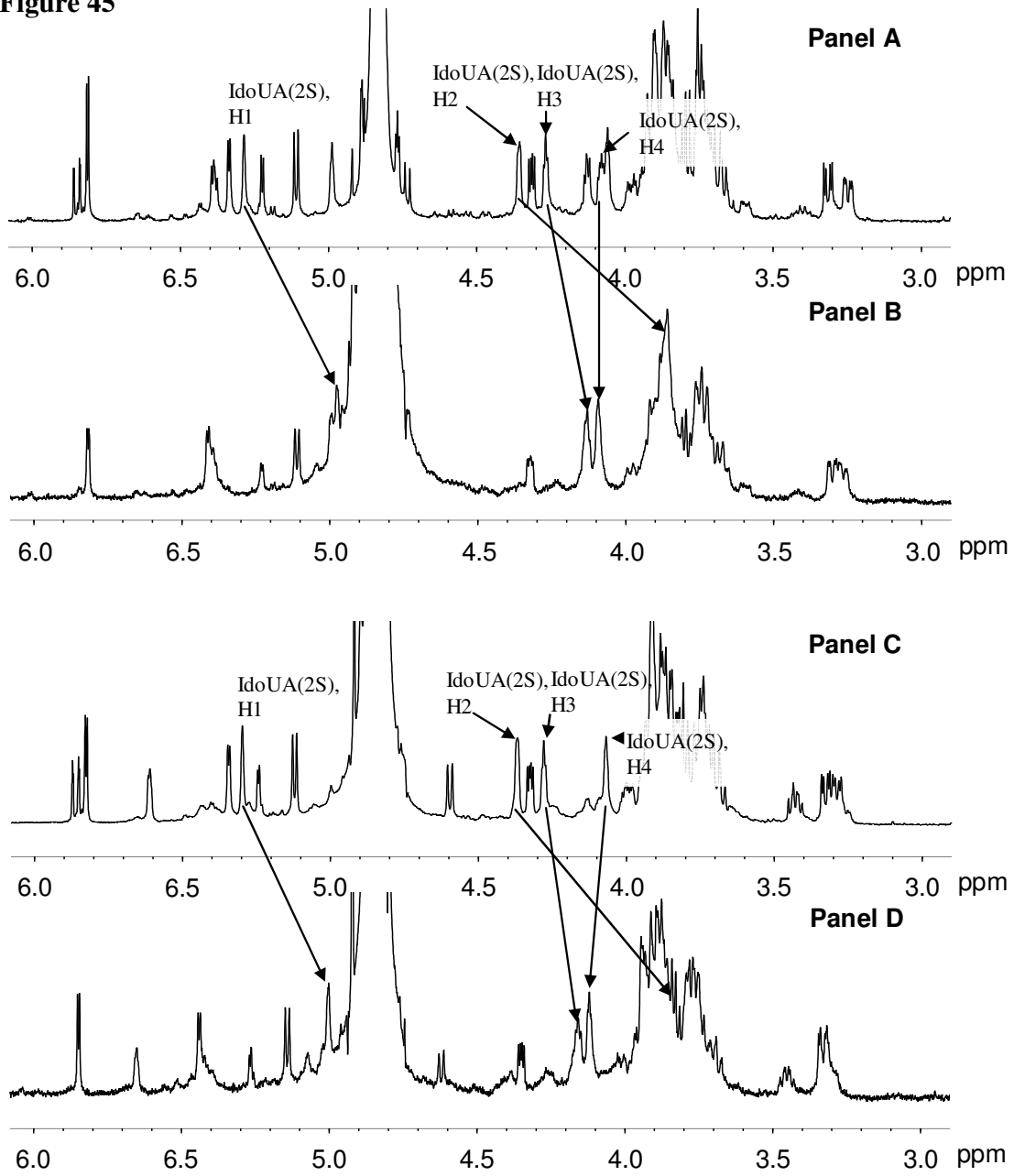


Table 8. ^1H chemical shifts assignments (ppm) for HS dp6-3 prior to and after undergoing de-2-O-sulphation.

The difference column refers to the difference between the protons original chemical shift and the chemical shift value after de-2-O- sulphation.

Residue		Original (ppm)		After de-2-O-sulphation (ppm)	Difference (ppm)
1. ΔUA	H1	5.122 (6.6)	H1	5.15(6.6)	+0.028
	H2	3.813 (6.0)	H2	3.836(6.0)	+0.023
	H3	4.325 (3.4)	H3	4.357(3.5)	+0.032
	H4	5.825	H4	5.854	+0.029
2. GlcNS	H1	5.343 (3.5)	H1	5.447(3.8)	+0.104
	H2	3.327 (10.7)	H2	3.345	+0.018
	H3	3.77	H3	3.738	-0.032
	H4	3.885	H4	3.894	+0.009
	H5	3.85	H5	n. d.	n. d.
	H6	3.797-3.966	H6	n. d.	n. d.
	H6'	3.797-3.966	H6'	n. d.	n. d.
3. IdoUA(2S) or IdoUA	H1	5.297 (1.6)	H1	5.009(2.5)	-0.288
	H2	4.368 (4.1)	H2	3.789(5.3)	-0.579
	H3	4.279 (3.15)	H3	4.164(4.4)	-0.115
	H4	4.068 (1.5)	H4	4.128(2.8)	+0.06
	H5	4.896	H5	4.842	-0.054
4. GlcNS	H1	5.613 (3.7)	H1	5.64(4.1)	+0.027
	H2	3.288 (10.1)	H2	3.326	+0.038
	H3	3.683	H3	3.686	+0.003
	H4	3.741	H4	3.748	+0.007
	H5	3.823	H5	3.863	+0.04
	H6	3.891-3.966	H6	n. d.	n. d.
	H6'	3.891-3.966	H6'	n. d.	n. d.
5. GlcUA	H1	4.596 (7.8)	H1	4.614(7.9)	+0.018
	H2	3.428	H2	3.439	+0.011
	H3	3.895	H3	3.909	+0.014
	H4	3.828	H4	3.862	+0.034
	H5	3.813	H5	n. d.	n. d.
6. GlcNAc (α)	H1	5.242 (3.1)	H1	5.27(3.2)	+0.028
	H2	3.905	H2	3.95	+0.045
	H3	3.934	H3	3.919	-0.045
	H4	3.726	H4	3.752	+0.026
	H5	3.996	H5	3.966	0
	H6	3.842-3.939	H6	n. d.	n. d.
	H6'	3.842-3.939	H6'	n. d.	n. d.
	NAc	2.075		2.091	+0.016

Coupling constants (in Hz) to proton directly below are given in brackets

n. d. Not determined

21. Molecular dynamic simulations of iduronate residues.

As mentioned in the aims section of the introduction, it was also sought to produce an improved set of coupling constants for iduronate conformational analysis. This was done using computational molecular dynamic modelling, in a manner similar to Forster and Mulloy (Forster and Mulloy 1993).

Molecular dynamic simulations were carried out for 1000ps on IdoUA and IdoUA(2S) monosaccharide models in explicit water at a temperature 300K and a pressure of 1atm (see experimental procedures). Over the course of the simulation, variations in the Cremer-Pople ring puckering parameters, θ and where warranted ϕ_2 (see introduction) monitored. Different simulations were started with models in each of the three low energy conformations of iduronate, 1C_4 , 2S_0 and 4C_1 (see introduction). Variations in the Cremer-Pople parameters during the simulations are shown in Figures 46 and 47.

For both IdoUA and IdoUA(2S) the 4C_1 and 2S_0 starting conformation were stable over the course of the simulation, the Cremer-Pople parameters averaging out at the following values:

IdoUA 4C_1	$\theta = 8.98^\circ$	
IdoUA(2S) 4C_1	$\theta = 9.17^\circ$	
IdoUA 2S_0	$\theta = 92.8^\circ$	$\phi_2 = 154.1^\circ$
IdoUA(2S) 2S_0	$\theta = 92.99^\circ$	$\phi_2 = 160.67^\circ$

The 2S_0 and 4C_1 constants shown in Table 9 were therefore averaged over all 1000 structures sampled during the MD simulation. For both IdoUA and IdoUA(2S) the 1C_4 conformation showed a single major conformational transition during the MD simulation. This occurred at 99ps for IdoUA (see Figure 46, Panels D and E) and 599ps for

IdoUA(2S) (see Figure 47, Panels D and E). The average values for θ and ϕ_2 after the transition occurred to the end of the simulation were 93.4° and 155° for IdoUA and 92.9° and 152.2° for IdoUA(2S). These average values established the conformational change to be in both cases from the 1C_4 conformation to the 2S_0 conformation. These conformational transitions validated that the glycam force field was able to make correct structural predictions with regard to the low energy conformations of the iduronate ring. This is important, because if transitions were made between conformations that were known to be of high energy and not exist experimentally, this would seriously undermine the credibility of any modelling studies undertaken using this force-field. The 1C_4 constants shown in Table 9 were averaged over 99 frames for IdoUA and 599 frames for IdoUA(2S).

Figure 46 Cremer-Pople ring puckering parameters for the methyl glycoside of IdoUA during a 1000ps molecular dynamic simulation.

Panel A, θ values for models starting in the 4C_1 conformation; Panel B, θ values for models starting in the 2S_0 conformation; Panel C, ϕ_2 values for models starting in the 2S_0 conformation; Panel D, θ values for models starting in the 1C_4 conformation; Panel E, ϕ_2 values for models starting in the 1C_4 conformation. See experimental procedures for experimental detail.

Figure 46

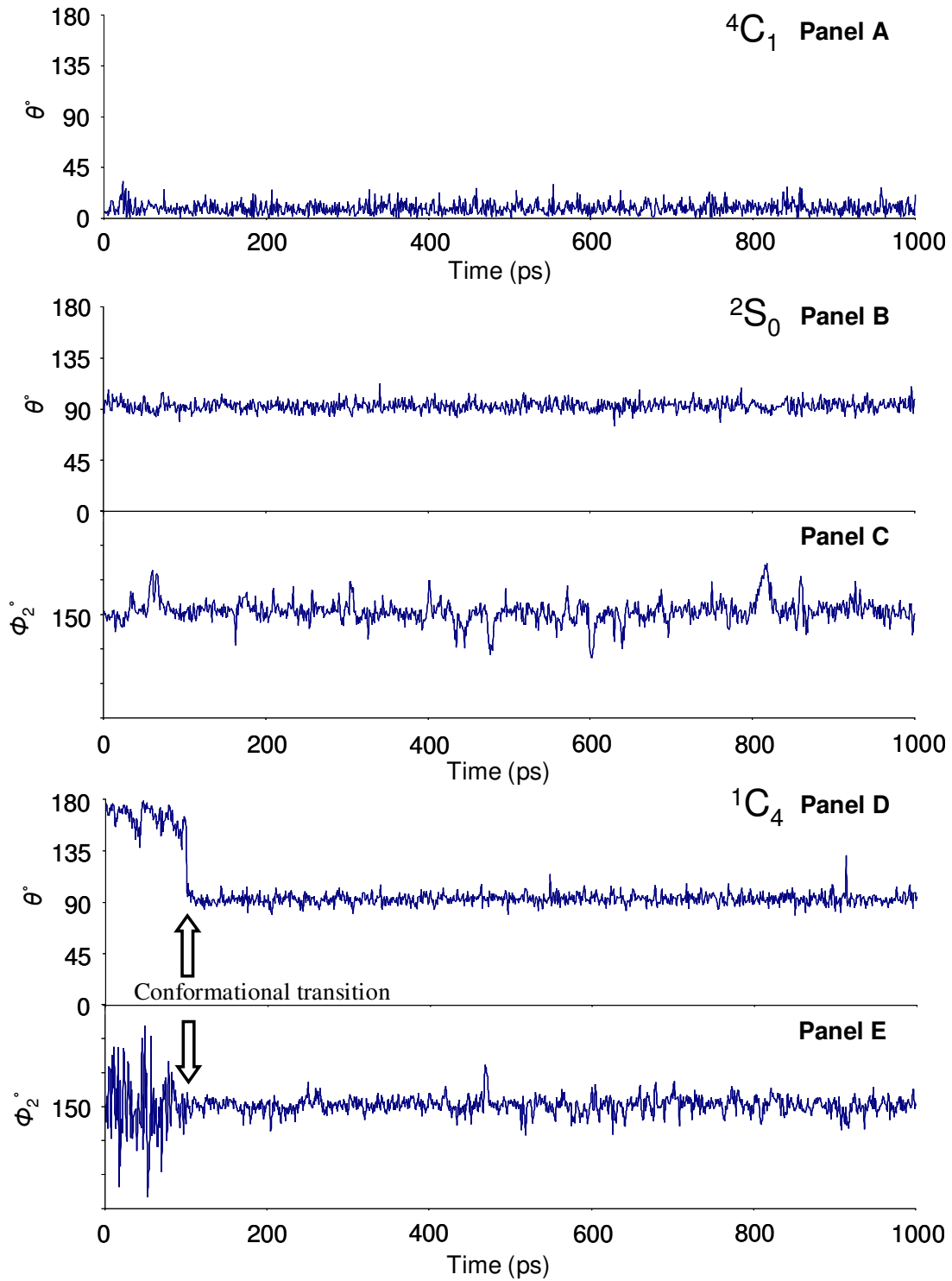


Figure 47 Cremer-Pople ring puckering parameters for the methyl glycoside of IdoUA(2S) during a 1000ps molecular dynamic simulation.

Panel A, θ values for models starting in the 4C_1 conformation; Panel B, θ values for models starting in the 2S_0 conformation; Panel C, ϕ_2 values for models starting in the 2S_0 conformation; Panel D, θ values for models starting in the 1C_4 conformation; Panel E, ϕ_2 values for models starting in the 1C_4 conformation. See experimental procedures for experimental detail.

Figure 47

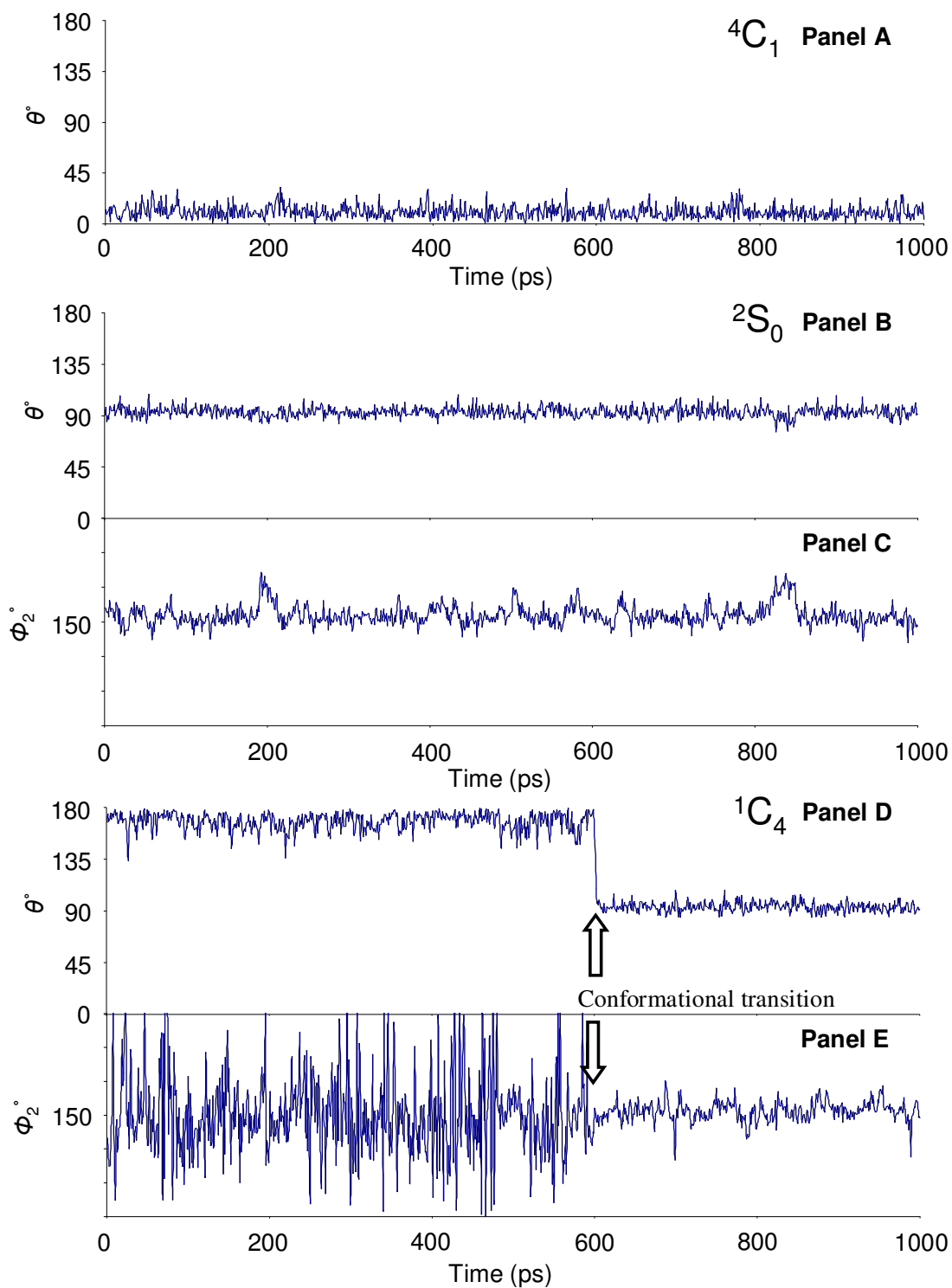


Table 9. The average value and standard deviation of the theoretical iduronate coupling constants as obtained by explicit water molecular dynamic simulations.

The program mdXvu was used to analyse the explicit water MD trajectories and extract theoretical coupling constants for the sugar in each frame according to the Haasnoot *et al* equation (Haasnoot *et al.* 1979). The 2S_0 and 4C_1 constants were averaged over 1000 MD trajectory frames. 1C_4 constants were averaged over 99 frames for IdoUA and 599 frames for IdoUA(2S) (see text).

	1C_4	2S_0	4C_1
${}^3J_{1,2}$	IdoUA 1.71 ± 0.53	IdoUA 4.36 ± 1.5	IdoUA 7.34 ± 0.64
	IdoUA(2S) 1.78 ± 0.53	IdoUA(2S) 3.99 ± 1.39	IdoUA(2S) 7.28 ± 0.69
${}^3J_{2,3}$	IdoUA 1.56 ± 0.99	IdoUA 9.25 ± 1.04	IdoUA 9.05 ± 1.02
	IdoUA(2S) 1.82 ± 0.94	IdoUA(2S) 9.26 ± 1.04	IdoUA(2S) 8.86 ± 1.1
${}^3J_{3,4}$	IdoUA 2.01 ± 0.93	IdoUA 5.78 ± 2.1	IdoUA 9.06 ± 0.79
	IdoUA(2S) 2.32 ± 0.99	IdoUA(2S) 6.36 ± 1.94	IdoUA(2S) 9.12 ± 0.77
${}^3J_{4,5}$	IdoUA 1.23 ± 0.7	IdoUA 4.26 ± 1.89	IdoUA 5.41 ± 1.06
	IdoUA(2S) 1.1 ± 0.71	IdoUA(2S) 4.73 ± 1.91	IdoUA(2S) 5.15 ± 1.09

22. Molecular modelling.

Static models were produced for all four HS derived oligosaccharides using the established strategy employed by Mulloy *et al.* (Mulloy *et al.* 1993) and Mikhailov *et al.* (Mikhailov *et al.* 1996; Mikhailov *et al.* 1997).

22.1 Static molecular modelling of HS dp6-3

In order to derive inter-proton distance restraints NOESY spectra were recorded on HS dp6-3 with mixing times of 50, 100, 200, 300, 400, 500 and 600ms. Inter-proton distance restraints are important, as frequently more than one low energy geometry is indicated on the basis of force field energy calculations alone. For HS dp6-3, the NOE intensities for symmetry related cross-peaks corresponding to the following enhancements were averaged:

UA H1 – GlcNS H4, GlcNS H1 – IdoUA(2S) H4, GlcNS H1 – IdoUA(2S) H3, IdoUA(2S) H1 – GlcNS H4, IdoUA(2S) H1 – GlcNS H6, GlcNS H1 – GlcUA H4 and GlcUA H1 – GlcNAc H4.

All of the above NOE cross-peaks span glycosidic linkages. It is glycosidic linkage geometry that mainly determines the overall 3-dimensional shape of an oligosaccharide (see introduction)

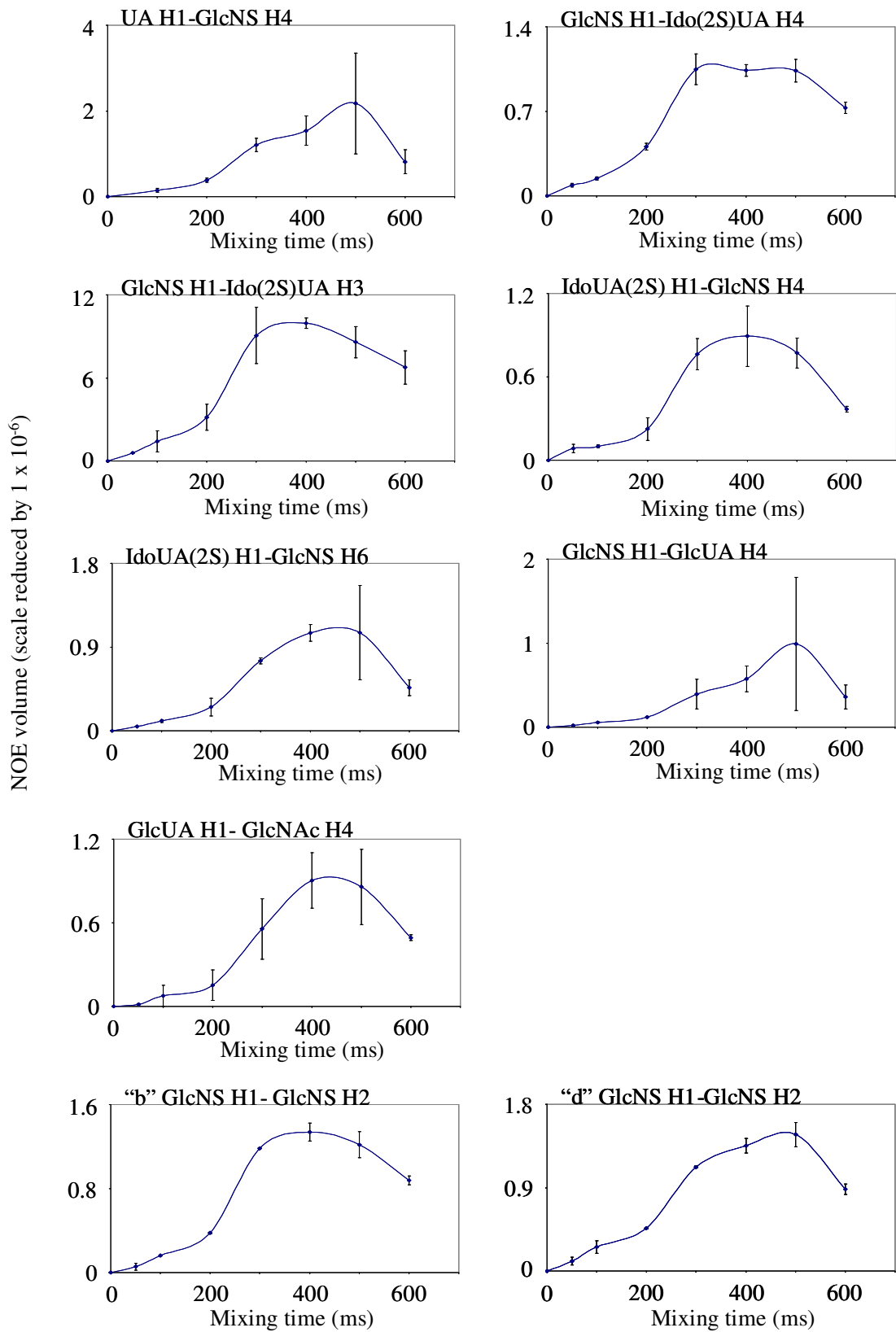
For distance calibration the intra-residue NOE GlcNS H1- GlcNS H2 of both glucosamines (residues “b” and “d”) was also assessed. Variations in the distance spanned by an intra-residue glucosamine H1-H2 NOE are limited by pyranose ring geometry.

These results are shown in Figure 48.

Figure 48 **NOE build up curves for all trans-glycosidic NOEs observed within oligosaccharide HS dp6-3.**

NOESY spectra were recorded using mixing times of 50, 100, 200, 300, 400, 500 and 600ms (see experimental procedures). Symmetry related cross-peaks were directly integrated using the Mest-ReC NMR analysis program. These volumes were then averaged to obtain the final volumes plotted. Error bars correspond to the standard error between the two NOE volumes averaged.

Figure 48



In contrast to the build up curves published by Mikhailov *et al.* (Mikhailov *et al.* 1996; Mikhailov *et al.* 1997) a considerable lag phase was observed in the build up profile of all the NOEs analysed (see Figure 48). Consequently a two-spin approximation analysis of the data was undertaken (see discussion). The build up rate over the initial lag phase, corresponding to the first 200ms of each profile (see Figure 48) was calculated and a value derived for the distance spanned by each NOE, using the formula

$$\text{Distance} = \text{reference distance} \times \left[\frac{\text{measured build up rate}}{\text{reference distance build up rate}} \right]^{-1/6}$$

(Neuhaus and Williamson 2000)

A theoretical GlcNS H1- GlcNS H2 reference distance of 2.4 Å was used as calibration.

The calculated approximate distances spanned by each NOE are listed below:

UA H1 – GlcNS H4	2.58 Å
GlcNS H1 – IdoUA(2S) H4	2.45 Å
GlcNS H1 – IdoUA(2S) H3	2.53 Å
IdoUA(2S) H1 – GlcNS H4	2.66 Å
IdoUA(2S) H1 – GlcNS H6	2.63 Å
GlcNS H1 – GlcUA H4	3.0 Å
GlcUA H1 – GlcNAc H4	2.9 Å

22.2 Calculation of potential energy surfaces.

Potential energy surfaces for each glycosidic linkage were calculated as described under experimental procedure. Potential energy surfaces can be used to locate low energy glycosidic linkage geometries, which can then be assessed against a set of inter-proton distance restraints. Low energy geometries in which a given inter-proton distance exceeds a distance restraint can be rejected as unlikely to contribute significantly to the solution conformation of a modeled oligosaccharide. The results for HS dp6-3 are presented as contour plots in Figure 49.

NOE derived distances were used to define maximum inter-proton distances that must be present in a final proposed model structure. The NOE derived distances were used to assess the low energy geometries across each glycosidic linkage, as indicated by the potential energy surface calculations.

Figure 49 Potential energy surfaces calculated for each glycosidic linkage within HS dp6-3.








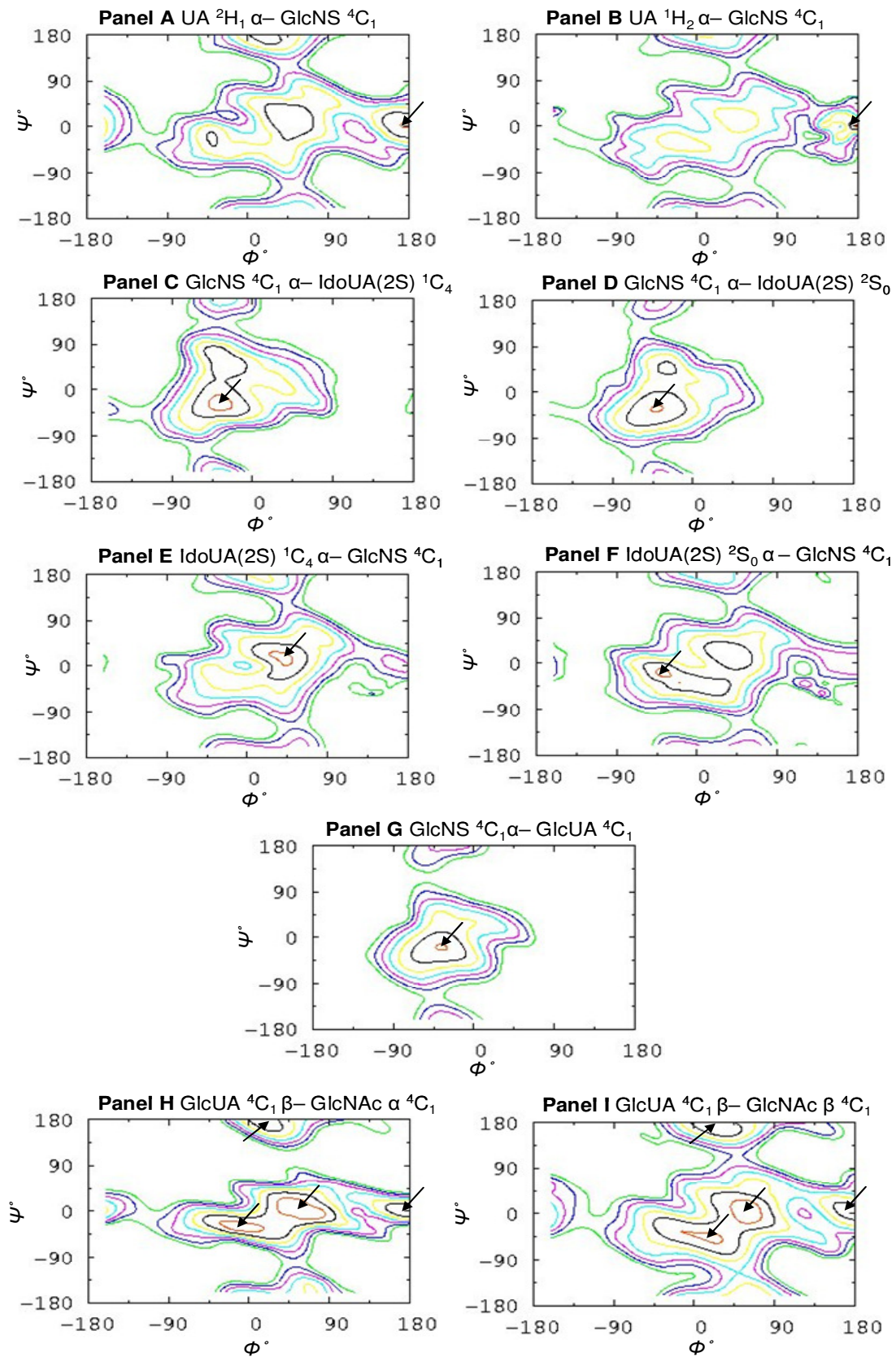
Contour lines are shown at the global energy minima () marked with an arrow, and at intervals of 3 kJ/mol above this, () global minima +3kJ/mol, () global minima +6kJ/mol, () global minima +9kJ/mol, () global minima +12 kJ/mol, () global minima +15 kJ/mol, () global minima +18 kJ/mol. The potential energy was evaluated using the glycam04 force field within the AMBER molecular modelling suite and partial charge calculation method 1 (see experimental methods).

Figure 49



22.3 Geometry of the UA-GlcNS linkage (monosaccharide “a”-monosaccharide “b”)

The potential energy surfaces indicated there to be a minimum energy geometry at dihedral angles of around $\Phi = 180^\circ$ and $\Psi = 0^\circ$, using either starting conformation of the uronic acid (see Figure 49, Panels A and B). On inspection of the energy minimised structures corresponding to $\Phi = 180^\circ$ and $\Psi = 0^\circ$, it was found that the ${}^1\text{H}_2$ conformation of UA had made a transition to the ${}^2\text{H}_1$ conformation during the calculation. The minima indicated on both maps were therefore the energy result calculated for the same structure UA ${}^2\text{H}_1$ - GlcNS ${}^4\text{C}_1$ $\Phi = 180^\circ$ and $\Psi = 0^\circ$. An overlay of the energy minimised structure calculated from the ${}^2\text{H}_1$ and a ${}^1\text{H}_2$ starting conformation of the uronic acid is given in Figure 50. The rmsd deviation between the two energy minimised final structures was 0.22 Å. As with iduronate (see results section 23) this conformational transition validates that the glycam force field is able to make correct structural predictions with regard to the low energy conformations of unsaturated uronic acid residues. Significantly a transition to an intermediate conformation between ${}^1\text{H}_2$ and ${}^2\text{H}_1$ as observed by Mikhailov *et al.* (Mikhailov *et al.* 1996) (see introduction, section 18.3) was not observed. This weighs further support to the argument that this residue was parameterised incorrectly by Mikhailov *et al.*

Figure 50

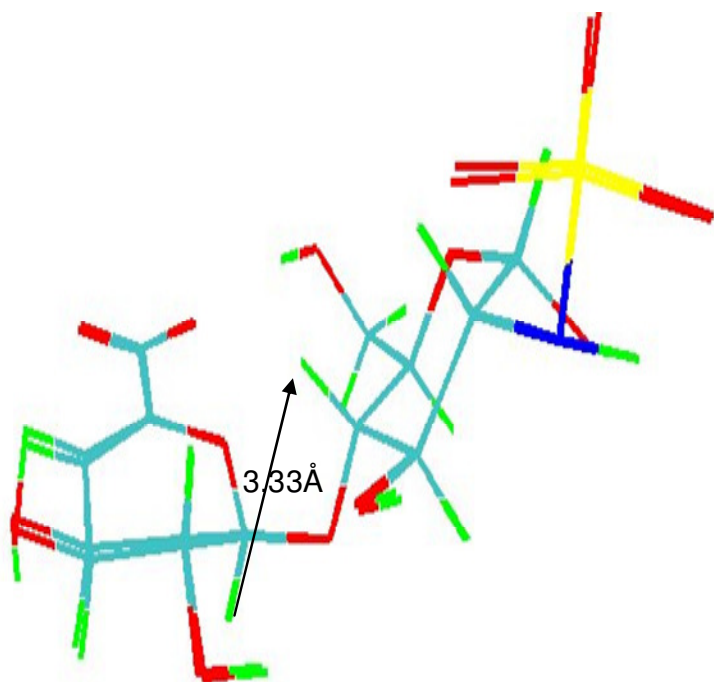


Figure 50 Energy minimised structures for the UA-GlcNS glycosidic linkage at $\Phi = 180^\circ$ and $\Psi = 0^\circ$.

An overlay of two final structures is shown, for one structure the UA starting conformation was ${}^2\text{H}_1$ for the other ${}^1\text{H}_2$. The distance from proton H1 of the uronic acid to proton H4 of the glucosamine was 3.33Å as indicated.

The UA H1 to GlcNS H4 distance of 3.33Å clearly largely exceeds the NOE maximum indicated distance of 2.58Å. Additionally, in this geometry a potential NOE between these two protons will be obstructed by the C1- glycosidic oxygen bond of the uronic acid. For these reasons it was thought unlikely that this geometry contributes significantly to the solution conformation across this glycosidic bond. Other low energy geometries were sought more consistent with the NOE data.

At the +3kJ/mol level three other low energy regions are indicated centred around $\Phi = 20^\circ - \Psi = 180^\circ$, $\Phi = 40^\circ - \Psi = 20^\circ$ and $\Phi = -40^\circ - \Psi = -20^\circ$ for a starting conformation of ${}^2\text{H}_1$ and $\Phi = 0^\circ - \Psi = 180^\circ$, $\Phi = 40^\circ - \Psi = 20^\circ$ and $\Phi = -20^\circ - \Psi = -20^\circ$ for a starting conformation of ${}^1\text{H}_2$ (see Figure 49, Panels A and B). For all these geometries the UA remained in its starting conformation throughout the calculation. The energy minimised structures are shown in Figure 51.

For both conformations of the uronic acid a significant contribution of the $\Phi = 20^\circ - \Psi = 180^\circ$ geometry (see Figure 49, Panels A and B) to the solution conformation, is not supported by the NOE evidence. UA H1 to GlcNS H4 distances of 3.33Å for ${}^1\text{H}_2$ and 3.35Å for ${}^2\text{H}_1$ are much greater than the 2.58Å distance derived from the intensity of the NOE between these two protons. Additionally a potential NOE between these two protons in this geometry will again be obstructed, here by the bond between C4 of GlcNS and the glycosidic oxygen.

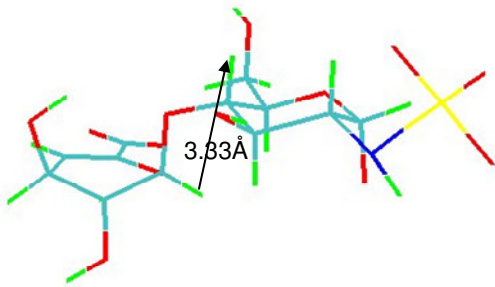
All other geometries shown in Figure 51 (Panels C-F) can not be ruled out as contributing to the solution conformation across the UA-GlcNS glycosidic bond. Movement between these geometries may occur in solution.

Figure 51 Energy minimised structures for the UA-GlcNS glycosidic in a number of low energy geometries, as indicated by the potential energy surface calculations.

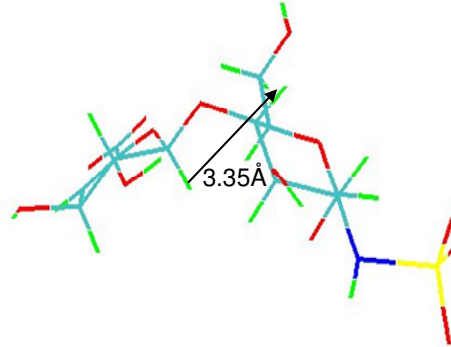
The distance from proton H1 of the uronic acid to proton H4 of the glucosamine is as indicated.

Figure 51

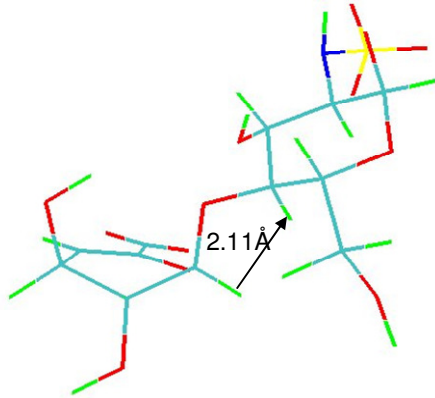
Panel A UA $^1\text{H}_2\text{-GlcNS } ^4\text{C}_1$ $\phi=20^\circ$ $\psi=180^\circ$



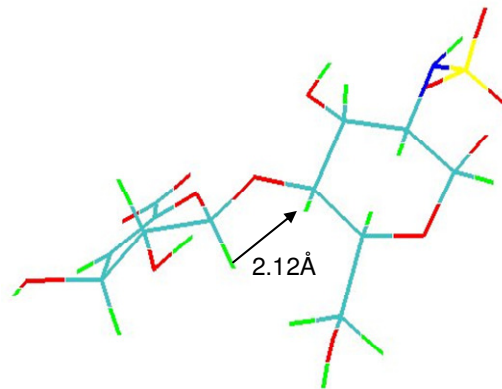
Panel B UA $^2\text{H}_1\text{-GlcNS } ^4\text{C}_1$ $\phi=20^\circ$ $\psi=180^\circ$



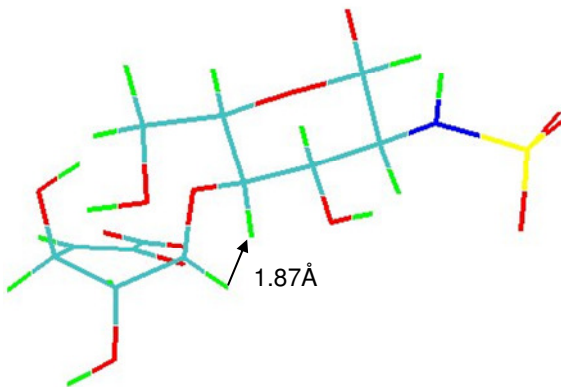
Panel C UA $^1\text{H}_2\text{-GlcNS } ^4\text{C}_1$ $\phi=40^\circ$ $\psi=20^\circ$



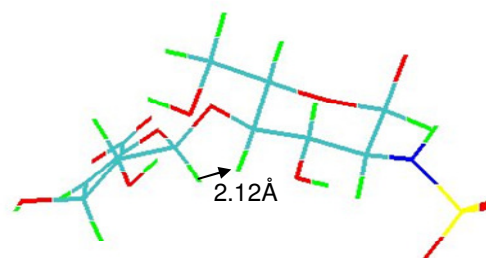
Panel D UA $^2\text{H}_1\text{-GlcNS } ^4\text{C}_1$ $\phi=40^\circ$ $\psi=20^\circ$



Panel E UA $^1\text{H}_2\text{-GlcNS } ^4\text{C}_1$ $\phi=-20^\circ$ $\psi=-20^\circ$



Panel F UA $^2\text{H}_1\text{-GlcNS } ^4\text{C}_1$ $\phi=-40^\circ$ $\psi=-20^\circ$



22.4 Geometry of the GlcNS-IdoUA(2S) linkage (monosaccharide “b”-monosaccharide “c”)

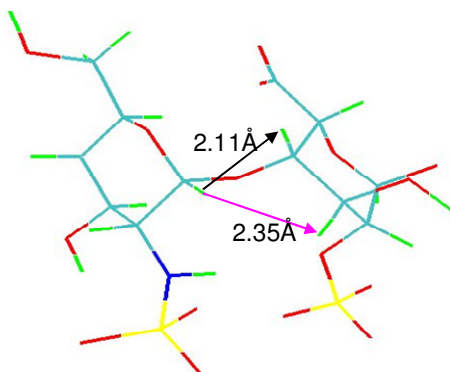
The potential energy surface calculations indicated there to be a minimum energy geometry at dihedral angles of around $\Phi = -40^\circ$ and $\Psi = -20^\circ$ for both IdoUA(2S) in the 1C_4 and 2S_0 conformations (see Figure 49, Panels C and D). The energy minimised structures at these geometries are shown in Figure 52.

For both conformations (1C_4 and 2S_0) of IdoUA(2S) the GlcNS H1 - IdoUA(2S) H4 distance at the geometries shown in Figure 52 are well within the NOE derived distance limit of 2.45Å. For the 1C_4 conformation the GlcNS H1- IdoUA(2S) H3 distance is also within the NOE limit of 2.53Å. However, this GlcNS H1- IdoUA(2S) H3 NOE derived distance is exceeded when IdoUA(2S) is in the 2S_0 conformation (see Figure 52, Panel B). Other low energy geometries were sought for this glycosidic linkage (see Figure 49, Panel D) that more consistent with the NOE data.

At the first energy level higher than the global minima, the +3kJ/mol level, the low energy structure corresponding to glycosidic dihedral angles $\Phi = -20^\circ$ and $\Psi = -40^\circ$ was examined. In this structure the distance actually increases to 3.8Å (data not shown). No low energy geometries could be found in which the H1 GlcNS- H3 IdoUA(2S) distance was not greater than the NOE derived maximum distance of 2.53Å, when IdoUA(2S) was in the 2S_0 conformation. This supports the conclusion that this IdoUA(2S) unit predominately occurs in the 1C_4 conformation.

Figure 52

Panel A GlcNS 4C_1 - IdoUA(2S) 1C_4
 $\Phi = -40^\circ$ $\Psi = -20^\circ$



Panel B GlcNS 4C_1 - IdoUA(2S) 2S_0
 $\Phi = -40^\circ$ $\Psi = -20^\circ$

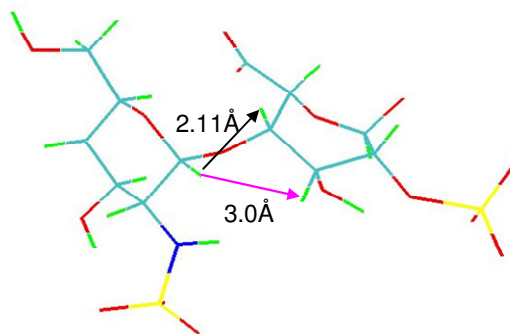


Figure 52 Energy minimised structures for the GlcNS-IdoUA(2S) (1C_4 and 2S_0) glycosidic linkage at $\Phi = -40^\circ$ and $\Psi = -20^\circ$.

The distance from proton H1 of the glucosamine to proton H4 of the iduronic acid is as indicated (\rightarrow). The distance from proton H1 of the glucosamine to proton H4 of the iduronic acid is as indicated (\rightarrow).

22.5 Geometry of the IdoUA(2S)-GlcNS linkage (monosaccharide “c”-monosaccharide “d”)

The potential energy surface calculations indicated there to be a minimum energy geometry at dihedral angles of around $\Phi = 40^\circ$ and $\Psi = 20^\circ$ for IdoUA(2S) in the 1C_4 and around $\Phi = -40^\circ$ and $\Psi = -20^\circ$ for IdoUA(2S) in the 2S_0 conformation. The energy minimised structures at these geometries are shown in Figure 53.

For IdoUA(2S) in the 1C_4 conformation both the IdoUA(2S) H1- GlcNS H4 and IdoUA H1- GlcNS H4 distances are well within the NOE distances of 2.66Å and 2.63Å respectively (see Figure 53, Panel A). However, for IdoUA(2S) in the 2S_0 conformation, while the IdoUA(2S) H1- GlcNS H4 distance is within the NOE distance of 2.63Å, the IdoUA(2S) H1- GlcNS H6 distance significantly exceeds the 2.66Å NOE distance (see Figure 53, Panel B). Other low energy geometries were sought for this glycosidic linkage (see Figure 49, Panel F) more consistent with the NOE data.

The potential energy surface calculated for IdoUA(2S) 2S_0 - GlcNS is more complex than any of the other maps calculated. The $\Phi = -40^\circ$ and $\Psi = -20^\circ$ minima lies at one end of an extended region of low energy structures bounded by a “global minima +3kJ/mol” contour. This region extends from around $\Phi = 20^\circ - \Psi = -40^\circ$ to enclose the global minima at $\Phi = -40^\circ$ and $\Psi = -20^\circ$ (see Figure 49, Panel F). Throughout the whole of this region the IdoUA(2S) H1- GlcNS H6 distance exceeds 3.5Å (data not shown). A second region of low energy structures is centred around $\Phi = 40^\circ - \Psi = 20^\circ$. For this geometry the IdoUA(2S) H1- GlcNS H6 distance falls to 2.25Å and is more compatible with the NOE derived distance limit. This structure is shown in Figure 54.

Figure 53

Panel A IdoUA(2S) 1C_4 - GlcNS 4C_1
 $\Phi = 40^\circ$ $\Psi = 20^\circ$

Panel B IdoUA(2S) 2S_0 - GlcNS 4C_1
 $\Phi = -40^\circ$ $\Psi = -20^\circ$

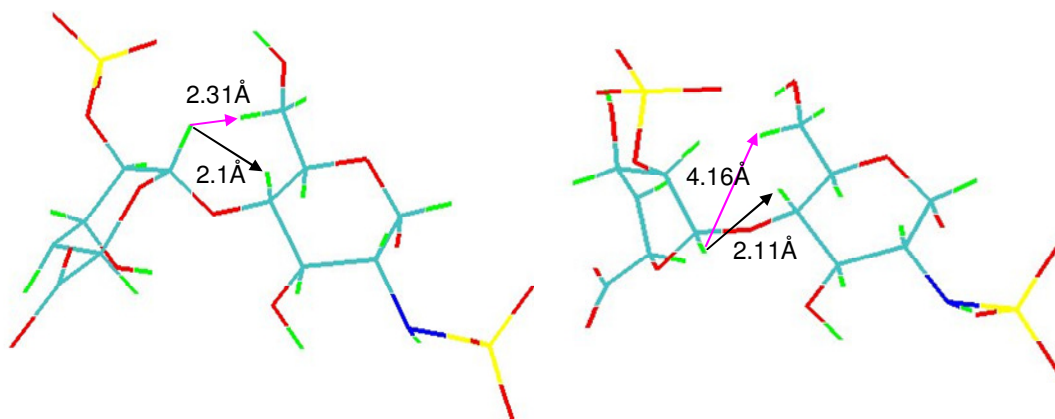


Figure 53 Energy minimised structures for the IdoUA(2S)-GlcNS glycosidic linkage at $\Phi = 40^\circ$ and $\Psi = 20^\circ$ for IdoUA(2S) in the 1C_4 conformation and $\Phi = -40^\circ$ and $\Psi = -20^\circ$ for IdoUA(2S) in the 2S_0 conformation.

The distance from proton H1 of the iduronic acid to proton H4 of the glucosamine is as indicated (\rightarrow). The distance from proton H1 of the iduronic acid to the closest glucosamine H6 proton is indicated (\rightarrow).

Figure 54

IdoUA(2S) 2S_0 - GlcNS 4C_1 $\Phi = 40^\circ$ $\Psi = 20^\circ$

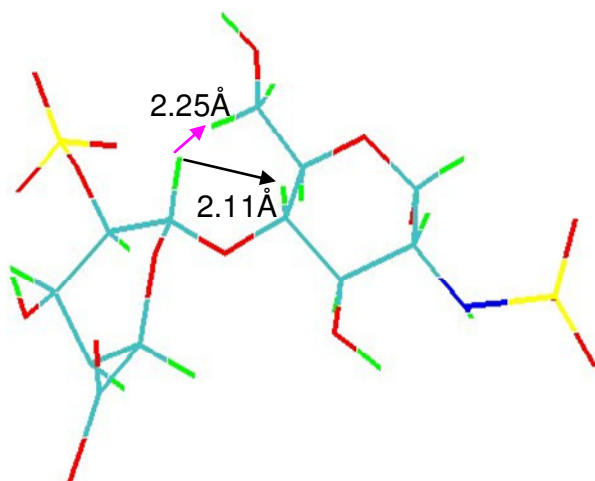


Figure 54 An energy minimised structure for the IdoUA(2S) 2S_0 - GlcNS glycosidic linkage at $\Phi = 40^\circ$ and $\Psi = 20^\circ$

The distance from proton H1 of the iduronic acid to proton H4 of the glucosamine is as indicated (\rightarrow). The distance from proton H1 of the iduronic acid to the closest glucosamine H6 proton is indicated (\rightarrow).

22.6 Geometry of the GlcNS- GlcUA linkage (monosaccharide “d”- monosaccharide “e”)

The potential energy surface indicated there to be a minimum energy geometry at dihedral angles centred around $\Phi = -40^\circ$ and $\Psi = -20^\circ$ (see Figure 49, Panel G). The energy minimised structure at this geometry is shown in Figure 55. The GlcNS H1- GlcUA H4 distance for the geometry shown in Figure 55 is well within the NOE derived distance limit of 3Å.

Figure 55

GlcNS 4C_1 - GlcUA 4C_1 $\Phi = -40^\circ$ $\Psi = -20^\circ$



Figure 55 The energy minimised structure for the GlcNS 4C_1 - GlcUA 4C_1 glycosidic linkage at $\Phi = -40^\circ$ and $\Psi = -20^\circ$

The distance from proton H1 of the glucosamine to proton H4 of the glucuronic acid is as indicated (\rightarrow).

22.7 Geometry of the GlcUA- GlcNAc (α or β) linkage (monosaccharide “e”-monosaccharide “f”)

The potential energy surface indicated there to be a number of geometries corresponding to the global minimal energy configuration of this linkage. As with the UA-GlcNS linkage, two global minima energy geometries were indicated centred around $\Phi = 180^\circ$ - $\Psi = 0^\circ$ and $\Phi = 20^\circ$ - $\Psi = 180^\circ$ (see Figure 49, Panels H and I). For both anomeric configurations and in both of these geometries, the GlcUA H1- GlcNAc H4 inter-proton distance significantly exceeds the NOE derived limit of 2.9Å (data not shown).

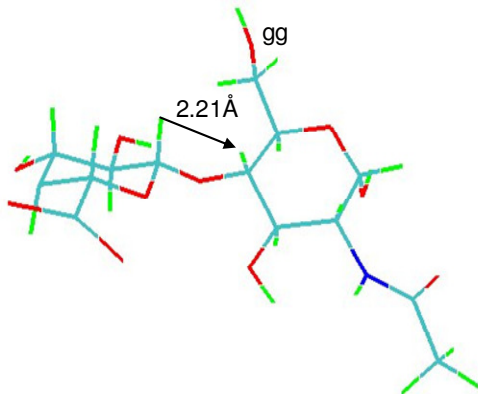
Two other global minima energy configurations are indicated centred around $\Phi = 60^\circ$ - $\Psi = 0^\circ$ and $\Phi = 0^\circ$ - $\Psi = -20^\circ$ for GlcNAc α and $\Phi = 60^\circ$ - $\Psi = 0^\circ$ and $\Phi = 20^\circ$ - $\Psi = -40^\circ$ for GlcNAc β (see Figure 49, Panels H and I). The energy minimised structures at these geometries are shown in Figure 56. For the geometries shown in Figure 56, in both anomeric configurations of GlcNAc the GlcUA H1- GlcNAc H4 distance is well within the NOE derived distance limit of 2.9Å (see Figure 56, Panel A-D). None of these geometries can be excluded as not occurring across this glycosidic linkage.

The C5-C6 rotamer conformations of both the GlcNAc α and β - anomers were initiated in the gg conformation for all potential energy surface calculations. Comparing the energy minimised structures shown in Figure 35, it is clear that a transition has been made during the calculation, by the GlcNAc β -anomer, from the starting gg conformation to the gt conformation. No such transition is observed for the α -anomer.

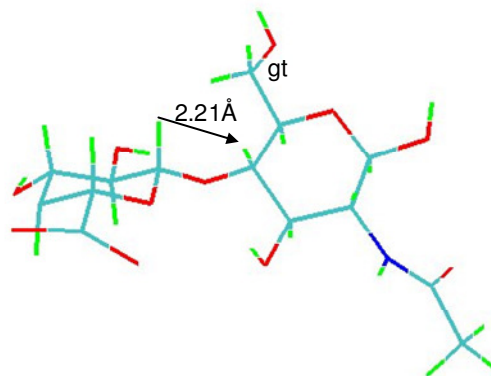
For all structures presented previous to Figure 35, the rotamer conformation remained in its initial conformation of gg throughout the calculation (see discussion).

Figure 56

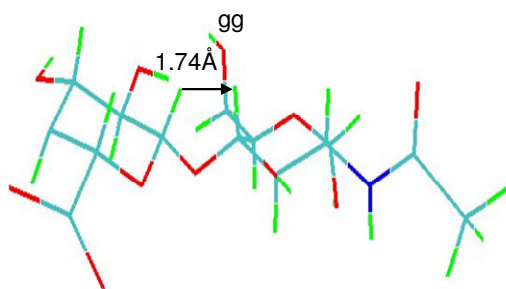
Panel A GlcUA ⁴C₁- GlcNAc α ⁴C₁ $\phi=60^\circ$ $\psi=0^\circ$



Panel B GlcUA ⁴C₁- GlcNAc β ⁴C₁ $\phi=60^\circ$ $\psi=0^\circ$



Panel C GlcUA ⁴C₁- GlcNAc α ⁴C₁ $\phi=0^\circ$ $\psi=-20^\circ$



Panel D GlcUA ⁴C₁- GlcNAc β ⁴C₁ $\phi=20^\circ$ $\psi=-40^\circ$

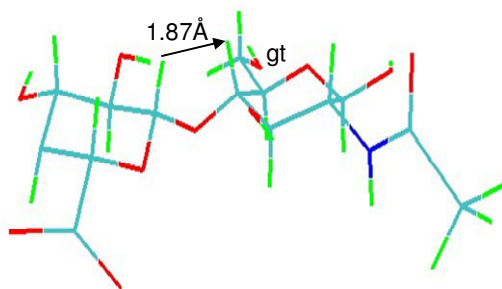


Figure 56 Energy minimised structures for the GlcUA-GlcNAc (α or β) glycosidic linkage in a number of low energy geometries, as indicated by the potential energy surface calculations.

The distance from proton H1 of the glucuronic acid to proton H4 of the glucosamine is as indicated (\rightarrow). Also labelled in each panel is the rotamer conformation of the GlcNAc C5-C6 bond (see text).

22.8 Summary for HS dp6-3

Selected Φ and Ψ angle combinations resulting in low energy geometries, across each glycosidic linkage of HS dp6-3 are given in Table 13. All of these geometries are supported by NOE evidence as possibly occurring within the solution conformation of this oligosaccharide (see sections 24.1 - 24.7).

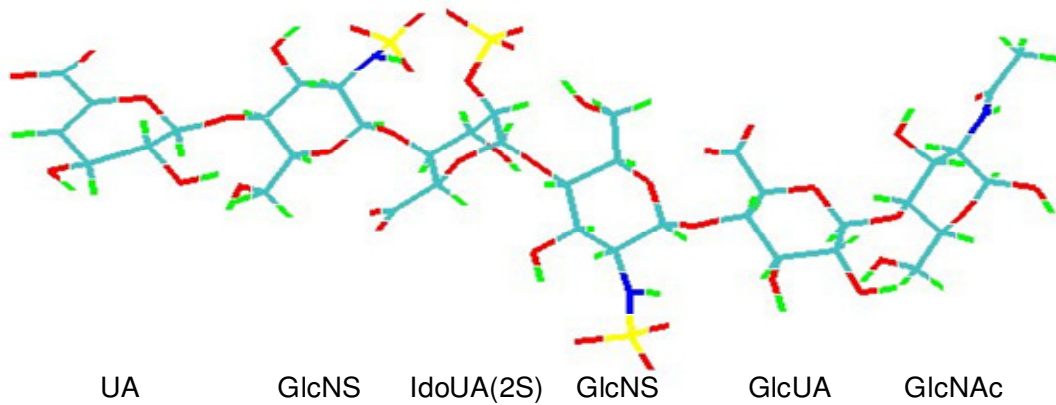
Two static models for HS dp6-3 were constructed. In each the UA monosaccharide was modelled in the predominant conformation indicated by the coupling constant analysis (see discussion) ($^2\text{H}_1$). The terminal GlcNAc monosaccharide was modelled in the predominant α anomeric configuration. Where more than one possible glycosidic geometry was indicated (see Table 10), geometries preferred as a result of the *exo*-anomeric effect were modelled (see introduction). Two models were constructed one in which the IdoUA(2S) monosaccharide was in the $^1\text{C}_4$ conformation and one in which it is in the $^2\text{S}_0$ conformation. GlcUA always occurs in the $^4\text{C}_1$ conformation (see introduction). Construction of two models containing alternate IdoUA(2S) conformations, follows the modelling procedure adopted by Mulloy *et al* (Mulloy *et al.* 1993). These models are shown in Figure 57.

Table 10 Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS **dp6-3**. All selected dihedral angles result in minimal energy geometries across the glycosidic linkage. These angles are also selected on the basis of NOE evidence as possibly contributing to the solution conformation of the oligosaccharide (see text, sections 24.1 – 24.7). Dihedral angles preferred as a result of the *exo*-anomeric effect are shown in red (see introduction).

Linkage and monosaccharide conformation	Dihedral angles	
	Φ	Ψ
a-b		
UA ¹ H ₂ - GlcNS ⁴ C ₁	40	20
	-20	-20
UA ² H ₁ - GlcNS ⁴ C ₁	40	20
	-40	-20
b-c		
GlcNS ⁴ C ₁ - IdoUA(2S) ¹ C ₄	-40	-20
GlcNS ⁴ C ₁ - IdoUA(2S) ² S ₀	-40	-20
c-d		
IdoUA(2S) ¹ C ₄ - GlcNS ⁴ C ₁	40	20
IdoUA(2S) ² S ₀ - GlcNS ⁴ C ₁	40	20
d-e		
GlcNS ⁴ C ₁ - GlcUA ⁴ C ₁	-40	-20
e-f		
GlcUA ⁴ C ₁ - GlcNAc α ⁴ C ₁	60	0
	0	-20
GlcUA ⁴ C ₁ - GlcNAc β ⁴ C ₁	60	0
	20	-40

Figure 57

Model A



Model B

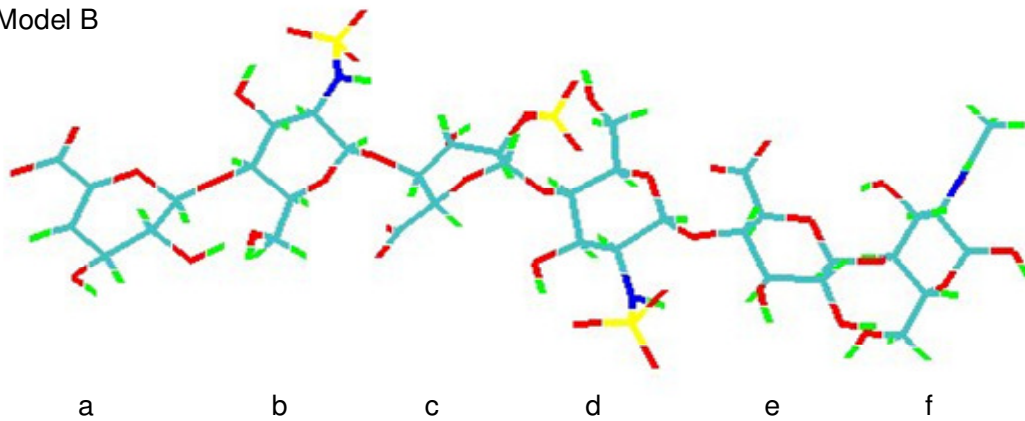


Figure 57 Two final proposed model structure of oligosaccharide HS dp6-3.

The two models differ in the conformation of the IdoUA(2S) monosaccharide at position “c”.

Model A: UA ²H₁-GlcNS ⁴C₁-IdoUA(2S) ¹C₄-GlcNS ⁴C₁-GlcUA ⁴C₁-GlcNAcα ⁴C₁

Model B: UA ²H₁-GlcNS ⁴C₁-IdoUA(2S) ²S₀-GlcNS ⁴C₁-GlcUA ⁴C₁-GlcNAcα ⁴C₁

23. NOE inter-proton distance restraints in the static molecular modelling of HS dp6-2, dp6-5 and dp6-6.

For the construction of molecular models for HS dp6-2, -5 and -6, trans-glycosidic NOEs observed at a 500ms mixing time were taken as occurring over a distance of $<3\text{\AA}$ (see discussion). Low energy geometries on each potential energy surface were then evaluated accordingly (see discussion).

23.1 Static molecular modelling of HS dp6-2.

Potential energy surfaces across many of the glycosidic linkages within HS dp6-2 were calculated as part of modelling of HS dp6-3 (see Figure 49). Additional energy surfaces were calculated for both GlcNS-IdoUA and IdoUA-GlcNAc(α or β) linkages present in HS dp6-2. These are shown in Figure 58. Table 11 lists the Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-2 supported as possible values by an NOE derived distance restraint of $<3\text{\AA}$ (see introduction to section 25 above and discussion).

In construction of the molecular models, shown in Figure 59, where several possible glycosidic geometries were indicated (see Table 11), again those preferred as a result of the *exo*-anomeric effect were modelled. In each model, the UA monosaccharide was modelled in the predominant conformation indicated by the coupling constant analysis ($^2\text{H}_1$). The terminal GlcNAc monosaccharide was modelled in the predominant α anomeric configuration. Two models were constructed one in which the IdoUA(2S) and IdoUA monosaccharides are in the $^1\text{C}_4$ conformation and one in which both are in the $^2\text{S}_0$ conformation (see Figure 59).

Note, it is entirely possible that these iduronate monosaccharides may be present in alternate conformations, i.e. one in the 1C_4 conformation and one in the 2S_0 conformation. There is no experimental evidence to suggest that changes between iduronate conformations occur in a concerted fashion along an oligosaccharide chain. The construction of only two models, with all iduronates in either one conformation, 1C_4 (Figure 59, model A), or another, 2S_0 (Figure 59, model B), follows the modelling procedure adopted by Mulloy *et al.* (Mulloy *et al.* 1993). The models shown in Figure 59 can then be directly compared to the previous heparin dodecasaccharide models constructed by Mulloy *et al.* (see introduction, section 17).

Figure 58 Potential energy surfaces calculated across the glycosidic linkages indicated, within HS dp6-2.








Contour lines are shown at the global energy minima () marked with an arrow, and at intervals of 3 kJ/mol above this, () global minima +3kJ/mol, () global minima +6kJ/mol, () global minima +9kJ/mol, () global minima +12 kJ/mol, () global minima +15 kJ/mol, () global minima +18 kJ/mol. The potential energy was evaluated using the glycam04 force field within the AMBER molecular modelling suite and partial charge calculation method 1 (see experimental methods).

Figure 58

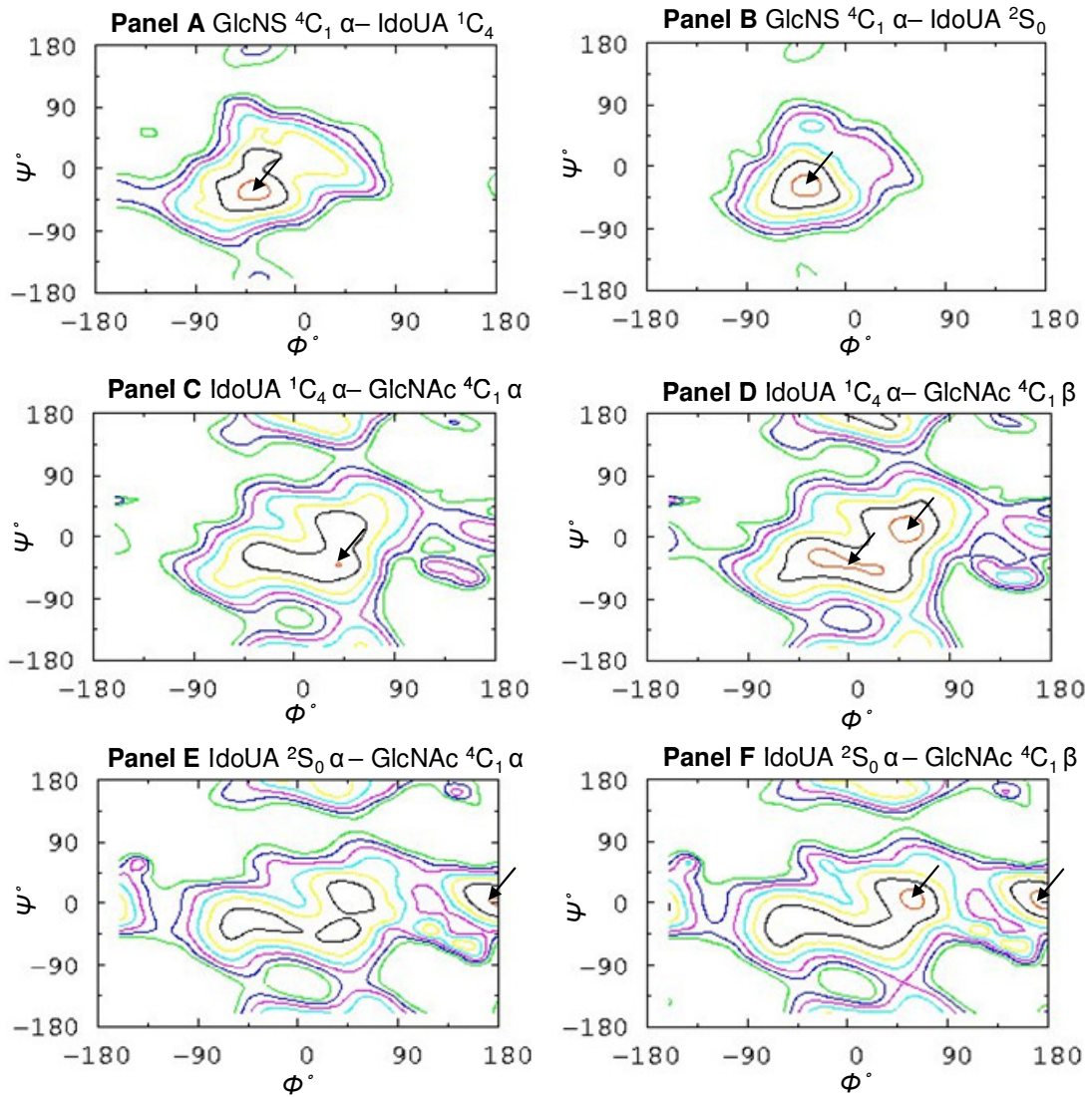
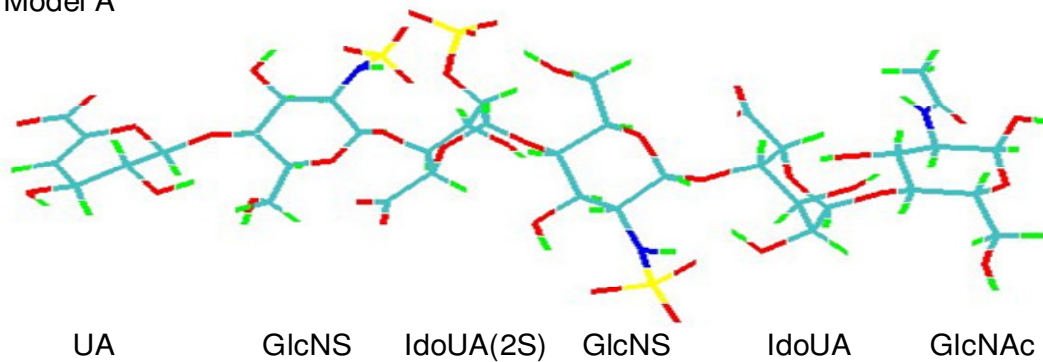


Table 11 Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS **dp6-2**. All selected dihedral angles result in minimal energy geometries across the glycosidic linkage. These angles are also selected on the basis of an NOE derived inter-proton distance restraint of $<3\text{\AA}$, as possibly contributing to the solution conformation of the oligosaccharide. Dihedral angles preferred as a result of the *exo*-anomeric effect are shown in red (see introduction).

Linkage and monosaccharide conformation	Dihedral angles	
	Φ	Ψ
a-b		
UA $^1\text{H}_2$ - GlcNS $^4\text{C}_1$	40	20
	-20	-20
UA $^2\text{H}_1$ - GlcNS $^4\text{C}_1$	40	20
	-40	-20
b-c		
GlcNS $^4\text{C}_1$ - IdoUA(2S) $^1\text{C}_4$	-40	-20
GlcNS $^4\text{C}_1$ - IdoUA(2S) $^2\text{S}_0$	-40	-20
c-d		
IdoUA(2S) $^1\text{C}_4$ - GlcNS $^4\text{C}_1$	40	20
IdoUA(2S) $^2\text{S}_0$ - GlcNS $^4\text{C}_1$	40	20
d-e		
GlcNS $^4\text{C}_1$ - IdoUA $^1\text{C}_4$	-40	-20
GlcNS $^4\text{C}_1$ - IdoUA $^2\text{S}_0$	-40	-20
e-f		
IdoUA $^1\text{C}_4$ - GlcNAc α $^4\text{C}_1$	40	-40
IdoUA $^1\text{C}_4$ - GlcNAc β $^4\text{C}_1$	50	10
	0	40
IdoUA $^2\text{S}_0$ - GlcNAc α $^4\text{C}_1$	50	10
	40	-40
	-40	-20
IdoUA $^2\text{S}_0$ - GlcNAc β $^4\text{C}_1$	50	10

Figure 59

Model A



Model B

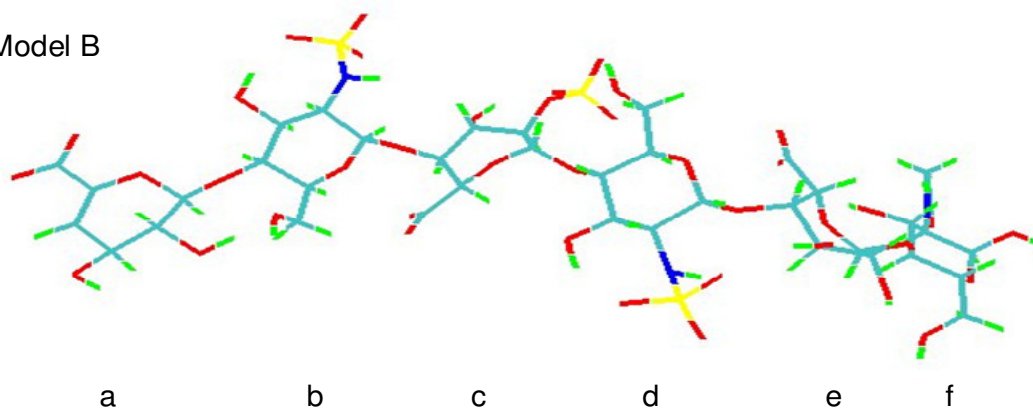


Figure 59 Two final proposed model structure of oligosaccharide HS dp6-2.

The two models differ in the conformation of the IdoUA(2S) monosaccharide at position “c” and in the conformation of the IdoUA monosaccharide at position “e”.

Model A: UA ²H₁-GlcNS ⁴C₁-IdoUA(2S) ¹C₄-GlcNS ⁴C₁-IdoUA ¹C₄-GlcNAcα ⁴C₁

Model B: UA ²H₁-GlcNS ⁴C₁-IdoUA(2S) ²S₀-GlcNS ⁴C₁-IdoUA ²S₀-GlcNAcα ⁴C₁

23.2 Static molecular modelling of HS dp6-5.

Potential energy surfaces were calculated across the UA-GlcNAc(6S), GlcNAc(6S)-GlcUA, GlcUA-GlcNS and IdoUA-GlcNAc (α or β) glycosidic linkages. The potential energy surface across the GlcNS-IdoUA linkage was previously calculated when modelling HS dp6-2. These results are shown in Figure 60. The subsequent table (Table 12) lists potential Φ and Ψ dihedral angles across each glycosidic linkage, supported as possible values by an NOE derived distance restraint of $<3\text{\AA}$ (see the previous introduction to section 25 and discussion).

For construction of the molecular models, shown in Figure 61, where several possible glycosidic geometries were indicated (see Table 12), again those preferred as a result of the *exo*-anomeric effect were modelled. In each model the UA monosaccharide was modelled in the predominant conformation indicated by the coupling constant analysis (see discussion) (${}^2\text{H}_1$). The terminal GlcNAc monosaccharide was modelled in the predominant α anomeric configuration. Two models were constructed one in which the IdoUA monosaccharides is in the ${}^1\text{C}_4$ conformation and one in which it is in the ${}^2\text{S}_0$ conformation (see Figure 61).

Figure 60 Potential energy surfaces calculated across the glycosidic linkages indicated, within HS dp6-5.

Contour lines are shown at the global energy minima (■) marked with an arrow, and at intervals of 3 kJ/mol above this, (■) global minima +3kJ/mol, (■) global minima +6kJ/mol, (■) global minima +9kJ/mol, (■) global minima +12 kJ/mol, (■) global minima +15 kJ/mol, (■) global minima +18 kJ/mol. The potential energy was evaluated using the glycam04 force field within the AMBER molecular modelling suite and partial charge calculation method 1 (see experimental methods).

Figure 60

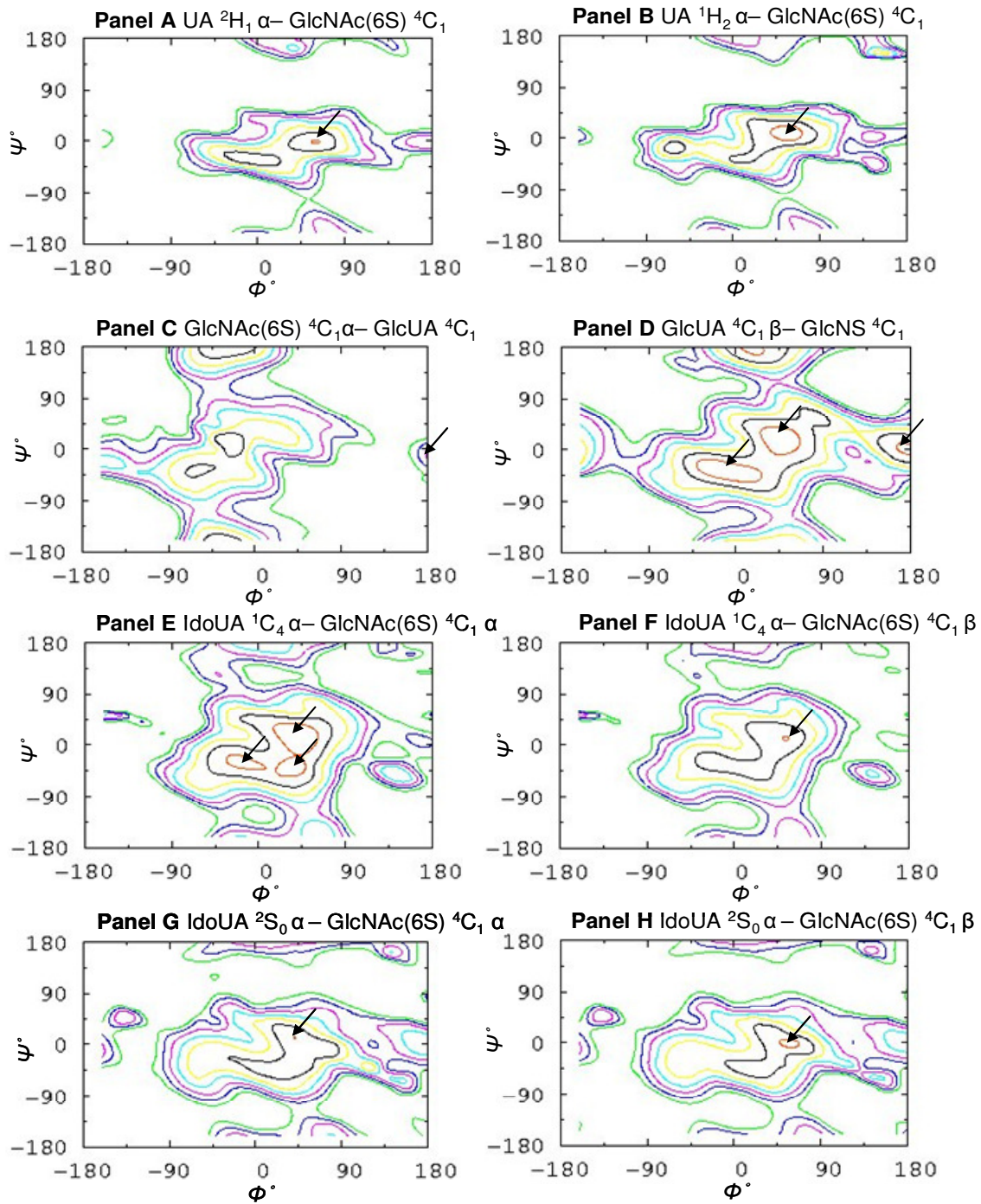


Table 12 Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-5.

All selected dihedral angles result in minimal energy geometries across the glycosidic linkage. These angles are also selected on the basis of an NOE derived inter-proton distance restraint of $<3\text{\AA}$, as possibly contributing to the solution conformation of the oligosaccharide. Dihedral angles preferred as a result of the *exo*-anomeric effect are shown in red (see introduction).

Linkage and monosaccharide conformation	Dihedral angles	
	Φ	Ψ
a-b		
UA $^1\text{H}_2$ - GlcNAc(6S) $^4\text{C}_1$	60	0
UA $^2\text{H}_1$ - GlcNAc(6S) $^4\text{C}_1$	60	0
b-c		
GlcNAc(6S) $^4\text{C}_1$ - GlcUA $^4\text{C}_1$	-60 -20	-40 10
c-d		
GlcUA $^4\text{C}_1$ - GlcNS $^4\text{C}_1$	40 0	20 -40
d-e		
GlcNS $^4\text{C}_1$ - IdoUA $^1\text{C}_4$	-40	-20
GlcNS $^4\text{C}_1$ - IdoUA $^2\text{S}_0$	-40	-20
e-f		
IdoUA $^1\text{C}_4$ - GlcNAc(6S) α $^4\text{C}_1$	40	20
	40	-40
	-20	-20
IdoUA $^1\text{C}_4$ - GlcNAc(6S) β $^4\text{C}_1$	50	10
	40	20
IdoUA $^2\text{S}_0$ - GlcNAc α $^4\text{C}_1$	40	20
IdoUA $^2\text{S}_0$ - GlcNAc β $^4\text{C}_1$	60	0

Figure 61

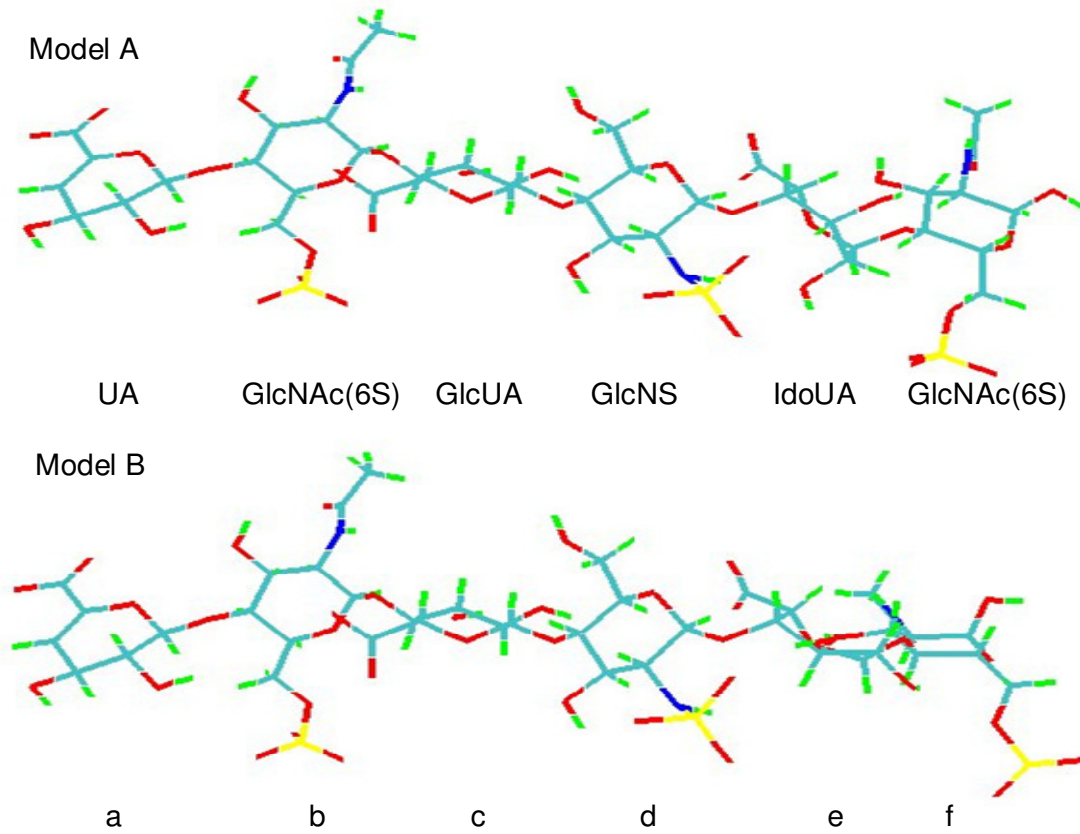


Figure 61 Two final proposed model structure of oligosaccharide HS dp6-5.

The two models differ in the conformation of the IdoUA monosaccharide at position “e”.

Model A: UA ²H₁-GlcNAc(6S) ⁴C₁-GlcUA ⁴C₁-GlcNS ⁴C₁-IdoUA ¹C₄-GlcNAc(6S)α ⁴C₁

Model B: UA ²H₁-GlcNAc(6S) ⁴C₁- GlcUA ⁴C₁ -GlcNS ⁴C₁-IdoUA ²S₀-GlcNAc(6S)α ⁴C₁

23.3 Static molecular modelling of HS dp6-6.

Potential energy surfaces were calculated across the IdoUA(2S)-GlcNS(6S) and GlcNS(6S)-IdoUA linkages. The potential energy surfaces across all other linkages were calculated when modelling HS dp6-2 and dp6-3 (see Figure 49, Panels A-B and Figure 58, Panels C-F). These results are shown in Figure 62. The subsequent table (Table 13) lists potential Φ and Ψ dihedral angles across each glycosidic linkage, supported as possible values by an NOE derived distance restraint of $<3\text{\AA}$ (see the previous introduction to section 25 and discussion).

For construction of the molecular models, shown in Figure 63, where several possible glycosidic geometries were indicated, again those preferred as a result of the *exo*-anomeric effect were modelled. In each model the UA monosaccharide was modelled in the predominant conformation indicated by the coupling constant analysis (see discussion) ($^2\text{H}_1$). The terminal GlcNAc monosaccharide was modelled in the predominant α anomeric configuration. Two models were constructed, one in which the IdoUA(2S) monosaccharide and the IdoUA monosaccharide are in the $^1\text{C}_4$ conformation, and one in which both are in the $^2\text{S}_0$ conformation (see note in section 25.1, on the reasoning behind the construction of only two models).

Figure 62

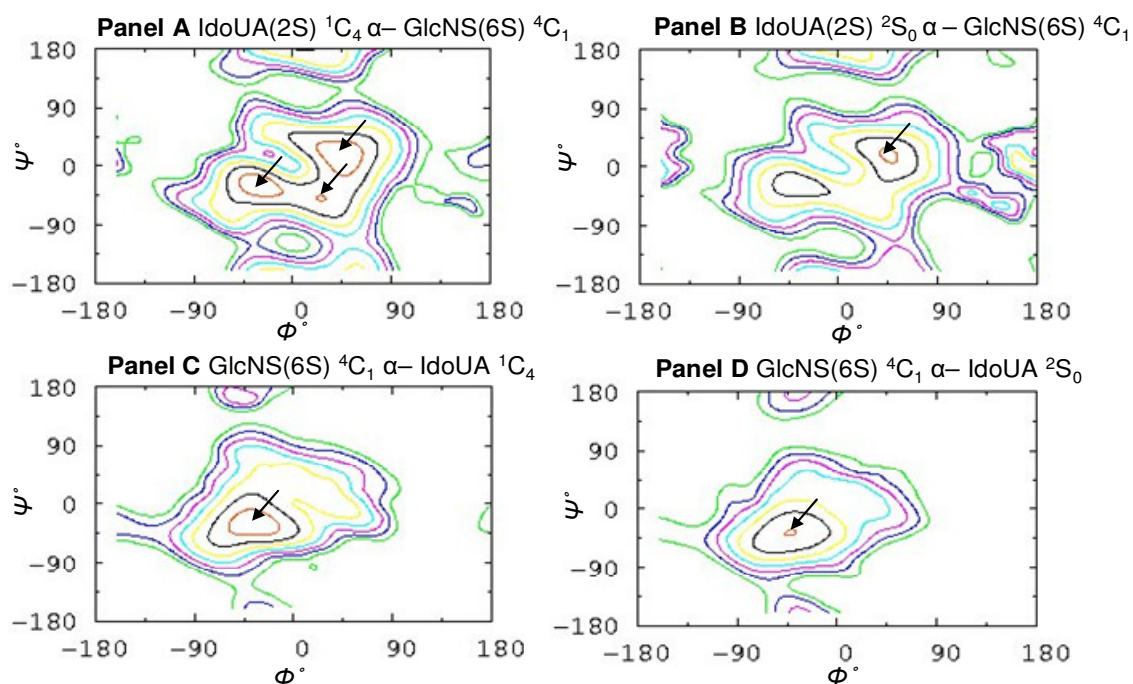


Figure 62 Potential energy surfaces calculated across the glycosidic linkages indicated, within HS dp6-6.

Contour lines are shown at the global energy minima (■) marked with an arrow, and at intervals of 3 kJ/mol above this, (■) global minima +3kJ/mol, (■) global minima +6kJ/mol, (■) global minima +9kJ/mol, (■) global minima +12 kJ/mol, (■) global minima +15 kJ/mol, (■) global minima +18 kJ/mol. The potential energy was evaluated using the glycam04 force field within the AMBER molecular modelling suite and partial charge calculation method 1 (see experimental methods).

Table 13 Selected Φ and Ψ dihedral angles across each glycosidic linkage of HS dp6-6.

All selected dihedral angles result in minimal energy geometries across the glycosidic linkage. These angles are also selected on the basis of an NOE derived inter-proton distance restraint of $<3\text{\AA}$, as possibly contributing to the solution conformation of the oligosaccharide. Dihedral angles preferred as a result of the *exo*-anomeric effect are shown in red (see introduction).

Linkage and monosaccharide conformation	Dihedral angles	
	Φ	Ψ
a-b		
UA $^1\text{H}_2$ - GlcNS $^4\text{C}_1$	40	20
	-20	-20
UA $^2\text{H}_1$ - GlcNS $^4\text{C}_1$	40	20
	-40	-20
b-c		
GlcNS $^4\text{C}_1$ - IdoUA(2S) $^1\text{C}_4$	-40	-20
GlcNS $^4\text{C}_1$ - IdoUA(2S) $^2\text{S}_0$	-40	-20
c-d		
IdoUA(2S) $^1\text{C}_4$ - GlcNS(6S) $^4\text{C}_1$	40	20
IdoUA(2S) $^2\text{S}_0$ - GlcNS(6S) $^4\text{C}_1$	40	20
d-e		
GlcNS(6S) $^4\text{C}_1$ - IdoUA $^1\text{C}_4$	-40	-20
GlcNS(6S) $^4\text{C}_1$ - IdoUA $^2\text{S}_0$	-40	-40
e-f		
IdoUA $^1\text{C}_4$ - GlcNAc α $^4\text{C}_1$	40	-40
IdoUA $^1\text{C}_4$ - GlcNAc β $^4\text{C}_1$	50	10
	0	40
IdoUA $^2\text{S}_0$ - GlcNAc α $^4\text{C}_1$	50	10
	40	-40
	-40	-20
IdoUA $^2\text{S}_0$ - GlcNAc β $^4\text{C}_1$	50	10

Figure 63

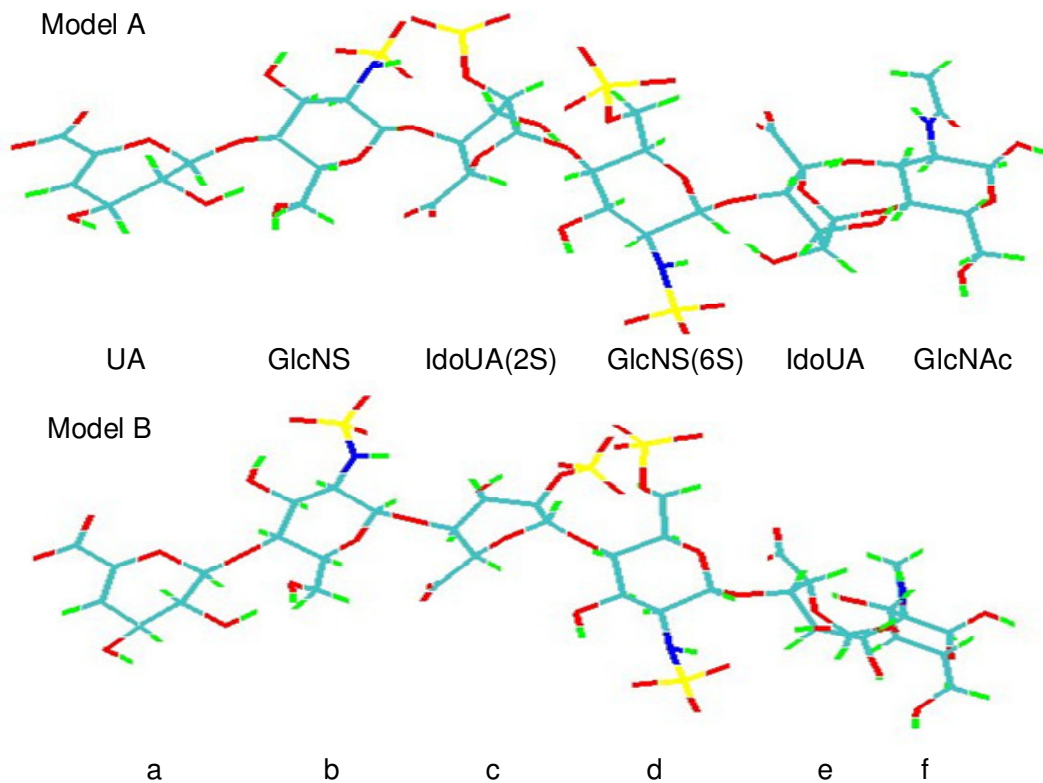


Figure 63 Two final proposed model structures of oligosaccharide HS dp6-6.

The two models differ in the conformation of the IdoUA(2S) monosaccharide at position “c” and the conformation of the IdoUA monosaccharide at position “e”..

Model A: UA ²H₁-GlcNS ⁴C₁-IdoUA(2S) ¹C₄-GlcNS(6S) ⁴C₁-IdoUA¹C₄-GlcNAcα ⁴C₁

Model B: UA ²H₁- GlcNS ⁴C₁-IdoUA(2S) ²S₀-GlcNS(6S) ⁴C₁-IdoUA ²S₀-GlcNAcα ⁴C₁

24. Molecular dynamic modelling

In reality molecules are not static structures and are constantly tumbling and moving in solution. As mentioned in the introduction, NMR derived structures represent an average of many structures that occur in solution. Throughout its lifetime in solution, for short periods of time a molecule may change shape and adopt or 'sample' an alternate conformation away from the average structure seen on its NMR analysis. This is particularly the case with oligosaccharide structures, which are known to be particularly flexible in solution (Imberty and Perez 2000). An idea can be gained as to the extent a modelled structure may deviate from its static structure, through the use of molecular dynamic modelling techniques. It was thought that dynamic simulations may also offer a new insight into any possible structural influences on oligosaccharide dynamics.

In contrast to the construction of potential energy surfaces, molecular dynamic simulations are more reliant on how through space interactions between atoms are modelled (Leach 2001). Partial charge values assigned to each atom play a crucial role in determining these through space interactions. As mentioned in the introduction, partial charges can not be measured experimentally but may be calculated using any number of computational techniques, reviewed in (Leach 2001). Therefore different partial charge sets assigned to the same molecular system, while equally valid, may produce different dynamic results. Using existing glycam04 force field parameters an evaluation of three different partial charge sets was made on HS dp6-2 and -3 structural models.

Variations of the two torsion angles Φ and Ψ were monitored over the course of a 1ns explicit water molecular dynamics simulation (see experimental procedures). For both oligosaccharides, static model A was dynamically modelled. Results for HS dp6-2 are

shown in Figure 64, and results for HS dp6-3 in Figure 65. Simulations on HS dp6-5 and -6 were carried out using solely charge set 1. These results are shown in Figures 66 and 67.

All iduronates and terminal uronic acid residues remained in their starting conformation throughout the course each simulation (data not shown). For iduronate residues this is in contrast to the molecular dynamic results obtained when modelling their monosaccharide structures in isolation (see results section 23). This result lends support to the hypothesis that the movement of C1 and C4 is severely restricted by the addition of flanking monosaccharide units to an otherwise free to move iduronate monosaccharide (Forster and Mulloy 1993). Hindrance of free motion at C1 and C4, severely decreasing the rate at which iduronate ring conformational transitions occur (Forster and Mulloy 1993). Angulo *et al.* (Angulo *et al.* 2003) ran 2ns dynamic simulations on a hexasaccharide structure, double the duration of the 1ns simulations undertaken in this study, and still observed no major iduronate conformational transitions.

24.1 HS dp6-2, -3, -5 and -6 molecular dynamics

With HS dp6-2 and -3, except for one notable exception, see panel C2 Figure 64 and discussion, all non-terminal glycosidic linkages over all charge sets remained close to their starting geometry. Several terminal glycosidic linkages moved to sample alternate geometries that were also suggested by the NMR structural refinement process (see Figures 64 and 65). For example A3 (Figure 64) was initiated at dihedral angles $\Phi = 40$, $\Psi = 20$ but also moved to sample the $\Phi = -40$, $\Psi = -20$ geometry (see Table 11). This involves a re-orientation of the terminal UA residue with respect to the rest of the

molecule; the shape of the rest and majority of the molecule remaining consistent. The results of these dynamic simulations were to some extent dependent on the partial charge model; compare Figure 65 panels 2A and 3A.

It is notable that only when intramolecular hydrogen bonds are dis-allowed in the calculation of partial charges, partial charge model 3, that the UA-GlcNS bond samples an alternate geometry; compare Figures 64 and 65 panels 3A with panels 1A and 2A. This suggests hydrogen bonds play a significant role in the stabilisation of the $\Phi = 40^\circ$, $\Psi = 20^\circ$ geometry. Hydrogen bond strength is effectively lowered by using partial charge calculation method 3. Therefore hydrogen bonding within the UA-GlcNS- disaccharide, appears to have a stronger influence in determining glycosidic linkage geometry than it does for other disaccharide structures. It is hard to determine if this is functionally significant or a modelling artefact. Further studies will need to address this issue.

Also notable, are dynamic differences across the reducing terminal IdoUA-GlcNAc and GlcUA-GlcNAc glycosidic linkages of HS dp6-2 and -3 respectively. For the IdoUA-GlcNAc linkage for all charge models (Figure 64, column E) the geometry of this linkage remains close to its initial value, with Φ at around 40° . In contrast, for the GlcUA-GlcNAc linkage a starting geometry of $\Phi = 60^\circ$, $\Psi = 0^\circ$ and an alternate geometry of $\Phi = 0^\circ$, $\Psi = -20^\circ$ are sampled using partial charge set 1 (Figure 65, panel E1) and partial charge set 3 (Figure 65, panel E3). It may be that the GlcUA-GlcNAc linkage has a higher degree of conformational flexibility over that of a IdoUA-GlcNAc linkage.

However, perhaps the key observation from these dynamic studies is the suggestion that 6-O-sulphate groups are able to exert an influence on the dynamics of adjacent glycosidic linkages. This is suggested by the fact that within HS dp6-6, the terminal IdoUA-GlcNAc

linkage appears to sample two additional conformations to the one in which the dynamic simulation was initiated. The simulation was initiated with IdoUA-GlcNAc glycosidic torsion angles at $\Phi = 40^\circ$, $\Psi = -40^\circ$ (see Table 13) and during the course of the simulation moves to sample a region of the potential energy surface around $\Phi = -20^\circ$, $\Psi = -20^\circ$ and $\Phi = -45^\circ$, $\Psi = -45^\circ$ (see Figure 67, panel E). A similarly modelled IdoUA-GlcNAc within HS dp6-2 (see Figure 64 panel E1) makes no such conformational transitions. Possibly the monosaccharide to which the IdoUA is linked on the opposite side, GlcNS(6S) for HS dp6-6 or GlcNS for HS dp6-2, is the main cause of this (see discussion).

With regard to the use of different partial charge set in the dynamic modelling of HS dp6-2 and HS dp6-3, these results represent in good start in the attempt to develop a new force field for the modelling of HS and heparin oligosaccharides. It was suspected that it may be the case that dynamics may vary as a result of charge model and for a number of glycosidic linkages this is indeed the case. A next step in the development of this force field may be to establish how well the new partial charges reproduce the quantum mechanical properties of a molecule in combination with the existing glycam04 torsion parameters. See discussion section 10, for the differences between quantum and mechanical molecular modelling. Quantum mechanical calculations take an extreme amount of computational time, however these calculations are currently underway both as a continuation of this thesis work and in the laboratory of Dr. Dupradeau, at the university of Amien, France.

Figure 64 Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation of HS dp6-2.

Structures were sampled every picosecond over the course of a 1ns TP5 explicit water molecular dynamic simulation (see experimental procedures). Φ and Ψ torsion angles across each glycosidic linkage for each of the 1000 structures sampled are plotted and connected in chronological order (■). Also plotted are contour lines representing the potential energy surface of each glycosidic linkage indicated (see Figure 49).

Results on line 1 were calculated using partial charge set 1, results on line 2 using partial charge set 2, and results on line 3 using partial charge set 3 (see experimental procedures).

Column A, UA $^2H_1 \alpha$ – GlcNS 4C_1

Column B, GlcNS $^4C_1 \alpha$ – IdoUA(2S) 1C_4

Column C, IdoUA(2S) $^1C_4 \alpha$ – GlcNS 4C_1

Column D, GlcNS $^4C_1 \alpha$ – IdoUA 1C_4

Column E, IdoUA $^1C_4 \alpha$ – GlcNAc $^4C_1 \alpha$

Figure 64

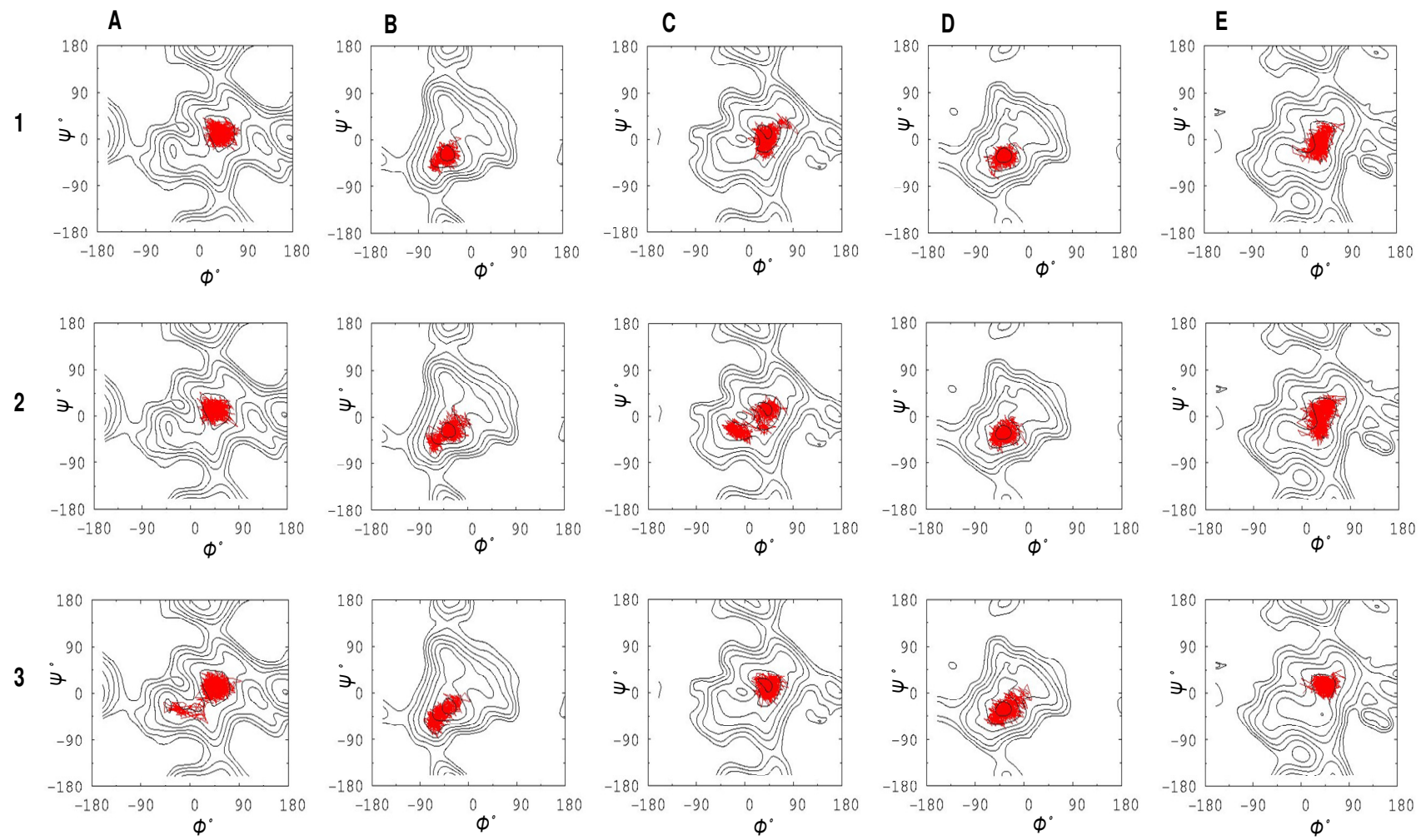


Figure 65 Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation of HS dp6-3.

Structures were sampled every picosecond over the course of a 1ns TP5 explicit water molecular dynamic simulation (see experimental procedures). Φ and Ψ torsion angles across each glycosidic linkage for each of the 1000 structures sampled are plotted and connected in chronological order (■). Also plotted are contour lines representing the potential energy surface of each glycosidic linkage indicated (see Figures 49 and 58).

Results on line 1 were calculated using partial charge set 1, results on line 2 using partial charge set 2, and results on line 3 using partial charge set 3 (see experimental procedures).

Column A, UA $^2H_1 \alpha$ – GlcNS 4C_1

Column B, GlcNS $^4C_1 \alpha$ – IdoUA(2S) 1C_4

Column C, IdoUA(2S) $^1C_4 \alpha$ – GlcNS 4C_1

Column D, GlcNS $^4C_1 \alpha$ – GlcUA 1C_4

Column E, GlcUA $^1C_4 \alpha$ – GlcNAc $^4C_1 \alpha$

Figure 65

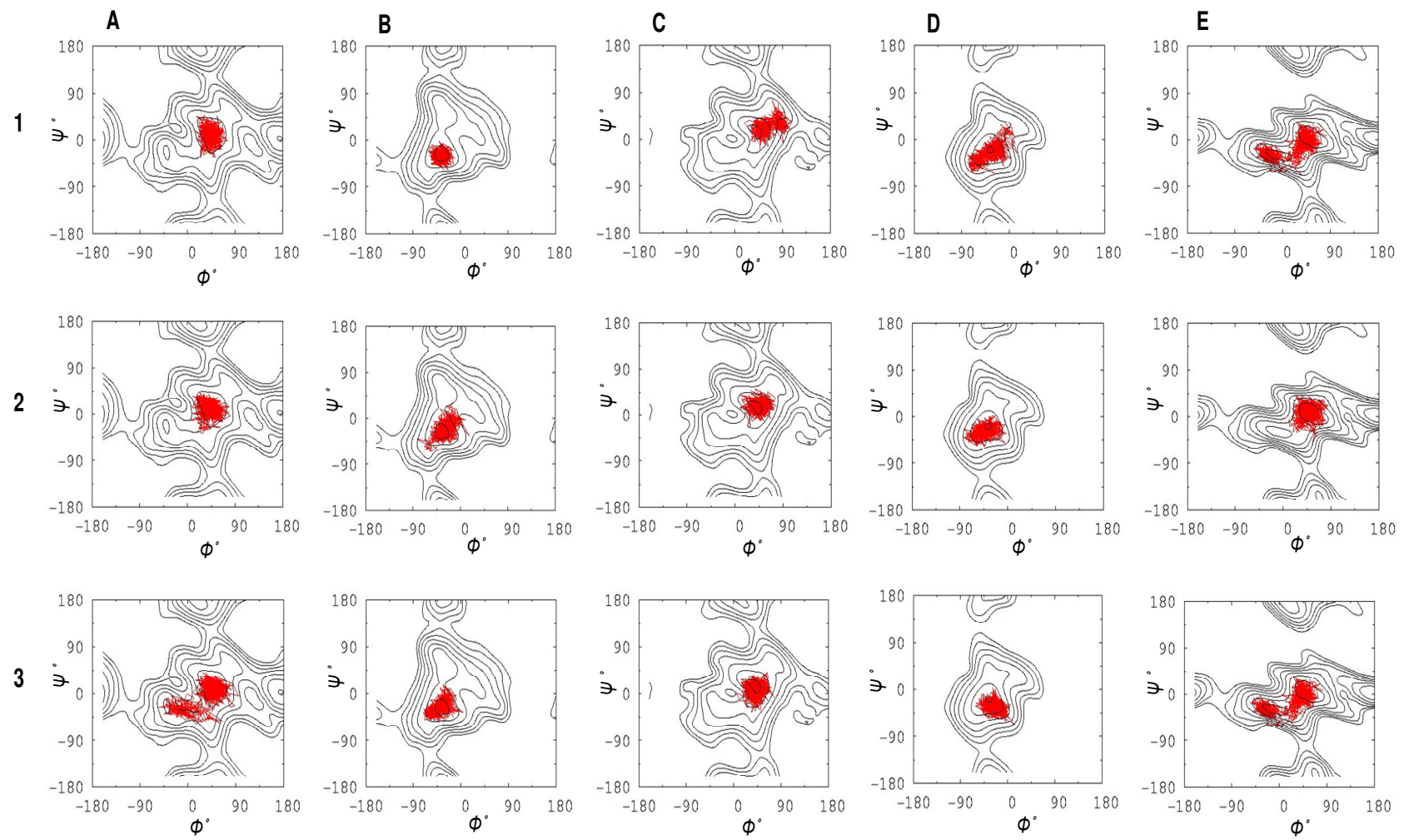


Figure 66

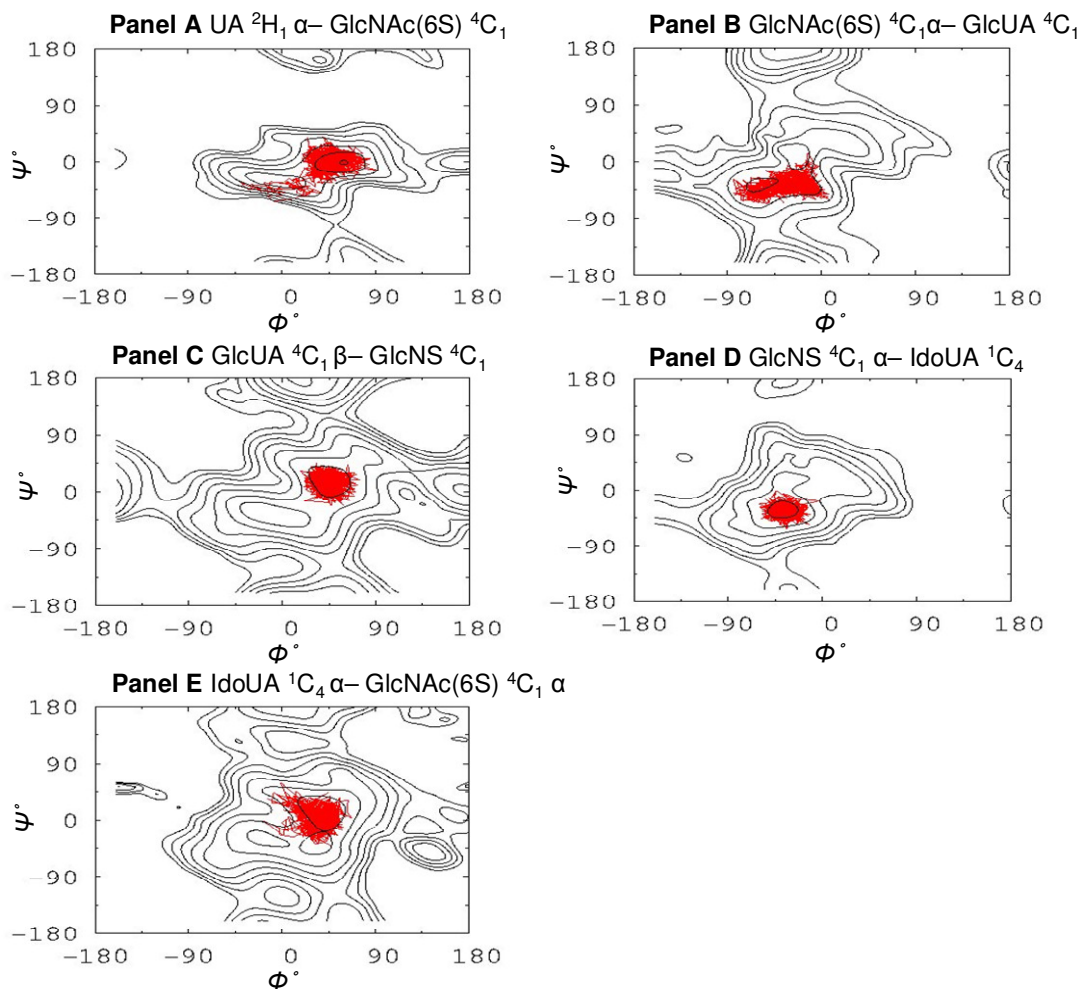


Figure 66 Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation, HS dp6-5, model A.

Structures were sampled every picosecond over the course of a 1ns TP5 explicit water molecular dynamic simulation (see experimental procedures). Φ and Ψ torsion angles across each glycosidic linkage for each of the 1000 structures sampled are plotted and connected in chronological order (■). Also plotted are contour lines representing the potential energy surface of each glycosidic linkage indicated (see Figures 58 and 60)

Figure 67

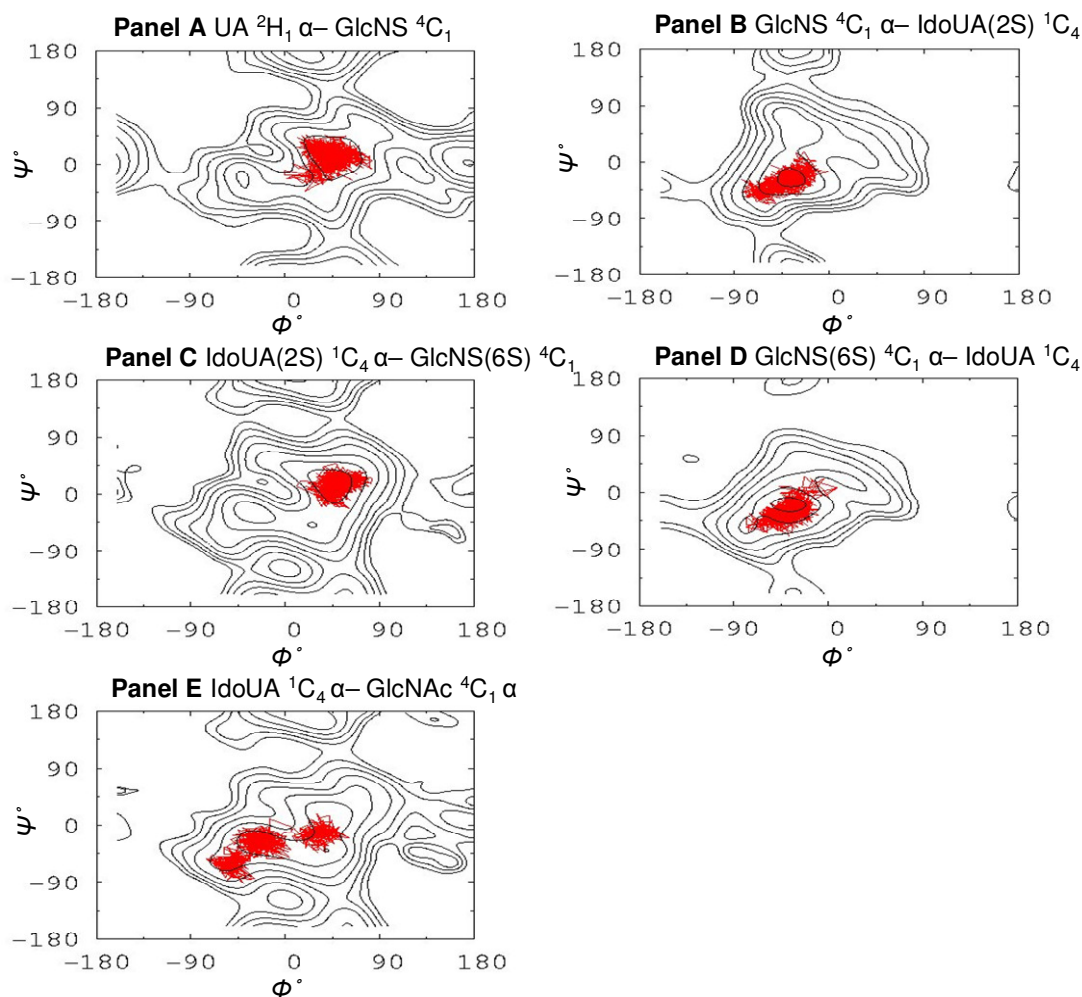


Figure 67 Variations in Φ and Ψ torsion angles over the course of a 1ns explicit water molecular dynamic simulation, HS dp6-6, model A.

Structures were sampled every picosecond over the course of a 1ns TP5 explicit water molecular dynamic simulation (see experimental procedures). Φ and Ψ torsion angles across each glycosidic linkage for each of the 1000 structures sampled are plotted and connected in chronological order (■). Also plotted are contour lines representing the potential energy surface of each glycosidic linkage indicated (see Figures 49, 58 and 62).

DISCUSSION

1. Initial studies into the degradation of HS by the heparinase III enzyme.

Many previous studies have used the heparinase III enzyme to degrade polymeric HS into shorter chain length oligosaccharides for structural and/or functional analysis (see introduction, NMR section). Frequently, the minimum quantity of enzyme required to bring about complete digestion of a known quantity of HS is not addressed.

Heparinase III cleavage of HS is through a β -eliminative mechanism, and generates oligosaccharides terminating at their non-reducing end in a $\Delta^4,5$ unsaturated uronic acid (see introduction). This absorbs U.V. light at a wavelength of 232nm and allows the progression of a heparinase III digest to be monitored. Typically, digestion is assumed to be complete when no further increase in absorbance at 232nm is detected (Pervin *et al.* 1995; Hileman *et al.* 1997). Due to the long incubation times used in the initial studies detailed in this study (see results section 1) spectrometer readings were subject to severe drift (data not shown). Therefore the progress of digestion was monitored through changes in the gel filtration profile of the sample.

A quantity of 40mU of heparinase III added in 10mU aliquots over the course of 88 hours was found sufficient to digest 200 mg of HS (100mg/ml) to completion (see Figure 20, panel E)

2. SAX-HPLC chromatography of sized HS oligosaccharides

SAX-HPLC analysis of the dp4 and dp6 oligosaccharide pools showed them to be considerably less heterogeneous in nature than the Heparinase III generated pools

analysed by Hileman *et al.* (Hileman *et al.* 1997) (see Figure 23). This could be due to three reasons:

- 1) Possible variations in HS start material, although porcine mucosal HS purchased from Celsus laboratories was used in this and the Hileman *et al.* study;
- 2) Improved gel filtration chromatography. There is a high probability that the samples analysed by Hileman *et al.* contain contaminating oligosaccharides of both a shorter and longer chain length. The presence of these contaminants in addition to the authentic chain length oligosaccharides increase sample heterogeneity;
- 3) The extent of enzyme digestion. Great care was taken in this thesis work to ensure all heparinase III enzyme digests were complete. A partial digest may result in a considerable increase in sample heterogeneity and more complicated SAX-HPLC profiles.

3. SAX-HPLC and PAGE analysis

It has been shown that the detection of a single peak on SAX-HPLC does not always correspond to the presence of a single oligosaccharide species (Vives *et al.* 2001; Goodger 2003). Likewise a split peak of the types labelled 1, 2 and 3 in Figure 23, Panel B, does not necessarily correspond to the presence of 2 oligosaccharide species. K5 polysaccharide derived oligosaccharides that are composed entirely of -GlcUA-GlcNAc-disaccharide units elute as similar split peaks on SAX-HPLC (Murphy 2003). For these reasons PAGE analysis was carried out on a number of dp6 and dp8 SAX-HPLC peaks.

A smear, rather than a number of clearly resolved bands was observed on PAGE analysis of dp6-1 (see Figure 24). It is thought that this peak may be composed of saccharide units

which remain covalently linked to a protein core amino acid sequence (see introduction, section 6.2, page 47, chain initiation).

The presence of two alternately charged oligosaccharides in dp6-2 and dp-3, suggested by their split SAX-HPLC peaks, was not supported by the PAGE analysis. Both stained for a single oligosaccharide band. Dp6-5 and dp6-6 also stained for a single band, however these peaks were not split on HPLC analysis.

On the basis of the dp6-5 and -6 results it might have been assumed that HS dp6-4 also contained a single oligosaccharide species. Likewise on the basis of the HS dp6-2 and -3 results it might have been assumed that dp6-7, a split peak, also contains one oligosaccharide species. However, at least two major species were detected in both of these samples. This demonstrates that each individual SAX-HPLC peaks must be assessed independently and that for a number of oligosaccharides, chromatographic separation is not achievable on the basis of charge by SAX-HPLC alone. This also applies to a number of dp8 oligosaccharides (see Figures. 23 and 24).

Split peaks of the type observed for dp6-2 and dp6-3, may be due to the presence of both α - and β -anomeric configurations for the reducing end terminal unit. With dp6-2 and -3 both α - and β -anomeric configurations were detected on NMR analysis for their terminal N-acetyl glucosamine residues (see Figure 33). It may be that one configuration (α or β) results in a slight increase in the affinity of the oligosaccharide for the HPLC stationary phase, possibly brought about by differences in the predominant rotamer conformations of GlcNAc α and GlcNAc β (discussed later).

4. PAGE and disaccharide analysis

The disaccharide analysis of HS dp6-2, -3, -5 and -6 supported the evidence from PAGE analysis (see Figure 24) that each is mainly composed of a single hexasaccharide structure (see Table 2). The disaccharide analysis of the dp8 subfractions -1, -3 and -7 however, did not support their PAGE analysis. Each was found to contain varying proportions of six different disaccharide units (see Table 2).

The probable reason for this discrepancy may be due to the quantity of dp8 subfractions material analysed by PAGE. In addition to the major oligosaccharide species stained, the disaccharide analysis results suggest these samples to also contain multiple additional oligosaccharides at low concentrations. In support of this, on overloading the lanes with dp8 material, a number of minor bands were also stained (data not shown). On the basis of the dp8 disaccharide and PAGE analyses, dp8 subfractions were not taken on for further NMR characterisation and remain stored for future study.

5. Variation between SAX-HPLC profiles

The reasons behind the observation of varying proportions of dp6-5 and -6 oligosaccharides in pooled dp6 material from the preparative size-exclusion column are unclear (see figure. 25). All digests were carried out on the same batch of HS purchased from Celsus laboratories, prior to addition to the size exclusion column. Several different batches of heparinase III enzyme were used, and initially it was thought that this may be the reason for variations. Slight variations did occur occasionally between batches, but frequently greater variations occurred when using the same batch of enzyme.

One possible explanation may be that for any one digest a particular 'selection' of polymeric heparan sulphate chains are present within the 200mg of material taken for digestion. Structural variations within this selection may then have an effect on the proportions of the various oligosaccharide sequences that result. In support of this, variations were also seen in the levels of dp6-2 and -3 however the most dramatic differences were observed with dp6-5 and -6.

An alternative explanation may be that subtle variations exist in the digestion rate of some disaccharide linkages. While a majority of susceptible disaccharide linkages have been cleaved by 40 mU of hep III over an 88 h period, a variable number of some resistant, but cleavable linkages may remain intact. It is possible that one of these linkages is present within HS dp6-5, leading to variations in the proportions of this oligosaccharide detected on apparent digestion completion. This may well be the result of the unusual –GlcNAc(6S)–GlcUA-GlcNS- trisaccharide sequence at the centre of HS dp6-5 (discussed later).

6. PAGE purification of dp6-5 and dp6-6

The further purification of dp6-5 and -6 by PAGE, resulted in significant sample losses. Approximately half the original material was lost on further purification. This was mainly due to the number of chromatographic steps that were then necessary to remove contaminating Azure A stain and polyacrylamide from these samples (see experimental procedures).

An alternative to Azure A staining may be the reverse staining technique as detailed by Hardy and Castellanos-Serra (Hardy and Castellanos-Serra 2004). A preliminary

evaluation of this technique towards the end of this study demonstrated that it could be used to clearly stain heparin derived oligosaccharides (data not shown). Bands could then be extracted from the gel without the presence of the difficult to remove Azure A organic dye.

The finding that polyacrylamide was present in these samples was unexpected as this was assumed to become efficiently chemically cross linked and insoluble on gel formation. Polyacrylamide itself is not thought to be toxic (King and Noss 1989), however the acrylamide monomer is a potent neurotoxin (Kupperman 1958; McCollister *et al.* 1964) and probable human carcinogen (Bull *et al.* 1984; Johnson *et al.* 1986). Vives *et al.* (Vives *et al.* 2001), reported that contaminants were a problem when gel-extracted oligosaccharides were tested for biological activity. Residual acrylamide monomer is likely an impurity in most polyacrylamide preparations, ranging from <1 ppm to 600 ppm (King and Noss 1989). Its removal by the HPLC method described by Vives *et al.*, or by the DEAE method described in this thesis, is crucial if gel purified oligosaccharides are to be tested for biological activity (Goodger 2003).

7. ^{13}C DEPT-135 and ^{13}C HSQC experiments.

Following assignment of a majority of the protons within all oligosaccharides, using COSY, TOCSY and NOESY spectra, additional ^{13}C DEPT-135 and ^{13}C HSQC spectra were recorded (where sample quantities permitted). A large amount of material is required for ^{13}C NMR studies (>1mg) due to the low natural presence of the ^{13}C nuclei. Spectra were recorded for HS dp6-2 and -3 and hep dp6-1 and -2 (see Appendix II, Figures 103-106). This was done because using COSY and TOCSY data, H6 glucosamine protons

proved particularly difficult to assign. Frequently very weak or no cross-peaks could be detected to these protons on TOCSY analysis. This was thought to be due to a splitting of the signal intensity between each of the two H6 protons, effectively halving each of their signal intensities.

A DEPT-135 pulse sequence results in a characteristic pattern of signal intensities for ^{13}C coupled to one, two or three hydrogens, CH, CH_2 or CH_3 . The CH_2 bonded carbons have negative signal intensity, while CH and CH_3 carbons have positive signal intensity (Becker 2000). Simply, HSQC spectra reveal which protons are coupled to which ^{13}C atoms through the occurrence of ^{13}C - ^1H cross-peaks (Becker 2000). The ^{13}C at glucosamine carbon position 6 is a CH_2 bonded carbon and has a negative DEPT-135 signal intensity. This negative intensity can easily be identified against the background of predominantly positive signals in all DEPT-135 recorded spectra (see appendix II). The large number of positive signals are the result of the large number of CH bonded carbons present in each oligosaccharide. On identifying signals arising from C6 carbons, HSQC spectra were then used to establish where the chemical shifts occurred for the two H6 protons.

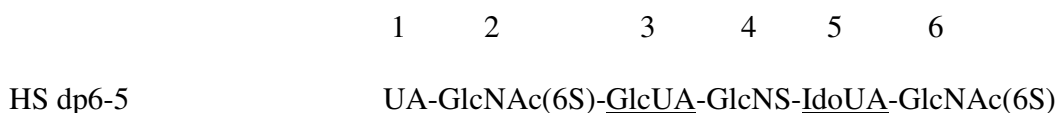
Where sample quantities did not permit the recording of ^{13}C spectra, a lack of TOCSY cross-peaks to a particular proton was assumed to indicate a H6 proton. This assumption was always confirmed by COSY analysis. H6-H5 COSY cross-peaks did not appear to suffer from the same loss in signal intensity as TOCSY cross-peaks (see Figure 69, panel A, in a later section of this discussion). Frequently it was possible to assign ring systems from H1-H5 using TOCSY and COSY spectra. Additional COSY cross-peaks from H5, i.e. not to the H4 resonance, were assigned to H5-H6.

8. Evidence for the conservation of HS monosaccharide sequences across species barriers

The three monosaccharide sequences, HS dp6-2, -3 and -6, have previously been isolated and sequenced from the cell surface of a cultured murine 3T3 fibroblast cell line (Merry *et al.* 1999). HS dp6-3 and -6 have also previously been isolated and sequenced from an alternate source of porcine HS (Vives *et al.* 1999). In the field of protein structure and functional studies, amino acid sequences are often found to be conserved across species and form domains within larger protein structures (Whitford 2005). These domains often have a conserved evolutionary function. It seems remarkably co-incidental that 3 out of the 4 HS oligosaccharides characterised in this study, and two from the study of Vives *et al.* (Vives *et al.* 1999) should be homologous with oligosaccharides isolated from a murine cell line (Merry *et al.* 1999). It is speculated that in common with amino acid sequence conservation, conservation of HS monosaccharide sequences may also occur across species barriers, to confer conservation of function.

9. HS dp6-5, a novel monosaccharide sequence

HS dp6-5 represents a unique unpublished monosaccharide sequence. However, it is very similar in sequence to a hexasaccharide previously characterised by Hileman *et al.* (Hileman *et al.* 1997) (see below).



The two sequences differ only in the order in which the internal GlcUA and IdoUA monosaccharides (underlined) occur within the sequence. For HS dp6-5, correctly establishing this order is key to marking this oligosaccharide out as containing a new, novel monosaccharide sequence. The NMR evidence for the positioning of GlcUA at position 3 and IdoUA at position 5 is clear and un-ambiguous.

A characteristic pattern of NOE cross-peaks within the NOESY spectrum of HS dp6-5 identifies the presence of a -GlcNS-IdoUA- disaccharide linkage (see Figure 68). Analyses of the 2D-COSY and TOCSY spectra recorded for HS dp6-5 clearly establish that the non-reducing terminal UA unit is linked to a GlcNAc(6S) monosaccharide (see Figure 69). This places GlcNAc(6S) at position 2. Consequently, the only monosaccharide positions available for which a -GlcNS-IdoUA- disaccharide sequence can occur are at positions 4 and 5. The analyses presented in Figures 68 and 69 confirm the results of a full sequential assignment of this saccharide, undertaken as detailed in the results section, for HS dp6-2 and -3.

Figure 68

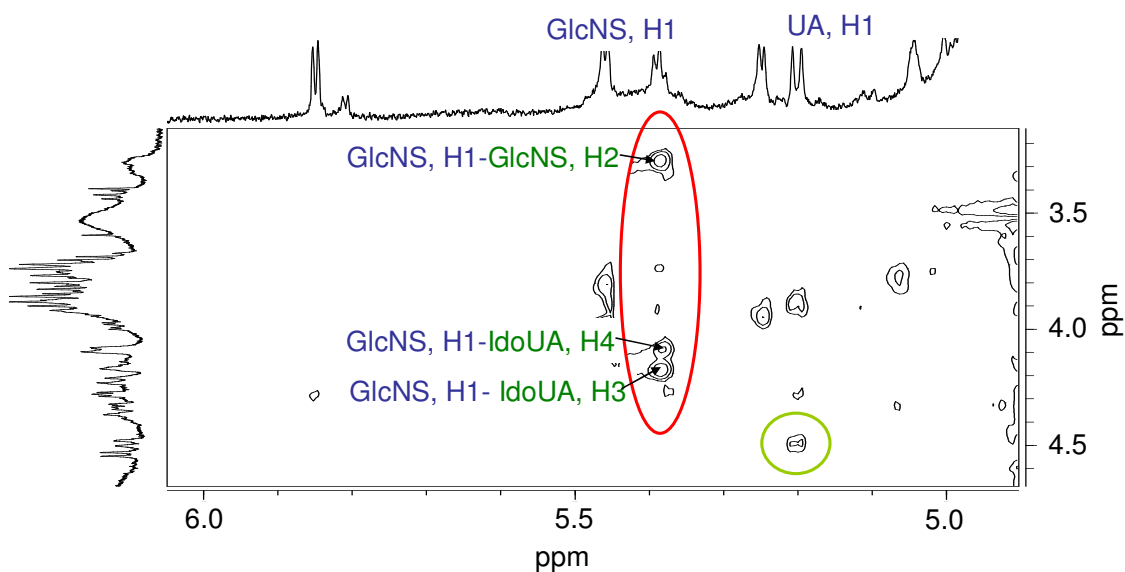


Figure 68 An enlargement of a section from a 2D-NOESY spectrum (500ms mixing time) recorded for HS dp6-5.

Circled in red are a pattern of three NOE cross-peaks that indicate the presence of the disaccharide sequence –GlcNS-IdoUA-. This same pattern of cross-peaks was detected in each of HS dp6-1 and -2 each containing a –GlcNS-IdoUA(2S)- disaccharide sequence.

Also circled in green is an NOE cross-peak between H1 of the terminal uronic acid unit and a proton in the preceding ring system, speculatively assigned as H6. This preceding ring system on this initial analysis may be either a GlcNAc(6S) monosaccharide or GlcNS monosaccharide unit.

Figure 69 Enlargements of sections from 2D- COSY and TOCSY spectra recorded for HS dp6-5.

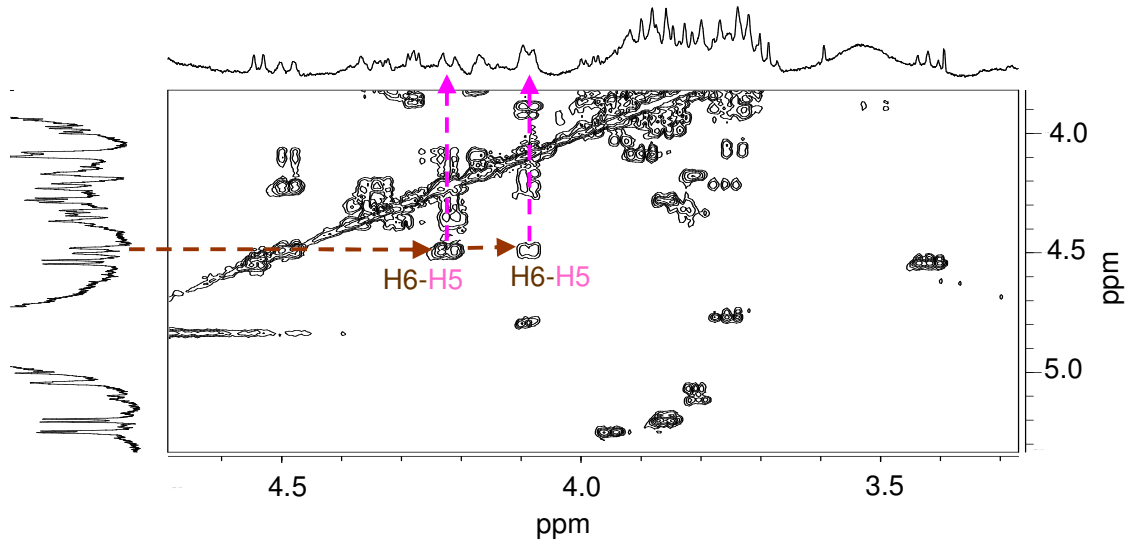
In the TOCSY spectrum no through bond connectivity's could be detected to the proton(s) speculatively assigned as H6, (see discussion main text, on H6 assignments).

In the COSY spectrum (panel A) two through bond connectivities were detected to protons at 4.219 ppm and 4.098 ppm, both were assigned as H5. In the TOCSY spectrum (panel B) through bond connectivity's could be detected to two anomeric protons at 5.459 ppm and 5.248 ppm.

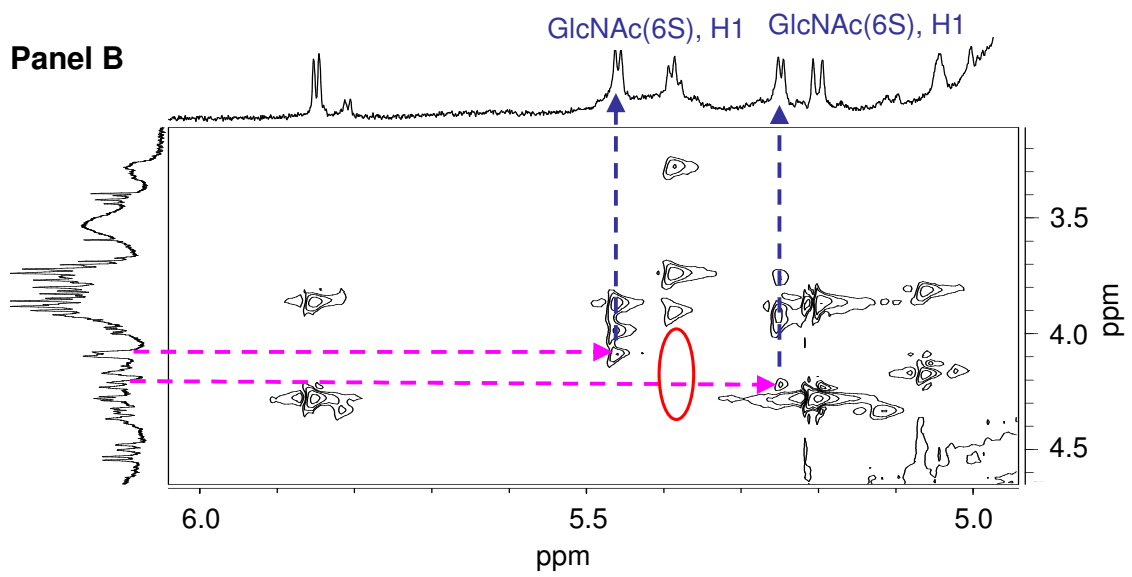
As highlighted by the circled area in panel B, no cross-peaks were detected to the previously assigned GlcNS anomeric proton (see Figure 68). This clearly established positions 1 and 2 of this oligosaccharide sequence to be UA and GlcNAc(6S) respectively.

Figure 69

Panel A



Panel B



10. Molecular modelling GlcNAc β , gg to gt transitions

All low energy geometries calculated for the GlcNAc β anomeric configuration contained the gt rotamer of GlcNAc β . This was despite its initiation in the gg conformation (see Figure 56). No other N-glucoamine monosaccharides made such transitions and remained in the gg rotamer state throughout the energy minimisation process. To investigate this, new quantum mechanical modelling calculations were carried out. Quantum mechanical techniques explicitly take into account the movement of electrons during calculations. Molecular modelling techniques e.g. the glycam04 forcefield, assume that adjustments in the movement of electrons in response to changes in the positioning of the atomic nuclei, are so rapid that they can be discounted from any modelling calculations. Geometry optimisation at the HF/6-31G* level of theory was performed on both GlcNAc β and GlcNAc α structures using the GAMESS quantum mechanical modelling program (data not shown). Simply, HF/6-31G* is quantum mechanical shorthand to describe how the energy of the system was calculated and is known as a 'basis set'. For its contribution to this discussion it is not thought necessary to unravel the complicated nomenclature of quantum mechanical basis sets. For such an explanation see (Leach 2001). HF/6-31G* is a well known and respected basis set with which to calculate quantum mechanical properties (Leach 2001). As with the molecular mechanical modelling calculations GlcNAc β made a gg to gt transition on quantum mechanical structure optimisation (equivalent to molecular mechanical energy minimisation). Similarly geometry optimised GlcNAc α remained in the gg rotamer state. Geometry optimisation using quantum mechanical modelling techniques ruled out the gg to gt transition as simply being a molecular mechanical partial charge effect.

It is possible that a gg to gt transition (see introduction, Figure 1 for a diagram of gg and gt conformational states) is the consequence of going against the endo-anomeric effect (see introduction). The β -anomeric configuration of GlcNAc (see introduction, Figure 19 for diagram) may be more energetically favoured by the gt rotamer state. Similar transitions have not been reported however for glucose monosaccharides. Suggesting the gg to gt transition may somehow involve the N-acetyl group.

It was noticeable that for all molecular and quantum mechanical energy minimised GlcNAc β structures, a hydrogen bond was present between the N-acetyl group oxygen and the hydroxyl group proton attached to carbon position 1. This was deduced on the basis of the distance and the angle between these two atoms upon completion of the modelling calculations (a distance of less than 2Å and an angle of less than 60°). Due to the change in C1 geometry, in the α anomeric configuration, formation of such a hydrogen bond is not possible. How or why this then results in a rotation from gg to gt around the C5-C6 bond is not clear. Computational predictions that the GlcNAc β anomer has a preference to exist in the gt rotamer state are an interesting apparently un-reported phenomenon. The possible biological implications of this discovery now remain for future study.

11. Derivation of molecular modelling inter-proton distance restraints

In theory, a full relaxation matrix analysis of NOE data can be used in order to account for multispin effects and derive more accurate interproton distances. Such an analysis was employed by Mulloy *et al.* through use of their NOEMOL program (Mulloy *et al.* 1993)

and by Mikhailov *et al.* using the MARDIGRAS program (Mikhailov *et al.* 1996; Mikhailov *et al.* 1997).

The severe lag phase observed in the build up curves of all NOEs analysed (see Figure 48) indicates that spin diffusion from other protons, additional to the two directly involved in the NOE, contributes significantly to the intensity observed. Spin diffusion is the result of NOE enhancements being transmitted along a chain of spin systems. The pyranose ring of each monosaccharide unit is just such a chain of spin systems. Indirect intensity enhancements as a result of spin diffusion, in-fact dominate all cross-relaxation pathways (see Figure 48). It has been shown that in cases such as this a full relaxation matrix analysis of the data, does not yield any more accurate inter-proton distances than a simple two-spin approximation analysis (Clore and Gronenborn 1989). Hence a full relaxation matrix analysis of the data was not undertaken. A simple two spin approximation analysis was used in the derivation of inter-proton distance restraints for HS dp6-3.

The NOE build up curves shown in Figure 48 are in marked contrast to the near perfect build up curves published by Mikhailov *et al.* (Mikhailov *et al.* 1996; Mikhailov *et al.* 1997). Possibly the increased sulphation level of the heparin derived hexasaccharides analysed by Mikhailov *et al.* reduces the effect of spin diffusion. However no evidence could be found for this in the literature and why this should be the case is unclear. The NOE kinetics observed by Mikhailov *et al.* are unusual, in that no evidence for the phenomenon of spin diffusion can be seen at all. Even at the shortest mixing time of 100ms used by them these data points form part of a text book perfect NOE build up curve. Given that all protons are part of a pyranose ring system, at least partial evidence

might be expected for the existence of spin diffusion. Of course, it is entirely possible that a lag phase is only observed in the NOE build up kinetics of heparin derived oligosaccharides at mixing times of less than 100ms. Again why this difference should occur between HS and heparin derived NOE kinetics is unclear. NOE build up curves were not produced for the heparin derived oligosaccharides produced as part of this thesis work, mainly because of the NMR instrument time involved.

Practically, construction of a NOE build up curves takes a considerable amount of NMR instrument time. Yamada *et al.* have previously stated that significant NOE connectivities were only observed over distances of less than 3Å within a range of 8 HS/heparin and chondroitin sulphate derived disaccharides (Yamada *et al.* 1992). Although how they arrived at this conclusion is not stated. An analysis of the molecular models constructed for HS dp6-3 confirms this is also the case with this hexasaccharide. On this basis a maximum inter-proton distance of 3Å was used to evaluate all low energy geometries in the construction of models for HS dp6-2, -5 and -6. NOE build up curves for these saccharides were not constructed.

Given the broad low energy regions of each potential energy surface, no advantage in the refinement process is gained using a 'build up rate' derived inter-proton distance over a value of 3Å. By 'broad low energy region' it is meant that the area bounded by low energy contour lines on each potential energy surface are not restricted to only a very tight range of Φ and Ψ values. The opposite, and an example of a narrow set of low energy regions would be seen if a potential energy surface were calculated for a individual H-C-C-H torsion angle in ethane. Three different very narrow ranges of Φ and Ψ values would be low energy, conforming to the three possible staggered conformations

of ethane. In all cases for HS dp6-3 refinement using the 'build up rate' derived distance and a distance of 3Å placed glycosidic linkages in the same low energy geometries. No 3Å geometries could be discounted on the use of the more accurate distance.

12. SAX-HPLC elution position, modelling insights and biological implications

Given their high negative charge density, electrostatic interactions between the positively charged stationary phase of the HPLC column and N- and O- sulphate groups of an oligosaccharide will be primarily responsible for determining its binding affinity. Secondary to this, electrostatic interactions between the less electro-negative carboxyl and acetyl groups may also play a role. It can also be imagined that interactions with the stationary phase of the HPLC column may to some extent also mimic binding interactions with the positively charged HS binding site of many protein ligands.

For the forthcoming sections of this discussion, the static models produced for each HS oligosaccharide in results sections 24-25 have been reformatted using the VMD molecular graphics viewer (Humphrey *et al.* 1996). Reformatting was done in order to emphasize the positioning of electronegative groups within each HS oligosaccharide. By comparing these re-formatted models, a number of new insights can be gained into how the positioning of electronegative groups may influence SAX-HPLC elution position. These insights, and biological implications where appropriate, are discussed in the forthcoming sections.

13. HS dp6-2

Reformatted models for HS dp6-2 (UA-GlcNS-IdoUA(2S)-GlcNS-IdoUA-GlcNAc) are shown in Figures 70 to 72. For Figures 70 and 71, to produce model E, model D was re-orientated so that the maximum number of electronegative groups are positioned along the visible surface of the molecule. It is probable that it is this surface that interacts with the stationary-phase of the HPLC column and HS binding sites at the surface of a protein ligand. In this orientation a total of 5 groups may contribute to a binding affinity for the stationary phase (labelled 1-5 in Figures 70-72).

The superimposition model (Figure 72) reveals that the positioning of electronegative groups 4 and 5 are largely unaffected by iduronate conformational changes. However, the positioning of electronegative groups 1 to 3 are re-organised. Similar re-orientations were reported by Mulloy *et al.* (Mulloy *et al.* 1993).

Figure 70 HS dp6-2 molecular models.

All internal iduronates are modelled in the 1C_4 conformation.

Model A, whole model, line representation,

Model B, whole model, solid surface representation,

Model C, whole model, wireframe surface representation,

Model D, as model C with N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (red);

Model E, a re-orientation of model D.

X, Y, Z, relative coordinate axis are shown for models B-D

A number of electronegative groups are numbered 1-5 in model E

1. UA, carboxyl group
2. GlcNS, N-sulphate group
3. IdoUA(2S), 2-O-sulphate group
4. IdoUA, carboxyl group
5. GlcNAc, N-acetate group

All models were formatted using the VMD molecular graphics program (Humphrey *et al.* 1996).

Figure 70

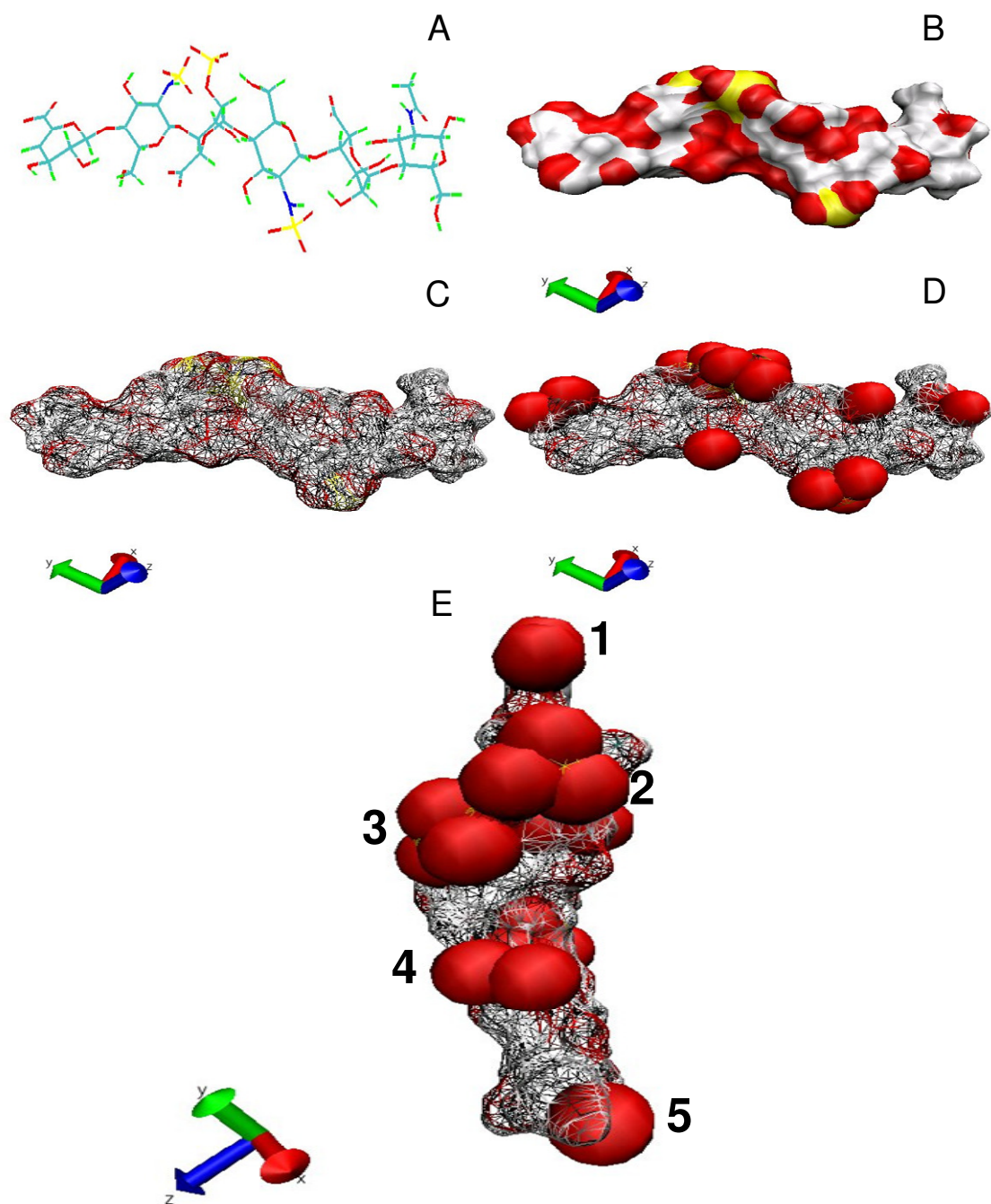


Figure 71 HS dp6-2 molecular models.

All internal iduronates are modelled in the 2S_0 conformation.

Model A, whole model, line representation,

Model B, whole model, solid surface representation,

Model C, whole model, wireframe surface representation,

Model D, as model C with N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (yellow);

Model E, a re-orientation of model D.

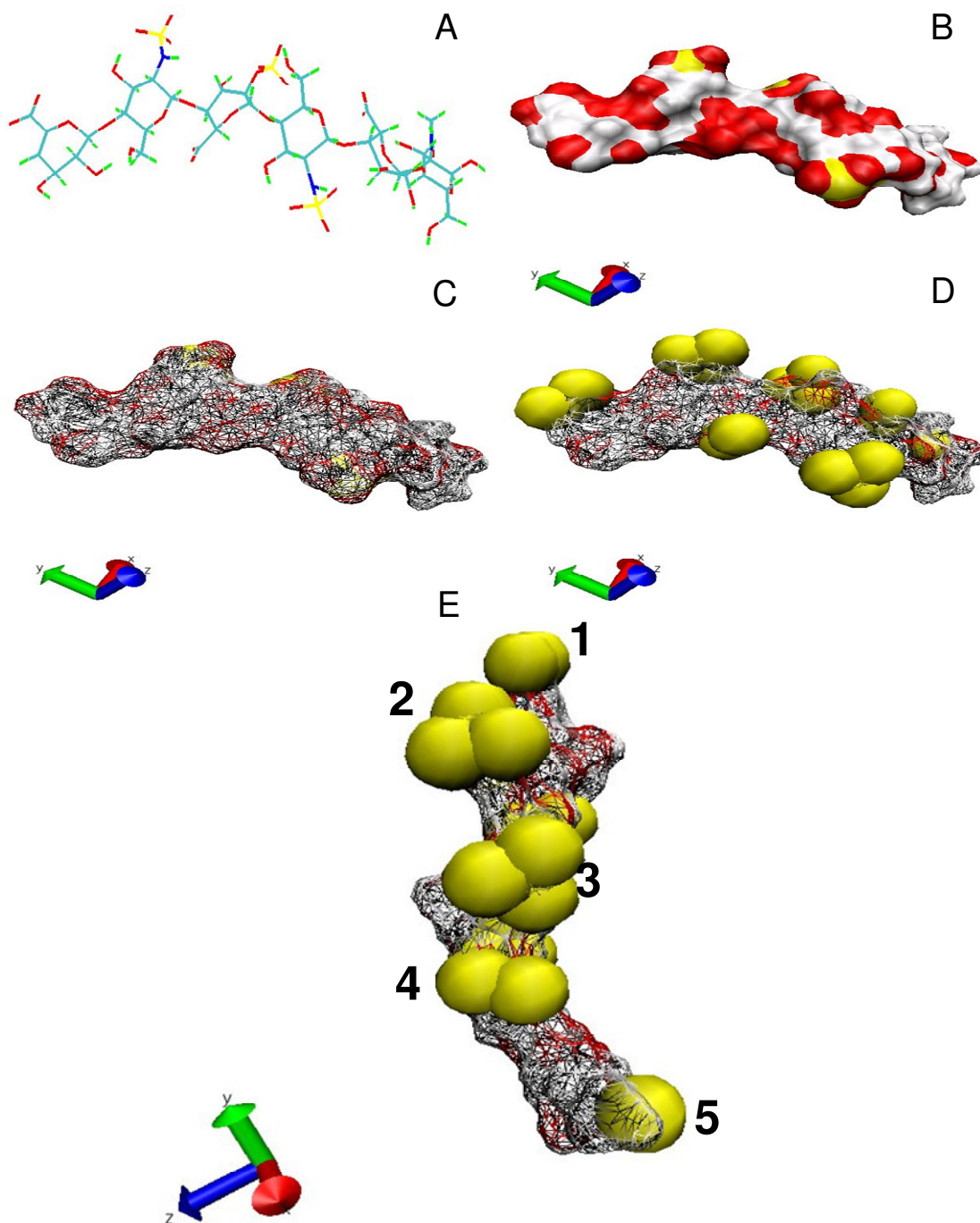
X, Y, Z, relative coordinate axis are shown for models B-D

A number of electronegative groups are numbered 1-5 in model E

1. UA, carboxyl group
2. GlcNS, N-sulphate group
3. IdoUA(2S), 2-O-sulphate group
4. IdoUA, carboxyl group
5. GlcNAc, N-acetyl group

All models were formatted using the VMD molecular graphics program (Humphrey *et al.* 1996).

Figure 71



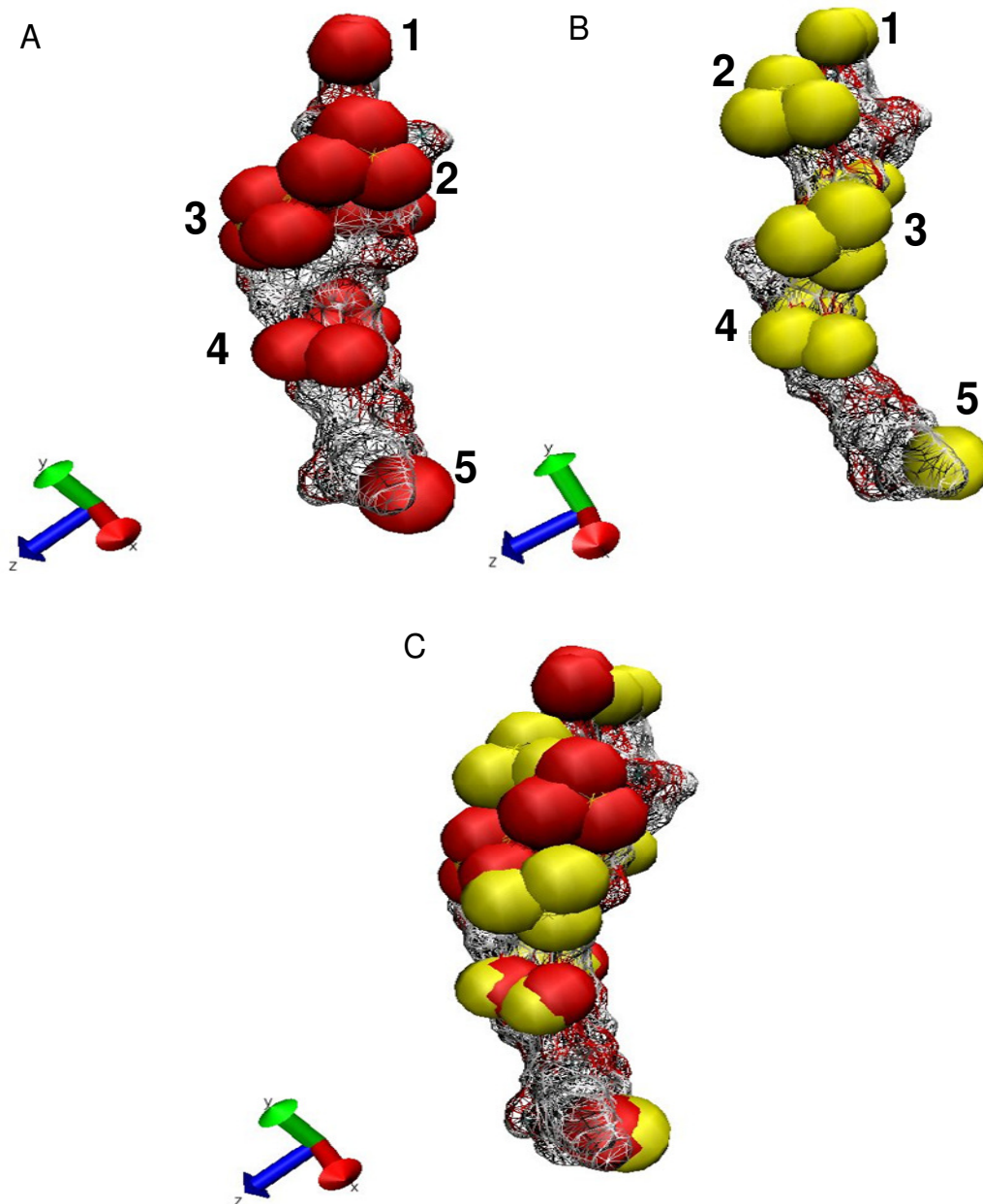
UA-GlcNS-IdoUA(2S)²S₀-GlcNS-IdoUA²S₀-GlcNAc

Figure 72 A superimposition model of HS dp6-2.

Two models of HS dp6-2, one in which both internal iduronates are in the 1C_4 conformation (model A), and one in which both internal iduronates are in the 2S_0 conformation (model B), have been superimposed to form model C.

Superimposition alignment was carried out using the rmsd extension of the VMD molecular graphics program (Humphrey *et al.* 1996). Model representations and colours are as in Figures 70 and 71.

Figure 72



N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (red); UA-GlcNS-IdoUA(2S)¹C₄-GlcNS-IdoUA¹C₄-GlcNAc

N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (yellow); UA-GlcNS-IdoUA(2S)²S₀-GlcNS-IdoUA²S₀-GlcNAc

14. HS dp6-3, differences to HS dp6-2 and possible biological implications

Reformatted models for HS dp6-3 are shown in Figure 73. In the orientation shown in Figure 73, as with HS dp6-2, again a total of 5 electronegative groups may contribute to a binding affinity for the stationary phase to the SAX column (labelled 1-5). Positional rearrangements are again observed on iduronate conformational changes within the oligosaccharide.

Both individual models, and superimposition models for HS dp6-2 and -3, are very similar, compare Figures 73 and 72, and it is not immediately clear why HS dp6-3 should have an increased binding affinity for the SAX-HPLC column. HS dp6-2 and HS dp6-3 are almost identical in their 3-dimensional shape and in the positioning of their electronegative substituents. If it can be understood why these two molecules interact differently with the stationary phase of the SAX-HPLC column, a greater insight may be gained into similar interactions at the protein surface.

Differences in their HPLC binding affinities may be due to a subtle conformational change that occurs between the two molecules around electronegative group 5. Models for both HS dp6-2 and -3 have been produced with the whole of each reducing terminal GlcNAc residue, containing group 5, in a Van der Waals representation, see Figure 74. In model B it can clearly be seen that the methyl group obscuring the acetyl oxygen (group 5) in model A is rotated. This places the acetyl oxygen in the same plane as electronegative groups 1-4, possibly now allowing it to interact with the SAX-HPLC stationary phase. Due to the steric hinderance of the methyl group in model A it is doubtful if the acetyl oxygen interacts with the HPLC stationary phase in the same plane as groups 1-4. The increased affinity of HS dp6-3 may therefore be due to a contribution

of the electronegative acetyl oxygen to electrostatic interactions with the HPLC stationary phase.

Acetyl oxygen groups are often overlooked as significant contributors to electrostatic interactions between oligosaccharides and proteins. This is perhaps because many binding studies use heparin or HS S-domain structures which contain few, if any, GlcNAc monosaccharide units. It is believed that the modelling and SAX-HPLC data presented in this thesis, represent the first evidence for the involvement of HS N-acetyl groups in oligosaccharide electrostatic interactions. Given that GlcNAc containing disaccharide units typically make up >50% of all disaccharide units in standard HS preparations, this observation may be of significant biological relevance. Previously, it has been suggested that the role of N-acetyl and NA/NS domains within HS is merely to space the S-domains correctly, so that more than one S-domain may make contact with more than one HS binding site at the protein surface (Mulloy and Forster 2000). It may also be that N-acetyl groups in the regions between the S-domains may also make additional stabilising electrostatic interactions with the protein surface.

Figure 73 A superimposition model of HS dp6-3.

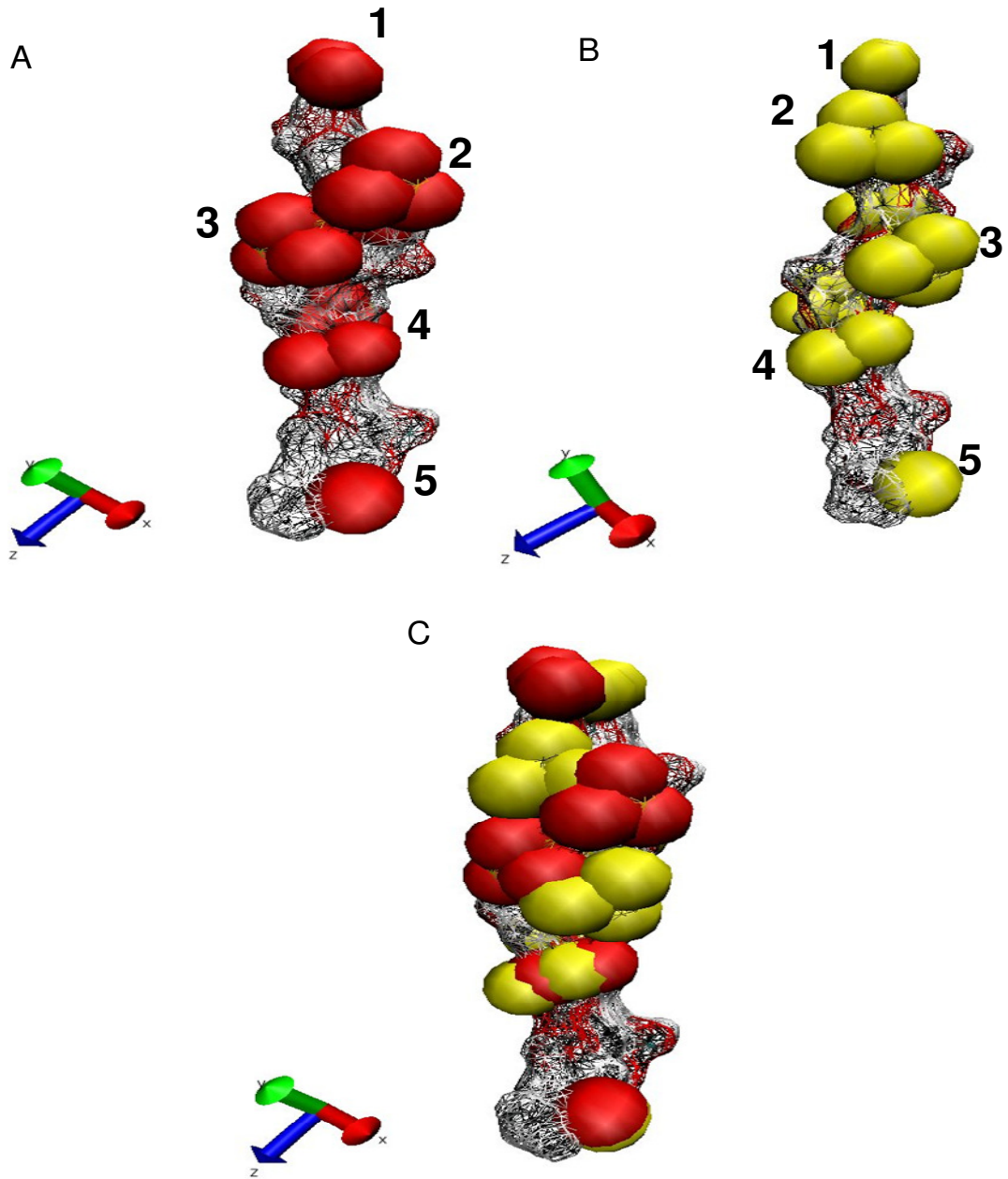
Two models of HS dp6-3, one in which the internal iduronate is in the 1C_4 conformation (model A), and one in which the internal iduronate is in the 2S_0 conformation (model B), have been superimposed to form model C.

Superimposition alignment was carried out using the rmsd extension of the VMD molecular graphics program (Humphrey *et al.* 1996). Model representations and colours are as in Figures 70 and 71.

A number of electronegative groups are numbered 1-5

1. UA, carboxyl group
2. GlcNS, N-sulphate group
3. IdoUA(2S), 2-O-sulphate group
4. GlcUA, carboxyl group
5. GlcNAc, N-acetate group

Figure 73



N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (red); UA-GlcNS-IdoUA(2S)¹C₄-GlcNS-GlcUA-GlcNAc

N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (yellow); UA-GlcNS-IdoUA(2S)²S₀-GlcNS-GlcUA-GlcNAc

Figure 74

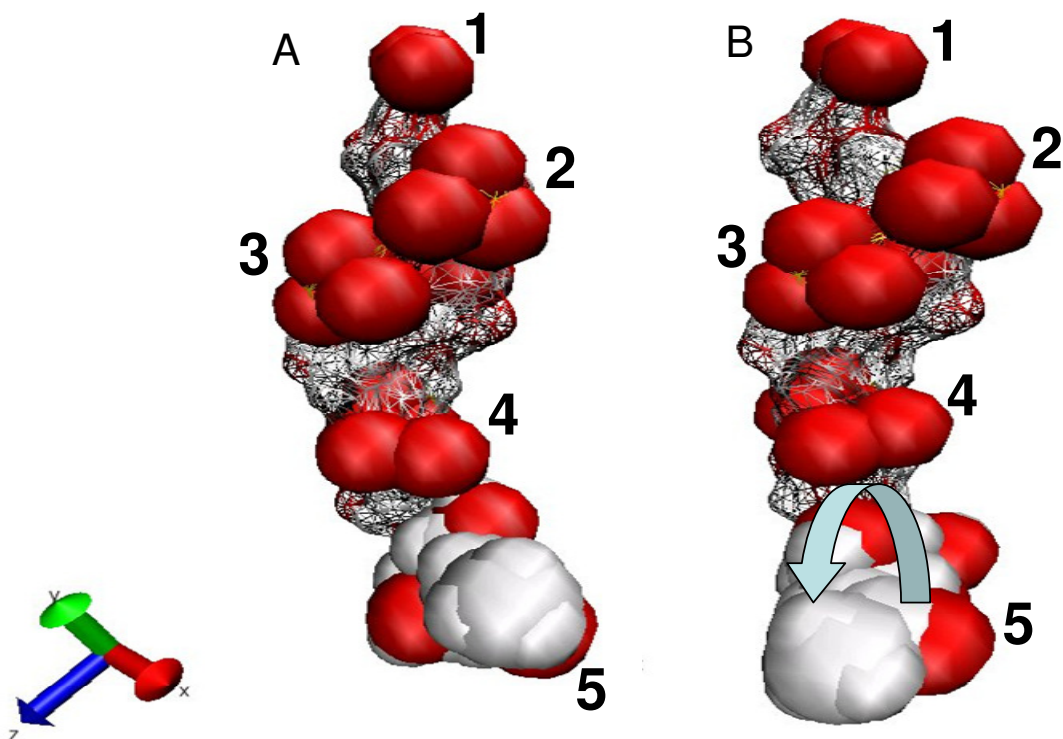


Figure 74 HS dp6-2 and HS dp6-3 molecular models.

HS dp6-2 is modelled with all internal iduronates in the 1C_4 conformation (Model A). HS dp6-2 is also modelled with its lone internal iduronate in the 1C_4 conformation (Model B).

Model representations and colours are as in Figures 72 and 73 with the addition that the whole reducing terminal GlcNAc residue of each model is also shown in a Van der Waals representation.

All models were formatted using the VMD molecular graphics program (Humphrey *et al.* 1996).

15. A molecular re-orientation for HS dp6-5

Reformatted models for HS dp6-5 are shown in Figure 75. In model A the oligosaccharide is shown in the same orientation as oligosaccharides dp6-2 and -3 in Figures 70-74. However, it is doubtful if the molecular surface shown here constitutes the binding surface for this oligosaccharide to the SAX-HPLC stationary phase, or to the surface of a potential protein ligand. Only two electronegative groups are present. It is more likely that the stationary phase binding surface occurs on the opposite side of the molecule as shown in model B, consisting of 4 electronegative groups. A superimposition model for HS dp6-5 in the orientation of model B is shown in Figure 76. From the superimposition model it can be seen that in this proposed binding orientation only 4 electronegative groups constitute the binding surface. Four electronegative groups are in contrast to the 5 electronegative groups potentially forming the binding surface of HS dp6-2 and -3. The increased SAX-HPLC binding affinity for HS dp6-5 may be due to the type of electronegative group rather than the number present at the binding surface. The proposed binding surface for HS dp6-5 is composed of three O-sulphate groups (groups 1, 3 and 4) and a single carboxyl group (group 2). The additional O-sulphate group may add an additional degree of affinity that outweighs the contributions from the carboxyl and acetyl groups of HS dp6-2 and -3.

The molecular re-orientation of HS dp6-5 relative to all other models demonstrates that ligand binding sites do not necessarily occur on the same side of a HS polymer, in a so called 'cis' binding orientation (Gallagher and Lyon 2000). It could be imagined that if the nonreducing terminal of HS dp6-5 were linked to the reducing terminal of HS dp6-3, where each then to bind a protein ligand, these ligands would be positioned on opposite

sides of the oligosaccharide chain, in a so called 'trans' binding orientation (Gallagher and Lyon 2000). The possibility for this 'trans' mode of binding was implied by regions of high electronegativity, formed by closely positioned N-, 2-O- and 6-O-sulphate groups, occurring consecutively on opposite sides of the dodecasaccharide chain in the heparin dodecasaccharide models of Mulloy *et al.* (Mulloy *et al.* 1993) (see introduction, Figure 17). Two proteins binding to two such consecutive regions of high electronegativity would then be positioned on opposite sides of the oligosaccharide chain. This mode of binding actually exists in the Pellegrini crystal structure of an FGF1:FGFR2c:heparin complex (Pellegrini *et al.* 2000) (see introduction, Figure 14). The molecular modelling and SAX-HPLC evidence in this study for the protein binding surface of HS dp6-5 to be on the opposite side of the molecule to HS dp6-2, -3 and -6, further re-inforces the existence of both 'trans' and 'cis' ligand binding modes within HS oligosaccharides.

Figure 75 HS dp6-5 molecular models.

HS dp6-5 is modelled with the internal iduronate in the 1C_4 conformation in both cases.

Model B is a re-orientation of model A (see text).

X, Y, Z, relative coordinate axis are shown for both models.

A number of electronegative groups are numbered 1-4 in both models.

1. UA, carboxyl group (model A), GlcNAc(6S), 6-O- sulphate group (model B),

2. IdoUA carboxyl group (model A), GlcUA carboxyl group (model B)

3. GlcNS, N-sulphate group,

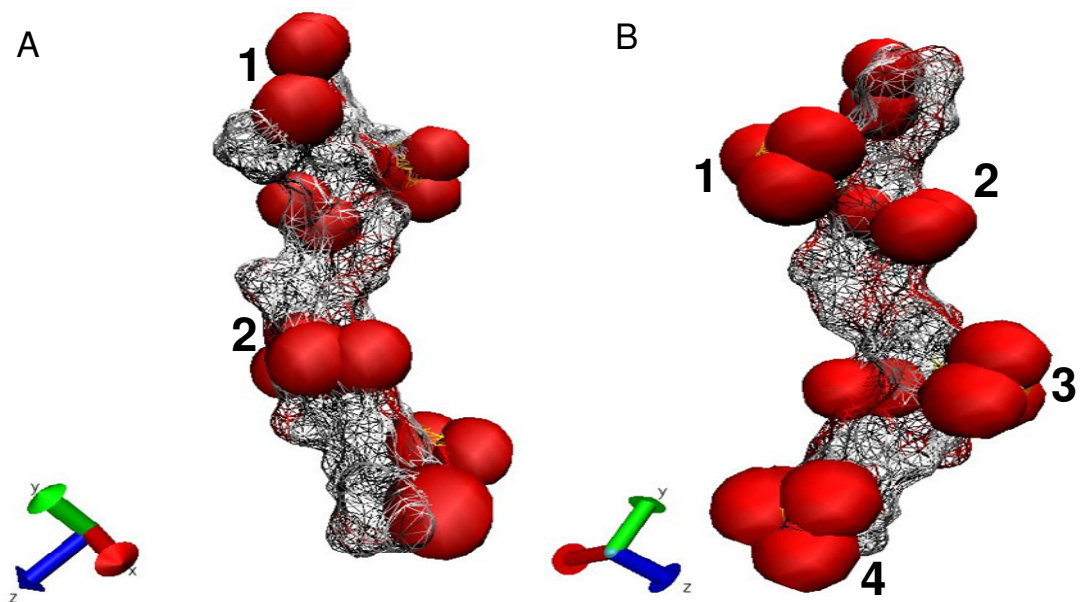
4. GlcNAc(6S), 6-O-sulphate group,

Model representations and colours are as in Figures 70 and 71.

All models were formatted using the VMD molecular graphics program (Humphrey *et al.*

1996)

Figure 75



UA-GlcNAc(6S)-GlcUA-GlcNS-IdoUA¹C₄-GlcNAc(6S)

Figure 76 A superimposition model of HS dp6-5.

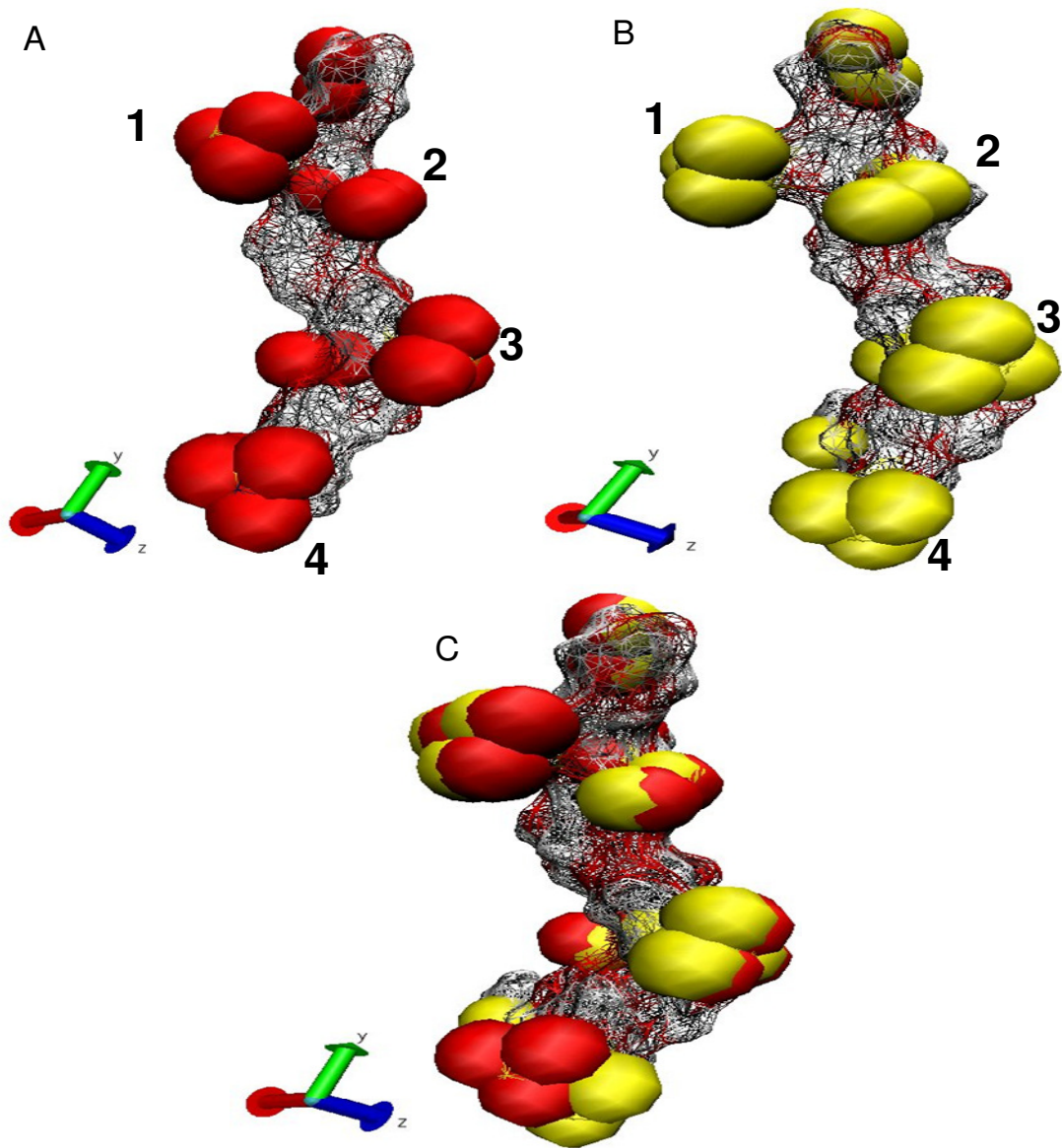
Two models of HS dp6-5, one in which the internal iduronate is in the 1C_4 conformation (model A), and one in which the internal iduronate is in the 2S_0 conformation (model B), have been superimposed to form model C.

Superimposition alignment was carried out using the rmsd extension of the VMD molecular graphics program (Humphrey *et al.* 1996). Model representations and colours are as in Figures 70 and 71.

A number of electronegative groups are numbered 1-4

1. GlcNAc(6S), 6-O- sulphate group,
2. GlcUA, carboxyl group,
3. GlcNS, N-sulphate group,
4. GlcNAc(6S), 6-O-sulphate group,

Figure 76



N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (red); UA-GlcNAc(6S)-GlcUA-GlcNS-IdoUA¹C₄-GlcNAc(6S)

N- and O-sulphate groups, carboxyl groups and acetate oxygens in Van der Waals representation (yellow); UA-GlcNAc(6S)-GlcUA-GlcNS-IdoUA²S₀-GlcNAc(6S)

16. De-2-O-sulphation analysis

The main purpose of generating de-2-O-sulphated oligosaccharides was to examine more closely the role 2-O-sulphation plays in influencing conformational equilibrium. An unexpected finding of the NMR analysis of the de-2-O-sulphated heparin hexasaccharides was the resistance of their terminal uronic acid residues to de-2-O-sulphation.

Iduronate chemical de-2-O-sulphation (see experimental procedures) is brought about by a base induced intramolecular displacement by the hydroxyl oxygen attached to carbon position 3, of the O-sulphate group attached to carbon position 2. The result of this displacement is the formation of an epoxide, 2,3-epoxigulonic acid. How this epoxide then decomposes determines the end product of de-2-O-sulphation (Jaseja *et al.* 1989; Piani *et al.* 1993). On lyophilisation to dryness the epoxide decomposes to leave a de-2-O-sulphated iduronate residue. It may be that the presence of a C4-C5 double bond within terminal uronate residues disrupts the efficient formation of the epoxide intermediate.

The results presented in sections 20 and 21 show that terminal uronates are not totally resistant to de-2-O-sulphation with on average 14.5% of the original total number of 2-O-sulphates removed on each treatment. It was shown that simulating multiple treatments by increasing the concentration of sodium hydroxide, does not increase the efficiency with which de-2-O-sulphation of these residues occurs. It is suggested that 7 passes of the procedure detailed under experimental procedures may be enough to totally remove all UA(2S) 2-O-sulphate groups. Four passes of the de-2-O-sulphation procedure was the maximum number undertaken as part of this thesis study. After 4 passes >40% of the UA(2S) 2-O-sulphate groups still remained *in-situ*.

16.1 Biological implications

A great number of studies have been published in which chemically de-2-O-sulphated polymeric heparin have been used in functional analyses. The following are just a selection (Maccarana *et al.* 1993; Coltrini *et al.* 1994; Ishihara *et al.* 1994; Ono *et al.* 1999; Chuang *et al.* 2000; Lundin *et al.* 2000; Lyon *et al.* 2000; Ostrovsky *et al.* 2002; Yabushita *et al.* 2002; Ricard-Blum *et al.* 2004; Robinson *et al.* 2005). More recently chemically de-sulphated polymeric HS has also been used (Robinson *et al.* 2006). The aim of many of these studies is to try and establish a minimal oligosaccharide sulphation requirement for ligand binding or biological activity. UA or UA(2S) residues at the non-reducing end of an oligosaccharide are usually the result of a deliberate heparinase or chemical β -eliminative degradation of a longer saccharide chain. Of the twelve studies cited above, all produced chemically de-2-O-sulphated heparin/HS for their functional studies from polymeric commercial grade heparin/HS. In no study had this material been deliberately treated with heparanase or undergone chemical β -eliminative degradation. Given this, it might be expected that the de-2-O-sulphated material used may contain none of the de-2-O-sulphation resistant UA(2S) monosaccharides. However, Jandik *et al.* (Jandik *et al.* 1996) have shown that under extreme conditions of high pH (1M sodium hydroxide) and elevated temperature (60°C) heparin may be degraded. Note that Jandik *et al.* specifically examined only the chemical stability of heparin, not HS. Given the strong structural similarity between HS and heparin it is probable that HS is also degraded under similar extreme conditions. Jandik *et al.* detected an appreciable increase in the number of non-reducing terminal UA(2S) residues present in a heparin sample incubated at 60°C in 1M sodium hydroxide over a 10h period (Jandik *et al.* 1996), (as detected by a rise in

the absorbance of the sample at 232nm). The original animal source for most commercially purchased heparin/HS is bovine lung or porcine intestine (Linhardt and Gunay 1999). Extraction and purification of commercial grade heparin/HS from these tissues involves just such an exposure to high temperatures and alkaline pH for prolonged periods of time (Linhardt and Gunay 1999). In the light of the stability study by Jandik *et al.* it is thought likely that all commercial heparin/HS preparations may contain variable quantities of UA(2S) residues, the precise number of which varying between suppliers as a result of differing manufacturing conditions. In support of this a number of resonances at around 105 ppm were detected in the ^{13}C spectra of porcine mucosal heparin samples, recorded separately by both Desai *et al.* and Guerrini *et al.* (Desai and Linhardt 1995; Guerrini *et al.* 2001). One of these resonances may correspond to the C4 resonance of a terminal UA(2S) residue (see Table 5). However, neither study assigned a UA(2S) C4 resonance. Guerrini *et al.* also recorded a ^1H spectrum for both porcine and bovine tissue extracted heparin. In these spectra a number of resonances can be seen that may correspond to chemical shifts for UA(2S) H1, H2 and H3 (see Table 5). This was not commented on by Guerrini *et al.* (Guerrini *et al.* 2001).

A key role for terminal residues and in particular UA(2S) residues in establishing biological activity is suggested by the crystal structure of an FGF2:FGFR1:heparin complex produced by Schlessinger *et al.* (Schlessinger *et al.* 2000) (see introduction). In the Schlessinger crystal complex, two terminal UA(2S) residues lie at the centre of a negatively charged cleft that runs between the two receptor molecules (see introduction). In this position, the 2-O-sulphate and carboxyl groups of both UA(2S) residues may make extensive stabilising electrostatic contacts with both receptor molecules. When it is

considered that a single HS chain may extend for up to more than 50 disaccharide units in length, it seems remarkable that the crystal complex of Schlessinger involves the ends of two different oligosaccharide molecules. Within polymeric HS/heparin, by far the majority of sequences do not involve any terminal units. Without the crystallographic evidence, a more logical FGF2:FGFR1:heparin complex might be expected not to involve terminal monosaccharide units at all. It may be that an FGF2:FGFR1 signalling complex, and possibly other biological signalling complexes, have a preference for formation around sequences containing a terminal monosaccharide residue such as the UA(2S). In further support of this it has been reported that only after its enzymatic degradation does polymeric HS gain the ability to activate FGF-2 mitogenicity (Kato *et al.* 1998). This may be correlated to an increase in the number of terminal oligosaccharide units around which active complexes may form. If this is the case the biological influence of a small number of unremoved 2-O-sulphate groups at the non-reducing end of oligosaccharide may be more exaggerated when it is assayed for biological activity. Therefore, great care needs to be taken in the interpretation of results from studies where it has been shown that de-2-O-sulphation of a saccharide does not influence its biological activity. The specific studies of (Ono *et al.* 1999; Ostrovsky *et al.* 2002; Robinson *et al.* 2006) are discussed further below. However, without any additional experimental evidence, solely using chemically de-2-O-sulphated heparin, Yabushita *et al.* (Yabushita *et al.* 2002), have concluded that 2-O-sulphate groups are not important for the biological activity of the chlamydial major outer membrane protein. In this case and in others, it may be that internal 2-O-sulphate groups are not important for biological activity but that terminal UA(2S) sulphate groups are.

The heparan sulphate 2-O-sulphotransferase enzyme (HS2ST) is responsible for the 2-O-sulphation of HS during its biosynthesis (see introduction). An alternative to chemical de-2-O-sulphation, is to produce cell lines in which this gene has been knocked out, so called HS2ST^{-/-} mutant cell lines. HS extracted from these cell lines when grown in culture contains absolutely zero 2-O-sulphate groups (Merry *et al.* 2001). A number of contradictions now exist in the literature on the requirements for 2-O-sulphate groups when studied using either chemically de-2-O-sulphated heparin/HS or HS extracted from HS2ST^{-/-} cell lines. Ostrovsky *et al.* (Ostrovsky *et al.* 2002) have shown that de-2-O-sulphated heparin is still biologically active in an FGF-1, FGFR2IIIb mitogenesis assay. However, Merry *et al.* (Merry *et al.* 2001) have shown using HS2ST^{-/-} HS, that 2-O-sulphate groups are required for the FGF-1 ligand to bind HS. This contradiction may be explained by the presence of none reducing terminal UA(2S) monosaccharides in the de-2-O-sulphated material prepared by Ostrovsky *et al.* In support of this, Guerrini *et al.* have reported a role for none reducing terminal monosaccharides in the formation of FGF-1 heparin complexes (Guerrini *et al.* 2002).

Ono *et al.* (Ono *et al.* 1999) have shown using de-2-O-sulphated heparin, that 2-O-sulphate groups are not required for the biological activity of the VEGF isoform 165 (VEGF₁₆₅). In support of this Robinson *et al.* (Robinson *et al.* 2006) have shown that HS2ST^{-/-} material does **not** have a reduced affinity for VEGF₁₆₅ when compared to normal HS. However, Robinson *et al.* also reported that chemically de-2-O-sulphated HS was able to compete with polymeric HS for binding interactions with VEGF₁₆₅. These two results by Robinson *et al.* seem contradictory as to the role of 2-O-sulphate groups in

VEGF₁₆₅ interactions. This contradiction may be explained by the presence of UA(2S) monosaccharides in the de-2-O-sulphated material prepared by Robinson *et al.*

The identified resistance of UA(2S) to chemical de-sulphation should cause a re-think on the use of chemically de-2-O-sulphated heparin/HS in ligand binding or biological assays. Most studies report 2-O-sulphate levels as being reduced by >80% after one round of sodium hydroxide treatment. However, it now cannot be ruled out that the critical positioning of the remaining 2-O-sulphate groups, a majority of which will form UA(2S) units are not an important factor in influencing ligand binding interactions and ultimately biological activity when assayed for.

17. The influence of sulphation pattern on iduronate and terminal uronic acid equilibria

The coupling constant reviews of Ferro *et al.* (Ferro *et al.* 1990) and Ragazzi *et al.* (Ragazzi *et al.* 1993) were expanded to also include all data published subsequent to 1990 and 1993 respectively. Using the theoretical coupling constants given in Table 9 for iduronate monosaccharides, and theoretical coupling constants taken from Ragazzi *et al.* for terminal uronic acid monosaccharides, Table 14 was constructed. Also included in Table 14 are populational data for HS dp6-2, -3, -5, de-2-O-HS dp6-3, Hep dp6-1 and de-2-O-hep dp6-1. A comparison of all data shown in Table 14 reveals new insights into the influence of sulphation pattern on conformational equilibria.

Structure A (see Table 14) was previously analysed as the terminal unit of an extended K5 polysaccharide (Ragazzi *et al.* 1993). The previous conclusion, drawn by Ragazzi *et al.*, that the attachment of a polysaccharide chain to the UA-GlcNAc disaccharide unit

(structure C in Table 14), causes a shift towards the ${}^2\text{H}_1$ conformation from the ${}^1\text{H}_2$ conformation, is confirmed by analysis of HS dp6-1 and HS dp6-2 (compare HS dp6-1, -2, and structures A to C in Table 14). The two additional disaccharide units added to UA-GlcNAc disaccharide in HS dp6-1 and -2, in contrast to the polymeric structure analysed by Ragazzi *et al.*, do not constitute a polysaccharide chain. By comparing structure C (see Table 14) to structure B (see Table 14) it can be concluded that rather than a polysaccharide chain, the addition of a single disaccharide unit is able to exert the same shift towards ${}^2\text{H}_1$.

Comparing structures C and E (see Table 14), N-sulphation as opposed to N-acetylation, to the reducing side of a terminal uronic acid has little influence on its conformational equilibrium. Structure D (see Table 14) is included to demonstrate the influence of solvent. Structure D was analysed by Bazin *et al.* (Bazin *et al.* 1998) at 25°C in deuteriochloroform. All other analyses cited in Table 14 were carried out at 25°C in buffered D₂O.

Comparing structures C and F, 6-O-sulphation of a linking glucosamine causes a shift towards ${}^1\text{H}_2$, addition of an N-sulphate group as in structure G has little further influence. None of the coupling constant data produced since 1993, or as part of this thesis work contradict the previous conclusions drawn by Ragazzi *et al.* (Ragazzi *et al.* 1993), that by far the greatest influence towards the ${}^1\text{H}_2$ conformation is 2-O-sulphation of UA itself (compare structures C, E, F and G to H). A shift back towards ${}^2\text{H}_1$ as seen for a UA-GlcNAc disaccharide is not seen when UA(2S)-GlcNS(6S) becomes part of a more extended oligosaccharide chain (compare structures H to I).

For non-sulphated iduronate monosaccharides, using structure J with which to make a comparison, neither N-sulphation of the reducing side glucosamine or 6-O-sulphation of the non-reducing side glucosamine is able to exert a significant conformational influence (compare structures J and K). It is only when the non-reducing side glucosamine becomes N-sulphated as in HS dp6-2 that a significant influence towards 1C_4 is exerted (compare HS dp6-2 to structures J and K).

Comparing HS dp6-2 and HS dp6-5 suggests that 6-O-sulphation of the non-reducing glucosamine exerts influence back towards 2S_0 , by 8%. However, comparison of HS dp6-5 to structure L and de-2-O-sulphated HS dp6-3 suggests it may exert a slight influence more towards 1C_4 , by 3% and 2% respectively. This highlights the issue that there is a limit at which coupling constants can be measured accurately, therefore little should be read into population differences of less than 10%. For this reason it is concluded that the main influence on non-sulphated iduronate conformation is N-sulphation of the non-reducing side glucosamine. This is present in all of saccharides HS dp6-5 through to structure N. None of these saccharides show a significant alteration in conformational balance above 10% when compared to HS dp6-2.

Ferro *et al.* (Ferro *et al.* 1990) concluded that 2-O-sulphation alters the balance of iduronate conformational equilibrium more towards 1C_4 . This is indeed the case when comparing HS dp6-3 (76% 1C_4 , see Table 14) and de-2-O-sulphated HS dp6-3 (51% 1C_4 , see Table 14). However, this is not the case when comparing Hep dp6-1 (51% 1C_4 , see Table 14) and de-2-O-sulphated Hep dp6-1 (57% 1C_4 , see Table 14). This suggests that

flanking 6-O-sulphation plays a major influence in altering IdoUA(2S) conformational equilibrium.

Comparing HS dp6-2, -3, structure O and structure P (see Table 14) it is only when both N-sulphation and 6-O-sulphation are present at the non-reducing side glucosamine that a significant influence towards the 2S_0 conformation is exerted, by an average of 18.6%. Comparing structure P to structure Q suggests that N-sulphation and 6-O-sulphation flanking IdoUA(2S) exert the same influence as N- and 6-O-sulphation solely occurring to the non-reducing side (see Table 14). A significant role is not suggested for a non-reducing side N-sulphate group. In all saccharides Hep dp6-1 through to structure U a non-reducing side N-sulphate group and a flanking and/or same side 6-O-sulphate group are present.

Structures S through to U also demonstrate that a 3-O-sulphate group positioned on either side of IdoUA(2S) is able to exert significant conformational influence towards 2S_0 . The same influence is not seen with IdoUA, compare the influence of 3-O-sulphation within structures S to U with that between structures M and N (see Table 14).

For structures V and W the 4C_1 conformation was calculated to contribute significantly to the conformational equilibrium of the two iduronates indicated. Iduronates within structures V and W occur within a synthetic oligosaccharide and it may be that the synthetic process used in its production has left a lasting contribution of the 4C_1 conformation (Lucas *et al.* 2003).

Table 14 Conformational distributions of iduronate and terminal uronic acid monosaccharides within oligosaccharides characterised in this study and a number of previously published studies.

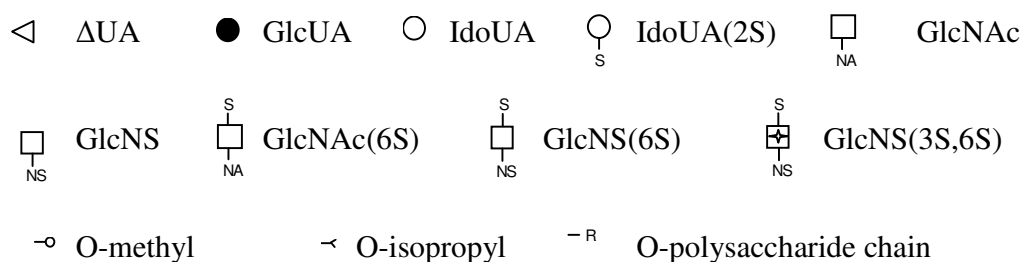
Theoretical coupling constants for iduronates were taken from Table 9, and theoretical coupling constants for terminal uronic acids from Ragazzi *et al.* (Ragazzi *et al.* 1993)

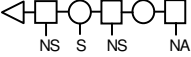
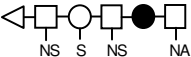

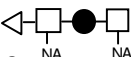






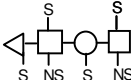
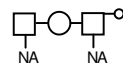
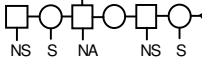
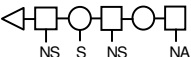
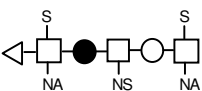
For iduronates, a three conformational least squares fit was performed to derive estimated conformer populations, for terminal uronic acids a two conformational least squares fit (see introduction).

The residual value for the sum of square differences after parameters a, b and c (or a, b for UA) had been optimised is given as a guide to fit accuracy (see introduction). Better fitting population estimates having a value closer to zero.

Where more than one residue of the type analysed exists within a sulphation sequence, the residue(s) analysed are indicated by arrows.

Symbols:



Δ UA	$^1\text{H}_2$ $^2\text{H}_1$ (%)	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	Residual sum of square differences	Reference for coupling constants
HS dp6-2	27 73	6.9	6.3	3.1	-	0.22	this thesis
							
HS dp6-3	31 69	6.6	6.0	3.4	-	0.03	this thesis
							
A	31 69	6.59	6.01	3.36	-	0.01	Ragazzi <i>et al.</i> (1993)
							
B	29 71	5.9	6.8	3.0	-	0.76	Hileman <i>et al.</i> (1997)
							
C	43 57	5.96	5.4	3.58	-	0.06	Ragazzi <i>et al.</i> (1993)
							
D	16 84	7.2	6.5	3	-	0.2	Bazin <i>et al.</i> (1998)
							
E	41 59	6.03	5.47	3.56	-	0.02	Ragazzi <i>et al.</i> (1993)
							
F	51 49	5.48	4.93	3.75	-	0.04	Ragazzi <i>et al.</i> (1993)
							
G	53 47	5.36	4.81	3.79	-	0.06	Ragazzi <i>et al.</i> (1993)
							
H	92 8	3.28	2.77	4.52	-	0.04	Ragazzi <i>et al.</i> (1993)
							
I	100 0	2.0	2.0	4.8	-	0.25	Pervin <i>et al.</i> (1995)
							
Non-sulphated IdoA	$^1\text{C}_4$ $^4\text{C}_1$ $^2\text{S}_0$ (%)	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$		
J	37 0 63	3.9	6.6	3.6	3.2	0.93	van_Boeckel <i>et al.</i> (1987)
							
K	34 1 65	3.5	6.6	4.5	3.3	0	Ojeda <i>et al.</i> (2002)
							
HS dp6-2	57 0 43	2.5	5.0	3.3	2.8	0.34	this thesis
							
HS dp6-5	49 0 51	2.5	5.3	4.1	3.4	0.76	this thesis
							

L		52	0	48	3.3	5.4	3.6	2.4	0.25	van_Boeckel <i>et al.</i> (1987)
de-2-O-HS dp6-3		51	1	48	2.5	5.3	4.4	2.8	0.55	this thesis
de-2-O-Hep dp6-1		57	0	43	1.6	5.3	3.4	2.8	1.87	this thesis
M		52	0	48	2.1	5.4	4.2	2.5	0.98	Lucas <i>et al.</i> (2003)
N		52	0	48	3.3	5.4	3.6	2.4	0.25	Ferro <i>et al.</i> (1990)
Sulphated IdoA		1C_4	4C_1	2S_0	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$		
		(%)			(Hz)					
HS dp6-2		74	0	26	1.6	4.1	3.15	2.2	0.76	this thesis
HS dp6-3		76	0	24	1.6	4.1	3.15	1.5	0.99	this thesis
O		83	0	17	1.4	3.2	2.8	2.1	0.77	van_Boeckel <i>et al.</i> (1987)
P		59	0	41	2.8	5.4	3.1	2.5	1.07	van_Boeckel <i>et al.</i> (1987)
Q		50	0	50	3	5.8	4	2.8	0.21	Ojeda <i>et al.</i> (2002)
Hep dp6-1		51	0	49	2.5	5.6	4.1	3.2	0.29	this thesis
R		59	0	41	2.8	5.4	3.1	2.5	1.07	van_Boeckel <i>et al.</i> (1987)
S		33	0	67	3.9	7.3	3.9	3.3	1.98	van_Boeckel <i>et al.</i> (1987)
T		26	0	74	3.5	8.3	4.2	3	2.8	Ferro <i>et al.</i> (1990)

U 6 0 96 5.3 9.4 4.4 4.4 5.4 van_Boeckel *et al.* (1987)

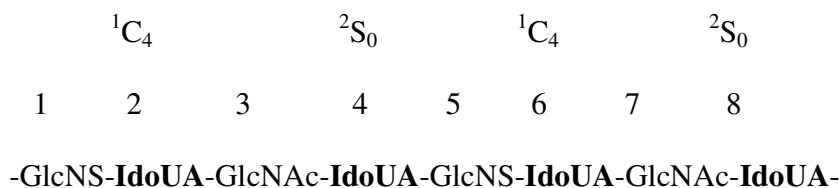
V 64 18 18 3.6 4.2 3.6 3.5 1.72 Lucas *et al.* (2003)

W 60 28 12 3.6 4.3 4.5 3.7 1.25 Lucas *et al.* (2003)

17.1 Biological implications

One biological implication is with regard to the conformational equilibrium of iduronate residues in the alternating NA/NS domains of HS (see introduction, section 8.2, page 65).

Iduronates in the following sequence:

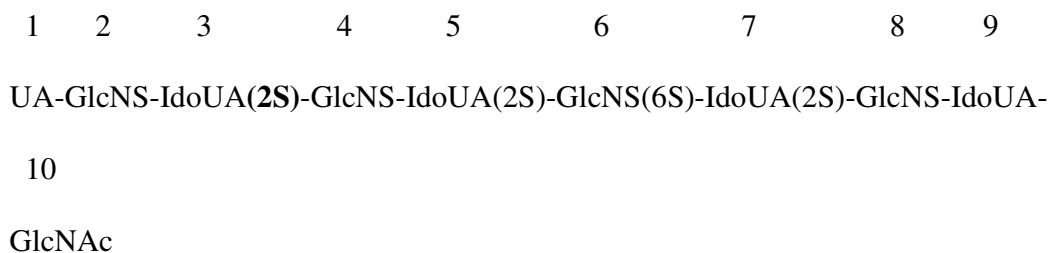


could be expected to be in predominantly different conformations; iduronates 2 and 6 predominantly in the 1C_4 conformation, iduronates 4 and 8 predominantly in the 2S_0 conformation, as indicated. Even with the addition of 6-O-sulphate groups to the glucosamines in this sequence, it can be concluded from the data presented in Table 14 that this pattern would not be expected to alter significantly.

NA/NS domains have been implicated in a number of ligand binding interactions (see in particular the discussion on endostatin in the introduction, section 12, page 77). This pattern of alternating iduronate conformations may be an important structural feature for biological activity.

Another important implication is with regard to the positioning of 6-O-sulphate groups. Within HS S-domains, 6-O-sulphation appears to be under tight regulatory control and either centrally positioned or positioned at the reducing terminal (Merry *et al.* 1999; Vives *et al.* 1999; Goodger 2003). Centrally positioned this will have the effect of

influencing the conformational equilibria of both proceeding and preceding IdoUA(2S) residues. For example, in the following sequence (sequence 10c (Merry *et al.* 1999)):



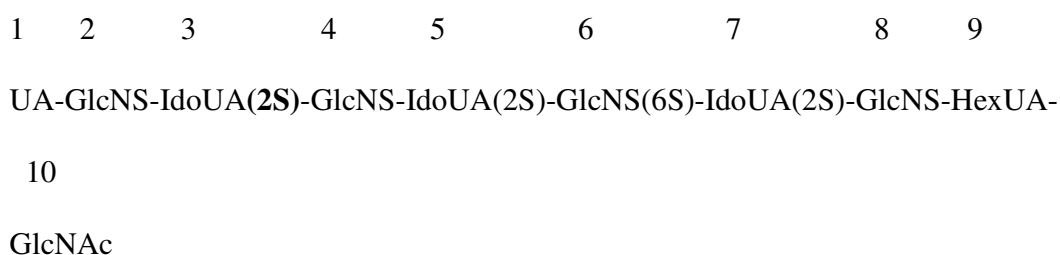
iduronates 5 and 7 may be expected to be present at 50% 1C_4 , 50% 2S_0 , whereas iduronate 3 will be predominantly present in the 1C_4 conformation (approx 75% 1C_4).

It could be suggested that the shift in the conformational equilibria of iduronates 5 and 7, towards 2S_0 , may be important for the potential biological activity of this oligosaccharide. A similar shift has been shown to be important within iduronate 5 of an antithrombin III binding synthetic pentasaccharide (see introduction, section 14.1, page 94). However, it may be that changes in the conformational equilibrium of the iduronate at monosaccharide position 3, towards the 1C_4 conformation that are biologically relevant, as detailed below.

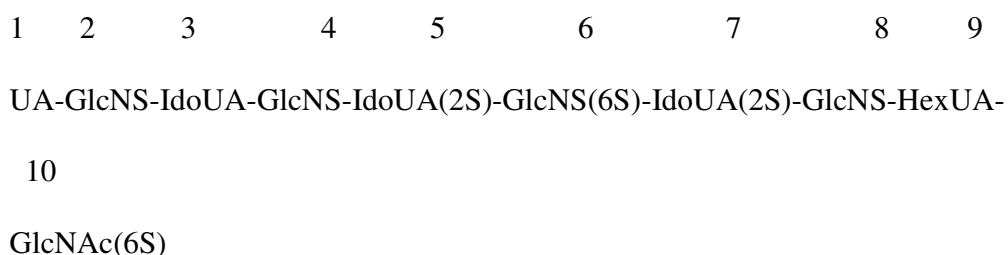
All decasaccharides, sequenced and tested for biological activity by Goodger, are given in Figure 77. Also shown in Figure 77 are the predominant conformations of all iduronates in these sequences, as deduced from Table 14. Where no conformation predominates these iduronates are marked as 50/50 (50% 1C_4 , 50% 2S_0). The biological activity of a decasaccharide almost identical in monosaccharide sequence to the one shown above was tested for biological activity by Goodger (Goodger 2003),

decasaccharide D, Figure 77. The only difference between the two saccharides being that Goodger was not able to establish the epimeric identity of the monosaccharide at sequence position 9, this may have been either IdoUA or GlcUA (see Figure 77). The decasaccharide assayed by Goodger **was** able to elicit a mitogenic response from FGF-2 in combination with the FGFR1 IIIc receptor. A similar decasaccharide (decasaccharide F, Figure 77) shown below **was not** able to elicit a biological response with FGF-2 in the same assay system.

Decasaccharide D, biologically active (see Figure 77)



Decasaccharide F, biologically inactive (see Figure 77)



Note that the 2-O-sulphate group at monosaccharide position 3 is absent from the biologically inactive decasaccharide shown above. The 6-O-sulphate group added at monosaccharide position 10, may or may not promote biological activity. Both inactive and active decasaccharides were found by Goodger that contained this terminal 6-O-

sulphate group (Goodger 2003), see Figure 77. It could be argued that the 2-O-sulphate group itself, removed from monosaccharide position 3, rather than the conformation of the iduronate is important for biological activity. However, out of the total of 4 biologically active deca-saccharides identified by Goodger two did not contain a 2-O-sulphate group at monosaccharide position 3, see Figure 77. These same two deca-saccharides also had a GlcUA at monosaccharide position 3; GlcUA being stable in the 4C_1 conformation. The possible significance of this is not clear until the sequence of the other two biologically active deca-saccharides is examined. Both contain IdoUA(2S) at monosaccharide position 3 flanked by two solely N-sulphated glucosamines. IdoUA(2S) in such a sequence could be predicted to be at >70% 1C_4 (see Table 14). It may be that the presence of a chair conformation (either 4C_1 or 1C_4) at monosaccharide position 3 of the deca-saccharides tested by Goodger, is one of the key features that bestow biological activity.

An analysis of the monosaccharide sequences shown in Figure 77 for conserved patterns of sulphate group positioning identifies that all active deca-saccharides carry N-sulphate groups at all non-terminal glucosamine sugars. A conserved 2-O-sulphate group also occurs at monosaccharide position 5. However, these two conserved features are also present in the two non-biologically active deca-saccharides 'E' and 'F', see Figure 77. This suggests other structural features contained within the active oligosaccharides may also be important for biological activity. In all active deca-saccharides, 6-O-sulphation at monosaccharide positions 5 and/or 10 in combination with a predominance of either the 1C_4 or 4C_1 conformation at monosaccharide position 3 also appears to be conserved. Given this, the question then arises as to why deca-saccharide E is not biologically active.

6-O- sulphation at monosaccharide position 8 within decasaccharide E appears to knock out the biological activity of what otherwise could be assumed to be an active saccharide (compare biological activities of decasaccharides E and D). Yet the same 6-O-sulphation at monosaccharide position 8 within decasaccharide A does not inactivate this oligosaccharide. The influence of 6-O-sulphation at position 8 to render some decasaccharide sequences inactive, whilst leaving the biological activity of others intact, implies that there may be another as yet unknown structural feature at work within these oligosaccharides. Discussed in the forthcoming sections are work that was done in order to try and establish other structural features that may play a role in influencing the biological activity of the Goodger decasaccharides.

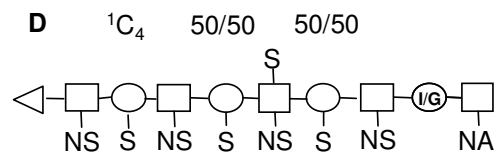
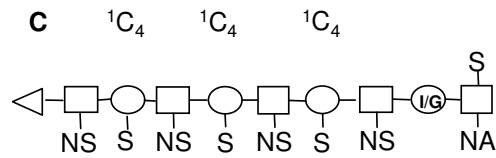
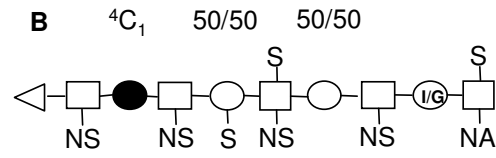
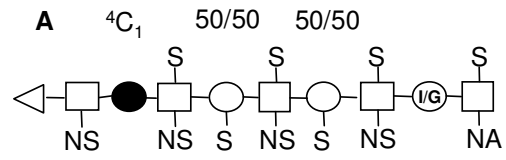
Figure 77 The monosaccharide sequences of 4 biologically active and 2 biologically inactive decasaccharides.

The decasaccharide sequences shown were assayed for activity in a FGF-2, FGFR1 IIIc Baf cell mitogenesis assay. This Figure was adapted from a similar Figure shown in the PhD thesis of Goodger (Goodger 2003). Monosaccharide positions within the decasaccharide are numbered 1-10. Also labelled, as deduced from Table 14, are the dominant conformation of each iduronate monosaccharide. The conformation of each glucuronic acid is also labelled as 4C_1 although this does not vary. For iduronates, where no conformation predominates this monosaccharide is labelled 50/50 (50% 1C_4 , 50% 2S_0) Symbols codes are as in Table 14, with the addition of the symbol (I/G) to signify a hexuronic acid (HexUA) of an unknown epimeric identity (either IdoUA or GlcUA).

Figure 77

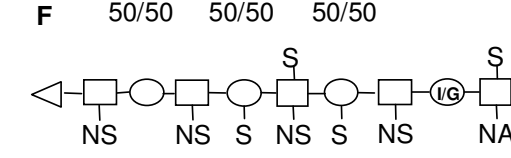
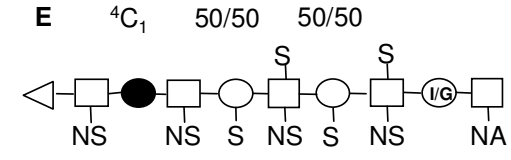
Active with FGF-2 and FGFR1 IIIc

1 2 3 4 5 6 7 8 9 10



Inactive with FGF-2 and FGFR1 IIIc

1 2 3 4 5 6 7 8 9 10



18. Structural features in addition to sulphation pattern and iduronate conformation may influence biological activity

In order to explain variations in the biological activity of the Goodger decasaccharides, the structures of all HS dp6 models were investigated in detail. For the purposes of this investigation, features such as the presence or absence of a particular sulphate group or a preference for iduronate in one conformation or another were dis-regarded, although they are undoubtedly important. It was thought that any additional structural features that occur within these dp6 oligosaccharides may also occur within the dp10 oligosaccharides of Goodger. The oligosaccharides of Goodger could not be modelled and studied directly as NMR data does not exist for these oligosaccharides with which to support possible structural conclusions.

19. A Comparison of HS dp6 structural models

It was thought that the differences in biological activity may be due to differences in the overall three dimensional shape of an oligosaccharide, changing as monosaccharide sequence changes. Figure 78 was constructed in order to assess the extent of such structural differences within the HS dp6 structural models. As can be seen in Figure 78, apart from obvious variations due to a presence or lack of a particular sulphate group and structural variations around the reducing terminal N-acetyl group of HS dp6-3 (previously discussed, see section 14), very little differences occur within the overall three dimensional shape of each oligosaccharide. This lead to the idea that molecular dynamic differences may exist between these oligosaccharides.

Figure 78 A comparison of of the molecular shape of all HS dp6 structural models.

Models in which electronegative groups are shown in red contain all iduronates in the 1C_4 conformation (models A-D). Models in which electronegative groups are shown in yellow contain all iduronates in the 2S_0 conformation (models E-H)

Models A and E, HS dp6-2,

Models B and F, HS dp6-3,

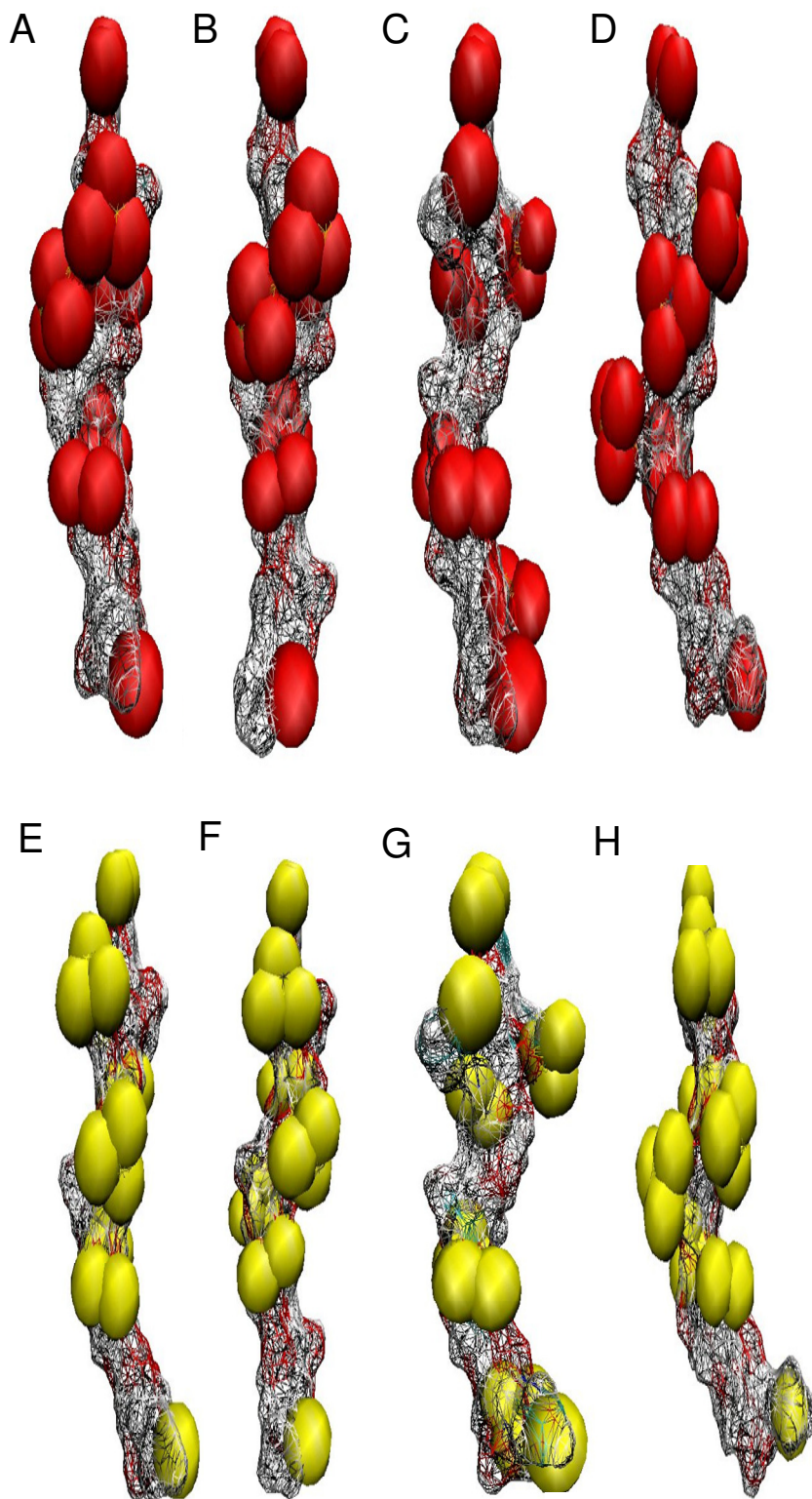
Models C and G, HS dp6-5,

Models D and H, HS dp6-5

Model representations are as in Figures 70 and 71.

All models were formatted using the VMD molecular graphics program (Humphrey *et al.* 1996).

Figure 78



20. Molecular dynamic studies

The construction of static models based on NMR data represents an average structure. This static structure does not allow for the fact that for a proportion of its time in solution, a molecule may exist or 'sample' a number of alternate conformations. As detailed in the results section, most dynamically modelled glycosidic linkage geometries remained close to their starting values or moved to 'sample' values that were also suggested by the structural refinement process. However, in one case glycosidic linkage geometries sampled an additional area of their potential energy surface that had previously been ruled out by the structural refinement process, see panel C2, Figure 64 (note, the structural refinement process only took into consideration IdoUA(2S) H1 NOE cross-peaks). The results from this panel show that when using charge set 2, the IdoUA(2S)–GlcNS internal disaccharide unit of HS dp-2, moves away from its initial geometry of $\Phi = 40$, $\Psi = 20$ to sample an area of its potential energy surface around the torsion angle $\Phi = -10$, $\Psi = -20$. For the purposes of this discussion the initial $\Phi = 40$, $\Psi = 20$ geometry is termed geometry 'a' and the $\Phi = -10$, $\Psi = -20$ geometry 'b'. If a molecular model is constructed for HS dp6-2 with the internal IdoUA(2S)–GlcNS linkage in geometry 'b' ($\Phi = -10$, $\Psi = -20$), all that seems to occur on a first analysis is a minor re-positioning of the carboxyl group at the non-reducing terminal (arrowed in Figure 79). If however, this hexasaccharide is extended in length, to a dodecasaccharide similar to that as modelled by Mulloy *et al.* (Mulloy *et al.* 1993), the supposed non trivial nature of this change becomes apparent. See Figure 80. A minor change in the local conformation of the oligosaccharide around the IdoUA(2S)-GlcNS linkage, arrowed in Figure 80, causes a small deviation or kink in the oligosaccharide chain away from its usual conformation.

As the chain increases in length away from the glycosidic linkage at which the kink occurs, its effects become more and more dramatic. Raman *et al.* have previously reported on the existence of such a chain kink within the trisaccharide sequence – GlcNS(6S)-IdoUA(2S)-GlcNS(6S)- present within all x-ray crystallography determined heparin-FGF1 and FGF-2 complexes. Given its occurrence in all complexes studied, Raman *et al.* speculated that this kink may be an essential feature of heparin/HS-FGF interactions (Raman *et al.* 2003). Raman *et al.* also concluded that after the occurrence of a kink the chain is restored to its original helical conformation, Figure 80 generally supports this conclusion. However, if no compensatory adjustments are made in the geometry of subsequent glycosidic linkages, the original course of this helical structure through three dimensional space, is altered by the occurrence of a chain kink. This has important consequences for the relative positioning of disaccharide units at opposite ends of an oligosaccharide chain. This may be an important piece of structural/functional information previously overlooked by Raman *et al.*

Experimental proof for the existence of geometry b is present in the NOESY spectra of HS dp6-2 and HS dp6-3. If geometry 'b' were to exist in solution IdoUA(2S) H2-GlcNS H6 and/or H4 NOE cross-peak(s) should be clearly detectable. On examination of the HS dp6-2 and HS dp6-3 NOESY spectra just such NOE cross-peaks are detected, see Figure 81. It was thought that similar kinks may occur within the decasaccharide structures of Goodger.

Figure 79

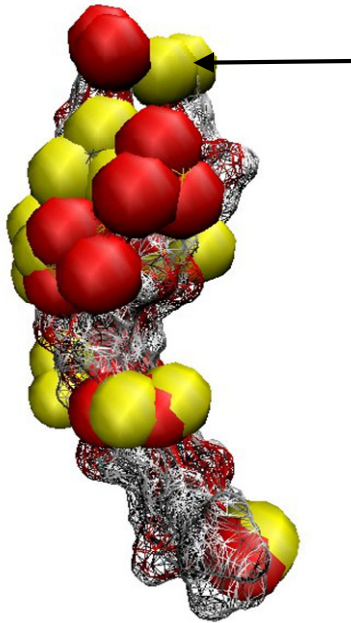


Figure 79 A second superimposition model of HS dp6-2.

Two models of HS dp6-2 are superimposed. In both models all internal iduronates are in the 1C_4 conformation. In one model (Van der Waals representations in red) the internal in IdoUA(2S) –GlcNS linkage is in a $\Phi= 40, \Psi= 20$ geometry. In the second model (Van der Waals representations in yellow) the internal in IdoUA(2S) –GlcNS linkage is in a $\Phi= -10, \Psi= -20$ geometry (geometry b).

Superimposition alignment was carried out using the rmsd extension of the VMD molecular graphics program (Humphrey *et al.* 1996).

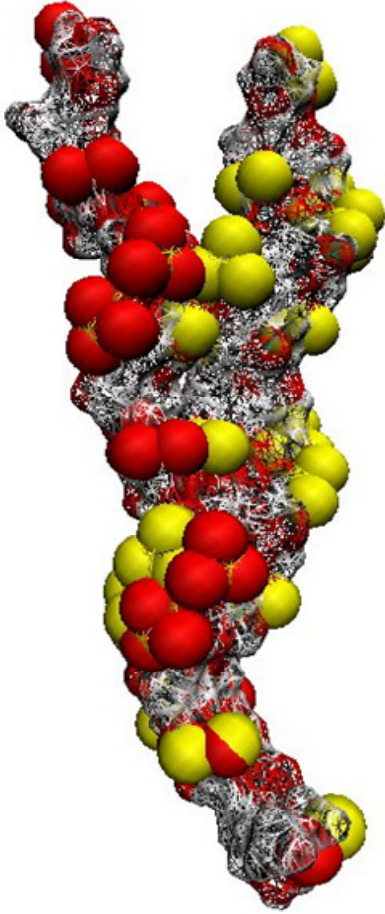
Figure 80 An extended dodecasaccharide superimposition model of HS dp6-2.

The terminal UA unit of the two dp6-2 models superimposed in Figure 79 were removed and the oligosaccharides extended by the addition of the sequence -IdoUA(2S)¹C₄-GlcNS-IdoUA(2S)¹C₄-GlcNS-IdoUA(2S)¹C₄-GlcNS-UA²H₁. All geometries for these linkages were taken from Table 11 and as such are consistent with NOE data. The two models were then superimposed, with structural alignment based on the first five residues of the original hexasaccharide models. In one model (Van der Waals representations in red, model A, whole molecule in red, model B) the internal IdoUA(2S) –GlcNS linkage arrowed in model B is in a $\Phi = 40$, $\Psi = 20$ geometry. In the second model (Van der Waals representations in yellow, model A, whole molecule in yellow, model B) the internal IdoUA(2S) –GlcNS linkage is in a $\Phi = -10$, $\Psi = -20$ geometry.

Superimposition alignment was carried out using the rmsd extension of the VMD molecular graphics program (Humphrey *et al.* 1996).

Figure 80

A



B

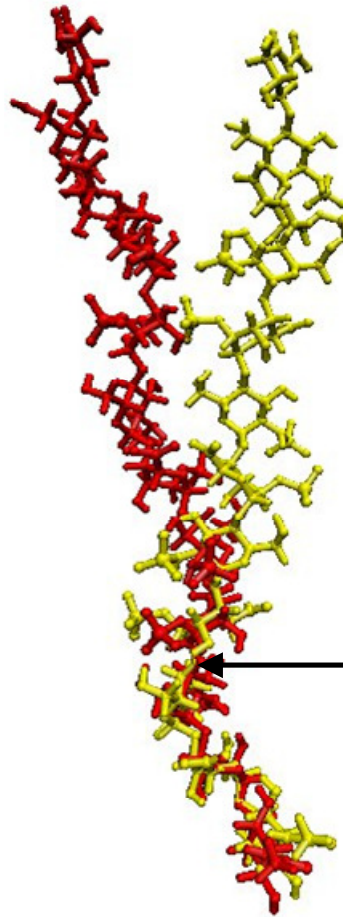


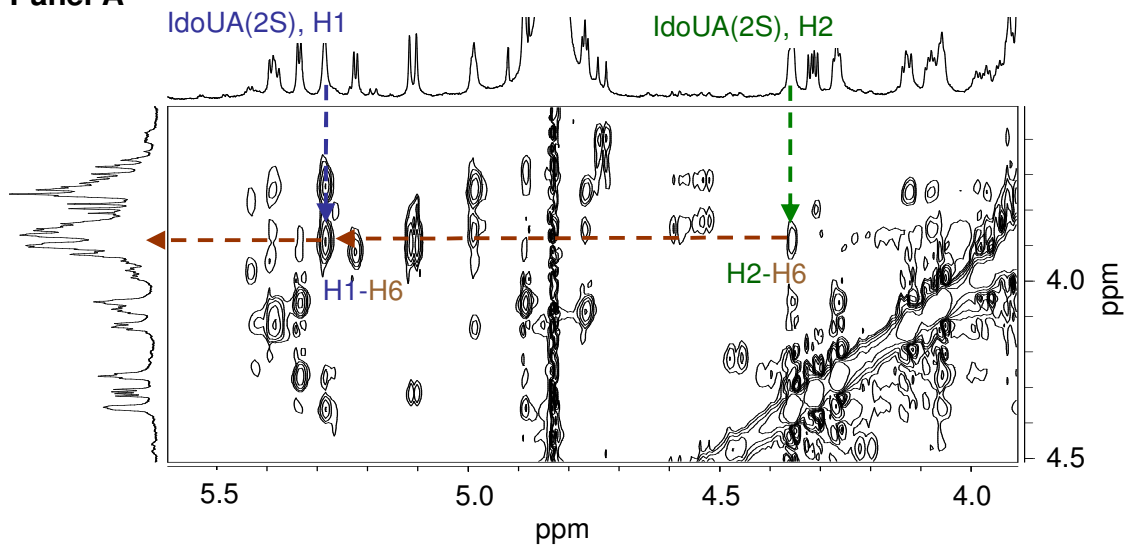
Figure 81 Enlarged sections from the NOESY spectra (500ms mixing time) of **HS dp6-2 and HS dp6-3.**

In both spectra trans-glycosidic NOE cross-peaks can be detected from both the anomeric proton of the IdoUA(2S) residue and proton H2 of this same ring system. For HS dp6-2 (panel A) a H2 cross-peak is only observed to a H6 proton of the preceding glucosamine. For HS dp6-3 (panel B) cross-peaks are observed between the H2 proton a H6 proton and a H4 proton of the same preceding glucosamine.

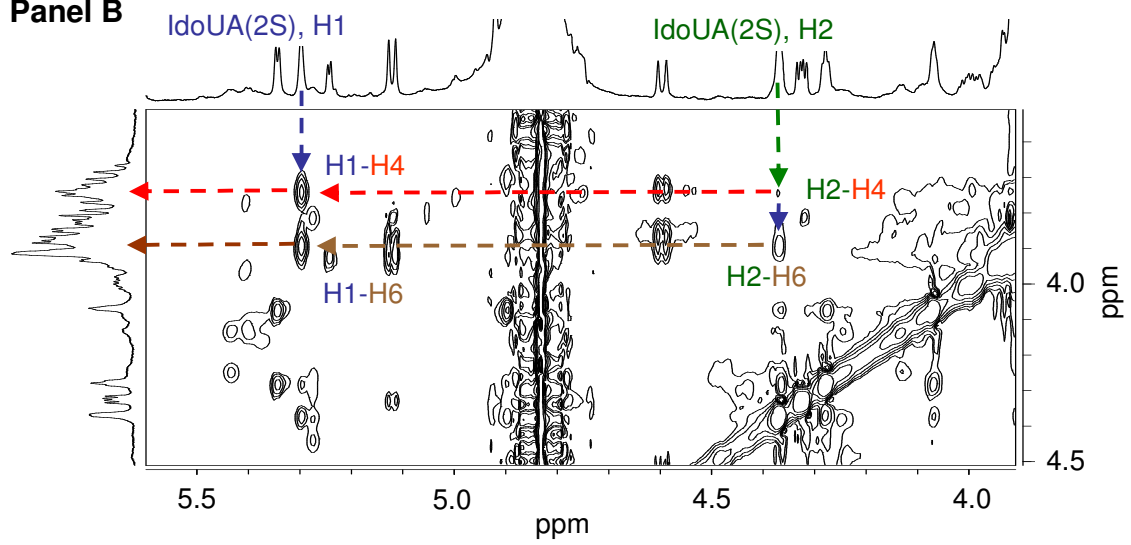
The colours and annotation scheme are as in results sections 10-11

Figure 81

Panel A



Panel B



21. Molecular dynamic evidence that internal 6-O-suphation may cause an increased sampling rate of geometry ‘b’

The molecular dynamic data was studied to try and establish structural features which may induce kink formation. Here after when oligosaccharides are referred to as containing a kink, it should be understood that this kink is a dynamic feature of the chain and that oligosaccharides may move in and out of this kinked conformation with differing frequencies in solution. Given that all heparin-FGF complexes contain a chain kink (Raman *et al.* 2003) (discussed in section 20), it may be the ability of certain structural features to increase the conformational lifetime of the kinked conformation over the conformational lifetime of the straight conformation that may be biologically relevant.

During dynamic simulations on HS dp6-6 the terminal IdoUA–GlcNAc linkage appeared to sample two glycosidic linkage geometries that would lead to a kinked oligosaccharide chain, geometry ‘b’ ($\Phi = -10^\circ$, $\Psi = -20^\circ$) and an additional geometry termed geometry ‘c’ $\Phi = -45^\circ$, $\Psi = -45^\circ$ (see Figure 67, panel E). A similarly modelled IdoUA–GlcNAc within HS dp6-2 (see Figure. 64 panel E1) made no such conformational transitions. The IdoUA-GlcNAc disaccharide forms part of a different terminal trisaccharide sequence in both HS dp6-6 and HS dp6-2, as shown below:

-GlcNS-IdoUA-GlcNAc	HS dp6-2
-GlcNS(6S)-IdoUA-GlcNAc	HS dp6-6

All molecular dynamic calculations carried out as part of this thesis work were deliberately set up so that each monosaccharide was only influenced by charges from its immediate neighbour

monosaccharides. This was done by using an electrostatic cut-off value of 6Å (see experimental procedures). Given this, the 6-O-sulphate group present in the trisaccharide sequence of HS dp6-6 is a strong candidate for altering the IdoUA-GlcNAc dynamics within HS dp6-6 when compared to HS dp6-2. The influence of this internal 6-O-sulphate group may indeed be direct, by through space interactions with atoms involved in the –IdoUA-GlcNAc linkage. However, given that all dynamics calculations were carried out with explicit water molecules, it may also interrupt local hydrogen bonding networks between the water molecules and the –IdoUA-GlcNAc disaccharide unit that may stabilise the $\Phi = 40^\circ$, $\Psi = -40^\circ$ conformation.

Unfortunately, due to the low amount of sample material available for HS dp6-6, it is not possible to verify if the alternate geometries exist experimentally. An IdoUA H2-GlcNAc H6 and/or H4 NOE should be clearly detectable for these geometries.

22. PAGE evidence for a link between internal 6-O-sulphation and a chain kink

Vives *et al.* have previously purified HS dp6-6 and HS dp6-2 from an alternate source of porcine HS (Vives *et al.* 2001). The migration position on PAGE analysis for both of these oligosaccharides was found to be significantly different, HS dp6-6 migrating at a slower rate through the gel than HS dp6-2. Similar results were obtained as part of this thesis work, see results, Figure 24.

The sequences of HS dp6-6 and HS dp6-2 are given below

	1	2	3	4	5	6
HS dp6-6	UA-GlcNS-IdoUA(2S)-GlcNS(6S)-IdoUA-GlcNAc					
HS dp6-2	UA-GlcNS-IdoUA(2S)-GlcNS-IdoUA-GlcNAc					

Ignoring the influence of molecular shape in the general theory of PAGE electrophoresis given in a standard biochemistry text such as (Mathews and van_Holde 1990), HS oligosaccharides with more negative charge density migrate through the gel at a faster rate than oligosaccharides with less charge density, the more negatively charged molecule experiencing a stronger pull towards the anode electrode. In this modified theory, it is a puzzle as to why HS dp6-6 migrates at a slower rate than HS dp6-2, HS dp6-6 does after all have a higher negative charge density. Bringing back in the influence of molecular shape on PAGE theory, asymmetric molecules encounter a higher degree of frictional resistance as they move through the gel than symmetrical molecules (Mathews and van_Holde 1990). The molecular dynamic data suggest that for a proportion of its conformational lifetime HS dp6-6 may contain a kink. This kink may then hinder its progress through the pores of the polyacrylamide gel. The NMR and computational data suggest HS dp6-2 may also kink but at a significantly reduced rate to that of the HS dp6-6 molecule. As a consequence of this, HS dp6-2 is able to migrate at a faster rate through the pores in the polyacrylamide gel. Interestingly, if 'kink-rate' can be related to PAGE migration distance in this way it can be predicted from Figure 24 that HS dp6-5 kinks at a much slower rate than either of HS dp6-2 or HS dp6-6 (compare Figure 24 migration distances for each oligosaccharide). This implies that the -GlcUA-GlcNS-disaccharide unit centrally positioned within HS dp6-5 may act as a stabilising feature against chain kinks induced by 6-O-sulphation.

23. Implication of the link between chain kinks and 6-O-sulphation for the biological activity of the Goodger deca-saccharides

From Figure 77, the fact that oligosaccharide 'C' is biologically active suggests an increased 'kink-rate' in free solution is not an essential feature for biological activity. Oligosaccharide 'C' does not contain an internal 6-O-sulphate group. With regard to the observation made by Raman *et al.*, that all heparin oligosaccharides in complex with FGF molecules contain a chain kink, the fact that oligosaccharide 'C' is biologically active suggests chain kinks, additional to their occurrence in free solution, may also be induced on protein binding. The question then is 'could a raised 'kink-rate' in the wrong part of an oligosaccharide chain prior to protein binding knock out its biological activity?' The prime target for such a kink to occur is between monosaccharides 9 and 10 of all the Goodger deca-saccharides. The 6-O-sulphation of monosaccharide 8 as in the inactive oligosaccharide 'E' may be expected to increase the rate at which a kink may occur in free solution between these two residues. However, 6-O-sulphation also occurs at monosaccharide 8 within deca-saccharide 'A' but yet this oligosaccharide is still biologically active. Two reasons are suggested for this

- 1) The reducing terminal unit (unit 10) of deca-saccharide 'A' carries a 6-O-sulphate group. In free solution, when this unit is 6-O-sulphated, it may be that 6-O-sulphation of monosaccharide 8 is able to exert less of a conformational influence on the dynamics of the glycosidic bond between monosaccharides 9 and 10. Further molecular dynamic studies will be able to ascertain if this is the case.
- 2) The epimeric identity of monosaccharide 9 has not been identified. It may be that this is a glucuronic acid residue. Implied evidence from the migration position of HS dp6-5 on PAGE suggests a -GlcUA-GlcNAc(6S) terminal disaccharide may display different

dynamic properties to a terminal –IdoUA-GlcNAc(6S) disaccharide. Again further molecular dynamic studies will ascertain if this is the case.

24. Experimental evidence for kinked heparin oligosaccharides in free solution?

Given the experimental evidence for the existence of kinks in free solution within HS oligosaccharides, the NOESY spectra of all heparin hexasaccharides produced as part of this thesis work were also examined for similar experimental evidence. Due to severe spectral overlap in the area of the spectrum where IdoUA(2S)-H2 protons and GlcNS(6S)-H6/H4 protons occur, it was not possible to establish if geometry ‘b’ exists within the heparin hexasaccharides produced as part of this study. Spectral overlap can be expected to be a problem for all heparin or heparin derived oligosaccharides where glucosamines are 6-O-sulphated. 6-O-sulphation shifts all glucosamine H6 and H5 proton resonances downfield to an area of the spectrum where the H2 resonance of IdoUA(2S) occurs. Perhaps for this reason experimental evidence for the existence of geometry ‘b’ and kinked heparin chains in free solution has never previously been shown. Trans-glycosidic NOEs were not analysed in the de-6-O-sulphated heparin study of Mulloy *et al.* (Mulloy *et al.* 1994). If they were this study may have picked up the existence of IdoUA(2S) H2- GlcNS H6/H4 cross peaks before they were detected within HS dp6-2 and -3 as part of this thesis work.

25. Overall summary

It is felt that the original aims of this study have been met and surpassed. It has been demonstrated that the considerable experimental difficulties encountered when working with

incredibly heterogeneous HS samples can be overcome and valuable structural data gathered for authentic HS derived oligosaccharides, in addition to heparin equivalents.

The NMR and molecular modelling data, has enabled significant new insights to be gained into the role sulphation patterns may play in influencing oligosaccharide conformation and more importantly function. For the first time a role has also been suggested for N-acetyl groups in HS-protein electrostatic interactions. N-acetyl groups are normally overlooked as a contributor to electrostatic interactions due to their low electronegativity in comparison to sulphate and carboxyl groups. The modelling studies undertaken as part of this study suggest they may be overlooked in error and appropriately positioned N-acetyl groups may also be an important contributor to HS-protein electrostatic interactions and biological activity.

It has been known for some time that the balance of terminal uronic acid and internal iduronate equilibria are influenced by the sulphation state of their neighbouring N-glucosamine residues. However, a number of significant gaps were present in the literature concerning the analysis of iduronate residues in sequences such as –GlcNS-IdoUA-GlcNS- and –GlcNS-IdoUA-GlcNS(6S)- and –GlcNS-IdoUA(2S)-GlcNS-. These are common saccharide sequences found within HS but not heparin. Iduronate conformational information gathered as part of this thesis work has now filled these gaps. Using Table 14 it is now possible for the first time to predict the probable conformational balance of iduronate residues within almost all sulphation sequences in which they may occur. Using a library of decasaccharide structures tested for biological activity (Goodger 2003), it has been shown how this knowledge can lead to new insights into iduronate conformation and oligosaccharide functional relationships.

Molecular dynamic modelling studies have revealed that glycosidic linkages may not be fixed permanently in geometries favoured as a result of the exo-anomeric effect. Static molecular modelling has shown that when alternate geometries occur the effect is to 'kink' the chain away from its usual 3-dimensional topology. NOE data supports the existence of just such a kink within HS dp6-2 and HS dp6-3. Molecular dynamic studies suggest that the rate at which kink formation occurs at a particular glycosidic bond can be related to the presence of a nearby 6-O-sulphate group. Finally and most importantly, it has been suggested how the existence of such kinks may be tolerated in certain parts of an oligosaccharide chain and not in others if an oligosaccharide is to possess biological activity.

APPENDIX I

Figure 82

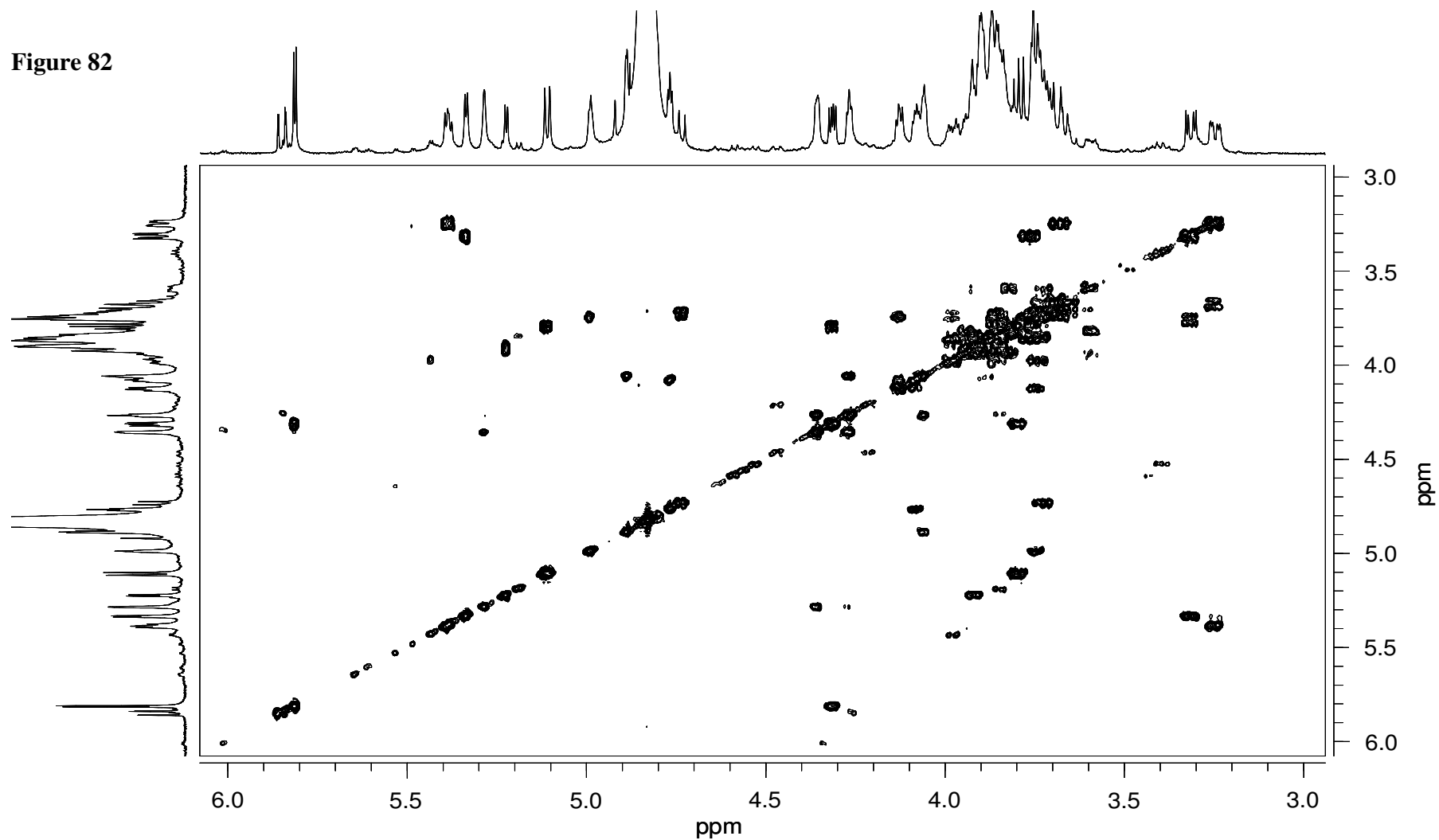


Figure 82 Two-dimensional 500 MHz ¹H-NMR COSY spectrum, HS dp6-2.

Figure 83

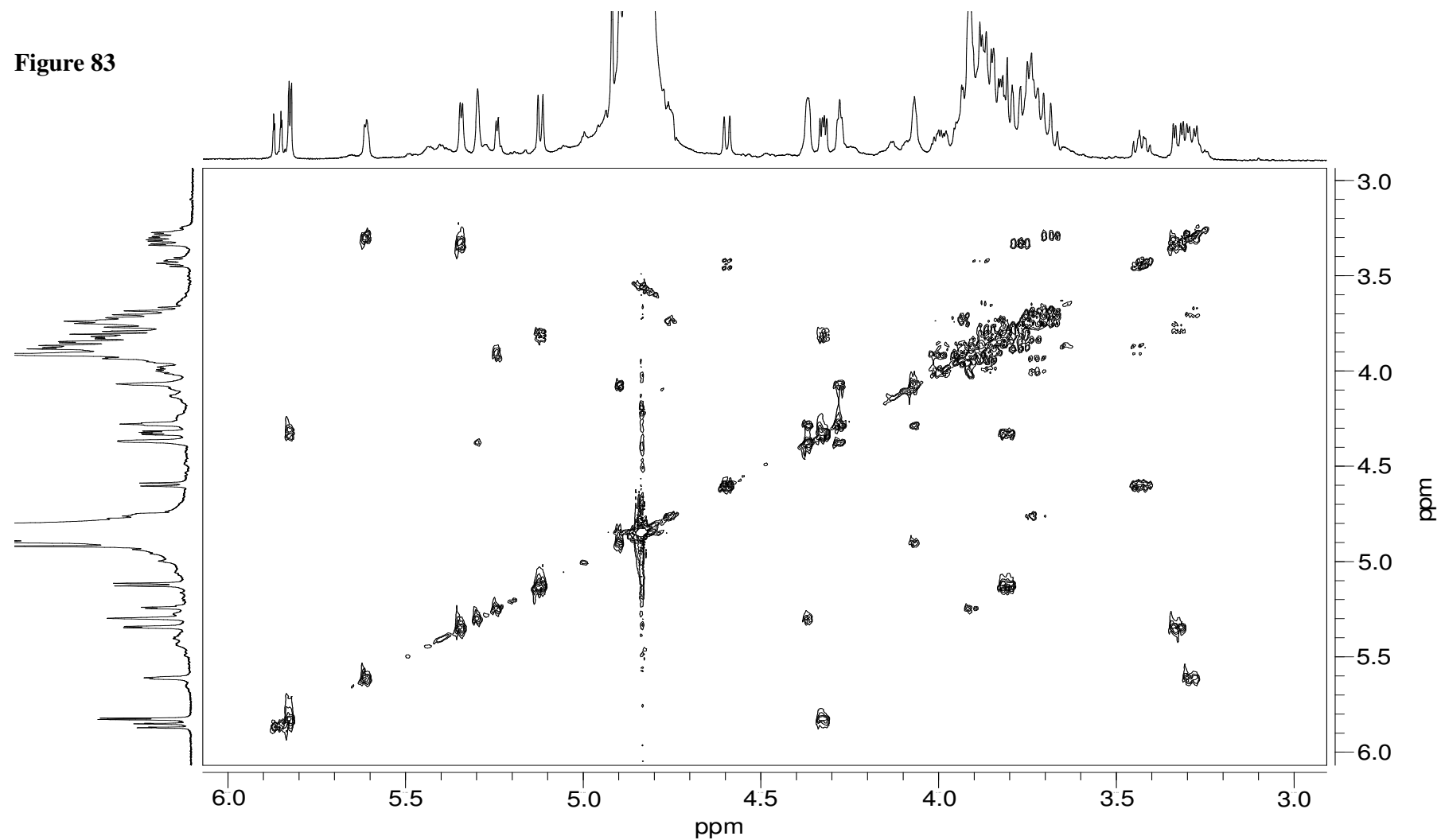


Figure 83 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, HS dp6-3.

Figure 84

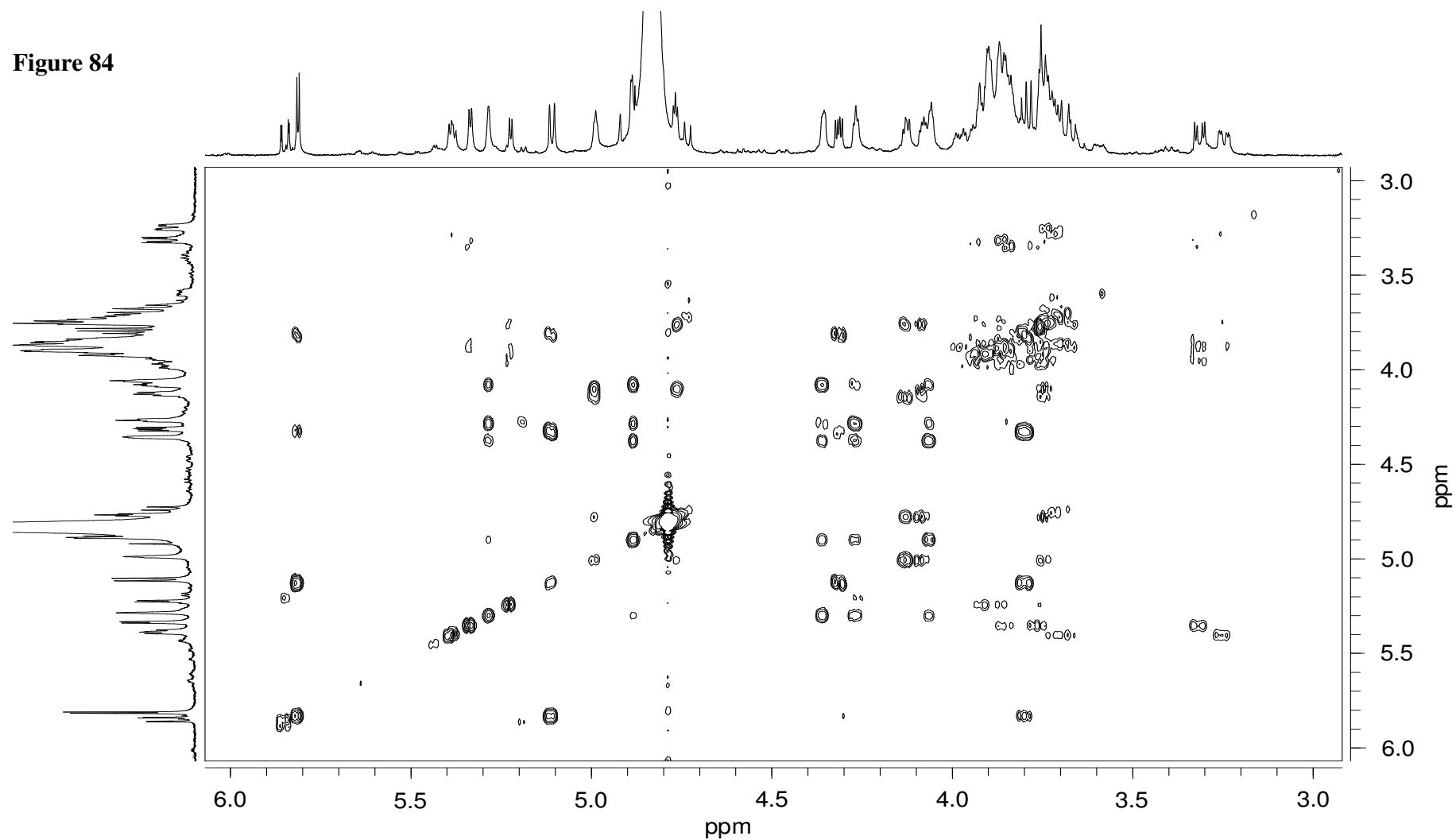


Figure 84 Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, HS dp6-2. Mixing time 120ms.

Figure 85

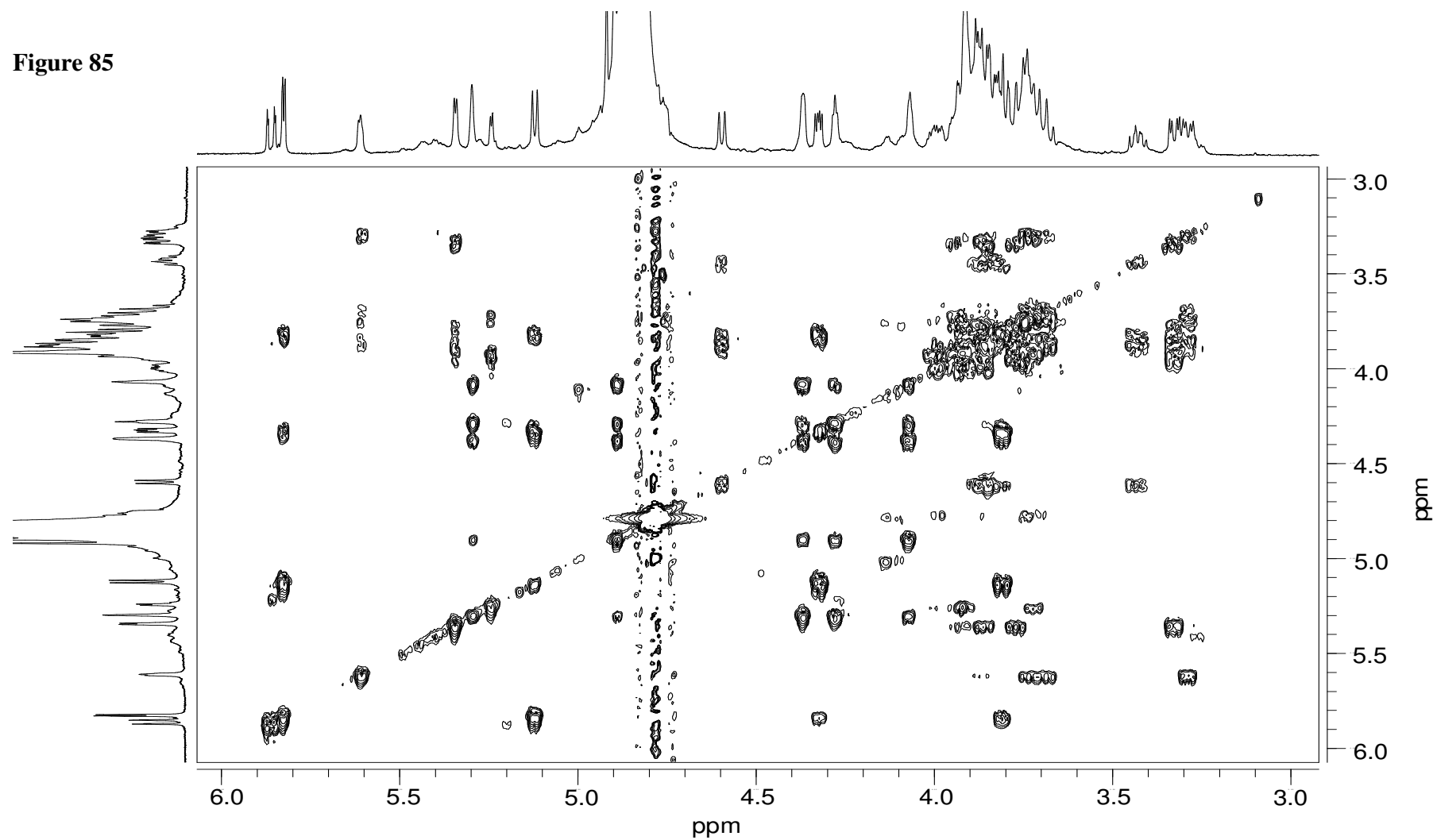


Figure 85 Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, HS dp6-3. Mixing time 120ms.

Figure 86

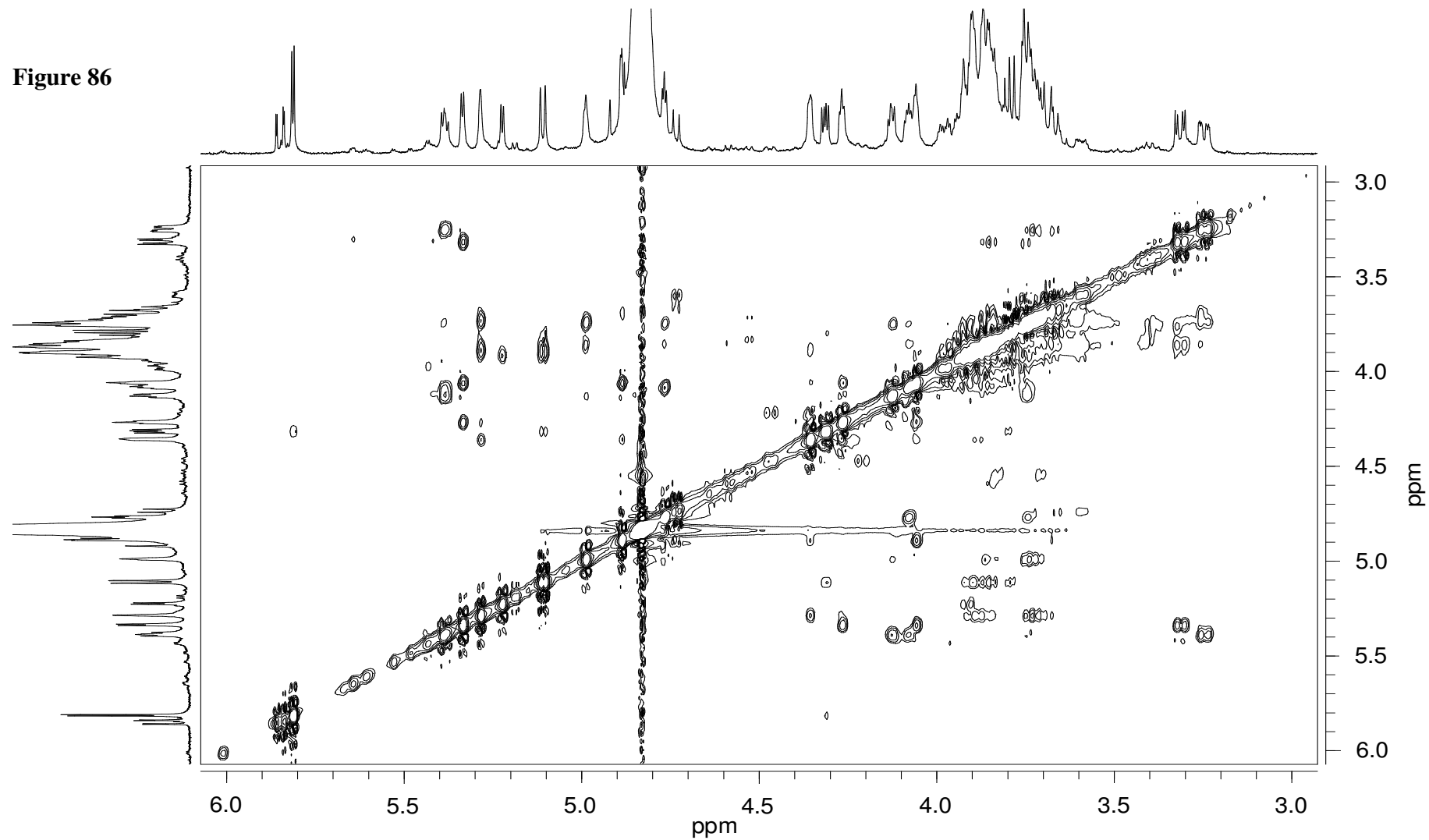


Figure 86 Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-2. Mixing time 500ms.

Figure 87

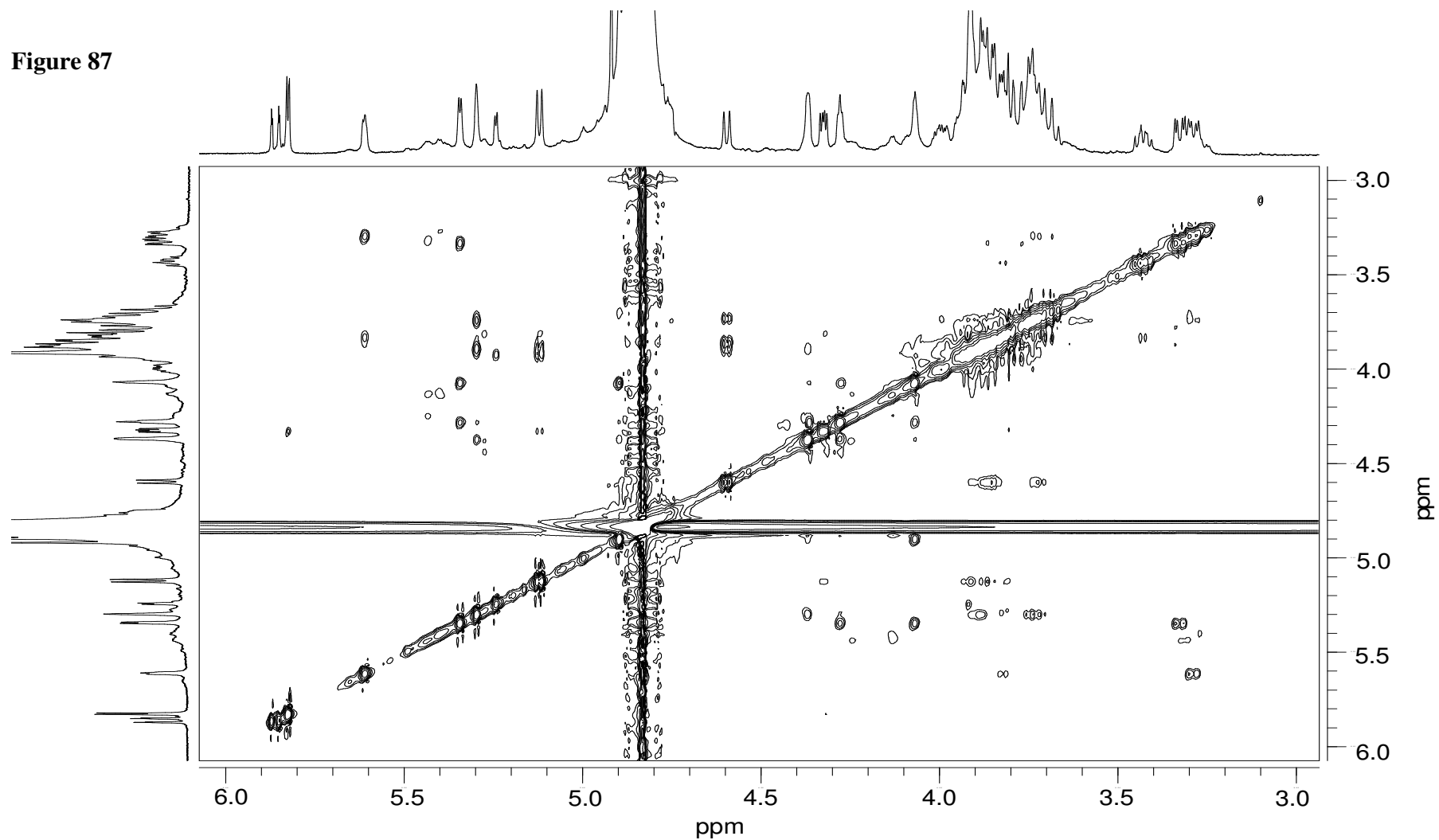


Figure 87 Two-dimensional 500 MHz ¹H-NMR NOESY spectrum, HS dp6-3. Mixing time 500ms.

Figure 88

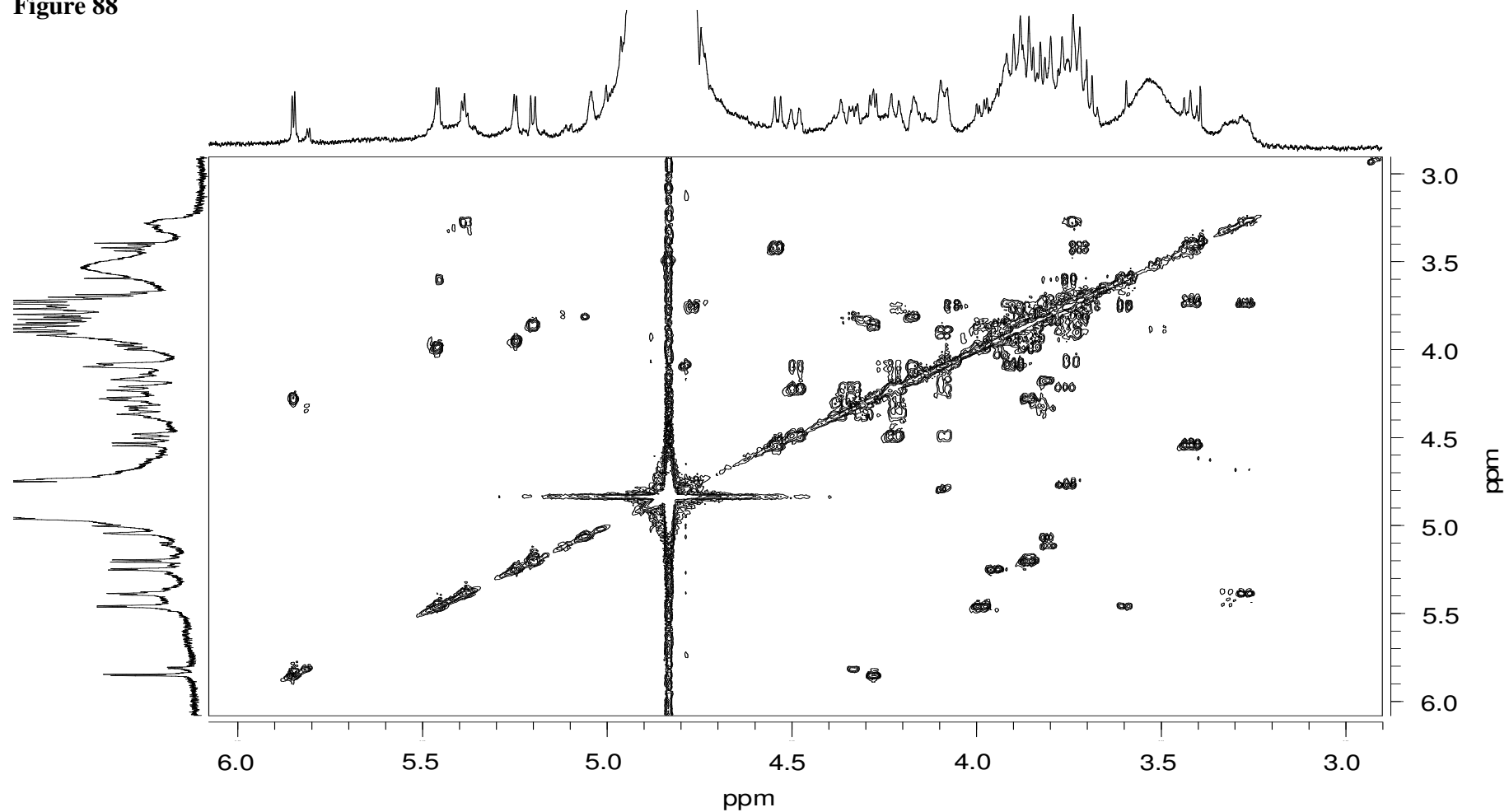


Figure 88 Two-dimensional 500 MHz ¹H-NMR COSY spectrum, HS dp6-5.

Figure 89

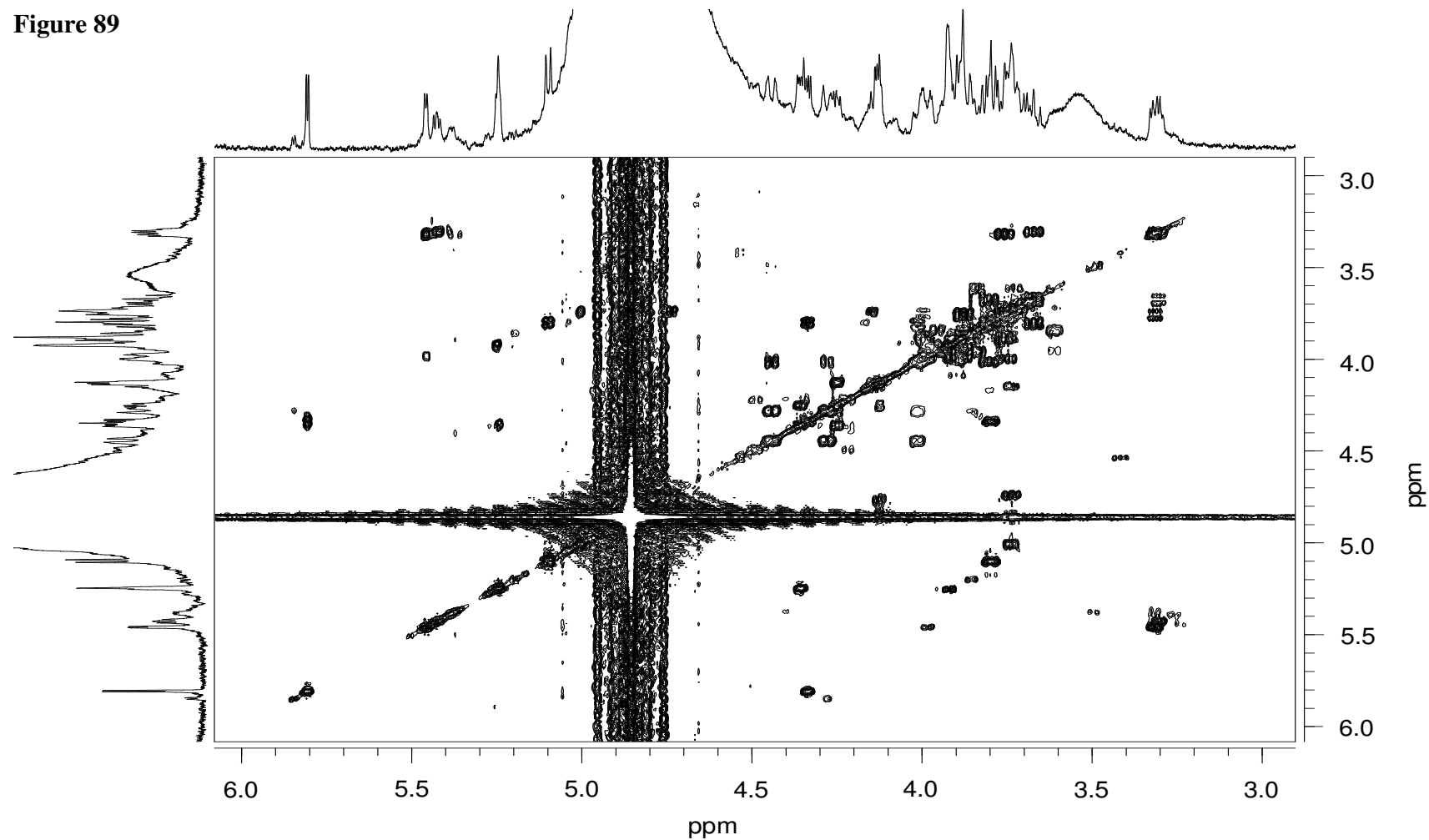


Figure 89 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, HS dp6-6.

Figure 90

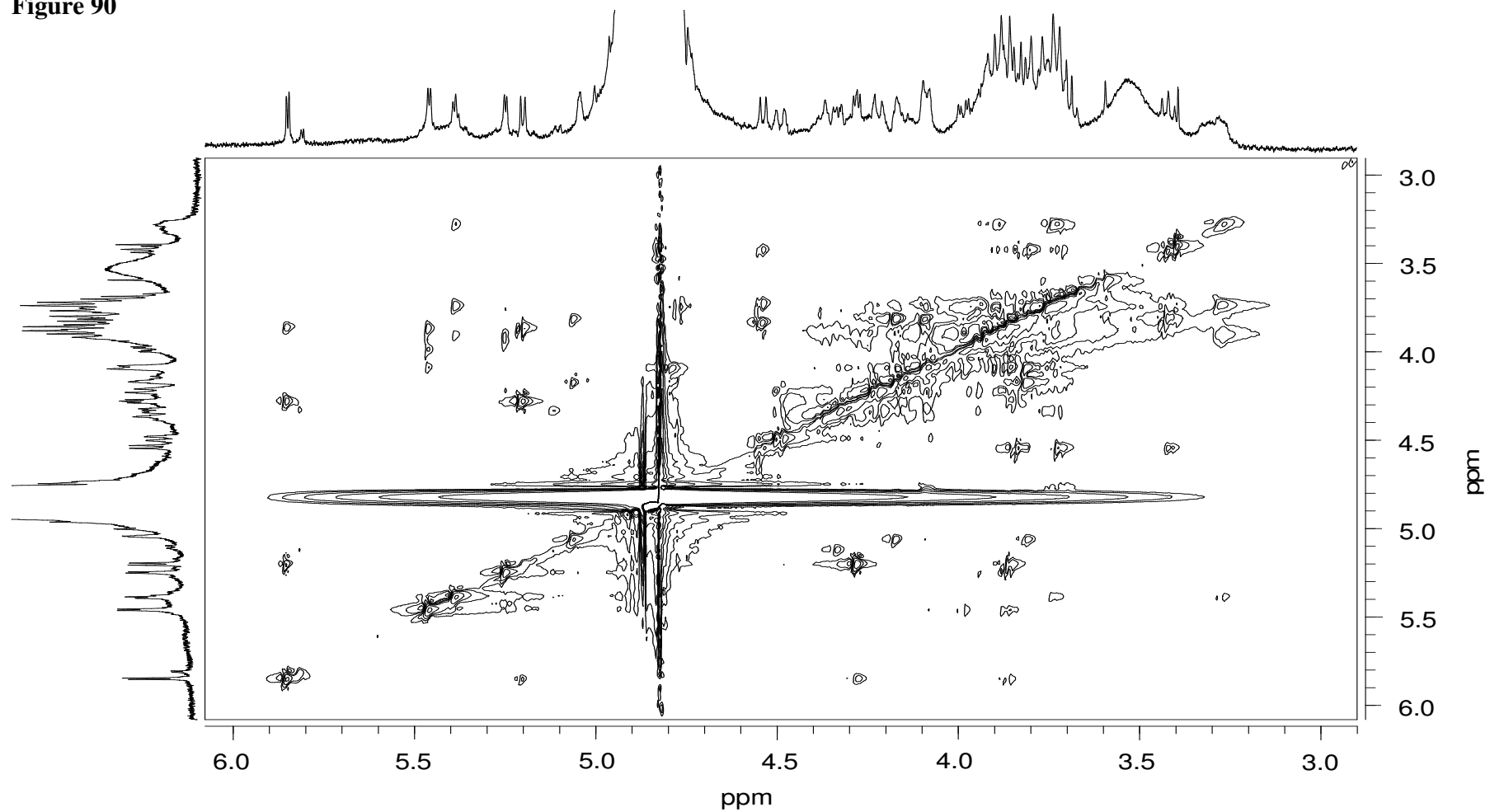


Figure 90 Two-dimensional 500 MHz ¹H-NMR TOCSY spectrum, HS dp6-5. Mixing time 80ms

Figure 91

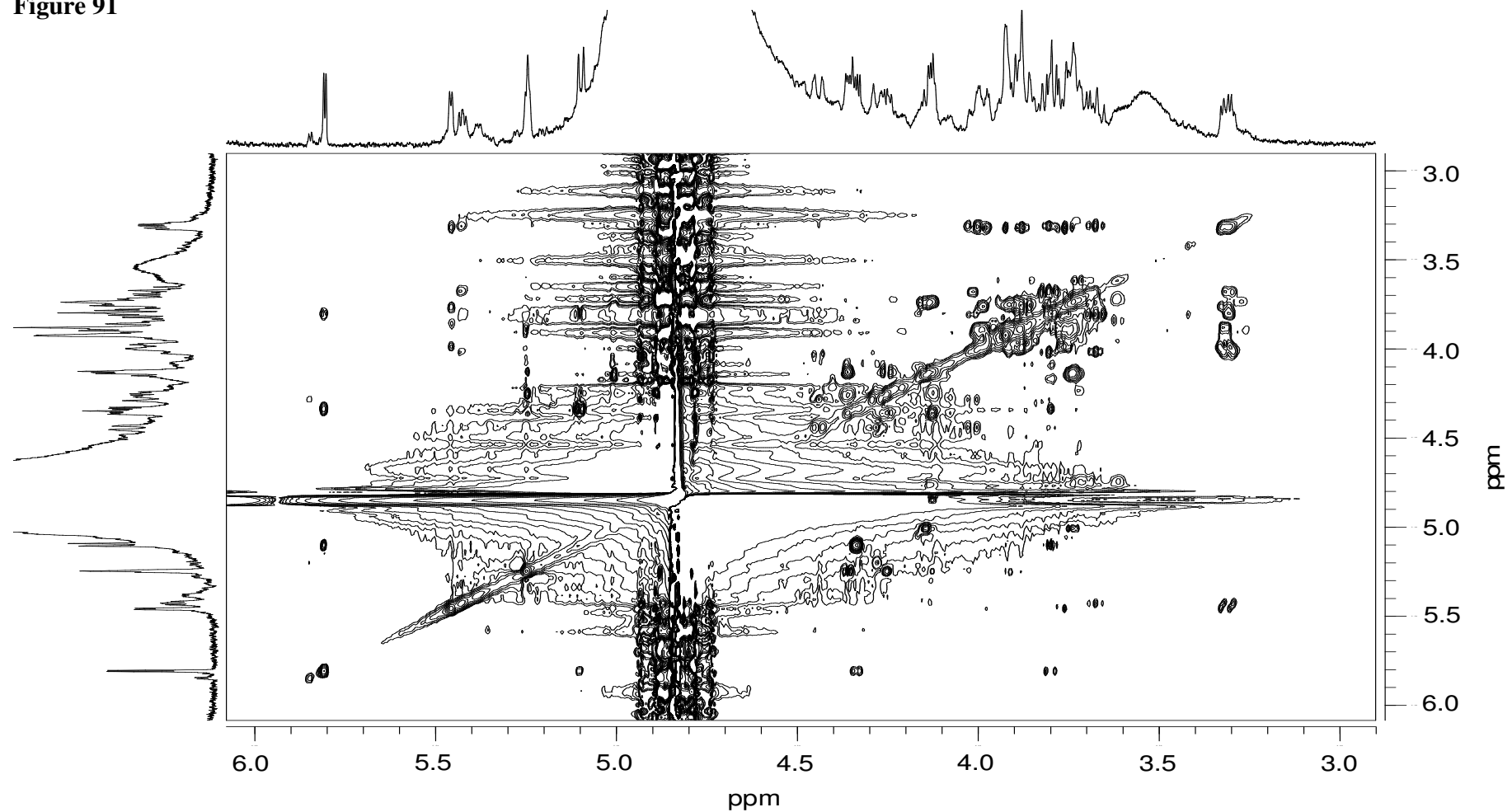


Figure 91 Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, HS dp6-6. Mixing time 80ms

Figure 92

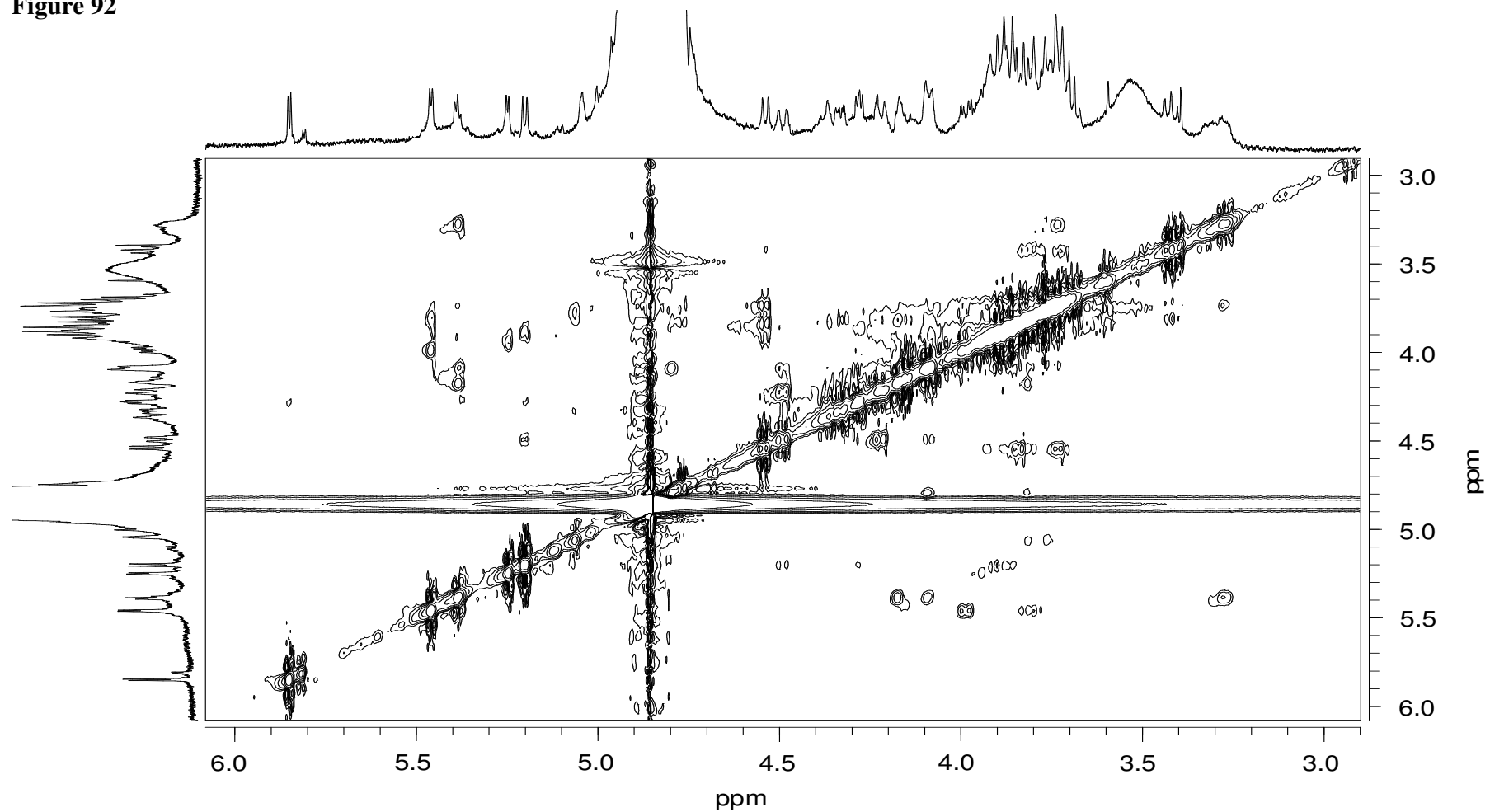


Figure 92 Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-5. Mixing time 500ms

Figure 93

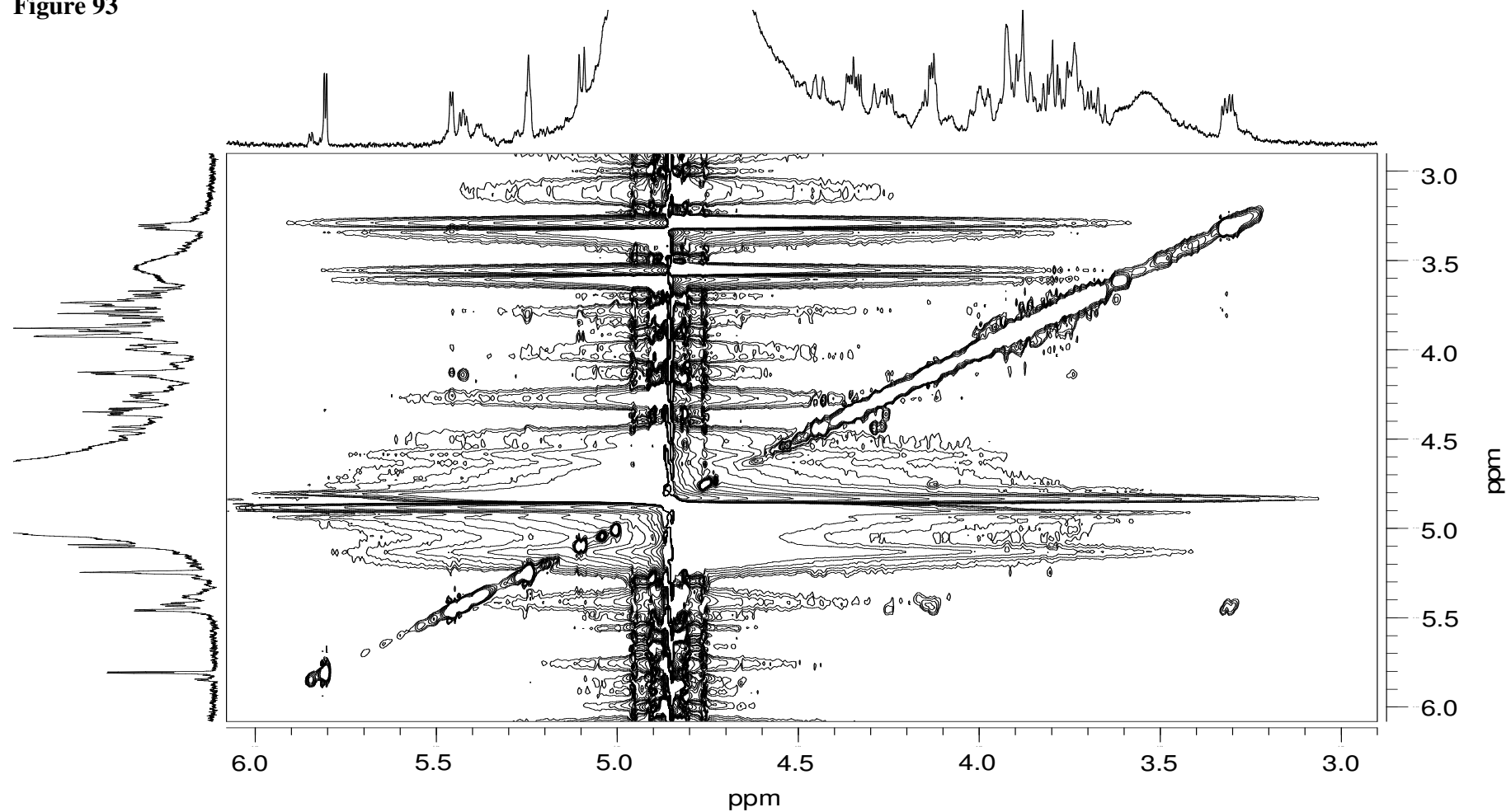


Figure 93 Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, HS dp6-6. Mixing time 500ms

Figure 94

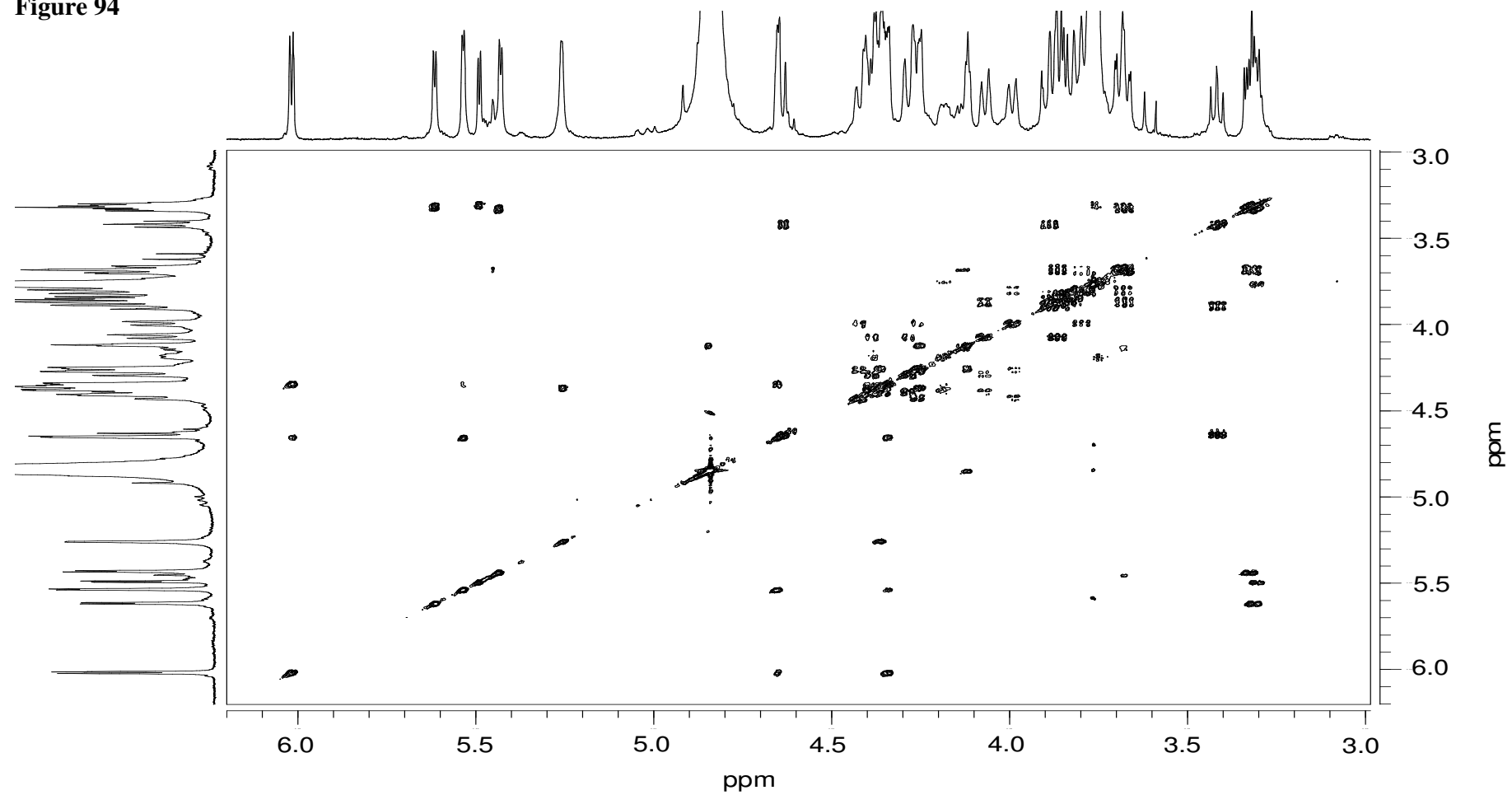


Figure 94 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, hep 6-1.

Figure 95

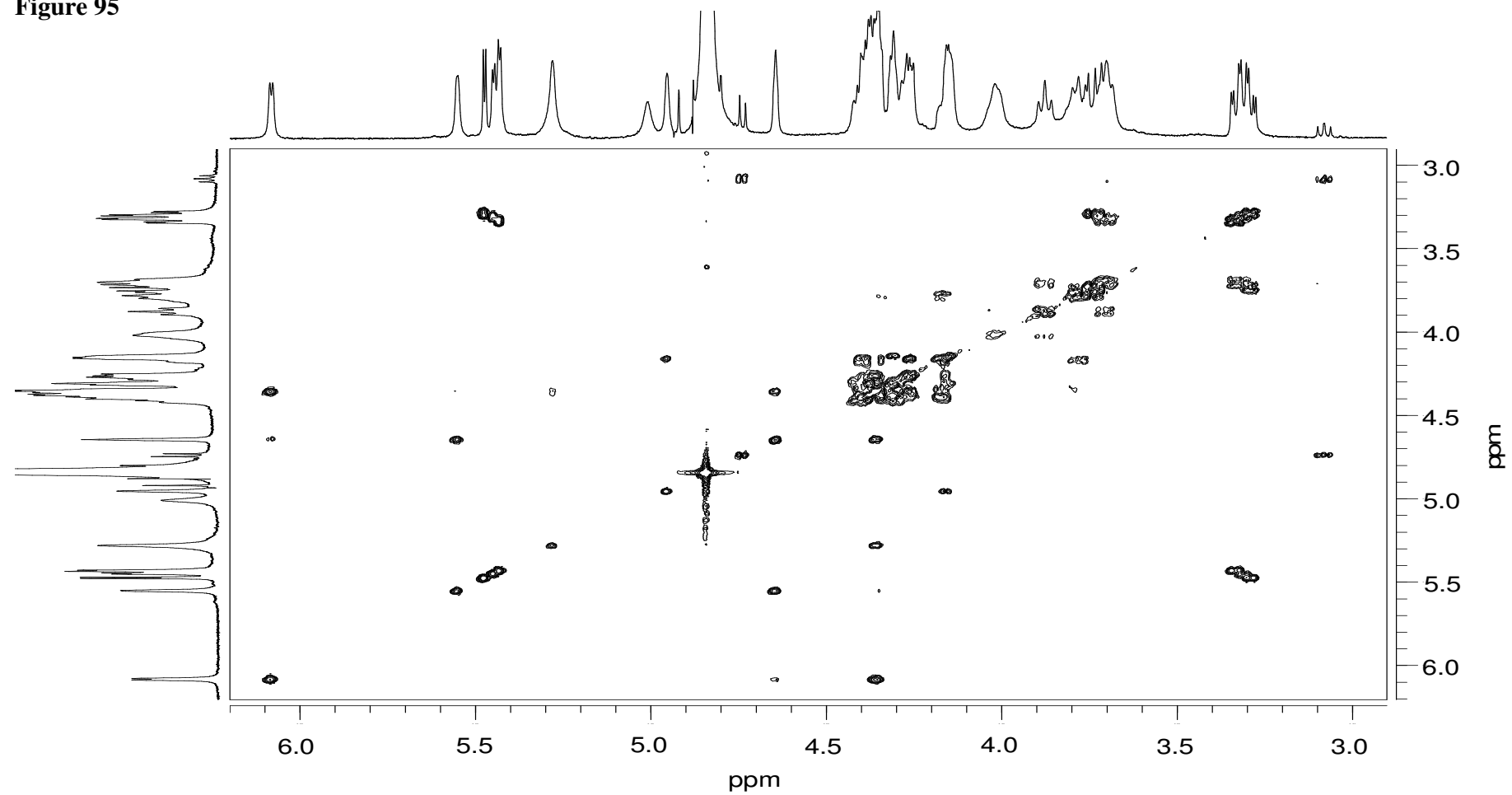


Figure 95 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, hep 6-2.

Figure 96

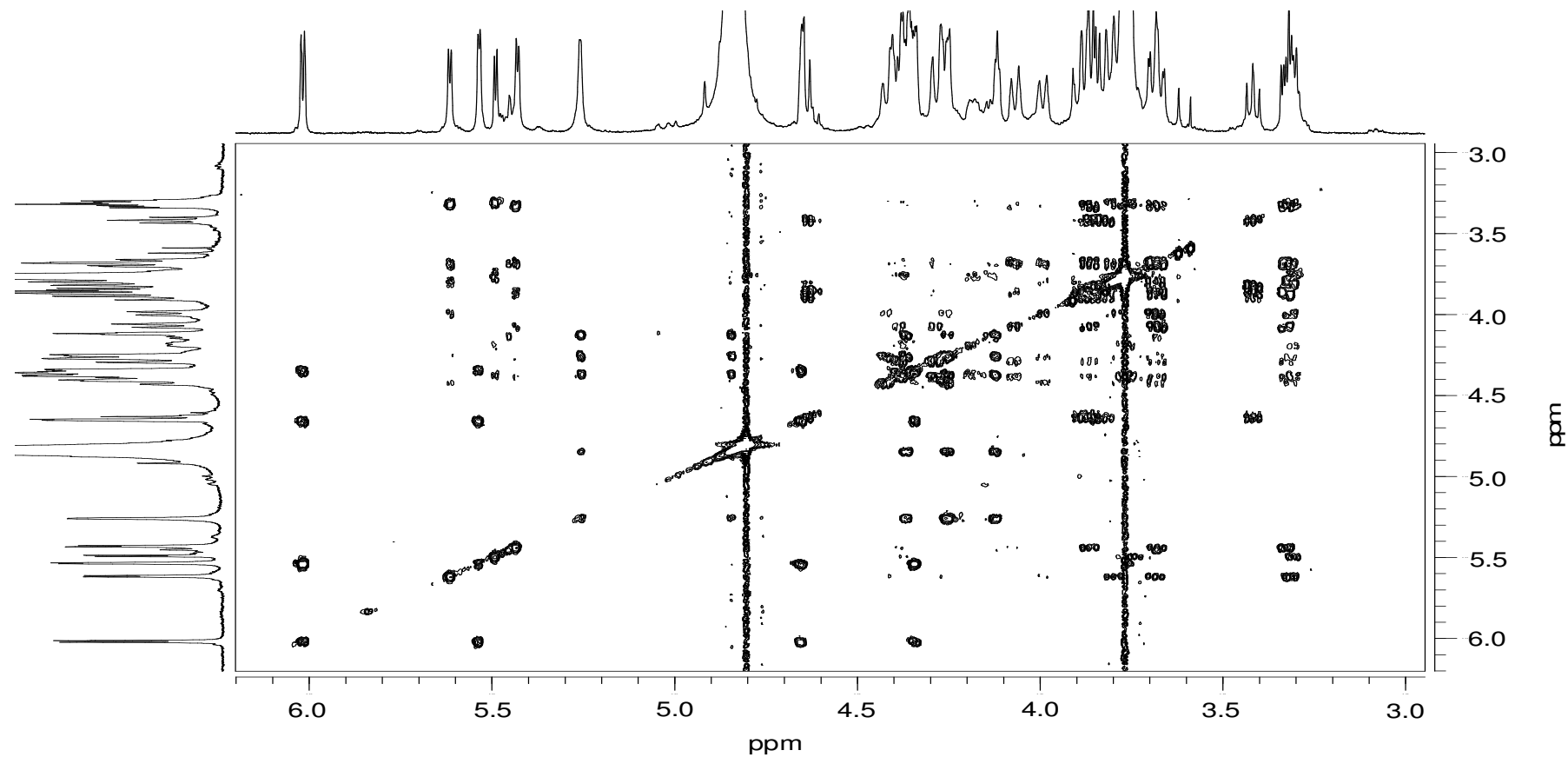


Figure 96 Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, hep 6-1. Mixing time 120ms

Figure 97

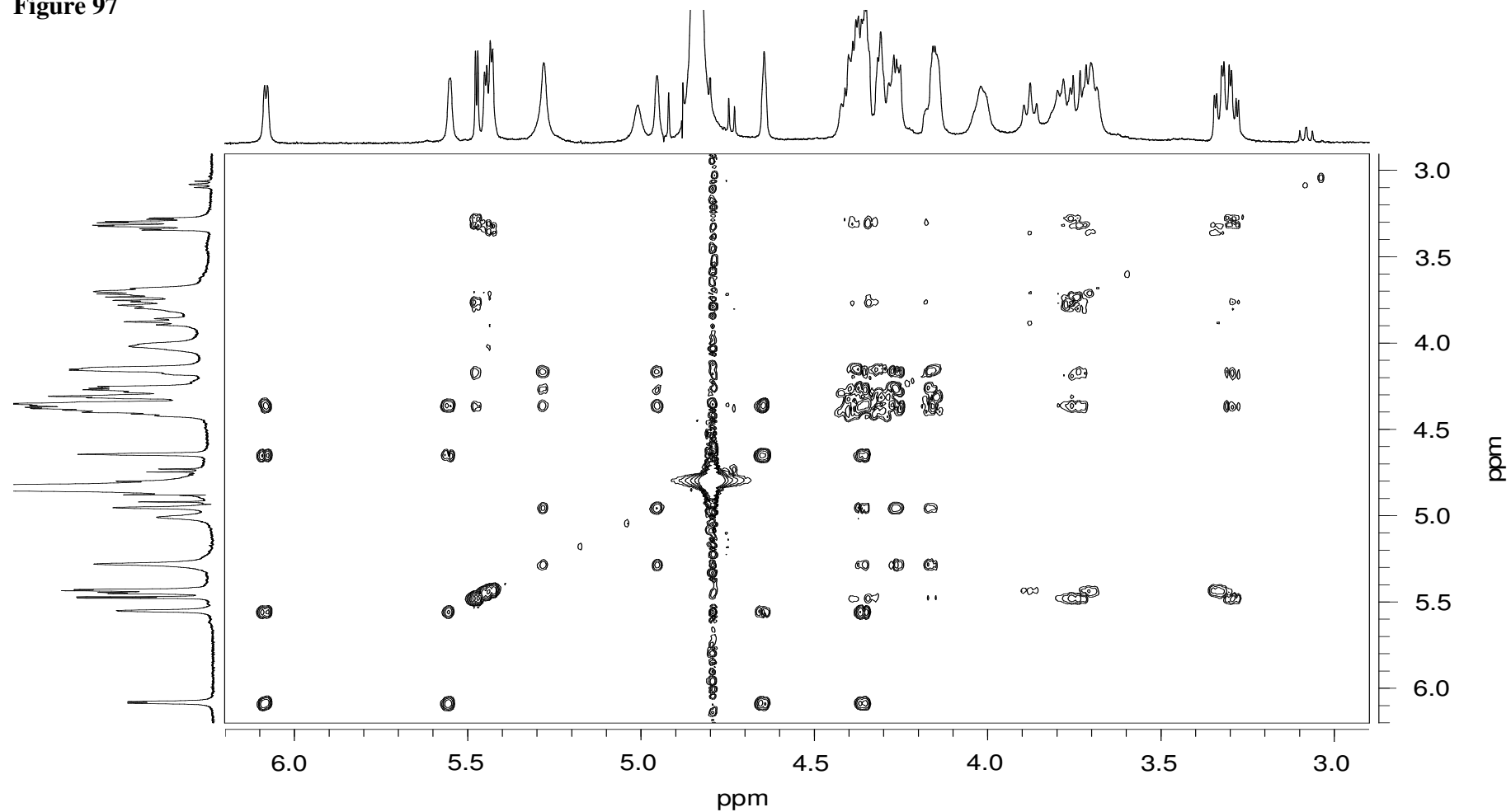


Figure 97 Two-dimensional 500 MHz ^1H -NMR TOCSY spectrum, hep 6-2. Mixing time 120ms

Figure 98

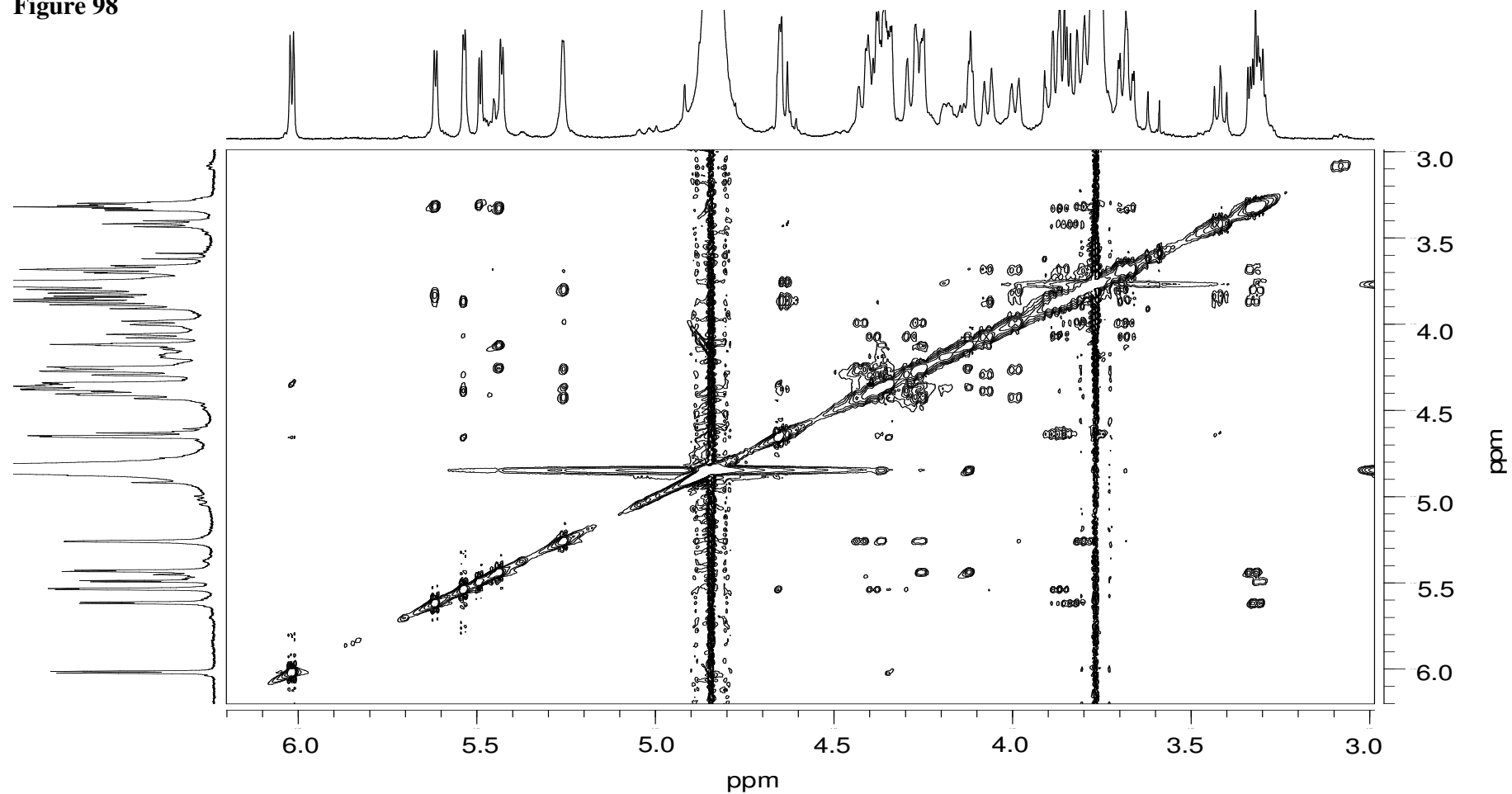


Figure 98 Two-dimensional 500 MHz ¹H-NMR NOESY spectrum, hep 6-1. Mixing time 500ms

Figure 99

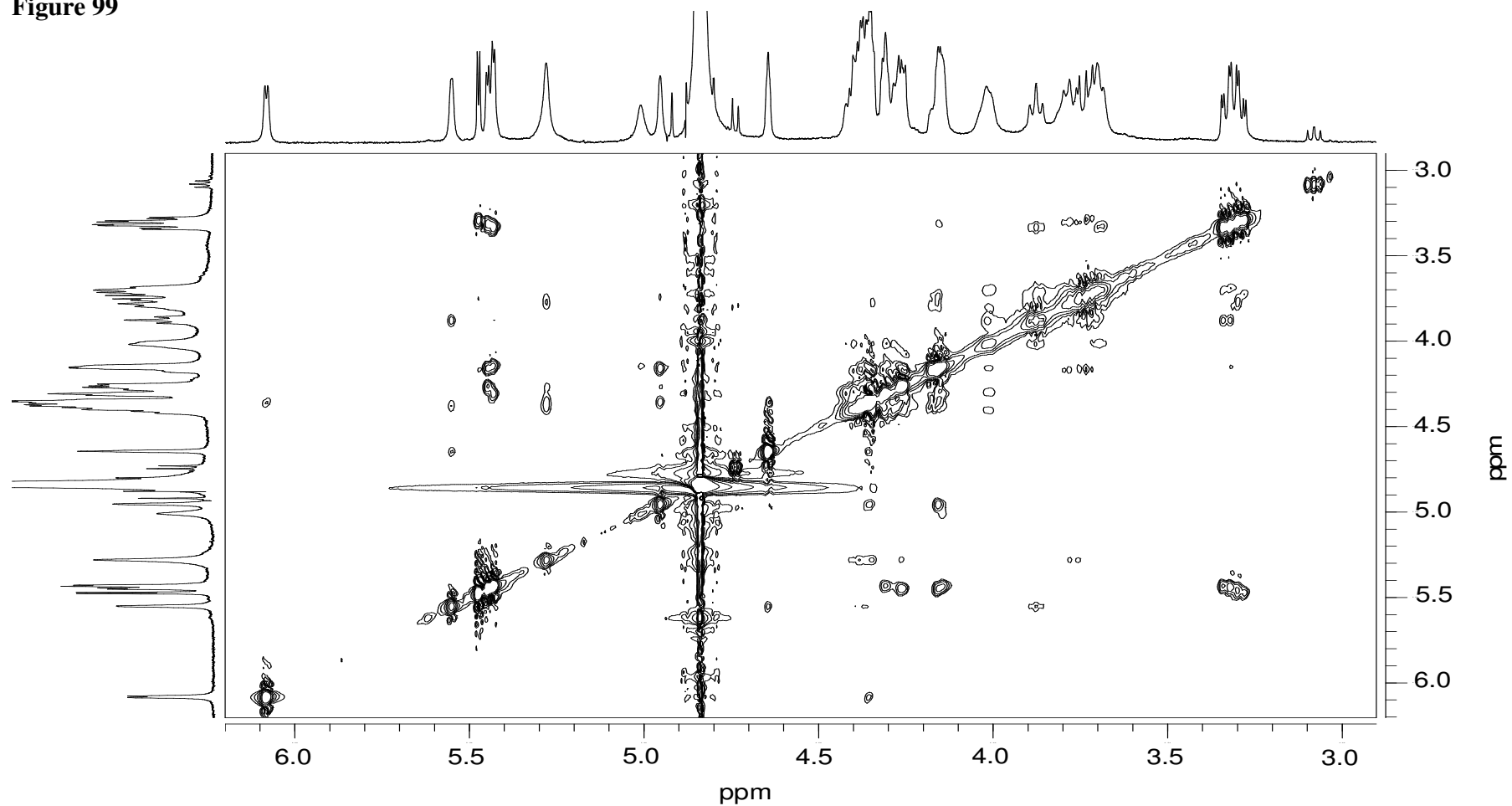


Figure 99 Two-dimensional 500 MHz ^1H -NMR NOESY spectrum, hep 6-2. Mixing time 500ms

Figure 100

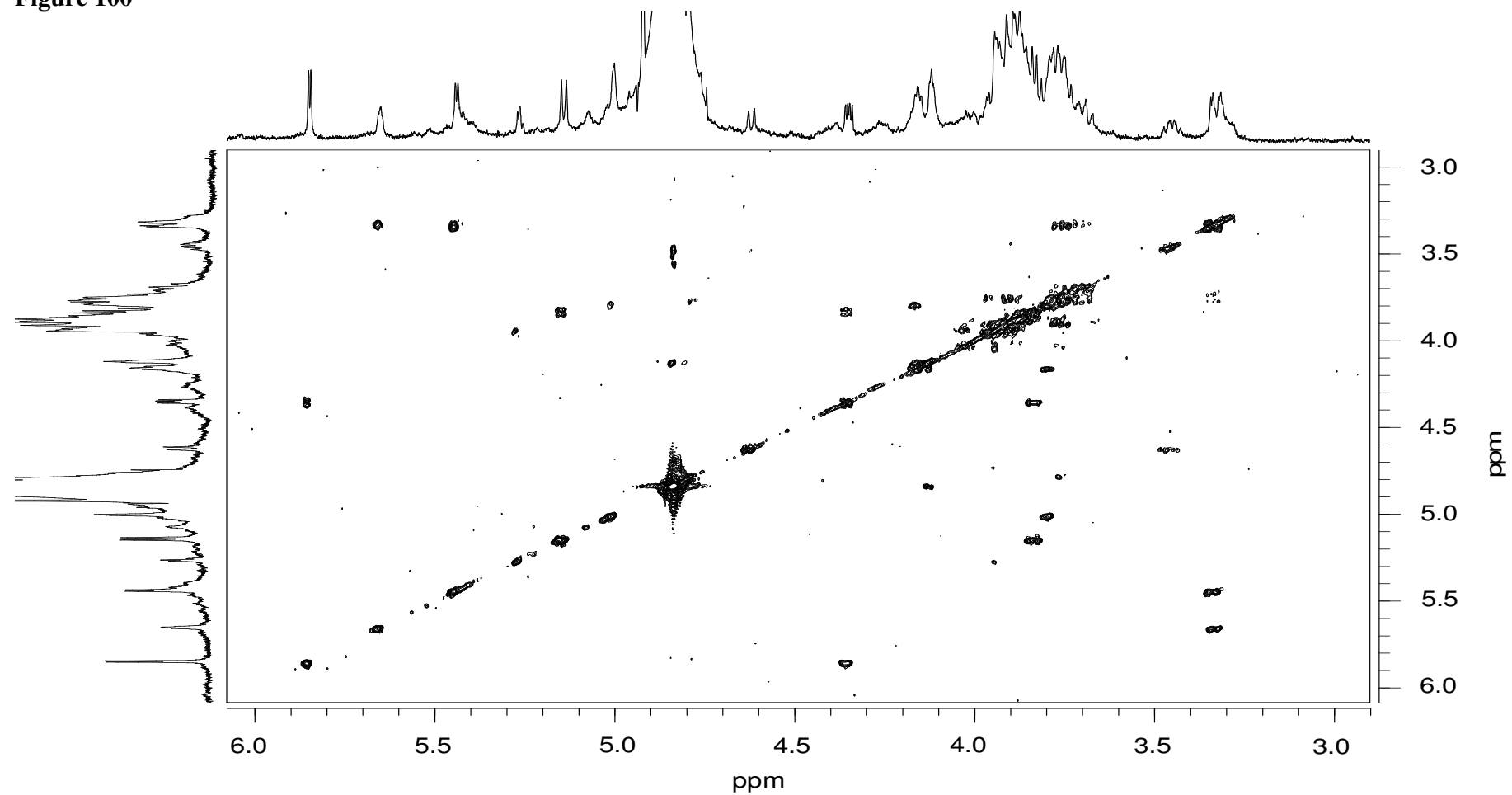


Figure 100 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, HS 6-3 after chemical de-2-O-sulphation.

Figure 101

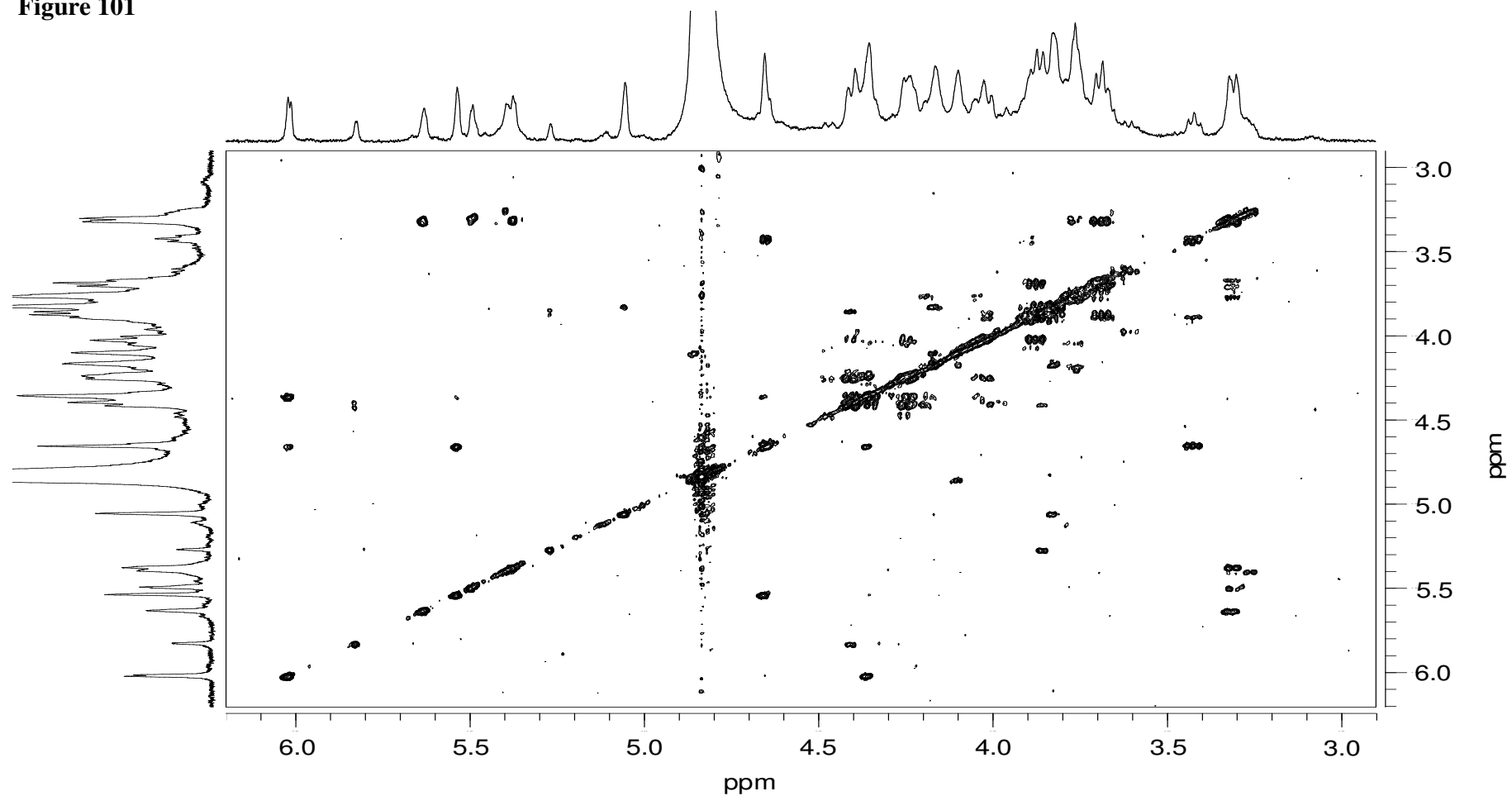


Figure 101 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, hep 6-1 after chemical de-2-O-sulphation.

Figure 102

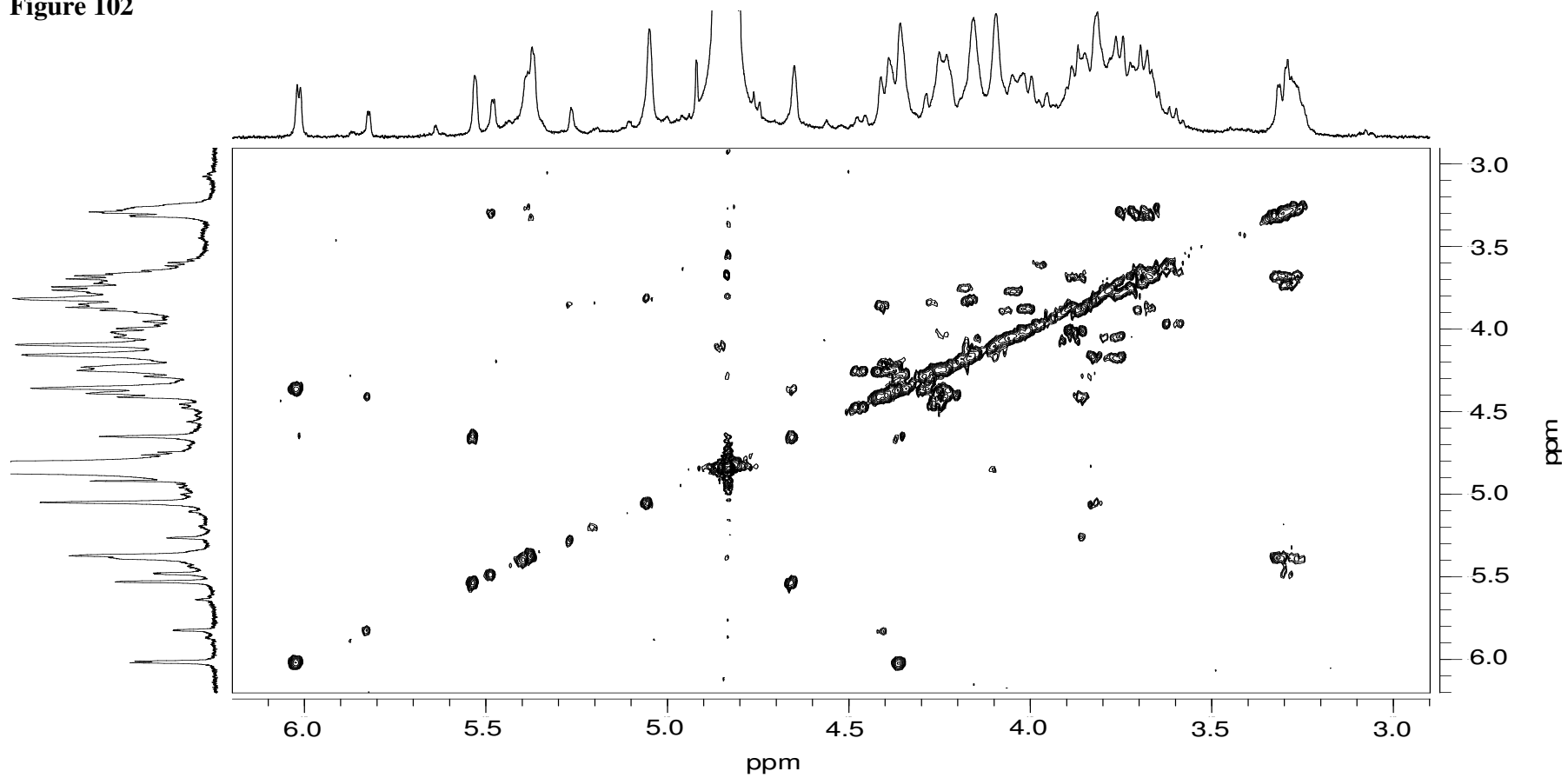


Figure 102 Two-dimensional 500 MHz ^1H -NMR COSY spectrum, hep 6-2 after undergoing two rounds of the chemical de-2-O-sulphation procedure

Appendix II

Figure 103

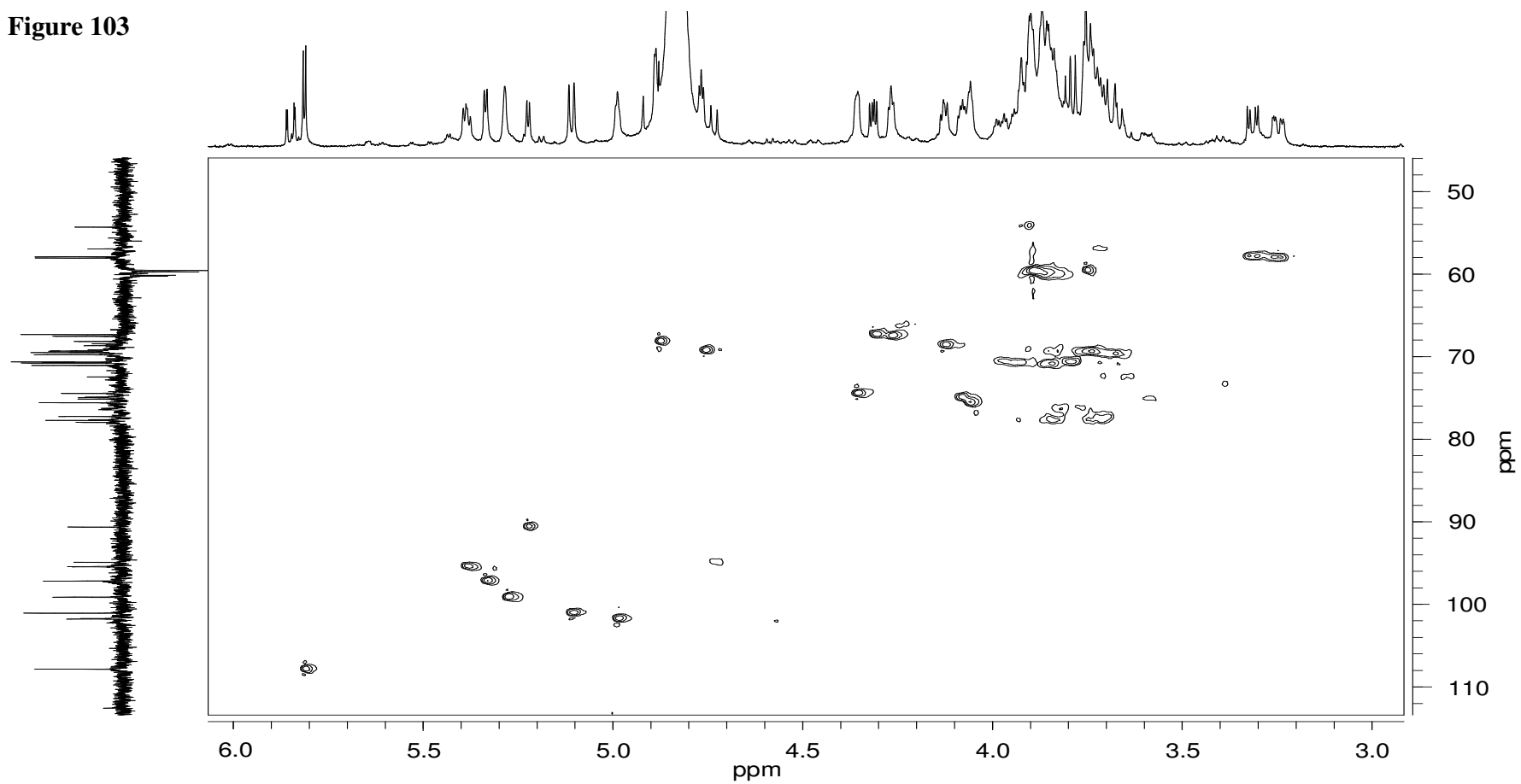


Figure 103 Two-dimensional 500 MHz HSQC spectrum, HS dp6-2. The vertical 1D spectrum was recorded using a DEPT-135 pulse sequence.

Figure 104

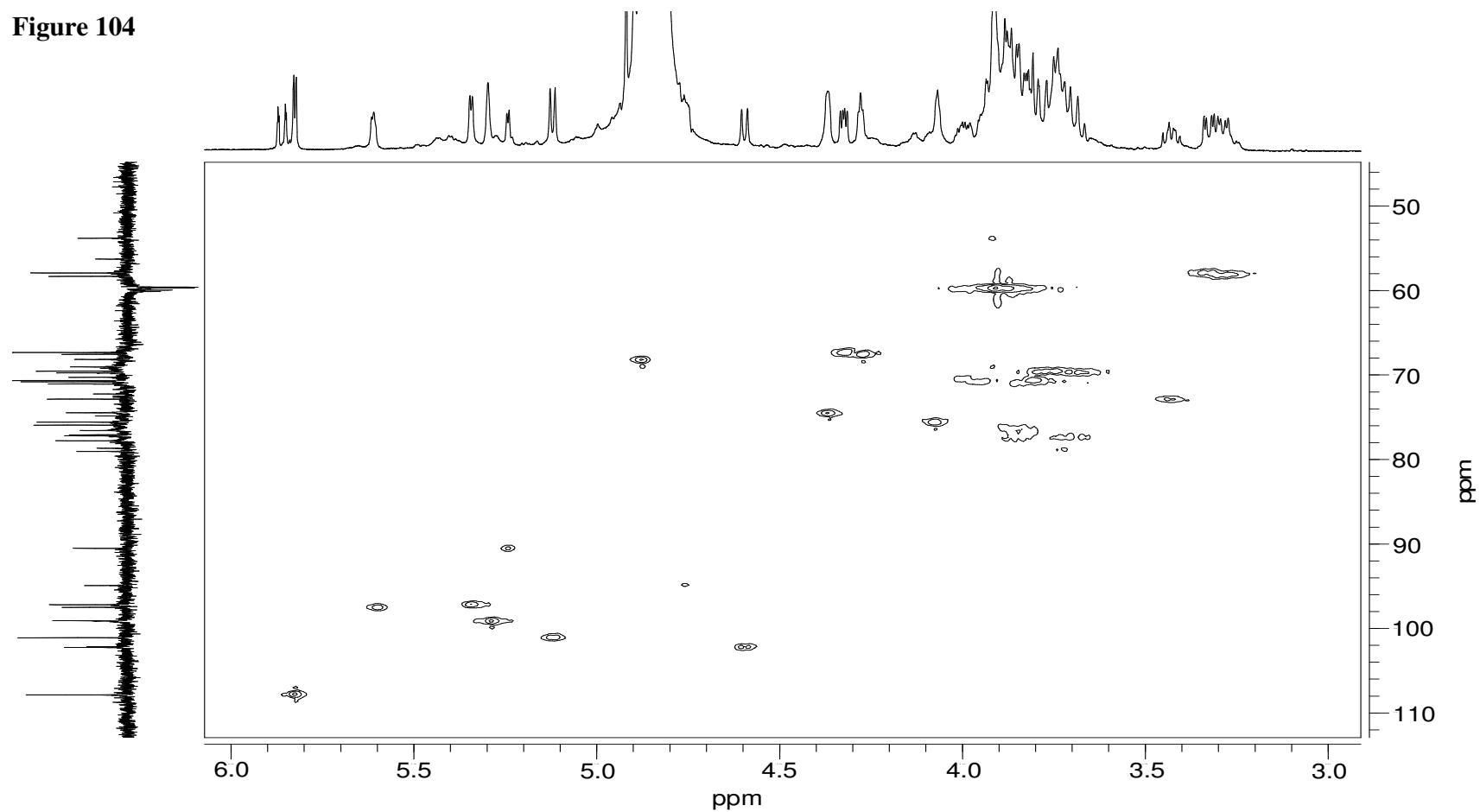


Figure 104 Two-dimensional 500 MHz HSQC spectrum, HS dp6-3. The vertical 1D spectrum was recorded using a DEPT-135 pulse sequence.

Figure 105

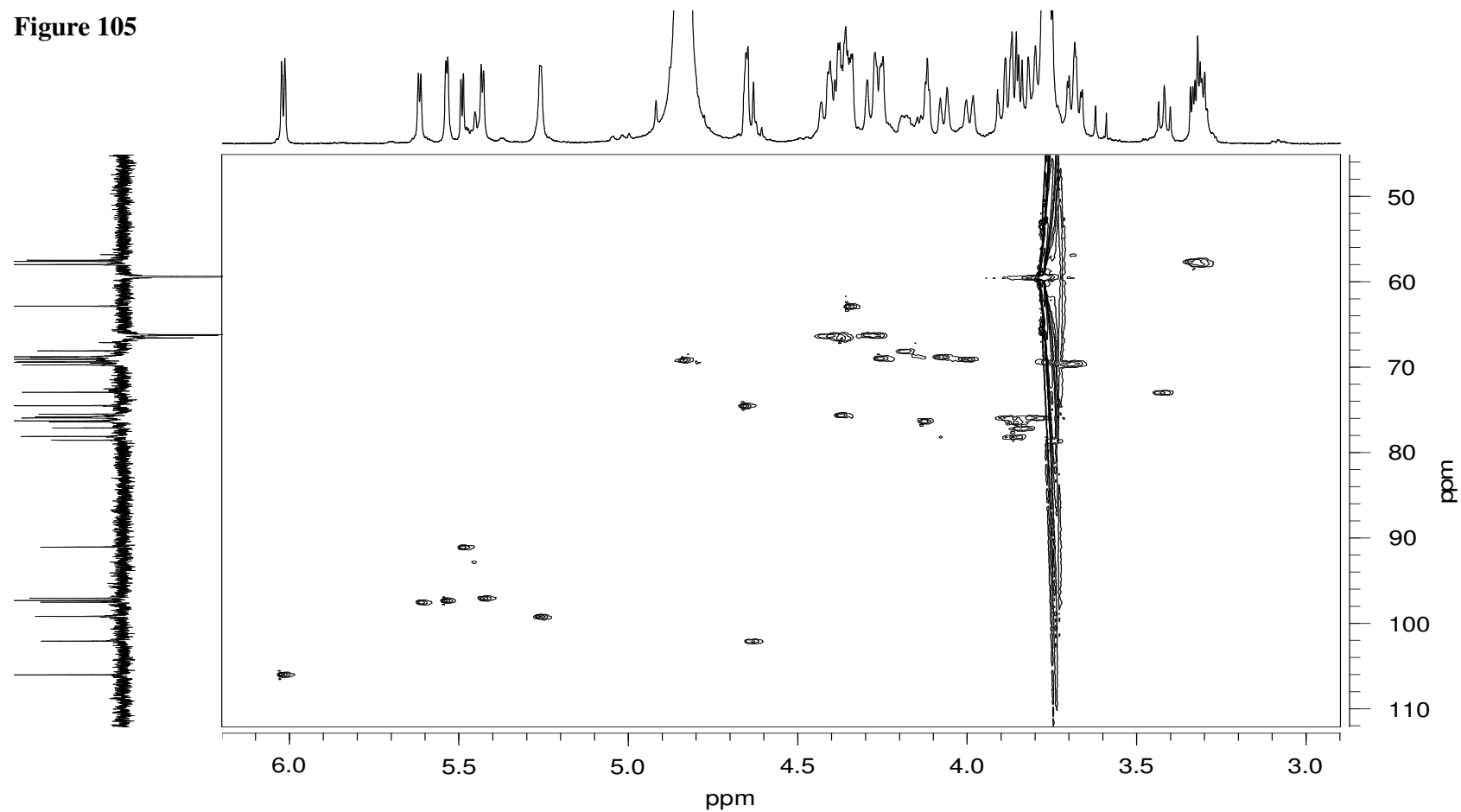


Figure 105 Two-dimensional 500 MHz HSQC spectrum, hep dp6-1. The vertical 1D spectrum was recorded using a DEPT-135 pulse sequence.

Figure 106

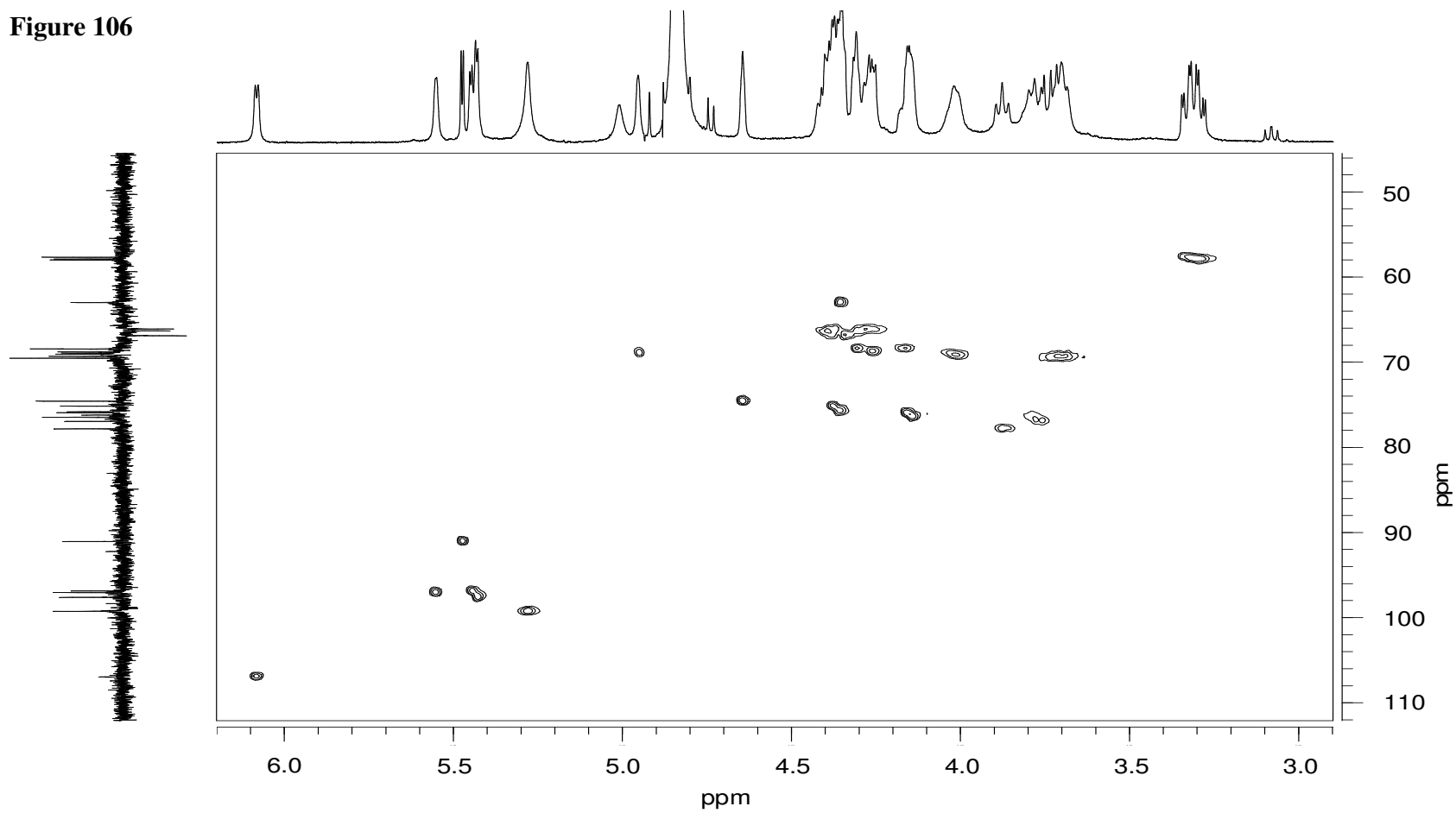


Figure 106 Two-dimensional 500 MHz HSQC spectrum, hep dp6-2. The vertical 1D spectrum was recorded using a DEPT-135 pulse sequence

- Abeijon, C., Mandon, E. C. and Hirschberg, C. B. (1997). Transporters of nucleotide sugars, nucleotide sulfate and ATP in the Golgi apparatus. *Trends. Biochem. Sci.* **22**(6): 203-7.
- Ai, X., Do, A., Kusche-Gullberg, M., Lindahl, U., Lu, K. and Emerson, J. C. P. (2006). Substrate specificity and domain functions of extracellular Heparan Sulfate 6-O-endosulfatases, QSulf1 and QSulf2. *J. Biol. Chem* **281**(8): 4969-4976.
- Ai, X., Do, A., Lozynska, O., Kusche_Gullberg, M., Lindahl, U. and P., E. J. C. (2003). QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. *J. Cell. Biol* **162**(2): 341-351.
- Aikawa, J. and Esko, J. D. (1999). Molecular cloning and expression of a third member of the heparan sulfate/heparin GlcNAc N-deacetylase/ N-sulfotransferase family. *J. Biol. Chem.* **274**(5): 2690-5.
- Aikawa, J., Grobe, K., Tsujimoto, M. and Esko, J. D. (2001). Multiple isozymes of heparan sulfate/heparin GlcNAc N-deacetylase/GlcN N-sulfotransferase. Structure and activity of the fourth member, NDST4. *J. Biol. Chem.* **276**(8): 5876-82.
- Andres, J. L., DeFalcis, D., Noda, M. and Massaue, J. (1992). Binding of two growth factor families to separate domains of the proteoglycan betaglycan. *J. Biol. Chem.* **267**: 5927-5930.
- Angulo, J., Hricovini, M., Gairi, M., Guerrini, M., de Paz, J. L., Ojeda, R., Martin-Lomas, M. and Nieto, P. M. (2005). Dynamic properties of biologically active synthetic heparin-like hexasaccharides. *Glycobiology* **15**(10): 1008-15.
- Angulo, J., Nieto, P. M. and Martin-Lomas, M. (2003). A molecular dynamics description of the conformational flexibility of the L-iduronate ring in glycosaminoglycans. *Chem. Commun*: 1512-1513.
- Ashikari-Hada, S., Habuchi, H., Kariya, Y., Itoh, N., Reddi, A. H. and Kimata, K. (2004). Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. *J Biol Chem* **279**(13): 12346-54.
- Asundi, V. K. and Carey, D., J. (1995). Self association of N-syndecan (syndecan-3) core protein is mediated by a novel structural motif in the transmembrane domain and ectodomain flanking regions. *J. Biol. Chem.* **270**(26404-26410).
- Baciu, P. C. and Goetinck, P. F. (1995). Protein Kinase C regulates recruitment of syndecan-4 into focal adhesions. *Mol. Biol. Cell.* **6**: 1503-1513.
- Baciu, P. C., Saonoella, S., Lee, S. H., Denhez, F., Leuthardt, D. and Goetinck, P. F. (2000). Syndesmos, a protein that interacts with the cytoplasmic domain of syndecan-4, mediates cell spreading and actin cytoskeletal organisation. *J. Cell. Sci.* **113**(2): 315-324.
- Bazin, H. G., Capila, I. and Linhardt, R. J. (1998). Conformational study of synthetic Δ^4 -uronate monosaccharides and glycosaminoglycan derived disaccharides. *Carbohydr. Res.* **309**: 135-144.
- Becker, C. F., Guimaraes, J. A. and Verli, H. (2005). Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution. *Carbohydr. Res.* **340**: 1499-1507.
- Becker, E. D. (2000). High Resolution NMR: Theory and Chemical Applications. San Diego, Academic Press.

- Bellin, R., Capila, I., Lincecum, J., Park, P. W., Reizes, O. and Bernfield, M. R. (2003). Unlocking the secrets of syndecans: Transgenic organisms as a potential key. *Glycoconj. J.* **19**: 295-304.
- Bernfield, M., Kokenyesi, R., Kato, M., Hinkes, M. T., Spring, J., Gallo, R. L. and Lose, E. J. (1992). Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. *Annu. Rev. Cell. Biol.* **8**: 365-93.
- Bjork, I. and Lindahl, U. (1982). Mechanism of the anticoagulant action of heparin. *Mol. Cell. Biochem.* **48**(3): 161-82.
- Blackhall, F. H., Merry, C. L., Lyon, M., Jayson, G. C., Folkman, J., Javaherian, K. and Gallagher, J. T. (2003). Binding of endostatin to endothelial heparan sulphate shows a differential requirement for specific sulphates. *Biochem. J.* **375**: 131-139.
- Bourdon, M. A., Oldberg, A., Pierschbacher, M. and Ruoslahti, E. (1985). Molecular cloning and sequence analysis of a chondroitin sulfate proteoglycan cDNA. *Proc. Natl. Acad. Sci. USA.* **82**(5): 1321-5.
- Brown, J. C., Sasaki, T., Gohring, W., Yamada, Y. and Timpl, R. (1997). The C-terminal domain V of perlecan promotes beta-1 integrin-mediated cell adhesion, binds heparin, nidogen and fibulin-2 and can be modified by glycosaminoglycans. *Eur. J. Biochem.* **250**: 39-46.
- Bubb, W. A. (2003). NMR spectroscopy in the study of carbohydrates: characterizing the structural complexity. *Concepts. Magn. Reson.* **19A**(1): 1-19.
- Bull, R. J., Robinson, M., Laurie, R. D., Stoner, E., Greisiger, J. R., Meier, J. R. and Strober, J. (1984). Carcinogenic effect of acrylamide in Sencar and A/J mice. *Cancer. Res.* **44**: 107-111.
- Cardin, A. D. and Weintraub, H. J. R. (1989). Molecular modelling of protein-glycosaminoglycan interactions. *Arteriosclerosis* **9**: 21-32.
- Carey, D., J., Conner, K., Asundi, V. K., O'Mahony, D. J., Stahl, R. C., Showalter, L. J., Cizmeci-Smith, G., Hartman, J. and Rothblum, L. I. (1997). cDNA cloning genomic organisation and in vivo expression of rat N-syndecan. *J. Biol. Chem.* **272**.
- Carey, D. J. (1997). Syndecans: multifunctional cell-surface co-receptors. *Biochem. J.* **327**: 1-16.
- Carey, F. A. (1992). Organic Chemistry, McGraw-Hill, Inc.
- Case, D. A., Cheatham, T. E. I., Darden, T., Gohlke, H., Luo, R., Merz, K. M. M. J., Onufriev, A., Simmerling, C., Wang, B. and Woods, R. J. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**: 1668-1688.
- Case, D. A., Darden, T. E., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., *et al.* (2004). AMBER 8. *University of California, San Francisco.*
- Cheung, W. F., Eriksson, I., Kusche_Gullberg, M., Lindhal, U. and Kjellen, L. (1996). Expression of the mouse mastocytoma glucosaminyl N-deacetylase/ N-sulfotransferase in human kidney 293 cells results in increased N-sulfation of heparan sulfate. *Biochemistry.* **35**(16): 5250-6.
- Chuang, W. L., Christ, M. D., Peng, J. and Rabenstein, D. L. (2000). An NMR and molecular modelling study of the site-specific binding of Histamine by Heparin,

- chemically modified Heparin and Heparin-derived oligosaccharides. *Biochemistry* **39**: 3542-3555.
- Chuang, W. L., Christ, M. D. and Rabenstein, D. L. (2001). Determination of the primary structures of heparin and heparan sulfate derived oligosaccharides using band selective homonuclear decoupled two dimensional ^1H NMR experiments. *Anal. Chem.* **73**: 2310-2316.
- Chuang, W. L., McAllister, H. and Rabenstein, D. L. (2002). Hexasaccharides from the histamine-modified depolymerization of porcine intestinal mucosal heparin. *Carbohydr. Res.* **337**(10): 935-45.
- Cifonelli, J. A. and King, J. A. (1977). Structural characteristics of heparan sulfates with varying sulfate contents. *Biochemistry.* **16**(10): 2137-41.
- Clamp, A., Blackhall, F. H., Henrioud, A., Jayson, G. C., Javaherian, K., Esko, J., Gallagher, J. T. and Merry, C. L. R. (2006). The morphogenic properties of oligomeric endostatin are dependent on cell surface Heparan Sulfate. *J. Biol. Chem.* **281**(21): 14813-14822.
- Clore, G. M. and Gronenborn, A. M. (1989). How accurately can interproton distances in macromolecules really be determined by full relaxation matrix analysis of nuclear overhauser enhancement data? *J. Magn. Reson.* **84**: 398-409.
- Cohen, A. R., Woods, D. F., Marfatia, S. M., Walther, Z., Chishti, A. H., Anderson, J. M. and Wood, D. F. (1998). Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. *J. Cell. Biol.* **142**(1): 129-38.
- Cole, G. J. and Halfter, W. (1996). Agrin: an extracellular matrix heparan sulfate proteoglycan involved in cell interactions and synaptogenesis. *Perspect. Dev. Neurobiol.* **3**(4): 359-371.
- Collins, P. and Ferrier, R. (1998). Monosaccharides, Their chemistry and their roles in natural products. Chichester, John Wiley & Sons.
- Coltrini, D., Rusnati, M., Zoppetti, G., Oreste, P., Grazioli, G., Naggi, A. and Presta, M. (1994). Differential effects of mucosal, bovine lung and chemically modified heparin on selected biological properties of fibroblast growth factor. *Biochem. J.* **303**: 583-590.
- Conrad, H. E. (1998). Heparin binding proteins. San Diego, Academic press.
- Coombe, D. R. and Kett, W. C. (2005). Heparan sulfate-protein interactions: therapeutic potential through structure-function insights. *Cell. Mol. Life Sci.* **62**(4): 410-424.
- Coutts, J. C. and Gallagher, J. T. (1995). Receptors for fibroblast growth factors. *Immun. Cell. Biol.* **73**(6): 584-9.
- Cremer, D. (1984). On the correct usage of the Cremer-Pople puckering parameters as quantitative descriptors of ring shapes - a reply to recent criticism by Petit, Dillen and Geise. *Acta. Cryst.* **B40**: 498-500.
- Cremer, D. and Pople, J. A. (1975). A general definition of ring puckering coordinates. *J. Am. Chem. Soc.* **97**: 1354-1358.
- Cros, S., Petitou, M., Sizun, P., Perez, S. and Imberty, A. (1997). Combined NMR and molecular modelling study of an Iduronic acid containing trisaccharide related to antithrombotic Heparin fragments. *Bioorg. Med. Chem.* **5**(7): 1301-1309.

- Danielsson, A., Raub, E., Lindahl, U. and Bjork, I. (1986). Role of ternary complexes, in which heparin binds both antithrombin and proteinase, in the acceleration of the reactions between antithrombin and thrombin or factor Xa. *J. Biol. Chem.* **261**(33): 15467-73.
- Das, S. K., Mallet, J. M., Esnault, J., Driguez, P. A., Duchaussoy, P., Sizun, P., Herault, J. P., Herbert, J. M., Petitou, M. and Sinay, P. (2001). Synthesis of conformationally locked carbohydrates: a skew-boat conformation of L-iduronic acid governs the antithrombotic activity of heparin. *Angew Chem Int Ed Engl* **40**: 1670-1673.
- David, G. (1993). Integral membrane heparan sulfate proteoglycans. *FASEB. J.* **7**(11): 1023-30.
- De Cat, B. and David, G. (2001). Developmental roles of the glypicans. *Sem. Cell. Develop. Biol.* **12**: 117-125.
- Deepa, S. S., Yamada, S., Zako, M., Goldberger, O. and Sugahara, K. (2004). Chondroitin sulfate chains on syndecan-1 and syndecan-4 from normal murine mammary gland epithelial cells are structurally and functionally distinct and cooperate with heparan sulfate chains to bind growth factors. A novel function to control binding of midkine, pleiotrophin, and basic fibroblast growth factor. *J. Biol. Chem* **279**: 37368-37376.
- Desai, U. R. and Linhardt, R. J. (1995). Molecular weight of Heparin using ¹³C nuclear magnetic resonance spectroscopy. *J. Pharm. Sci.* **84**(2): 212-215.
- Desai, U. R., Wang, H. M. and Linhardt, R. J. (1993). Specificity studies on the heparin lyases from *Flavobacterium heparinum*. *Biochemistry.* **32**(32): 8140-5.
- Dhoot, G. K., Gustafsson, M. K., Ai, X., Sun, W., Standiford, D. M. and Emerson, J. C. P. (2001). Regulation of Wnt signalling and embryo patterning by an extracellular sulfatase. *Science* **293**: 1663-1666.
- Dong, S., Cole, G. J. and Halfter, W. (2003). Expression of collagen XVIII and localization of its glycosaminoglycan attachment sites. *J. Biol. Chem.* **278**(3): 1700-7.
- Duchesne, L., Tissot, B., Rudd, T. R., Dell, A. and Fernig, D. G. (2006). N-glycosylation of fibroblast growth factor receptor 1 regulates ligand and heparan sulfate co-receptor binding. *J. Biol. Chem.* **281**(37): 27178-27189.
- Edge, A. S. and Spiro, R. G. (1985). Structural elucidation of glycosaminoglycans through characterization of disaccharides obtained after fragmentation by hydrazine-nitrous acid treatment. *Arch. Biochem. Biophys.* **240**(2): 560-72.
- Eriksson, I., Sandback, D., Ek, B., Lindahl, U. and Kjellen, L. (1994). cDNA cloning and sequencing of mouse mastocytoma glucosaminyl N-deacetylase/N-sulfotransferase, an enzyme involved in the biosynthesis of heparin. *J. Biol. Chem.* **269**(14): 10438-43.
- Esko, J. D. and Selleck, S. B. (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* **71**: 435-71.
- Ethell, I. M., Hagihara, K., Miura, Y., Irie, F. and Yamaguchi, Y. (2000). Synbindin, a novel syndecan-2-binding protein in neuronal dendritic spines. *J. Cell. Biol.* **151**(1): 53-68.

- Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J. and Rees, D. C. (1996). Heparin structure and interactions with basic fibroblast growth factor. *Science*. **271**(5252): 1116-20.
- Ferro, D. R., Provasoli, A., Ragazzi, M., Casu, B., Torri, G., Bossennec, V., Perly, B. and Sinay, P. (1990). Conformer populations of L-Iduronic acid residues in glycosaminoglycan sequences. *Carbohydr. Res.* **195**: 157-167.
- Filmus, J. (2001). Glypicans in the growth control of cancer. *Glycobiology*. **11**(3): 19R-23R.
- Forsberg, E., Pejler, G., Ringvall, M., Lunderius, C., Tomasini_Johansson, B., Kusche_Gullberg, M., Eriksson, I., Ledin, J., Hellman, L. and Kjellen, L. (1999). Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme. *Nature*. **400**(6746): 773-6.
- Forster, M. J. and Mulloy, B. (1993). Molecular dynamics study of Iduronate ring conformation. *Biopolymers* **33**: 575-588.
- Fritz, T. A., Gabb, M. M., Wei, G. and Esko, J. D. (1994). Two N-acetylglucosaminyltransferases catalyze the biosynthesis of heparan sulfate. *J. Biol. Chem.* **269**(46): 28809-14.
- Gallagher, J. T. and Lyon, M. (2000). Molecular structure of heparan sulfate and interactions with growth factors and morphogens. *Proteoglycans: structure, biology and molecular interactions*. R. V. Iozzo. New York, Marcel Dekker Inc. New York: 27-59.
- Gallagher, J. T. and Walker, A. (1985). Molecular distinctions between heparan sulphate and heparin. Analysis of sulphation patterns indicates that heparan sulphate and heparin are separate families of N-sulphated polysaccharides. *Biochem. J.* **230**(3): 665-74.
- Gallihier, P. M., Cooney, C. L., Langer, R. and Linhardt, R. J. (1981). Heparinase production by *Flavobacterium heparinum*. *Appl. Environ. Microbiol.* **41**(2): 360-5.
- Gao, Y., M., L., Chen, W. and Simons, M. (2000). Synectin, syndecan-4 cytoplasmic domain binding PDZ protein, inhibits cell migration. *J. Cell. Physiol.* **184**: 373-379.
- Goodger, S. J. (2003). The role and structure of heparan sulfate in the regulation of FGF-1 and FGF-2. School of Biological Sciences. Manchester, University of Manchester. **PhD**: 260.
- Griffin, C. C., Linhardt, R. J., Van Gorp, C. L., Toida, T., Hileman, R. E., Schubert, R. L. and Brown, S. E. (1995). Isolation and characterisation of heparan sulfate from crude porcine intestinal mucosal peptidoglycan heparin. *Carbohydr. Res.* **276**: 183-197.
- Groffen, A. J., Buskens, C. A., van_Kuppevelt, T. H., Veerkamp, J. H., Monnens, L. A. and van_den_Heuvel, L. P. (1998). Primary structure and high expression of human agrin in basement membranes of adult lung and kidney. *Eur. J. Biochem.* **254**(1): 123-8.
- Guerrini, M., Agulles, T., Bisio, A., Hricovini, M., Lay, L., Naggi, A., Poletti, L., Sturiale, L., Torri, G. and Casu, B. (2002). Minimal heparin/heparan sulfate

- sequences for binding to fibroblast growth factor-1. *Biochem. Biophys. Res. Commun.* **292**(1): 222-30.
- Guerrini, M., Bisio, A. and Torri, G. (2001). Combined quantitative (1)H and (13)C nuclear magnetic resonance spectroscopy for characterization of heparin preparations. *Semin. Thromb. Hemost.* **27**(5): 473-82.
- Guimond, S., Maccarana, M., Olwin, B. B., Lindahl, U. and Rapraeger, A. C. (1993). Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2, and FGF-4. *J. Biol. Chem.* **268**(32): 23906-14.
- Guimond, S. E. and Turnbull, J. E. (1999). Fibroblast growth factor receptor signalling is dictated by specific heparan sulphate saccharides. *Curr. Biol.* **9**: 1343-1346.
- Haasnoot, C. A. G., de Leeuw, F. A. A. M. and Altona, C. (1979). The relationship between proton-proton NMR coupling-constants and substituent electronegativities. 1. An empirical generalization of the Karplus equation. *Tetrahedron* **36**: 2783-2792.
- Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K. and Kimata, K. (2000). The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine. *J. Biol. Chem.* **275**(4): 2859-68.
- Hagner Mcwhirter, A., Lindahl, U. and Li, J. (2000). Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5. *Biochem. J.* **347 Pt 1**: 69-75.
- Hakansson, S. and Caffrey, M. (2003). Structural and dynamic properties of the HIV-1 Tat transduction domain in the free and heparin bound states. *Biochemistry* **42**: 8999-9006.
- Halfter, W., Dong, S., Schurer, B. and Cole, G. J. (1998). Collagen XVIII is a basement membrane heparan sulfate proteoglycan. *J. Biol. Chem.* **273**(39): 25404-12.
- Hardy, E. and Castellanos-Serra, L. R. (2004). Reverse-staining of biomolecules in electrophoresis gels: analytical and micropreparative applications. *Anal. Biochem.* **328**: 1-13.
- Hashimoto, Y., Orellana, A., Gil, G. and Hirschberg, C. B. (1992). Molecular cloning and expression of rat liver N-heparan sulfate sulfotransferase. *J. Biol. Chem.* **267**(22): 15744-50.
- Hileman, R. E., Smith, A. E., Toida, T. and Linhardt, R. J. (1997). Preparation and structure of heparin lyase-derived heparan sulfate oligosacchrides. *Glycobiology.* **7**(2): 231-239.
- Hohenester, E., Sasaki, T., Olsen, B. R. and Timpl, R. (1998). Crystal structure of the angiogenesis inhibitor endostatin at 1.5Å resolution. *EMBO J.* **17**(6): 1656-1664.
- Holmborn, K., Ledin, J., Smeds, E., Eriksson, I., Kusche-Gullberg, M. and Kjellen, L. (2004). Heparan Sulfate Synthesized by Mouse Embryonic Stem Cells Deficient in NDST1 and NDST2 is 6-O-Sulfated but Contains No N-Sulfate Groups. *J. Biol. Chem.* **279**(41): 42355-42358.
- Hook, M., Lindahl, U., Hallen, A. and Backstrom, G. (1975). Biosynthesis of heparin. Studies on the microsomal sulfation process. *J. Biol. Chem.* **250**(15): 6065-71.

- Hopf, M., Gohring, W., Kohfeldt, E., Yamada, Y. and Timpl, R. (1999). Recombinant domain IV of perlecan binds to nidogens, laminin-nidogen complex, fibronectin, fibulin-2 and heparin. *Eur. J. Biochem.* **259**: 917-925.
- Horowitz, A. and Simons, M. (1998). Phosphorylation of the cytoplasmic domain of syndecan-4 regulates activation of protein kinase C alpha. *J. Biol. Chem.* **273**: 25548-25551.
- Hricovini, M., Guerrini, M. and Bisio, A. (1999). Structure of heparin derived tetrasacchride complexed to the plasma protein antithrombin derived from NOEs, *J*-couplings and chemical shifts. *Eur. J. Biochem* **261**: 789-801.
- Hricovini, M., Guerrini, M., Bisio, A., Torri, G., Naggi, A. and Casu, B. (2002). Active conformations of Glycosaminoglycans. NMR determination of the conformation of Heparin sequences complexed with Antithrombin and Fibroblast Growth Factors in solution. *Semin. Thromb. Hemost.* **28**(4): 325-334.
- Hricovini, M., Guerrini, M., Bisio, A., Torri, G., Petitou, M. and Casu, B. (2001). Conformation of heparin pentasaccharide bound to antithrombin III. *Biochem. J.* **359**: 265-272.
- Hsueh, Y. P. and Sheng, M. (1999). Regulated expression and subcellular localization of syndecan heparan sulfate proteoglycans and the syndecan-binding protein CASK/LIN-2 during rat brain development. *J. Neurosci.* **19**(17): 7415-25.
- Huige, C. J. M. and Altona, C. (1995). Force field parameters for sulfates and sulfamates based on *ab initio* calculations: extensions of AMBER and CHARMM fields. *J. Comput. Chem.* **16**(1): 56-79.
- Humphrey, W., Dulke, A. and Schulten, K. (1996). VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **14**: 33-38.
- Imberty, A. and Perez, S. (2000). Structure, conformation and dynamics of bioactive oligosaccharides: Theoretical approaches and experimental validations. *Chem. Rev.* **100**: 4567-4588.
- Ishihara, M. (1994). Structural requirements in heparin for binding and activation of FGF-1 and FGF-4 are different from that for FGF-2. *Glycobiology.* **4**(6): 817-24.
- Ishihara, M., Shaklee, P. N., Yang, Z., Liang, W., Wei, Z., Stack, R. J. and Holme, K. (1994). Structural features in heparin which modulate specific biological activities mediated by basic fibroblast growth factor. *Glycobiology* **4**(4): 451-458.
- Ishihara, M., Takano, R., Kanda, T., Hayashi, K., Hara, S., Kikuchi, H. and Yoshida, K. (1995). Importance of 6-O-sulfate groups of glucosamine residues in heparin for activation of FGF-1 and FGF-2. *J. Biochem. (Tokyo)* **118**(6): 1255-60.
- Jackson, D. G., Bell, J. I., Dickinson, R., TImans, J., Shields, J. and Whittle, N. (1995). Proteoglycan forms of lymphocyte homing receptor CD44 are alternatively spliced variants containing the V3 exon. *J. Cell. Biol.* **128**: 673-685.
- Jacobsson, I., Lindahl, U., Jensen, J. W., Roden, L., Prihar, H. and Feingold, D. S. (1984). Biosynthesis of heparin. Substrate specificity of heparosan N-sulfate D-glucuronosyl 5-epimerase. *J. Biol. Chem.* **259**(2): 1056-63.
- Jakallan, A., Jack, D. B. and Bayly, C. I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterisation and validation. *J. Comput. Chem.* **23**: 1623-1641.

- Jandik, K. A., Kruep, D., Cartier, M. and Linhardt, R. J. (1996). Accelerated stability studies of heparin. *J. Pharm. Sci.* **85**(1): 45-51.
- Jaseja, M., Rabindra, N. R., Sauriol, F. and Perlin, A. S. (1989). Novel regio- and stereoselective modifications of heparin in alkaline solution. Nuclear magnetic resonance spectroscopic evidence. *Can. J. Chem.* **67**: 1449-1456.
- Jastrebova, N., Vanwildemeersch, M., Rapraeger, A. C., Gimenez-Gallego, G., Lindahl, U. and Spillmann, D. (2006). Heparan sulfate-related oligosaccharides in ternary complex formation with fibroblast growth factors 1 and 2 and their receptors. *J. Biol. Chem.* **281**(37): 26884-92.
- Johnson, K. A., Gorzinski, S. J., Bodner, K. M., Campbell, R., Wolf, C., Friedman, M. and Mast, R. (1986). Chronic toxicity and oncogenicity study on acrylamide incorporated in the drinking water of Fisher 344 rats. *Toxicol. App. Pharmacol.* **85**: 154-168.
- Kamimura, K., Koyama, T., Habuchi, H., Ueda, R., Masu, M., Kimata, K. and Nakato, K. (2006). Specific and flexible roles of heparan sulfate modifications in Drosophila FGF signaling. *J. Cell. Biol.* **174**(6): 773-778.
- Karplus, M. (1959). Contact electron-spin interactions of nuclear magnetic moments. *J. Chem. Phys.* **30**: 11-15.
- Karumanchi, S. A., Jha, V., Ramchandran, R., Karihaloo, A., Tsiokas, L., Chan, B., Dhanabal, M., Hanai, J. I., Venkataraman, G., Shriver, Z., *et al.* (2001). Cell surface glypicans are low-affinity endostatin receptors. *Mol. Cell.* **7**(4): 811-22.
- Kato, M., Wang, H., Bernfield, M., Gallagher, J. T. and Turnbull, J. E. (1994). Cell surface syndecan-1 on distinct cell types differs in fine structure and ligand binding of its heparan sulfate chains. *J. Biol. Chem.* **269**: 18881-18890.
- Kato, M., Wang, H., Kainulainen, V., Fitzgerald, M. L., Ledbetter, S., Ornitz, D. M. and Bernfield, M. (1998). Physiological degradation converts the soluble syndecan-1 ectodomain from an inhibitor to a potent activator of FGF-2. *Nat. Med.* **4**(6): 691-7.
- Kim, B. T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U. and Sugahara, K. (2001). Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4- N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/ heparin biosynthesis. *Proc. Natl. Acad. Sci.* **98**(13): 7176-81.
- Kim, C., Goldberger, O., Gallo, R. and Bernfield, M. (1994). Members of the syndecan family of heparan sulfate proteoglycans are expressed in distinct cell-, tissue-, and developmental specific patterns. *Mol. Biol. Cell.* **5**: 797-805.
- King, D. J. and Noss, R. R. (1989). Toxicity of polyacrylamide and acrylamide. *Rev. Environ. Health* **8**(1-4): 3-16.
- Kirschner, K. N. and Woods, R. J. (2001). Solvent interactions determine carbohydrate conformation. *Proc. Natl. Acad. Sci.* **98**(19): 10541-10545.
- Knox, S., Melrose, J. and Whitelock, J. (2001). Electrophoretic, biosensor, and bioactivity analyses of perlecan of different cellular origins. *Proteomics.* **1**: 1534-1541.

- Knox, S., Merry, C., Stringer, S., Melrose, J. and Whitelock, J. (2002). Not all perlecan are created equal: interactions with fibroblast growth factor (FGF) 2 and FGF receptors. *J. Biol. Chem.* **277**(17): 14657-65.
- Kokenyesi, R. and Bernfield, M. (1994). Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1. *J. Biol. Chem.* **269**(16): 12304-12309.
- Kreuger, J., Jemth, P., Sanders_Lindberg, E., Eliahu, L., Ron, D., Basilico, C., Salmivirta, K. and Lindahl, U. (2005). Fibroblast growth factors share binding sites in Heparan Sulphate. *Biochem. J.* **389**: 145-150.
- Kreuger, J., Matsumoto, T., Vanwildemeersch, M., Sasaki, T., Timpl, R., Claesson_Welsh, L., Spillmann, D. and Lindahl, U. (2002). Role of heparan sulfate domain organization in endostatin inhibition of endothelial cell function. *EMBO J.* **21**(23): 6303-11.
- Kreuger, J., Salmivirta, M., Sturiale, L., Gimenez_Gallego, G. and Lindahl, U. (2001). Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2. *J. Biol. Chem.* **276**(33): 30744-52.
- Kreuger, J., Spillmann, D., Li, J.-P. and Lindahl, U. (2006). Interactions between heparan sulfate and proteins: the concept of specificity. *J. Cell. Biol.* **174**(3): 323-327.
- Kuo, C. J., LaMontagne, K. R., Jr., Garcia-Cardena, G., Ackley, B. D., Kalman, D., Park, S., Christofferson, R., Kamihara, J., Ding, Y. H., Lo, K. M., *et al.* (2001). Oligomerization-dependent regulation of motility and morphogenesis by the collagen XVIII NC1/endostatin domain. *J. Cell. Biol.* **152**(6): 1233-46.
- Kupperman, A. S. (1958). Effects of acrylamide on the central nervous system of the cat. *J. Pharmacol. Exp. Ther.* **123**: 180-192.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.
- Lander, A. D. and Selleck, S. B. (2000). The elusive functions of proteoglycans: *in vivo veritas*. *J. Cell. Biol.* **148**(2): 227-32.
- Larnkjaer, A., Hansen, S. H. and Ostergaard, P. B. (1995). Isolation and characterisation of hexasaccharides derived from heparin. Analysis by HPLC and elucidation of structure by ¹H NMR. *Carbohydr. Res.* **266**: 37-52.
- Leach, A. R. (2001). Molecular modelling principles and applications. Harlow, Pearson education limited.
- Lin, X., Wei, G., Shi, Z., Dryer, L., Esko, J. D., Wells, D. E. and Matzuk, M. M. (2000). Disruption of gastrulation and heparan sulfate biosynthesis in EXT1-deficient mice. *Dev. Biol.* **224**(2): 299-311.
- Lind, T., Tufaro, F., McCormick, C., Lindahl, U. and Lidholt, K. (1998). The putative tumor suppressors EXT1 and EXT2 are glycosyltransferases required for the biosynthesis of heparan sulfate. *J. Biol. Chem.* **273**(41): 26265-8.
- Lindahl, U., Backstrom, G., Jansson, L. and Hallen, A. (1973). Biosynthesis of heparin. II. Formation of sulfamino groups. *J. Biol. Chem.* **248**(20): 7234-41.
- Lindahl, U., Kusche-Gullberg, M. and Kjellen, L. (1998). Regulated diversity of heparan sulfate. *J. Biol. Chem.* **273**: 24979-24982.

- Lindahl, U., Thunberg, L., Backstrom, G., Riesenfeld, J., Nordling, K. and Bjork, I. (1984). Extension and structural variability of the antithrombin-binding sequence in heparin. *J. Biol. Chem.* **259**(20): 12368-76.
- Linhardt, R. J. and Gunay, N. S. (1999). Production and chemical processing of low molecular weight heparins. *Semin. Thromb. Hemost.* **25 Suppl 3**: 5-16.
- Linhardt, R. J., Rice, K. G., Kim, Y. S., Lohse, D. L., Wang, H. M. and Loganathan, D. (1988). Mapping and quantification of the major oligosaccharide components of heparin. *Biochem. J.* **254**(3): 781-7.
- Linhardt, R. J., Turnbull, J. E., Wang, H. M., Loganathan, D. and Gallagher, J. T. (1990). Examination of the substrate specificity of heparin and heparan sulfate lyases. *Biochemistry.* **29**(10): 2611-7.
- Linker, A. and Hovingh, P. (1972). Isolation and characterization of oligosaccharides obtained from heparin by the action of heparinase. *Biochemistry.* **11**(4): 563-8.
- Liu, D., Shriver, Z., Venkataraman, G., El_Shabrawi, Y. and Sasisekharan, R. (2002). Tumor cell surface heparan sulfate as cryptic promoters or inhibitors of tumor growth and metastasis. *Proc. Natl. Acad. Sci.* **99**(2): 568-73.
- Liu, J., Shriver, Z., Blaiklock, P., Yoshida, K., Sasisekharan, R. and Rosenberg, R. D. (1999). Heparan sulfate D-glucosaminyl 3-O-sulfotransferase-3A sulfates N-unsubstituted glucosamine residues. *J. Biol. Chem.* **274**(53): 38155-62.
- Liu, J., Shworak, N. W., Sinay, P., Schwartz, J. J., Zhang, L., Fritze, L. M. and Rosenberg, R. D. (1999). Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities. *J. Biol. Chem.* **274**(8): 5185-92.
- Loganathan, D., Wang, H. M., Mallis, L. M. and Linhardt, R. J. (1990). Structural variation in the Antithrombin III binding site region and its occurrence in Heparin from different sources. *Biochemistry* **29**: 4362-4368.
- Lortat-Jacob, H. and Grimaud, J. A. (1991). Interferon-gamma binds to heparan sulfate by a cluster of amino acids located in the C-terminal part of the molecule. *FEBS* **280**(1): 152-154.
- Lortat_Jacob, H., Grosdidier, A. and Imberty, A. (2002). Structural diversity of heparan sulfate binding domains in chemokines. *Proc. Natl. Acad. Sci.* **99**(3): 1229-34.
- Lortat_Jacob, H., Turnbull, J. E. and Grimaud, J. A. (1995). Molecular organization of the interferon gamma-binding domain in heparan sulphate. *Biochem. J.* **310**: 497-505.
- Lucas, R., Angulo, J., Nieto, P. M. and Martin-Lomas, M. (2003). Synthesis and structural study of two new heparin-like hexasaccharides. *Org. Biomol. Chem.* **1**: 2253-2266.
- Lundin, L., Larsson, H., Kreuger, J., Kanda, S., Lindahl, U., Salmivirta, M. and Claesson_Welsh, L. (2000). Selectively desulfated heparin inhibits fibroblast growth factor-induced mitogenicity and angiogenesis. *J. Biol. Chem.* **275**(32): 24653-60.
- Lyon, M., Rushton, G., Askari, J. A., Humphries, M. J. and Gallagher, J. T. (2000). Elucidation of the structural features of heparan sulfate important for interaction with the hep-2 domain of fibronectin. *J Biol Chem* **275**(7): 4599-4606.

- Maccarana, M., Casu, B. and Lindahl, U. (1993). Minimal sequence in heparin/heparan sulfate required for binding of basic fibroblast growth factor. *J. Biol. Chem.* **268**(32): 23898-905.
- Maccarana, M., Sakura, Y., Tawada, A., Yoshida, K. and Lindahl, U. (1996). Domain structure of heparan sulfates from bovine organs. *J. Biol. Chem.* **271**(30): 17804-10.
- Marcum, J. A., McKenny, J. B., Galli, S. J., Jackman, R. W. and Rosenberg, R. D. (1986). Anticoagulant active heparin-like molecules from mast cell-deficient mice. *Am. J. Physiol.* **250**: H879-H888.
- Marneros, A. G. and Olsen, B. R. (2005). Physiological role of collagen XVIII and endostatin. *FASEB J.* **19**: 716-728.
- Mathews, C. K. and van_Holde, K. E. (1990). Biochemistry, Benjamin/Cummings publishing company, Inc.
- McCollister, D. D., Oyen, F. and Rowe, V. K. (1964). Toxicology of acrylamide. *Toxicol. App. Pharmacol.* **6**: 172-181.
- McCormick, C., Duncan, G., Goutsos, K. T. and Tufaro, F. (2000). The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. *Proc. Natl. Acad. Sci.* **97**(2): 668-73.
- McDonnell, K. M. and Grow, W. A. (2004). Reduced glycosaminoglycan sulfation diminishes the agrin signal transduction pathway. *Dev. Neurosci.* **26**(1): 1-10.
- Merry, C. L., Bullock, S. L., Swan, D. C., Backen, A. C., Lyon, M., Beddington, R. S., Wilson, V. A. and Gallagher, J. T. (2001). The molecular phenotype of heparan sulfate in the Hs2st^{-/-} mutant mouse. *J. Biol. Chem.* **276**(38): 35429-34.
- Merry, C. L., Lyon, M., Deakin, J. A., Hopwood, J. J. and Gallagher, J. T. (1999). Highly sensitive sequencing of the sulfated domains of heparan sulfate. *J. Biol. Chem.* **274**(26): 18455-62.
- Mikhailov, D., Linhardt, R. J. and Mayo, K. H. (1997). NMR solution conformation of heparin derived hexasaccharide. *Biochem. J.* **328**: 51-61.
- Mikhailov, D., Mayo, K. H., Pervin, A. and Linhardt, R. J. (1996). ¹³C-NMR relaxation study of heparin-disaccharide interactions with tripeptides GRG and GKG. *Biochem. J.* **315**: 447-454.
- Mikhailov, D., Mayo, K. H., Vlahov, I. R., Toida, T., Pervin, A. and Linhardt, R. J. (1996). NMR solution conformation of heparin derived tetrasacchride. *Biochem. J.* **318**: 93-102.
- Mulloy, B. and Forster, M. J. (2000). Conformation and dynamics of heparin and heparan sulfate. *Glycobiology* **10**(11): 1147-56.
- Mulloy, B., Forster, M. J., Jones, C. and Davies, D. B. (1993). NMR and molecular modelling studies of the solution conformation of heparin. *Biochem. J.* **293**: 849-858.
- Mulloy, B., Forster, M. J., Jones, C., Drake, A. F., Johnson, E. A. and Davies, D. B. (1994). The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study. *Carbohydr. Res.* **255**: 1-26.

- Murphy (2003). K5 lyase, a new enzyme for the analysis of Heparan Sulphate structure and function. Faculty of Medicine, Dentistry, Nursing and Pharmacy. Manchester, University of Manchester. **M. Phil:** 105.
- Murphy, K. J., Merry, C. L. R., Lyon, M., Thompson, J. E., Roberts, I. S. and Gallagher, J. T. (2004). A new model for the domain structure of Heparan Sulfate based on the novel specificity of K5 lyase. *J. Biol. Chem.* **279**(26): 27239-27245.
- Nakato, H., Futch, T. A. and Selleck, S. B. (1995). The division abnormally delayed (dally) gene: A putative integral membrane proteoglycan required for cell division patterning during postembryonic development of the nervous system in *Drosophila*. *Development.* **121**: 3687-3702.
- Neuhaus, D. and Williamson, M. P. (2000). The Nuclear Overhauser Effect in structural and conformational analysis. New York, Wiley-VCH, Inc.
- Norgard_Sumnicht, K. and Varki, A. (1995). Endothelial heparan sulfate proteoglycans that bind to L-selectin have glucosamine residues with unsubstituted amino groups. *J. Biol. Chem.* **270**(20): 12012-24.
- O'Reilly, M. S., Boeham, T., Shing, Y., Fukai, N., Vasios, G., Lane, W. S., Flynn, E., Birkhead, J. R., Olsen, B. R. and Folkman, J. (1997). Endostatin: an endogenous inhibitor of angiogenesis and tumour growth. *Cell* **88**: 277-285.
- Oh, E., Woods, A. and Couchman, J. R. (1997). Syndecan-4 proteoglycan regulates the distribution and activity of protein kinase C. *J. Biol. Chem.* **272**(13): 8133-8139.
- Oh, E. S., Woods, A. and Couchman, J. R. (1997). Multimerisation of the cytoplasmic domain of syndecan-4 is required for its ability to activate protein kinase C. *J. Biol. Chem.* **272**: 11805-11811.
- Ojeda, R., Angulo, J., Nieto, P. M. and Martin-Lomas, M. (2002). The activation of fibroblast growth factors by heparin: Synthesis and structural study of rationally modified heparin-like oligosaccharides. *Can. J. Chem.* **80**: 917-936.
- Ono, K., Hattori, H., Takeshita, S., Kurita, A. and Ishihara, M. (1999). Structural features in heparin that interact with VEGF165 and modulate its biological activity. *Glycobiology* **9**(7): 705-11.
- Ornitz, D. M. and Itoh, N. (2001). Fibroblast growth factors. *Genome. Biol.* **2**(3): Reviews 3005.
- Ornitz, D. M., Xu, J., Colvin, J. S., McEwan, D. G., MacArthur, C. A., Coulier, F., Gao, G. and Goldfarb, M. (1996). Receptor specificity of the fibroblast growth factor family. *J. Biol. Chem.* **271**(25): 15292-15297.
- Ostrovsky, O., Berman, B., Gallagher, J., Mulloy, B., Fernig, D. G., Delehedde, M. and Ron, D. (2002). Differential effects of heparin saccharides on the formation of specific fibroblast growth factor (FGF) and FGF receptor complexes. *J. Biol. Chem.* **277**(4): 2444-53.
- Pellegrini, L., Burke, D. F., Von_Delft, F., Mulloy, B. and Blundell, T. L. (2000). Crystal structure of fibroblast growth factor receptor 2 ectodomain bound to ligand and heparin. *Nature* **407**: 1029-1034.
- Perrimon, N. and Bernfield, M. (2000). Specificities of heparan sulphate proteoglycans in developmental processes. *Nature.* **404**: 725-728.

- Pervin, A., Gallo, C., Jandik, K. A., Han, X. J. and Linhardt, R. J. (1995). Preparation and structural characterization of large heparin-derived oligosaccharides. *Glycobiology* **5**(1): 83-95.
- Piani, S., Casu, B., Marchi, E. G., Torri, G. and Ungarelli, F. (1993). Alkali-induced optical rotational changes in heparins and heparan sulfates, and their relation to iduronic acid-containing sequences. *J. Carbohydr. Chem.* **12**: 507-521.
- Pilia, G., Huges-Benzie, R. M., MacKenzie, A., Baybayan, P., Chen, E. Y., Huber, R., Neri, G., Cao, A., Forabosco, A. and Schlessinger, D. (1996). Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nat. Genet.* **12**: 241-247.
- Prasthofer, T., Ek, B., Ekman, P., Owens, R., Hook, M. and Johansson, S. (1995). Protein kinase C phosphorylates two of the four known syndecan domains *in vitro*. *Biochem. Mol. Biol. Int.* **36**: 793-802.
- Pye, D. A., Vives, R. R., Hyde, P. and Gallagher, J. T. (2000). Regulation of FGF-1 mitogenic activity by heparan sulfate oligosaccharides is dependent on specific structural features: differential requirements for the modulation of FGF-1 and FGF-2. *Glycobiology.* **10**(11): 1183-92.
- Pye, D. A., Vives, R. R., Turnbull, J. E., Hyde, P. and Gallagher, J. T. (1998). Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity. *J. Biol. Chem.* **273**(36): 22936-42.
- Ragazzi, M., Ferro, D. R. and Provasoli, A. (1986). A force field study of the conformational characteristics of the Iduronate ring. *J. Comput. Chem.* **7**(2): 105-112.
- Ragazzi, M., Ferro, D. R., Provasoli, A., Pumilia, P., Cassinari, A., Torri, G., Guerrini, M., Casu, B., Nader, H. B. and Dietrich, C. P. (1993). Conformation of the unsaturated uronic acid residues of glycosaminoglycan disaccharides. *J. Carbohydr. Chem.* **12**: 523-535.
- Raman, R., Venkataraman, G., Ernst, S., Sasisekharan, V. and Sasisekharan, R. (2003). Structural specificity of heparin binding in the fibroblast growth factor family of proteins. *Proc. Natl. Acad. Sci.* **100**(5): 2357-2362.
- Rapraeger, A. C. (2001). Molecular interactions of syndecans during development. *Sem. Cell. Develop. Biol.* **12**: 107-116.
- Reiland, J., Ott, V. L., Lebakken, C. S., Yeaman, C., McCarthy, J. and Rapraeger, A. C. (1996). Pervanadate activation of intracellular kinase leads to phosphorylation and shedding of syndecan-1. *Biochem. J.* **319**: 39-47.
- Ricard-Blum, S., Feraud, O., Lortat-Jacob, H., Rencurosi, A., Fukai, N., Dkhissi, F., Vittet, D., Imberty, A., Olsen, B. R. and Van_der_Rest, M. (2004). Characterization of endostatin binding to heparin and heparan sulfate by plasmon resonance and molecular modelling. *J. Biol. Chem* **279**(4): 2927-2936.
- Robinson, C. J., Harmer, N. J., Goodger, S. J., Blundell, T. L. and Gallagher, J. T. (2005). Cooperative dimerization of fibroblast growth factor 1 (FGF1) upon a single heparin saccharide may drive the formation of 2:2:1 FGF1-FGFR2c-Heparin ternary complexes. *J. Biol. Chem.* **280**(51): 42274-42282.

- Robinson, C. J., Mulloy, B., Gallagher, J. T. and Stringer, S. E. (2006). VEGF165-binding sites within heparan sulfate encompass two highly sulfated domain and can be liberated by K5 lyase. *J Biol Chem* **281**(3): 1731-1740.
- Robinson, C. J. and Stringer, S. E. (2001). The splice variants of vascular endothelial growth factor (VEGF) and their receptors. *J. Cell. Sci.* **114**(Pt 5): 853-65.
- Rohrmann, K., Niemann, R. and Buddecke, E. (1985). Two N-acetylgalactosaminyltransferase are involved in the biosynthesis of chondroitin sulfate. *Eur. J. Biochem* **148**(3): 463-9.
- Rong, J., Habuchi, H., Kimata, K., Lindahl, U. and Kusche_Gullberg, M. (2001). Substrate specificity of the heparan sulfate hexuronic acid 2-O-sulfotransferase. *Biochemistry.* **40**(18): 5548-55.
- Sadir, R., Forest, E. and Lortat_Jacob, H. (1998). The heparan sulfate binding sequence of interferon-gamma increased the on rate of the interferon-gamma-interferon-gamma receptor complex formation. *J. Biol. Chem.* **273**(18): 10919-25.
- Sanderson, R. D., Turnbull, J. T., Gallagher, J. T. and Lander, A. D. (1994). Fine structure of heparan sulfate regulates syndecan-1 function and cell behaviour. *J. Biol. Chem.* **269**: 13100-13106.
- Sasaki, T., Fukai, N., Mann, K., Gohring, W., Olsen, B. R. and Timpl, R. (1998). Structure, function and tissue forms of the C-terminal globular domain of collagen XVIII containing the angiogenesis inhibitor endostatin. *EMBO J.* **17**(15): 4249-56.
- Sasaki, T., Larsson, H., Kreuger, J., Salmivirta, M., Claesson-Welsh, L., Lindahl, U., Hohenester, E. and Timpl, R. (1999). Structural basis and potential role of heparin/heparan sulfate binding to the angiogenesis inhibitor endostatin. *EMBO J.* **18**(22): 6240-8.
- Saunders, S., Paine-Saunders, S. and Lander, A. D. (1997). Expression of the cell surface proteoglycan glypican-5 is developmentally regulated in kidney, limb and brain. *Dev. Biol.* **190**: 78-93.
- Schlessinger, J., Plotnikov, A. N., Ibrahimi, O. A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., Linhardt, R. J. and Mohammadi, M. (2000). Crystal structure of a ternary FGF-FGFR-Heparin complex reveals a dual role for heparin in FGFR binding and dimerisation. *Mol. Cell.* **6**: 743-750.
- Schroeder, K., Hertzog, P. J., Ravasi, T. and Hume, D. A. (2004). Interferon-gamma: an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75**: 163-189.
- Shively, J. E. and Conrad, H. E. (1976). Formation of anhydrosugars in the chemical depolymerisation of heparin. *Biochemistry* **15**(18): 3932-3941.
- Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D. and Spear, P. G. (1999). A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell.* **99**(1): 13-22.
- Shworak, N. W., Liu, J., Petros, L. M., Zhang, L., Kobayashi, M., Copeland, N. G., Jenkins, N. A. and Rosenberg, R. D. (1999). Multiple isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. Isolation, characterization, and expression of human cDNAs and identification of distinct genomic loci. *J. Biol. Chem.* **274**(8): 5170-84.

- Shworak, N. W., Motoaki, S., Mulligan, R. C. and Rosenberg, R. D. (1994). Characterisation of ryudocan glycosaminoglycan acceptor sites. *J. Biol. Chem.* **269**(33): 21204-21214.
- Silbert, J. E. (1967). Biosynthesis of heparin. 3. Formation of a sulfated glycosaminoglycan with a microsomal preparation from mast cell tumors. *J. Biol. Chem.* **242**(21): 5146-52.
- Silbert, J. E. (1967). Biosynthesis of heparin. IV. N-Deacetylation of a precursor glycosaminoglycan. *J. Biol. Chem.* **242**(21): 5153-7.
- Sisu, E., Tripathy, S., Mallet, J. M., Driguez, P. A., Herault, J. P., Sizun, P., Herbert, J. M., Petitou, M. and Sinay, P. (2003). Synthesis of new conformationally constrained pentasaccharides as molecular probes to investigate the biological activity of heparin. *Biochimie* **85**: 91-99.
- Sleeman, M., Fraser, J., McDonald, M., Yuan, S., White, D., Grandison, P., Kumble, K., Watson, J. D. and Murison, W. J. G. (2001). Identification of a new fibroblast growth factor receptor. FGFR5. *Gene* **271**: 171-182.
- Spillmann, D., Witt, D. and Lindahl, U. (1998). Defining the Interleukin-8-binding domain of heparan sulfate. *J. Biol. Chem.* **273**(25): 15487-15493.
- Stickens, D., Zak, B. M., Rougier, N., Esko, J. and Werb, Z. (2005). Mice deficient in Ext2 lack heparan sulfate and develop exostoses. *Development* **132**(22): 5055-5068.
- Stipp, C. S., Litwack, E. D. and Lander, A. D. (1994). Cerebroglycan: an integral membrane heparan sulfate proteoglycan that is unique to the developing nervous system and expressed specifically during neuronal development. *J. Cell. Biol.* **124**: 149-160.
- Stringer, S. E., Forster, M. J., Mulloy, B., Bishop, C. R., Graham, G. J. and Gallagher, J. T. (2002). Characterization of the binding site on heparan sulfate for macrophage inflammatory protein 1alpha. *Blood*. **100**(5): 1543-50.
- Stringer, S. E. and Gallagher, J. T. (1997). Specific binding of the chemokine platelet factor 4 to heparan sulfate. *J. Biol. Chem.* **272**(33): 20508-14.
- Thiel, D. J., Le-Du, M. H., Walter, R. L., D'arcy, A., Chene, C., Fountoulakis, M., Garotta, G., Winkler, F. K. and Ealick, S. E. (2000). Observation of an unexpected third receptor molecule in the crystal structure of human interferon-gamma receptor complex. *Structure* **8**(9): 927-936.
- Toida, T., Yoshida, H., Toyoda, H., Koshiishi, I., Imanari, T., Hileman, R. E., Fromm, J. R. and Linhardt, R. J. (1997). Structural differences and the presence of unsubstituted amino groups in heparan sulphates from different tissues and species. *Biochem. J.* **322** (Pt 2): 499-506.
- Tsuda, H., Yamada, S., Yamane, Y., Yoshida, K., Hopwood, J. J. and Sugahara, K. (1996). Structures of five sulfated hexasaccharides prepared from porcine intestinal heparin using bacterial heparinase. *J. Biol. Chem.* **271**(18): 10495-502.
- Tumova, S., Woods, A. and Couchman, J. R. (2000). Heparan sulfate chains from glypicans and syndecans bind the Hep II domain of fibronectin similarly despite minor structural differences. *J. Biol. Chem.* **275**: 9410-9417.

- Valla, S., Li, J., Ertesvag, H., Barbeyron, T. and Lindahl, U. (2001). Hexuronyl C5-epimerases in aliginat and glycosaminoglycan biosynthesis. *Biochimie*. **83**: 819-830.
- van_Boeckel, C. A. A., van_Aelst, S. F., Wagenaars, G. N., Mellema, J. R., Paulsen, H., Peters, T., Pollex, A. and Sinnwell, V. (1987). Conformational analysis of synthetic heparin-like oligosaccharides containing α -L-idopyranosyluronic acid. *Recl. Trav. Chim. Pay-Bas*. **106**: 19-29.
- Vanhaverbeke, C., Simorre, J., Sadir, R., Gans, P. and Lortat_Jacob, H. (2004). NMR characterization of the interaction between the C-terminal domain of interferon- γ and heparin derived oligosaccharides. *Biochem. J*. **384**: 93-99.
- Verli, H. and Guimaraes, J. (2004). Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr. Res*. **339**: 281-290.
- Vives, R. R., Goodger, S. J. and Pye, D. A. (2001). Combined strong anion-exchange HPLC and PAGE approach for the purification of heparan sulphate oligosaccharides. *Biochem. J*. **354**: 141-147.
- Vives, R. R., Pye, D. A., Salmivirta, M., Hopwood, J. J., Lindahl, U. and Gallagher, J. T. (1999). Sequence analysis of heparan sulphate and heparin oligosaccharides. *Biochem. J*. **339 (Pt 3)**: 767-73.
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. and Case, D. A. (2004). Development and testing of a general Amber force field. *J. Comput. Chem*. **25(9)**: 1157-1174.
- Westling, C. and Lindahl, U. (2002). Location of N-unsubstituted glucosamine residues in heparan sulfate. *J Biol Chem* **277(51)**: 49247-49255.
- Whitelock, J. M., Graham, L. D., Melrose, J., Murdoch, A. D., Iozzo, R. V. and Underwood, P. A. (1999). Human perlecan immunoprecipitated from different endothelial cell sources has different adhesive properties to vascular cells. *Matrix. Biol*. **18**: 163-178.
- Whitford, D. (2005). Proteins, structure and function. Chichester, John Wiley and Sons, Ltd.
- Woods, A. and Couchman, J. R. (1994). Syndecan 4 heparan sulfate proteoglycan is a selectively enriched and widespread focal adhesion component. *Mol. Biol. Cell*. **5(2)**: 183-92.
- Yabushita, H., Noguchi, Y., Habuchi, H., Ashikari, S., Nakabe, K., Fujita, M., Noguchi, M., Esko, J. and Kimata, K. (2002). Effects of chemically modified heparin on Chlamydia trachomatis serovar L2 infection of eukaryotic cells in culture. *Glycobiology* **12(5)**: 345-351.
- Yamada, S., Yamane, Y., Tsuda, H., Yoshida, K. and Sugahara, K. (1998). A major common trisulfated hexasacchride core sequence, Hexuronic acid (2-sulfate)-Glucosamine (N-sulfate)-Iduronic acid-N-acetylglucosamine-Glucuronic acid-Glucosamine (N-sulfate), isolated from the low sulfated irregular region of porcine intestinal heparin. *J. Biol. Chem*. **273(4)**: 1863-1871.
- Yamada, S., Yoshida, K., Sugiura, M. and Sugahara, K. (1992). One and two dimensional $^1\text{H-NMR}$ characterisation of two series of sulfated disaccharides prepared from Chondroitin Sulfate and Heparan Sulfate/Heparin by bacterial eliminase digestion. *J. Biochem*. **112**: 440-447.

- Yamada, S., Yoshida, K., Sugiura, M., Sugahara, K., Khoo, K. H., Morris, H. R. and Dell, A. (1993). Structural studies on the bacterial lyase resistant tetrasaccharides derived from the antithrombin III binding site of porcine intestinal heparin. *J. Biol. Chem.* **268**(7): 4780-4787.
- Yates, E. A., Santini, F., de_Cristofano, B., Payre, N., Cosentino, C., Guerrini, M., Naggi, A., Torri, G. and Hricovini, M. (2000). Effect of substitution pattern on ^1H , ^{13}C chemical shifts and $^1\text{J}_{\text{CH}}$ coupling constants in heparin derivatives. *Carbohydr. Res.* **329**: 239-247.
- Yates, E. A., Santini, F., Guerrini, M., Naggi, A., Torri, G. and Casu, B. (1996). ^1H and ^{13}C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydr. Res.* **294**: 15-27.
- Zako, M., Dong, J., Goldberger, O., Bernfield, M., Gallagher, J. and Deakin, J. A. (2003). Syndecan-1 and -4 synthesised simultaneously by mouse mammary gland epithelial cells bear heparan sulfate chains that are apparently structurally indistinguishable. *J. Biol. Chem.* **278**(15): 13561-13569.
- Zhang, L., Lawrence, R., Schwartz, J. J., Bai, X., Wei, G., Esko, J. D. and Rosenberg, R. D. (2001). The effect of precursor structures on the action of glucosaminyl 3-O-sulfotransferase-1 and the biosynthesis of anticoagulant heparan sulfate. *J. Biol. Chem.* **276**(31): 28806-13.
- Zimmerman, P. and David, G. (1999). The syndecans, tuners of transmembrane signalling. *FASEB. J.* **13 Suppl**: S91-S100.
- Zimmermann, P., Tomatis, D., Rosas, M., Grootjans, J., Leenaerts, I., Degeest, G., Reekmans, G., Coomans, C. and David, G. (2001). Characterization of syntenin, a syndecan-binding PDZ protein, as a component of cell adhesion sites and microfilaments. *Mol. Biol. Cell.* **12**(2): 339-50.