



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 10:21 am GMT

PDB ID : 2B5T  
Title : 2.1 Angstrom structure of a nonproductive complex between antithrombin, synthetic heparin mimetic SR123781 and two S195A thrombin molecules  
Authors : Johnson, D.J.; Li, W.; Luis, S.A.; Carrell, R.W.; Huntington, J.A.  
Deposited on : 2005-09-29  
Resolution : 2.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

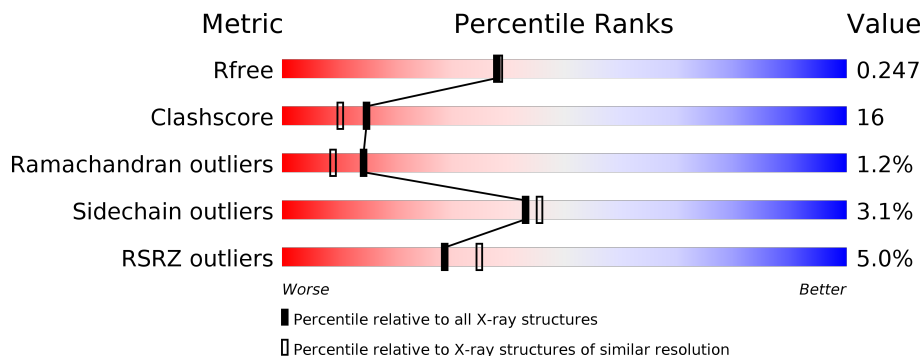
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	4608 (2.10-2.10)
Clashscore	122126	5109 (2.10-2.10)
Ramachandran outliers	120053	5059 (2.10-2.10)
Sidechain outliers	120020	5060 (2.10-2.10)
RSRZ outliers	108989	4497 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	49	<div> <div>2%</div> <div> <div></div> <div>65%</div> <div>31%</div> <div>•</div> </div> </div>
1	C	49	<div> <div>6%</div> <div> <div></div> <div>80%</div> <div>14%</div> <div>6%</div> </div> </div>
2	B	259	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>20%</div> <div>• •</div> </div> </div>
2	D	259	<div> <div>3%</div> <div> <div></div> <div>78%</div> <div>19%</div> <div>•</div> </div> </div>
3	I	432	<div> <div>8%</div> <div> <div></div> <div>60%</div> <div>33%</div> <div>• •</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	GU6	I	434	X	-	-	-
13	GU8	I	437	-	-	X	-
13	GU8	I	439	-	-	-	X
19	MAN	I	702	-	-	-	X
4	GOL	A	1102	-	-	X	-
4	GOL	B	1103	-	-	X	-
5	NAG	B	501	-	-	X	-
5	NAG	I	502	-	-	-	X
5	NAG	I	902	-	-	-	X
6	FUC	B	602	-	-	X	X
6	FUC	D	602	-	-	-	X
8	NDG	I	501	-	-	X	-

## 2 Entry composition

There are 20 unique types of molecules in this entry. The entry contains 9026 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	47	Total	C	N	O	S	0	1	0
			394	245	65	83	1			
1	C	46	Total	C	N	O	S	0	0	0
			372	233	60	78	1			

- Molecule 2 is a protein called Thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	252	Total	C	N	O	S	0	3	0
			2065	1317	368	366	14			
2	D	252	Total	C	N	O	S	0	0	0
			2022	1290	353	365	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	ALA	SER	ENGINEERED	UNP P00734
D	195	ALA	SER	ENGINEERED	UNP P00734

- Molecule 3 is a protein called Antithrombin-III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	415	Total	C	N	O	S	0	0	0
			3233	2065	540	609	19			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	137	ALA	SER	ENGINEERED	UNP P01008
I	317	CYS	VAL	ENGINEERED	UNP P01008
I	401	CYS	THR	ENGINEERED	UNP P01008

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



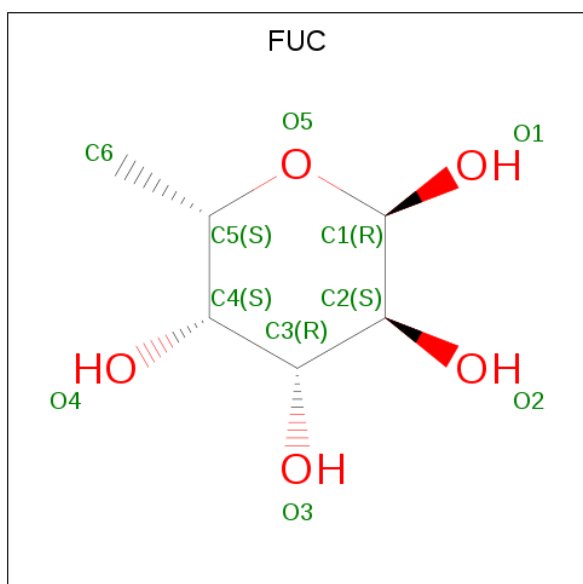
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	I	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



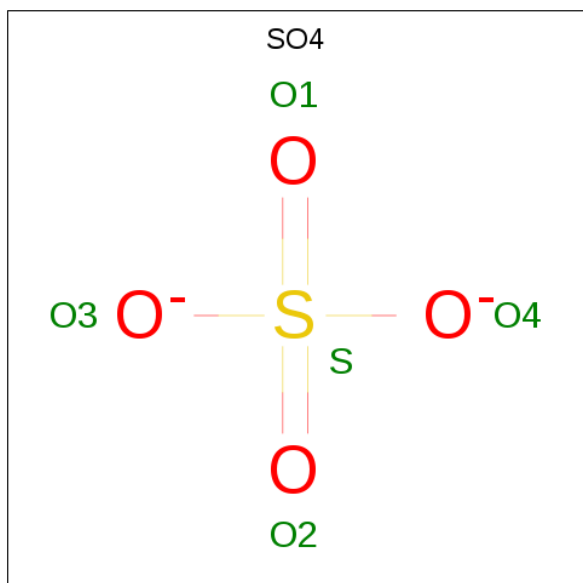
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	I	1	Total	C	N	O	0	0
			14	8	1	5		
5	I	1	Total	C	N	O	0	0
			14	8	1	5		
5	I	1	Total	C	N	O	0	0
			14	8	1	5		
5	I	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula:  $C_6H_{12}O_5$ ).



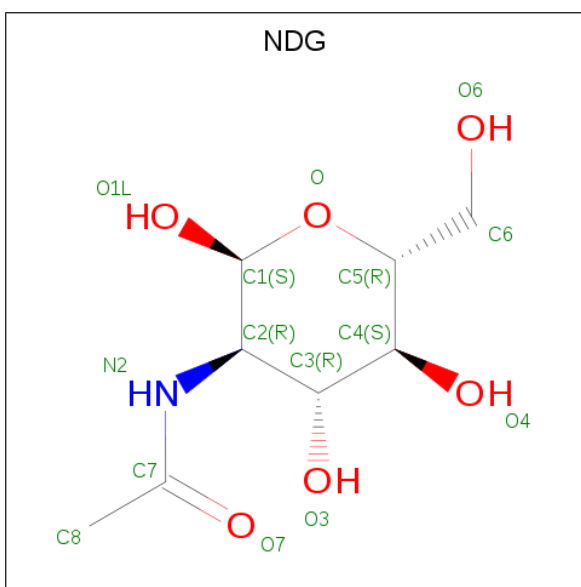
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			10	6	4		
6	D	1	Total	C	O	0	0
			10	6	4		
6	I	1	Total	C	O	0	0
			10	6	4		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



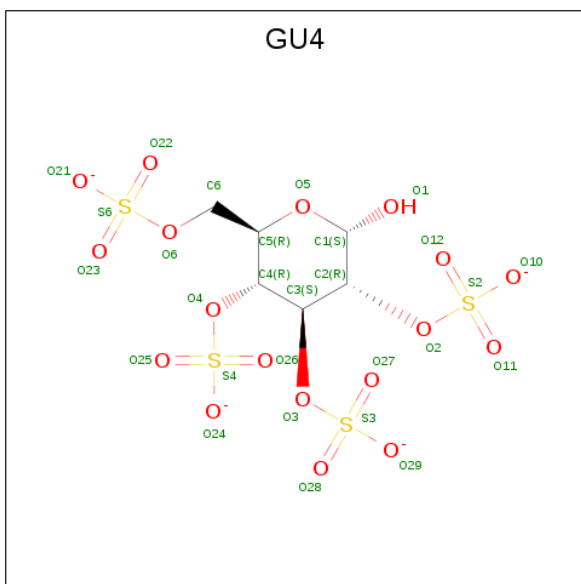
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	O	S	0	0
			5	4	1		
7	B	1	Total	O	S	0	0
			5	4	1		
7	B	1	Total	O	S	0	0
			5	4	1		
7	B	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	I	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	D	1	Total	C	N	O	0	0
			14	8	1	5		
8	I	1	Total	C	N	O	0	0
			14	8	1	5		

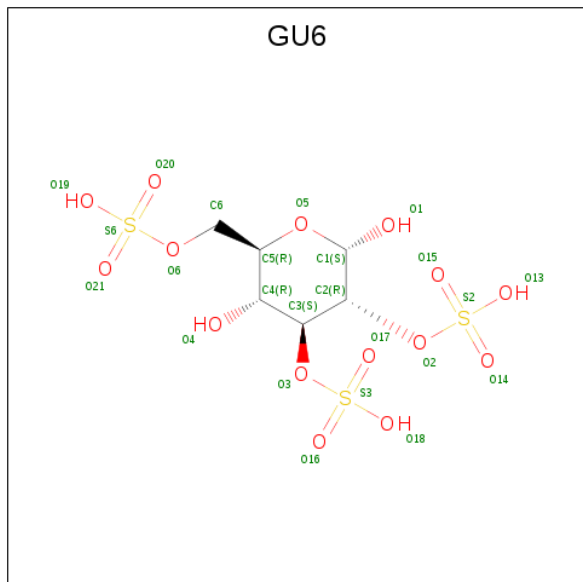
- Molecule 9 is 2,3,4,6-TETRA-O-SULFONATO-ALPHA-D-GLUCOPYRANOSE (three-letter code: GU4) (formula:  $C_6H_8O_{18}S_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	I	1	Total	C	O	S	0	0
			28	6	18	4		

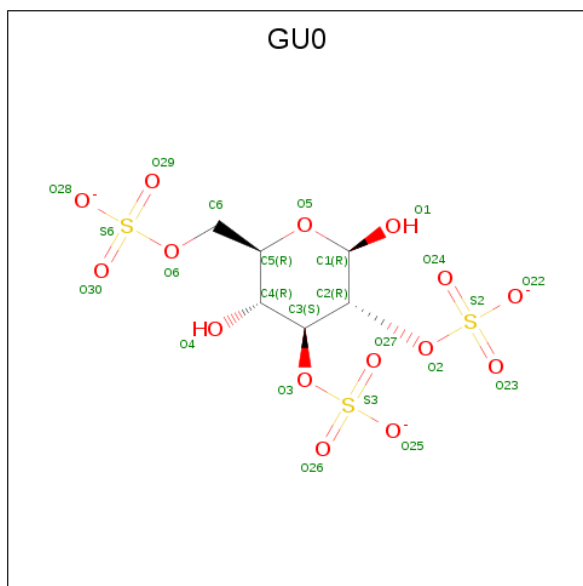


- Molecule 10 is 2,3,6-TRI-O-SULFONATO-ALPHA-D-GLUCOPYRANOSE (three-letter code: GU6) (formula:  $C_6H_{12}O_{15}S_3$ ).



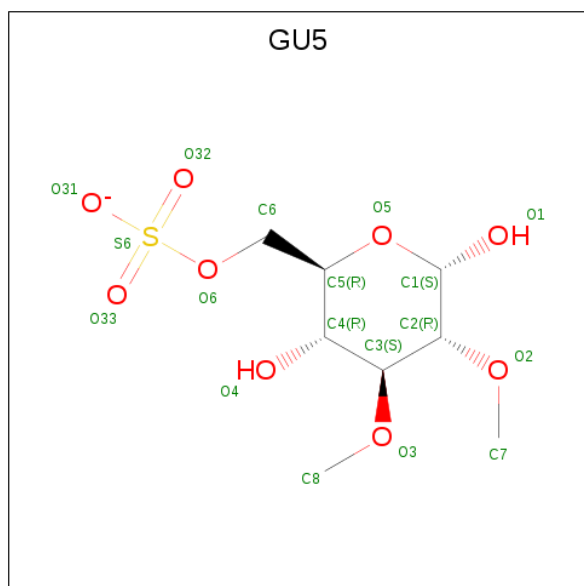
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	I	1	Total	C	O	S	0	0
			23	6	14	3		
10	I	1	Total	C	O	S	0	0
			23	6	14	3		

- Molecule 11 is 2,3,6-TRI-O-SULFONATO-ALPHA-L-GALACTOPYRANOSE (three-letter code: GU0) (formula:  $C_6H_9O_{15}S_3$ ).



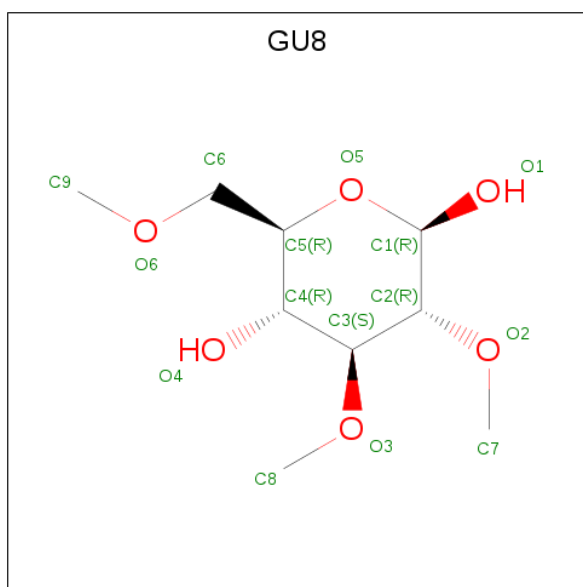
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	I	1	Total	C	O	S	0	0
			23	6	14	3		

- Molecule 12 is 2,3-DI-O-METHYL-6-O-SULFONATO-ALPHA-D-GLUCOPYRANOSE (three-letter code: GU5) (formula: C<sub>8</sub>H<sub>15</sub>O<sub>9</sub>S).



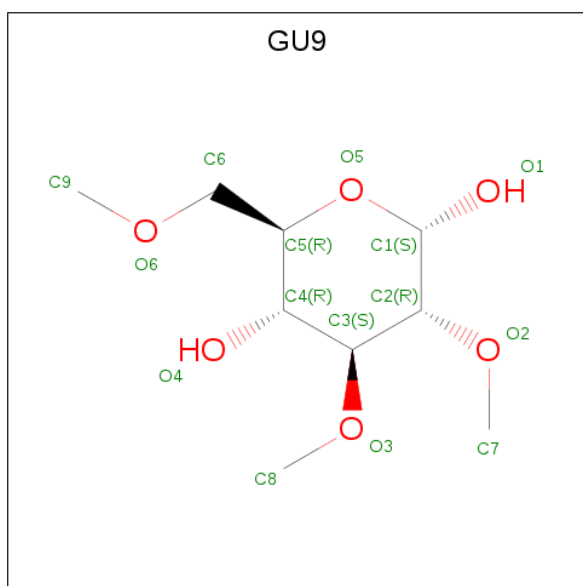
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
12	I	1	Total	C	O	S	0	0
			17	8	8	1		
12	I	1	Total	C	O	S	0	0
			17	8	8	1		

- Molecule 13 is 2,3,6-TRI-O-METHYL-BETA-D-GLUCOPYRANOSE (three-letter code: GU8) (formula: C<sub>9</sub>H<sub>18</sub>O<sub>6</sub>).



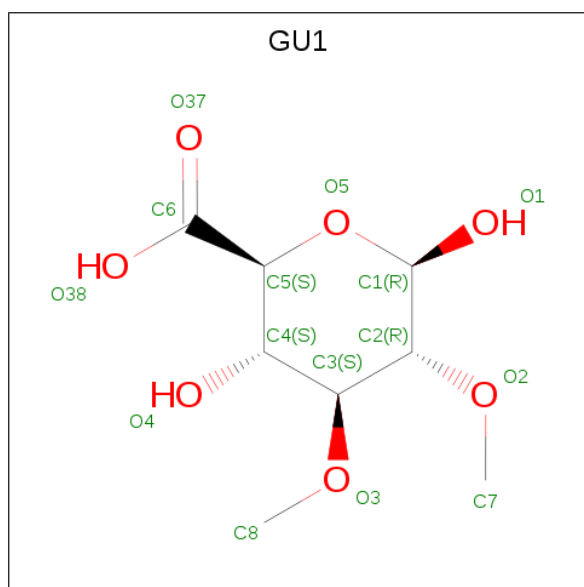
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	I	1	Total	C	O	0	0
			14	9	5		
13	I	1	Total	C	O	0	0
			14	9	5		
13	I	1	Total	C	O	0	0
			14	9	5		
13	I	1	Total	C	O	0	0
			14	9	5		

- Molecule 14 is 2,3,6-TRI-O-METHYL-ALPHA-D-GLUCOPYRANOSE (three-letter code: GU9) (formula: C<sub>9</sub>H<sub>18</sub>O<sub>6</sub>).



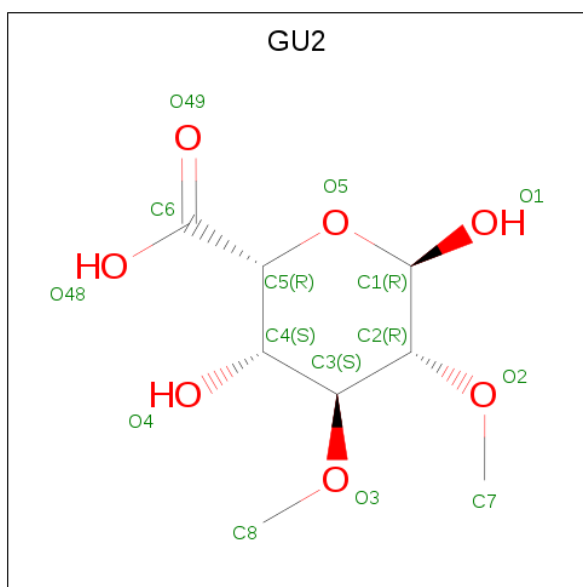
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
14	I	1	Total	C	O	0	0
			14	9	5		
14	I	1	Total	C	O	0	0
			14	9	5		
14	I	1	Total	C	O	0	0
			14	9	5		

- Molecule 15 is 2,3-DI-O-METHYL-BETA-D-GLUCOPYRANURONIC ACID (three-letter code: GU1) (formula: C<sub>8</sub>H<sub>14</sub>O<sub>7</sub>).



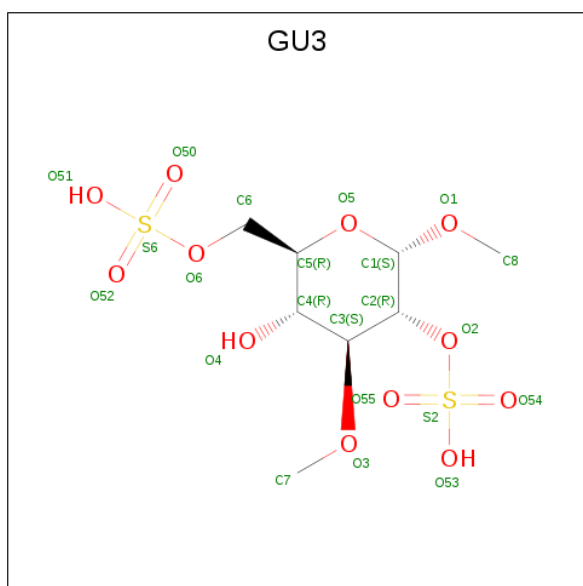
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
15	I	1	Total	C	O	0	0
			14	8	6		

- Molecule 16 is 2,3-DI-O-METHYL-ALPHA-L-IDOPYRANURONIC ACID (three-letter code: GU2) (formula: C<sub>8</sub>H<sub>14</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
16	I	1	Total	C	O	0	0
			14	8	6		

- Molecule 17 is methyl 3-O-methyl-2,6-di-O-sulfo-alpha-D-glucopyranoside (three-letter code: GU3) (formula:  $C_8H_{16}O_{12}S_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
17	I	1	Total	C	O	S	0	0
			21	8	11	2		

- Molecule 18 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
18	I	1	Total	C	O	0	0
			11	6	5		

- Molecule 19 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
19	I	1	Total	C	O	0	0
			11	6	5		
19	I	1	Total	C	O	0	0
			11	6	5		
19	I	1	Total	C	O	0	0
			10	6	4		

- Molecule 20 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
20	A	29	Total 29	O 29	0	0
20	B	132	Total 132	O 132	0	0
20	C	24	Total 24	O 24	0	0
20	D	117	Total 117	O 117	0	0
20	I	87	Total 87	O 87	0	0

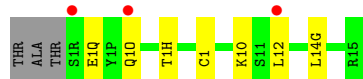
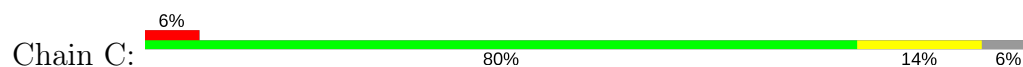
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

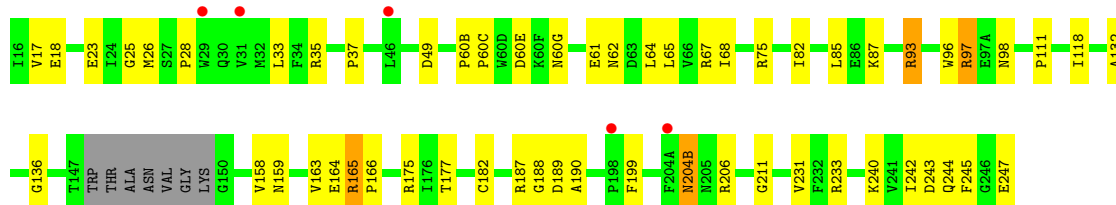
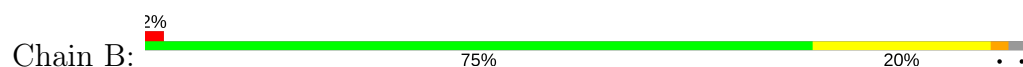
#### • Molecule 1: Thrombin



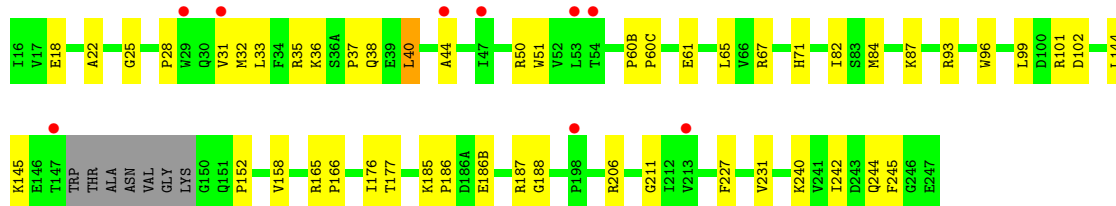
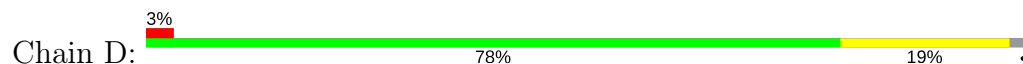
#### • Molecule 1: Thrombin



#### • Molecule 2: Thrombin



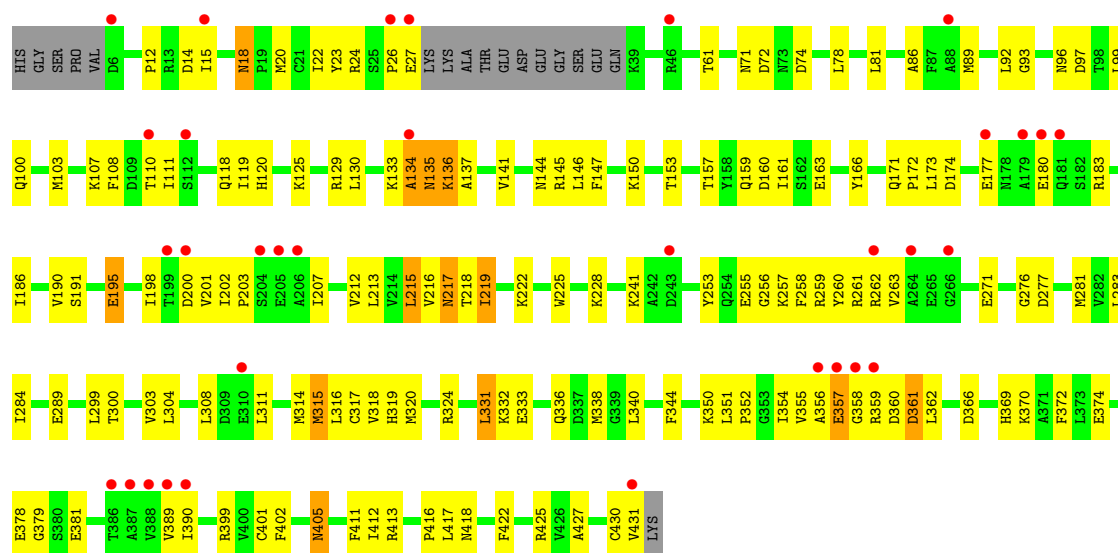
#### • Molecule 2: Thrombin



#### • Molecule 3: Antithrombin-III







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.28Å 59.53Å 108.41Å 77.55° 78.94° 68.08°	Depositor
Resolution (Å)	35.70 – 2.10 35.72 – 2.10	Depositor EDS
% Data completeness (in resolution range)	92.4 (35.70-2.10) 92.5 (35.72-2.10)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.10Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.209 , 0.247 0.210 , 0.247	Depositor DCC
$R_{free}$ test set	2937 reflections (4.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.6	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 61.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.014 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9026	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, BMA, NAG, NDG, GU3, GU2, GU1, GU0, SO4, GU5, GU4, GU9, GU8, GU6, FUC, MAN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.40	0/401	0.66	0/535
1	C	0.42	0/379	0.66	0/506
2	B	0.36	0/2118	0.66	0/2859
2	D	0.35	0/2074	0.64	0/2805
3	I	0.39	0/3299	0.65	2/4470 (0.0%)
All	All	0.38	0/8271	0.65	2/11175 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	402	PHE	N-CA-C	-5.51	96.14	111.00
3	I	360	ASP	N-CA-C	-5.47	96.24	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	394	0	369	15	0
1	C	372	0	346	5	0
2	B	2065	0	2030	52	0
2	D	2022	0	1965	36	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	I	3233	0	3143	124	0
4	A	12	0	16	5	0
4	B	30	0	40	4	0
4	D	24	0	32	4	0
4	I	6	0	8	2	0
5	B	14	0	12	8	0
5	I	56	0	50	2	0
6	B	10	0	10	6	0
6	D	10	0	10	4	0
6	I	10	0	10	5	0
7	B	20	0	0	2	0
7	D	5	0	0	0	0
7	I	5	0	0	0	0
8	D	14	0	12	4	0
8	I	14	0	11	8	0
9	I	28	0	7	2	0
10	I	46	0	17	2	0
11	I	23	0	6	1	0
12	I	34	0	24	6	0
13	I	56	0	60	21	0
14	I	42	0	45	15	0
15	I	14	0	10	3	0
16	I	14	0	10	0	0
17	I	21	0	14	2	0
18	I	11	0	8	0	0
19	I	32	0	27	0	0
20	A	29	0	0	0	0
20	B	132	0	0	0	0
20	C	24	0	0	0	0
20	D	117	0	0	1	0
20	I	87	0	0	4	0
All	All	9026	0	8292	274	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 16.

The worst 5 of 274 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:93[A]:ARG:HB3	2:B:93[A]:ARG:HH11	1.09	1.14
3:I:129:ARG:HD3	3:I:417:LEU:HD11	1.35	1.03

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:358:GLY:HA2	3:I:361:ASP:OD1	1.66	0.94
2:D:18:GLU:HG3	2:D:187:ARG:HG3	1.50	0.94
3:I:361:ASP:HB3	20:I:2052:HOH:O	1.68	0.92

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	46/49 (94%)	45 (98%)	1 (2%)	0	100	100
1	C	44/49 (90%)	42 (96%)	2 (4%)	0	100	100
2	B	251/259 (97%)	244 (97%)	7 (3%)	0	100	100
2	D	248/259 (96%)	239 (96%)	9 (4%)	0	100	100
3	I	411/432 (95%)	378 (92%)	21 (5%)	12 (3%)	5	1
All	All	1000/1048 (95%)	948 (95%)	40 (4%)	12 (1%)	14	9

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	134	ALA
3	I	136	LYS
3	I	195	GLU
3	I	241	LYS
3	I	357	GLU

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	43/43 (100%)	43 (100%)	0	100	100
1	C	40/43 (93%)	39 (98%)	1 (2%)	50	55
2	B	220/224 (98%)	210 (96%)	10 (4%)	30	29
2	D	214/224 (96%)	210 (98%)	4 (2%)	60	66
3	I	345/382 (90%)	331 (96%)	14 (4%)	33	33
All	All	862/916 (94%)	833 (97%)	29 (3%)	43	41

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	93	ARG
3	I	97	ASP
3	I	366	ASP
2	D	101	ARG
3	I	141	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	205	ASN
2	D	239	GLN
3	I	217	ASN
2	D	71	HIS
2	D	156	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

48 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	GOL	A	1102	-	5,5,5	0.72	0	5,5,5	0.48	0
4	GOL	A	1108	-	5,5,5	0.74	0	5,5,5	0.50	0
4	GOL	B	1101	-	5,5,5	0.80	0	5,5,5	0.44	0
4	GOL	B	1103	-	5,5,5	0.76	0	5,5,5	0.42	0
4	GOL	B	1104	-	5,5,5	0.62	0	5,5,5	0.41	0
4	GOL	B	1105	-	5,5,5	0.68	0	5,5,5	0.41	0
4	GOL	B	1109	-	5,5,5	0.70	0	5,5,5	0.43	0
7	SO4	B	2001	-	4,4,4	0.32	0	6,6,6	0.12	0
7	SO4	B	2002	-	4,4,4	0.33	0	6,6,6	0.08	0
7	SO4	B	2003	-	4,4,4	0.31	0	6,6,6	0.15	0
7	SO4	B	2006	-	4,4,4	0.38	0	6,6,6	0.08	0
5	NAG	B	501	2,6	14,14,15	0.65	0	17,19,21	0.94	1 (5%)
6	FUC	B	602	5	9,10,11	0.59	0	13,14,16	0.76	0
4	GOL	D	1106	-	5,5,5	0.62	0	5,5,5	0.40	0
4	GOL	D	1107	-	5,5,5	0.62	0	5,5,5	0.34	0
4	GOL	D	1110	-	5,5,5	0.73	0	5,5,5	0.42	0
4	GOL	D	1111	-	5,5,5	0.75	0	5,5,5	0.46	0
7	SO4	D	2004	-	4,4,4	0.31	0	6,6,6	0.11	0
8	NDG	D	501	2,6	14,14,15	0.71	0	17,19,21	0.86	1 (5%)
6	FUC	D	602	8	9,10,11	0.55	0	13,14,16	0.78	0
4	GOL	I	1112	-	5,5,5	0.67	0	5,5,5	0.38	0
7	SO4	I	2005	-	4,4,4	0.38	0	6,6,6	0.11	0
9	GU4	I	433	10	28,28,28	2.04	4 (14%)	31,45,45	0.96	1 (3%)
10	GU6	I	434	9,11	23,23,24	3.37	9 (39%)	24,36,38	3.62	7 (29%)
11	GU0	I	435	10,12	23,23,24	2.02	4 (17%)	24,36,38	0.91	2 (8%)
12	GU5	I	436	11,13	17,17,18	1.40	1 (5%)	20,24,26	0.82	0
13	GU8	I	437	12,14	14,14,15	0.50	0	16,18,20	0.76	0
14	GU9	I	438	13	14,14,15	0.59	0	16,18,20	0.94	1 (6%)
13	GU8	I	439	14	14,14,15	0.53	0	16,18,20	0.77	0
14	GU9	I	440	13	14,14,15	0.52	0	16,18,20	0.76	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
13	GU8	I	441	14	14,14,15	0.53	0	16,18,20	0.68	0
14	GU9	I	442	13	14,14,15	0.55	0	16,18,20	0.66	0
13	GU8	I	443	12,14	14,14,15	0.50	0	16,18,20	0.80	0
12	GU5	I	444	13,15	17,17,18	1.41	1 (5%)	20,24,26	0.88	1 (5%)
15	GU1	I	445	10,12	11,14,15	0.52	0	11,19,21	0.65	0
10	GU6	I	446	15,16	23,23,24	2.03	4 (17%)	24,36,38	1.14	3 (12%)
16	GU2	I	447	10,17	11,14,15	0.53	0	11,19,21	0.59	0
17	GU3	I	448	16	21,21,22	1.77	3 (14%)	25,31,33	1.15	1 (4%)
8	NDG	I	501	3,5,6	14,14,15	0.66	0	17,19,21	0.84	1 (5%)
5	NAG	I	502	8,18	14,14,15	0.57	0	17,19,21	0.85	1 (5%)
6	FUC	I	602	8	9,10,11	0.52	0	13,14,16	0.78	0
18	BMA	I	701	19,5	11,11,12	0.60	0	15,15,17	0.46	0
19	MAN	I	702	18	11,11,12	0.57	0	15,15,17	0.72	0
19	MAN	I	703	19,18	11,11,12	0.67	0	15,15,17	0.68	0
19	MAN	I	704	19	10,10,12	0.54	0	12,13,17	0.56	0
5	NAG	I	801	3	14,14,15	0.55	0	17,19,21	0.68	0
5	NAG	I	901	3,5	14,14,15	0.62	0	17,19,21	0.94	1 (5%)
5	NAG	I	902	5	14,14,15	0.55	0	17,19,21	0.63	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	1102	-	-	0/4/4/4	0/0/0/0
4	GOL	A	1108	-	-	0/4/4/4	0/0/0/0
4	GOL	B	1101	-	-	0/4/4/4	0/0/0/0
4	GOL	B	1103	-	-	0/4/4/4	0/0/0/0
4	GOL	B	1104	-	-	0/4/4/4	0/0/0/0
4	GOL	B	1105	-	-	0/4/4/4	0/0/0/0
4	GOL	B	1109	-	-	0/4/4/4	0/0/0/0
7	SO4	B	2001	-	-	0/0/0/0	0/0/0/0
7	SO4	B	2002	-	-	0/0/0/0	0/0/0/0
7	SO4	B	2003	-	-	0/0/0/0	0/0/0/0
7	SO4	B	2006	-	-	0/0/0/0	0/0/0/0
5	NAG	B	501	2,6	-	0/6/23/26	0/1/1/1
6	FUC	B	602	5	-	0/0/17/20	0/1/1/1
4	GOL	D	1106	-	-	0/4/4/4	0/0/0/0
4	GOL	D	1107	-	-	0/4/4/4	0/0/0/0
4	GOL	D	1110	-	-	0/4/4/4	0/0/0/0

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	D	1111	-	-	0/4/4/4	0/0/0/0
7	SO4	D	2004	-	-	0/0/0/0	0/0/0/0
8	NDG	D	501	2,6	-	0/6/23/26	0/1/1/1
6	FUC	D	602	8	-	0/0/17/20	0/1/1/1
4	GOL	I	1112	-	-	0/4/4/4	0/0/0/0
7	SO4	I	2005	-	-	0/0/0/0	0/0/0/0
9	GU4	I	433	10	-	0/21/41/41	0/1/1/1
10	GU6	I	434	9,11	1/1/7/8	0/16/32/36	0/1/1/1
11	GU0	I	435	10,12	-	0/16/32/36	0/1/1/1
12	GU5	I	436	11,13	-	0/10/26/30	0/1/1/1
13	GU8	I	437	12,14	-	0/7/23/27	0/1/1/1
14	GU9	I	438	13	-	0/7/23/27	0/1/1/1
13	GU8	I	439	14	-	0/7/23/27	0/1/1/1
14	GU9	I	440	13	-	0/7/23/27	0/1/1/1
13	GU8	I	441	14	-	0/7/23/27	0/1/1/1
14	GU9	I	442	13	-	0/7/23/27	0/1/1/1
13	GU8	I	443	12,14	-	0/7/23/27	0/1/1/1
12	GU5	I	444	13,15	-	0/10/26/30	0/1/1/1
15	GU1	I	445	10,12	-	0/4/24/28	0/1/1/1
10	GU6	I	446	15,16	-	0/16/32/36	0/1/1/1
16	GU2	I	447	10,17	-	0/4/24/28	1/1/1/1
17	GU3	I	448	16	-	0/15/31/35	0/1/1/1
8	NDG	I	501	3,5,6	-	0/6/23/26	0/1/1/1
5	NAG	I	502	8,18	-	0/6/23/26	0/1/1/1
6	FUC	I	602	8	-	0/0/17/20	0/1/1/1
18	BMA	I	701	19,5	-	0/2/19/22	0/1/1/1
19	MAN	I	702	18	-	0/2/19/22	0/1/1/1
19	MAN	I	703	19,18	-	0/2/19/22	0/1/1/1
19	MAN	I	704	19	-	0/2/16/22	0/1/1/1
5	NAG	I	801	3	-	0/6/23/26	0/1/1/1
5	NAG	I	901	3,5	-	0/6/23/26	0/1/1/1
5	NAG	I	902	5	-	0/6/23/26	0/1/1/1

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	I	434	GU6	O3-C3	-6.61	1.36	1.47
10	I	434	GU6	C4-C3	-6.43	1.39	1.52
10	I	434	GU6	O6-S6	-5.13	1.43	1.56
12	I	436	GU5	O6-S6	-4.97	1.43	1.56
17	I	448	GU3	O6-S6	-4.97	1.43	1.56

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	I	448	GU3	C2-O2-S2	-3.79	111.67	118.97
10	I	446	GU6	C2-O2-S2	-3.11	112.99	118.97
10	I	434	GU6	C4-C3-C2	-2.90	104.88	110.60
14	I	438	GU9	C4-C5-C6	-2.71	107.33	112.66
5	I	901	NAG	C2-N2-C7	-2.25	119.66	122.94

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
10	I	434	GU6	C5

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	I	447	GU2	C1-C2-C3-C4-C5-O5

32 monomers are involved in 74 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1102	GOL	5	0
4	B	1103	GOL	4	0
7	B	2002	SO4	1	0
7	B	2006	SO4	1	0
5	B	501	NAG	8	0
6	B	602	FUC	6	0
4	D	1106	GOL	1	0
4	D	1107	GOL	1	0
4	D	1110	GOL	1	0
4	D	1111	GOL	1	0
8	D	501	NDG	4	0
6	D	602	FUC	4	0
4	I	1112	GOL	2	0
9	I	433	GU4	2	0
11	I	435	GU0	1	0
12	I	436	GU5	4	0
13	I	437	GU8	8	0
14	I	438	GU9	4	0
13	I	439	GU8	5	0
14	I	440	GU9	5	0
13	I	441	GU8	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	I	442	GU9	6	0
13	I	443	GU8	2	0
12	I	444	GU5	2	0
15	I	445	GU1	3	0
10	I	446	GU6	2	0
17	I	448	GU3	2	0
8	I	501	NDG	8	0
5	I	502	NAG	1	0
6	I	602	FUC	5	0
5	I	901	NAG	1	0
5	I	902	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	47/49 (95%)	-0.11	1 (2%) 63 68	22, 34, 47, 66	1 (2%)
1	C	46/49 (93%)	0.02	3 (6%) 19 24	28, 38, 51, 58	0
2	B	252/259 (97%)	0.06	5 (1%) 65 69	14, 31, 50, 63	3 (1%)
2	D	252/259 (97%)	0.13	9 (3%) 42 49	18, 33, 47, 64	0
3	I	415/432 (96%)	0.48	33 (7%) 12 16	34, 51, 77, 89	0
All	All	1012/1048 (96%)	0.24	51 (5%) 29 35	14, 41, 68, 89	4 (0%)

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	388	VAL	7.0
3	I	358	GLY	6.1
3	I	386	THR	5.6
3	I	431	VAL	5.6
3	I	134	ALA	4.4

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	B	501	14/15	0.46	0.38	73,77,84,88	0
19	MAN	I	702	11/12	0.51	0.67	115,116,117,118	0
5	NAG	I	902	14/15	0.59	0.45	102,104,105,106	0
19	MAN	I	704	10/12	0.66	0.34	103,104,105,105	0
18	BMA	I	701	11/12	0.70	0.33	103,105,110,113	0
19	MAN	I	703	11/12	0.71	0.32	99,101,102,103	0
6	FUC	D	602	10/11	0.72	0.47	83,84,85,85	0
4	GOL	B	1101	6/6	0.73	0.35	49,55,57,60	0
5	NAG	I	901	14/15	0.74	0.23	92,94,96,99	0
8	NDG	D	501	14/15	0.74	0.31	59,66,74,80	0
14	GU9	I	438	14/15	0.75	0.37	81,86,89,90	0
4	GOL	I	1112	6/6	0.78	0.28	80,81,81,82	0
13	GU8	I	439	14/15	0.79	0.48	88,92,93,94	0
4	GOL	B	1103	6/6	0.80	0.20	63,65,65,66	0
6	FUC	B	602	10/11	0.80	0.43	90,91,91,91	0
5	NAG	I	502	14/15	0.80	0.42	98,100,101,103	0
4	GOL	D	1110	6/6	0.81	0.23	72,75,76,77	0
8	NDG	I	501	14/15	0.83	0.29	83,88,94,98	0
6	FUC	I	602	10/11	0.84	0.47	101,103,104,104	0
4	GOL	B	1105	6/6	0.85	0.25	95,95,95,95	0
4	GOL	A	1108	6/6	0.86	0.26	73,76,77,78	0
4	GOL	B	1109	6/6	0.86	0.17	62,62,64,64	0
14	GU9	I	442	14/15	0.87	0.24	69,75,80,81	0
4	GOL	D	1111	6/6	0.88	0.26	63,64,64,66	0
4	GOL	D	1106	6/6	0.88	0.13	66,67,67,68	0
7	SO4	B	2003	5/5	0.89	0.21	85,86,87,88	0
5	NAG	I	801	14/15	0.89	0.21	66,68,73,74	0
4	GOL	A	1102	6/6	0.89	0.27	54,57,58,60	0
14	GU9	I	440	14/15	0.90	0.23	84,84,86,87	0
13	GU8	I	437	14/15	0.90	0.22	54,59,65,70	0
13	GU8	I	441	14/15	0.91	0.25	79,83,84,85	0
16	GU2	I	447	14/15	0.92	0.17	61,64,69,70	0
17	GU3	I	448	21/22	0.94	0.14	56,61,66,67	0
13	GU8	I	443	14/15	0.94	0.11	60,63,69,69	0
7	SO4	B	2001	5/5	0.94	0.12	85,85,86,87	0
4	GOL	D	1107	6/6	0.95	0.13	32,41,44,45	0
10	GU6	I	446	23/24	0.95	0.10	47,61,65,67	0
4	GOL	B	1104	6/6	0.95	0.13	30,35,37,46	0
7	SO4	B	2002	5/5	0.96	0.12	89,89,90,91	0
12	GU5	I	444	17/18	0.96	0.10	52,57,59,59	0
15	GU1	I	445	14/15	0.96	0.09	51,54,57,57	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	SO4	B	2006	5/5	0.97	0.19	56,56,58,59	0
11	GU0	I	435	23/24	0.98	0.10	28,33,41,43	0
12	GU5	I	436	17/18	0.98	0.09	38,42,47,50	0
7	SO4	D	2004	5/5	0.98	0.10	61,63,64,64	0
10	GU6	I	434	23/24	0.98	0.11	25,30,32,37	0
7	SO4	I	2005	5/5	0.98	0.09	63,63,64,64	0
9	GU4	I	433	28/28	0.99	0.09	22,26,30,33	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.