



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 5, 2018 – 09:24 AM EDT

PDB ID : 3B9F
Title : 1.6 Å structure of the PCI-thrombin-heparin complex
Authors : Li, W.; Adams, T.E.; Huntington, J.A.
Deposited on : 2007-11-05
Resolution : 1.60 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

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 : rb-20031021
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) : rb-20031021

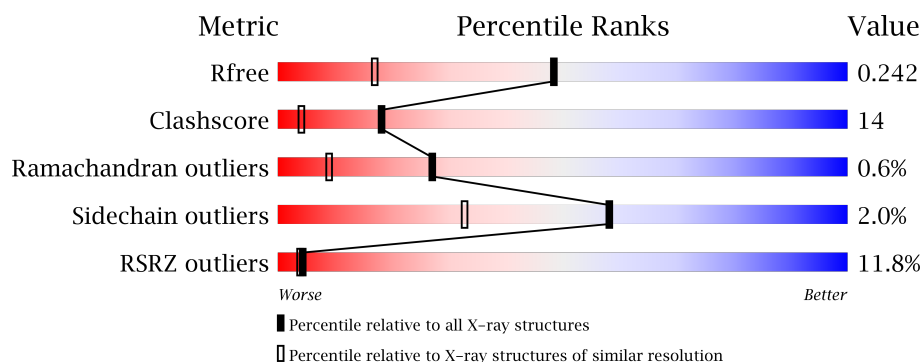
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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	2957 (1.60-1.60)
Clashscore	122126	3202 (1.60-1.60)
Ramachandran outliers	120053	3117 (1.60-1.60)
Sidechain outliers	120020	3116 (1.60-1.60)
RSRZ outliers	108989	2883 (1.60-1.60)

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 ≥ 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	L	49	<div> <div>6%</div> <div> <div></div> <div>82%</div> <div>8%</div> <div>10%</div> </div> </div>
2	H	259	<div> <div>5%</div> <div> <div></div> <div>71%</div> <div>26%</div> <div>..</div> </div> </div>
3	I	395	<div> <div>16%</div> <div> <div></div> <div>68%</div> <div>22%</div> <div>• 10%</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
6	SGN	H	1	-	-	-	X
7	IDS	H	2	-	-	-	X
9	GOL	I	390	-	-	X	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 6092 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 Prothrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	44	Total	C	N	O	S	0	2	0
			377	236	61	79	1			

- Molecule 2 is a protein called Prothrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	253	Total	C	N	O	S	0	34	0
			2294	1456	412	410	16			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	195	ALA	SER	engineered mutation	UNP P00734

- Molecule 3 is a protein called Plasma serine protease inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	356	Total	C	N	O	S	5	9	0
			2798	1796	461	524	17			

There are 25 discrepancies between the modelled and reference sequences:

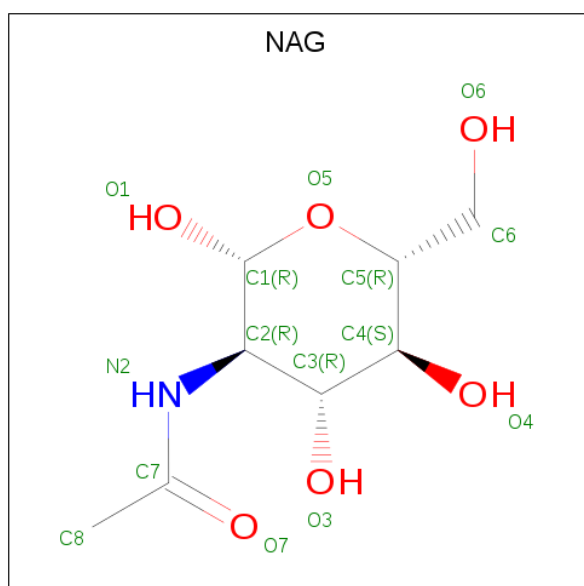
Chain	Residue	Modelled	Actual	Comment	Reference
I	-7	MET	-	expression tag	UNP P05154
I	-6	GLY	-	expression tag	UNP P05154
I	-5	HIS	-	expression tag	UNP P05154
I	-4	HIS	-	expression tag	UNP P05154
I	-3	HIS	-	expression tag	UNP P05154
I	-2	HIS	-	expression tag	UNP P05154
I	-1	HIS	-	expression tag	UNP P05154
I	0	HIS	-	expression tag	UNP P05154
I	1	HIS	-	expression tag	UNP P05154

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	2	HIS	-	expression tag	UNP P05154
I	3	HIS	-	expression tag	UNP P05154
I	4	HIS	-	expression tag	UNP P05154
I	5	SER	-	expression tag	UNP P05154
I	6	SER	-	expression tag	UNP P05154
I	7	GLY	-	expression tag	UNP P05154
I	8	HIS	-	expression tag	UNP P05154
I	9	ILE	-	expression tag	UNP P05154
I	10	ASP	-	expression tag	UNP P05154
I	11	ASP	-	expression tag	UNP P05154
I	12	ASP	-	expression tag	UNP P05154
I	13	ASP	-	expression tag	UNP P05154
I	14	LYS	-	expression tag	UNP P05154
I	15	HIS	-	expression tag	UNP P05154
I	16	MET	-	expression tag	UNP P05154
I	45	ASN	SER	see sequence details	UNP P05154

- Molecule 4 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



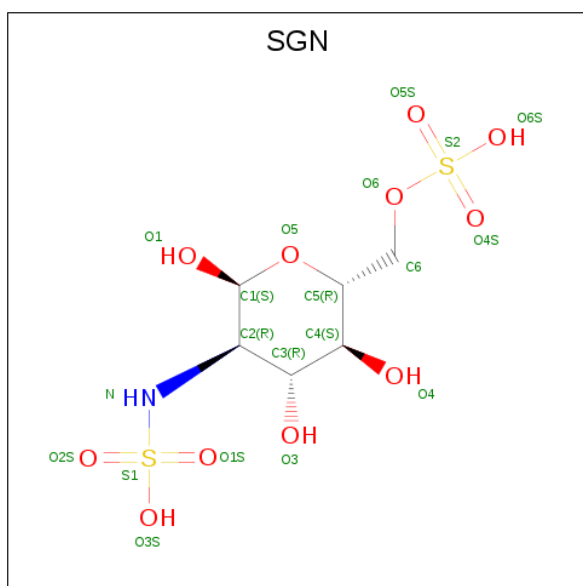
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	H	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: $C_6H_{12}O_5$).



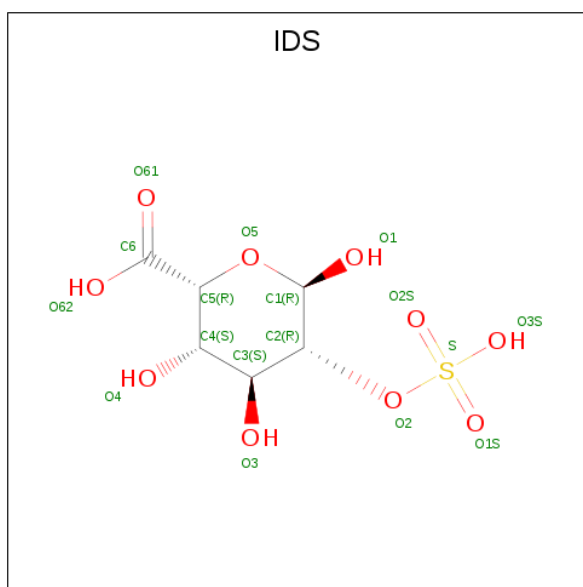
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	H	1	Total	C	O	0	0
			10	6	4		

- Molecule 6 is N,O6-DISULFO-GLUCOSAMINE (three-letter code: SGN) (formula: $C_6H_{13}NO_{11}S_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	H	1	Total	C	N	O	S	0	0
			15	6	1	7	1		

- Molecule 7 is 2-O-sulfo-alpha-L-idopyranuronic acid (three-letter code: IDS) (formula: $C_6H_{10}O_{10}S$).



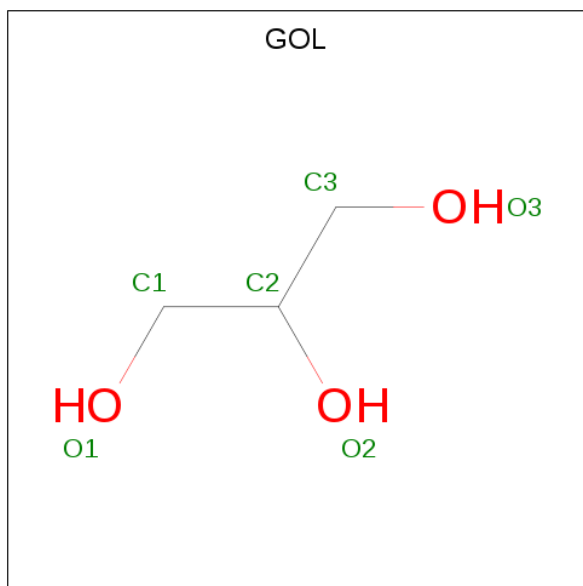
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	H	1	Total	C	O	S	0	0
			16	6	9	1		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	H	1	Total	O	S	0	0
			5	4	1		
8	I	1	Total	O	S	0	0
			5	4	1		
8	I	1	Total	O	S	0	0
			5	4	1		

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	H	1	Total C O 6 3 3	0	0
9	H	1	Total C O 6 3 3	0	0
9	H	1	Total C O 6 3 3	0	0
9	I	1	Total C O 6 3 3	0	0
9	I	1	Total C O 6 3 3	0	0

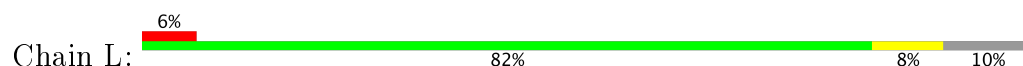
- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	L	83	Total O 83 83	0	0
10	H	272	Total O 272 272	0	0
10	I	168	Total O 168 168	0	0

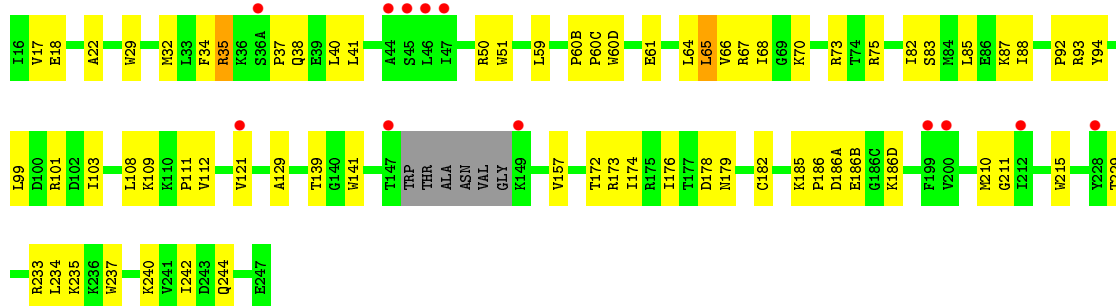
3 Residue-property plots [i](#)

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.

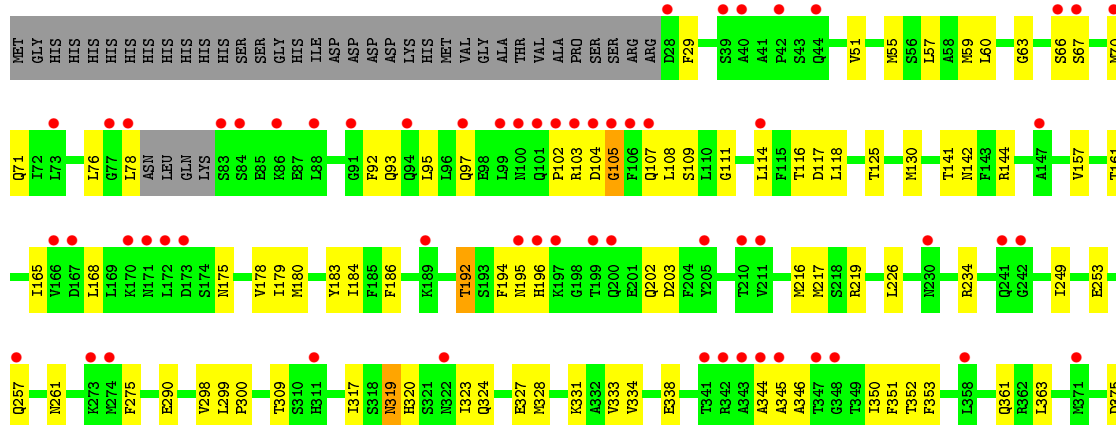
• Molecule 1: Prothrombin



• Molecule 2: Prothrombin



• Molecule 3: Plasma serine protease inhibitor





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	44.05Å 48.83Å 97.85Å 78.72° 81.52° 77.69°	Depositor
Resolution (Å)	30.95 – 1.60 30.95 – 1.60	Depositor EDS
% Data completeness (in resolution range)	93.4 (30.95-1.60) 93.6 (30.95-1.60)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.94 (at 1.60Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.208 , 0.234 0.218 , 0.242	Depositor DCC
R_{free} test set	2928 reflections (3.05%)	wwPDB-VP
Wilson B-factor (Å ²)	19.0	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 61.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6092	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.18% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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, NAG, IDS, FUC, SGN, SO4

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	L	0.35	0/384	0.63	0/514
2	H	0.31	0/2347	0.70	0/3169
3	I	0.29	0/2851	0.56	0/3860
All	All	0.30	0/5582	0.63	0/7543

There are no bond length outliers.

There are no bond angle outliers.

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	L	377	0	352	3	0
2	H	2294	0	2222	84	0
3	I	2798	0	2718	76	0
4	H	14	0	12	1	0
5	H	10	0	10	0	0
6	H	15	0	10	6	0
7	H	16	0	6	3	0
8	H	5	0	0	0	0
8	I	10	0	0	0	0
9	H	18	0	24	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	I	12	0	16	7	0
10	H	272	0	0	5	0
10	I	168	0	0	6	0
10	L	83	0	0	1	0
All	All	6092	0	5370	158	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 14.

The worst 5 of 158 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:H:185[A]:LYS:HB2	2:H:186[A]:PRO:C	1.77	1.05
2:H:185[A]:LYS:HB2	2:H:186(A)[A]:ASP:N	1.73	1.02
2:H:139:THR:HG22	2:H:157[A]:VAL:HG22	1.54	0.89
2:H:67[B]:ARG:HG2	2:H:82[B]:ILE:HD12	1.59	0.84
3:I:179:ILE:HD12	3:I:327:GLU:HB2	1.64	0.78

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	L	44/49 (90%)	43 (98%)	1 (2%)	0	100	100
2	H	282/259 (109%)	275 (98%)	7 (2%)	0	100	100
3	I	361/395 (91%)	337 (93%)	20 (6%)	4 (1%)	16	3
All	All	687/703 (98%)	655 (95%)	28 (4%)	4 (1%)	27	9

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	104	ASP
3	I	346	ALA
3	I	105	GLY
3	I	102	PRO

5.3.2 Protein sidechains ⓘ

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	L	42/43 (98%)	42 (100%)	0	100	100
2	H	241/224 (108%)	237 (98%)	4 (2%)	63	40
3	I	296/347 (85%)	289 (98%)	7 (2%)	52	25
All	All	579/614 (94%)	568 (98%)	11 (2%)	58	36

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	78	LEU
3	I	192	THR
3	I	203	ASP
2	H	182	CYS
3	I	196	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
3	I	195	ASN
3	I	196	HIS
3	I	319	ASN
3	I	175	ASN
3	I	294	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 ⓘ

12 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$
6	SGN	H	1	7	14,15,20	3.77	3 (21%)	17,22,31	1.38	1 (5%)
7	IDS	H	2	6	13,16,17	1.66	3 (23%)	13,24,26	0.61	0
9	GOL	H	3	-	5,5,5	0.28	0	5,5,5	0.27	0
9	GOL	H	4	-	5,5,5	0.30	0	5,5,5	0.23	0
4	NAG	H	778	2,5	14,14,15	0.51	0	17,19,21	0.71	0
5	FUC	H	780	4	9,10,11	0.43	0	13,14,16	0.40	0
8	SO4	H	781	-	4,4,4	0.34	0	6,6,6	0.08	0
9	GOL	H	782	-	5,5,5	0.30	0	5,5,5	0.15	0
8	SO4	I	388	-	4,4,4	0.34	0	6,6,6	0.07	0
8	SO4	I	389	-	4,4,4	0.34	0	6,6,6	0.07	0
9	GOL	I	390	-	5,5,5	0.28	0	5,5,5	0.27	0
9	GOL	I	391	-	5,5,5	0.29	0	5,5,5	0.27	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
6	SGN	H	1	7	-	0/6/23/31	0/1/1/1
7	IDS	H	2	6	-	0/5/25/29	1/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	GOL	H	3	-	-	0/4/4/4	0/0/0/0
9	GOL	H	4	-	-	0/4/4/4	0/0/0/0
4	NAG	H	778	2,5	-	0/6/23/26	0/1/1/1
5	FUC	H	780	4	-	0/0/17/20	0/1/1/1
8	SO4	H	781	-	-	0/0/0/0	0/0/0/0
9	GOL	H	782	-	-	0/4/4/4	0/0/0/0
8	SO4	I	388	-	-	0/0/0/0	0/0/0/0
8	SO4	I	389	-	-	0/0/0/0	0/0/0/0
9	GOL	I	390	-	-	0/4/4/4	0/0/0/0
9	GOL	I	391	-	-	0/4/4/4	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	H	2	IDS	O2-S	-4.72	1.44	1.56
7	H	2	IDS	O2S-S	2.16	1.53	1.45
7	H	2	IDS	O1S-S	2.17	1.53	1.45
6	H	1	SGN	S1-N	3.47	1.63	1.59
6	H	1	SGN	O1S-S1	9.52	1.53	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	1	SGN	O2S-S1-O1S	-4.24	109.29	120.24

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	H	2	IDS	C1-C2-C3-C4-C5-O5

6 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	H	1	SGN	6	0
7	H	2	IDS	3	0
9	H	4	GOL	1	0
4	H	778	NAG	1	0
9	I	390	GOL	6	0
9	I	391	GOL	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, 95th 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 ²)	Q<0.9
1	L	44/49 (89%)	-0.27	3 (6%) 17 16	12, 18, 28, 45	2 (4%)
2	H	253/259 (97%)	0.14	12 (4%) 31 29	9, 17, 35, 53	23 (9%)
3	I	355/395 (89%)	1.07	62 (17%) 1 1	14, 37, 59, 69	4 (1%)
All	All	652/703 (92%)	0.62	77 (11%) 4 4	9, 28, 54, 69	29 (4%)

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	343	ALA	12.6
3	I	345	ALA	10.9
3	I	102	PRO	7.2
3	I	103	ARG	7.0
3	I	106	PHE	6.8

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, 95th 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(\AA^2)	Q<0.9
6	SGN	H	1	15/20	0.30	0.63	65,72,76,76	0
7	IDS	H	2	16/17	0.54	0.43	71,72,76,76	0
9	GOL	I	391	6/6	0.66	0.17	60,60,61,61	0
4	NAG	H	778	14/15	0.67	0.23	50,54,59,63	0
9	GOL	H	4	6/6	0.76	0.21	55,55,55,55	2
8	SO4	I	389	5/5	0.82	0.26	76,76,76,76	2
5	FUC	H	780	10/11	0.84	0.48	66,68,68,69	0
9	GOL	H	3	6/6	0.84	0.18	48,49,49,50	3
8	SO4	H	781	5/5	0.87	0.14	77,77,77,77	3
9	GOL	I	390	6/6	0.87	0.12	44,45,45,46	2
9	GOL	H	782	6/6	0.91	0.12	20,24,25,27	3
8	SO4	I	388	5/5	0.96	0.10	70,70,70,71	3

6.5 Other polymers [i](#)

There are no such residues in this entry.