



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 08:35 PM BST

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 Full wwPDB X-ray Structure Validation 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

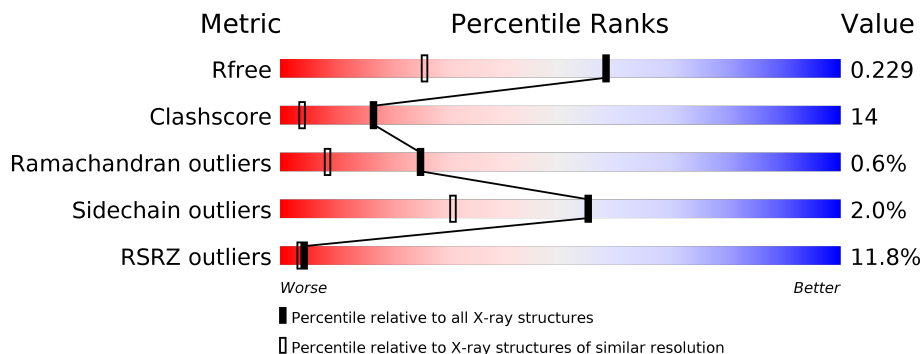
# 1 Overall quality at a glance

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))
$R_{free}$	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	
2	H	259	
3	I	395	
4	A	2	
5	B	2	

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
5	SGN	B	1	-	-	-	X
5	IDS	B	2	-	-	-	X
7	GOL	I	390	-	-	X	-

## 2 Entry composition i

There are 8 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
			Total	C	N	O	S			
1	L	44	377	236	61	79	1	0	2	0

- Molecule 2 is a protein called Prothrombin.

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

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
			Total	C	N	O	S			
3	I	356	2798	1796	461	524	17	5	9	0

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...

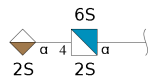
Chain	Residue	Modelled	Actual	Comment	Reference
I	2	HIS	-	expression tag	UNP P05154
I	3	HIS	-	expression tag	UNP P05154
I	4	HIS	-	expression tag	UNP P05154
I	5	SER	-	expression tag	UNP P05154
I	6	SER	-	expression tag	UNP P05154
I	7	GLY	-	expression tag	UNP P05154
I	8	HIS	-	expression tag	UNP P05154
I	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 an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	A	2	24	14	1	9	0	0	0

- Molecule 5 is an oligosaccharide called 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose.



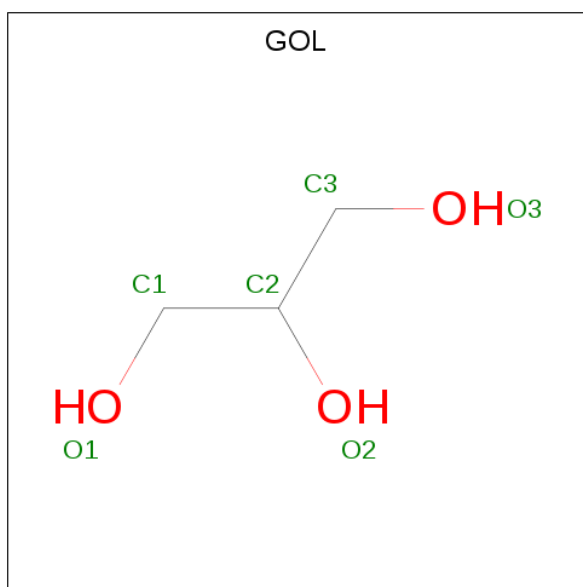
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	2	31	12	1	16	2	0	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

*Continued on next page...*

*Continued from previous page...*

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

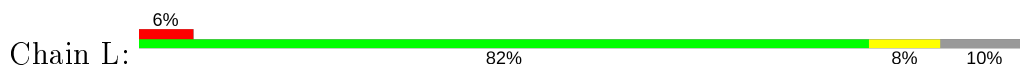
- Molecule 8 is water.

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

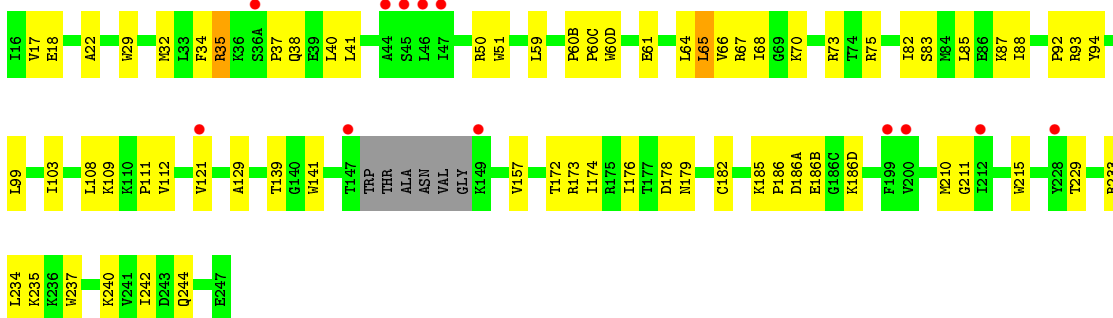
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

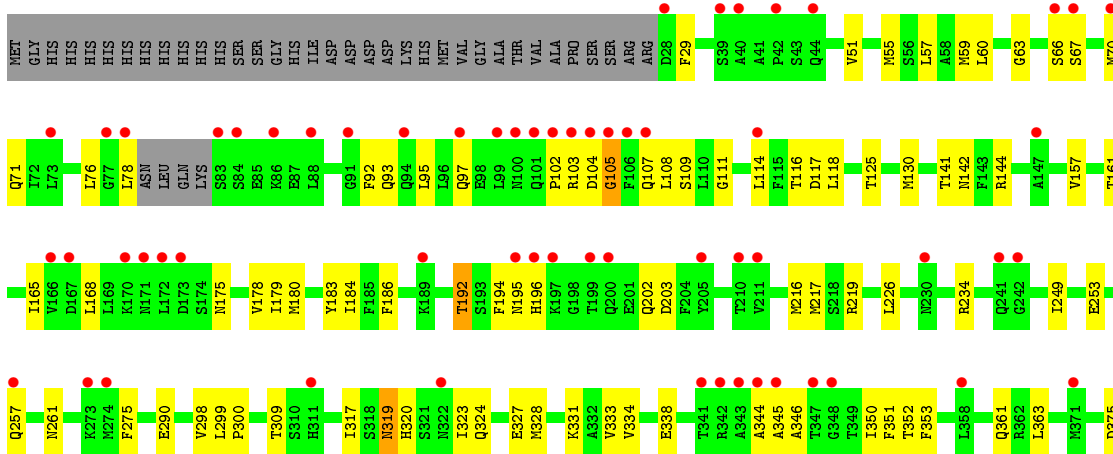
- Molecule 1: Prothrombin



- Molecule 2: Prothrombin



- Molecule 3: Plasma serine protease inhibitor







- Molecule 4: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain A: 50% 50%



- Molecule 5: 2-O-sulfo-alpha-L-idopyranuronic acid-(1-4)-2-deoxy-6-O-sulfo-2-(sulfoamino)-alpha-D-glucopyranose

Chain B: 100%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 1.60Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.208 , 0.234 0.202 , 0.229	Depositor DCC
$R_{free}$ test set	2928 reflections (3.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.0	Xtrriage
Anisotropy	0.068	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 61.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6092	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage'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.*

<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, 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	83	0
3	I	2798	0	2718	76	0
4	A	24	0	22	1	0
5	B	31	0	13	7	0
6	H	5	0	0	0	0
6	I	10	0	0	0	0
7	H	18	0	24	1	0
7	I	12	0	16	7	0
8	H	272	0	0	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	I	168	0	0	6	0
8	L	83	0	0	1	0
All	All	6092	0	5367	156	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.

All (156) 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
3:I:59:MET:HG2	3:I:130[A]:MET:SD	2.25	0.77
3:I:234:ARG:NH2	3:I:361:GLN:HE21	1.83	0.76
2:H:185[A]:LYS:HG3	2:H:186[A]:PRO:HA	1.70	0.73
2:H:185[A]:LYS:HB2	2:H:186[A]:PRO:CA	2.20	0.72
2:H:61:GLU:HG2	2:H:87:LYS:HA	1.71	0.71
2:H:237:TRP:HB2	5:B:1:SGN:H61	1.76	0.68
3:I:353:PHE:H	7:I:390:GOL:C1	2.06	0.68
2:H:185[A]:LYS:HB3	2:H:186(B)[A]:GLU:HB2	1.76	0.67
2:H:35:ARG:HB3	2:H:41[A]:LEU:HD11	1.76	0.67
2:H:59:LEU:HD13	2:H:88[B]:ILE:HG21	1.77	0.66
2:H:61:GLU:CG	2:H:87:LYS:HA	2.25	0.66
2:H:68:ILE:HD12	2:H:112[A]:VAL:HG21	1.77	0.65
2:H:108:LEU:HD13	2:H:112[A]:VAL:HG23	1.79	0.65
2:H:174[B]:ILE:CD1	3:I:351:PHE:HB3	2.28	0.64
3:I:55[B]:MET:HG2	3:I:92:PHE:CD1	2.33	0.64
3:I:157:VAL:HG21	3:I:168:LEU:HD22	1.80	0.63
2:H:34:PHE:CE1	2:H:67[A]:ARG:HD3	2.33	0.63
2:H:179:ASN:ND2	5:B:1:SGN:O3S	2.29	0.62
2:H:64:LEU:HD12	2:H:88[B]:ILE:HD11	1.82	0.62
3:I:161:THR:HG21	3:I:165:ILE:HB	1.80	0.62
3:I:353:PHE:H	7:I:390:GOL:H12	1.64	0.61
3:I:202:GLN:HB2	3:I:216[B]:MET:HG3	1.83	0.61
2:H:32[A]:MET:SD	2:H:70:LYS:HD3	2.40	0.60
1:L:1(Q):GLU:HG2	8:L:55:HOH:O	2.01	0.60
2:H:92:PRO:HG2	5:B:2:IDS:O6A	2.01	0.60
3:I:234:ARG:HH21	3:I:361:GLN:HE21	1.48	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:93:GLN:O	3:I:97:GLN:HG3	2.02	0.59
2:H:237:TRP:HD1	5:B:2:IDS:H3	1.69	0.58
3:I:178:VAL:HG21	3:I:323:ILE:HD11	1.85	0.58
2:H:93[A]:ARG:HG3	5:B:2:IDS:O6A	2.03	0.58
2:H:174[A]:ILE:HD11	3:I:350:ILE:HB	1.85	0.58
2:H:185[A]:LYS:HE2	2:H:186(D)[A]:LYS:O	2.04	0.58
2:H:173:ARG:NH2	3:I:350:ILE:HG23	2.20	0.57
3:I:116:THR:HG22	3:I:178:VAL:HG22	1.86	0.56
2:H:237:TRP:HB2	5:B:1:SGN:C6	2.35	0.56
2:H:61:GLU:CD	2:H:61:GLU:H	2.09	0.56
3:I:111:GLY:HA3	3:I:183:TYR:CE1	2.40	0.56
2:H:185[A]:LYS:HB3	2:H:186(B)[A]:GLU:N	2.21	0.55
2:H:60(D):TRP:CH2	7:I:390:GOL:H2	2.41	0.55
3:I:344:ALA:HB3	8:I:528:HOH:O	2.07	0.55
2:H:34:PHE:HB2	2:H:65:LEU:HD22	1.87	0.54
3:I:180:MET:HE2	3:I:328:MET:HG3	1.89	0.54
2:H:185[A]:LYS:CB	2:H:186(A)[A]:ASP:N	2.61	0.54
3:I:352:THR:HA	7:I:390:GOL:H11	1.90	0.54
2:H:185[A]:LYS:CB	2:H:186[A]:PRO:CA	2.86	0.53
3:I:118:LEU:HD11	3:I:142:ASN:HB2	1.90	0.53
3:I:67:SER:O	3:I:71:GLN:HG3	2.08	0.53
8:H:1014:HOH:O	4:A:1:NAG:H61	2.07	0.53
3:I:103:ARG:O	3:I:105:GLY:N	2.41	0.53
3:I:109:SER:O	3:I:184:ILE:HA	2.07	0.53
2:H:75[B]:ARG:HG2	2:H:75[B]:ARG:HH11	1.74	0.52
2:H:99:LEU:HD12	2:H:215:TRP:HB3	1.89	0.52
3:I:66:SER:O	3:I:70:MET:HG3	2.10	0.52
2:H:240[B]:LYS:CG	2:H:244[B]:GLN:HE21	2.23	0.52
1:L:14(C):GLU:O	1:L:14(G):LEU:HD23	2.10	0.51
3:I:161:THR:HG21	3:I:165:ILE:HD12	1.91	0.51
3:I:217:MET:SD	3:I:338:GLU:HG3	2.50	0.51
2:H:34:PHE:HB2	2:H:65:LEU:CD2	2.40	0.51
3:I:114:LEU:HD13	3:I:130[A]:MET:HE1	1.93	0.51
2:H:22:ALA:HB2	2:H:157[A]:VAL:HG23	1.91	0.51
3:I:108:LEU:C	3:I:108:LEU:HD23	2.31	0.51
3:I:353:PHE:H	7:I:390:GOL:H11	1.74	0.51
2:H:174[B]:ILE:HD12	2:H:215:TRP:CZ3	2.46	0.51
2:H:111:PRO:HD3	7:H:4:GOL:H32	1.92	0.51
3:I:165:ILE:HG23	3:I:331:LYS:HD3	1.93	0.51
3:I:63:GLY:HA3	3:I:317:ILE:HG13	1.93	0.50
3:I:344:ALA:O	3:I:345:ALA:HB2	2.12	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:85[A]:LEU:HD22	2:H:85[A]:LEU:N	2.26	0.50
3:I:125[A]:THR:HG23	8:I:490:HOH:O	2.10	0.50
3:I:253:GLU:HG2	8:I:520:HOH:O	2.12	0.50
3:I:29:PHE:HB3	7:I:391:GOL:O1	2.12	0.50
2:H:50[B]:ARG:HG3	2:H:51:TRP:CD1	2.47	0.49
2:H:35:ARG:NH1	2:H:37:PRO:HD2	2.27	0.49
2:H:64:LEU:HD12	2:H:88[B]:ILE:CD1	2.41	0.49
2:H:40[A]:LEU:C	2:H:40[A]:LEU:HD23	2.33	0.49
3:I:192:THR:OG1	3:I:344:ALA:CB	2.60	0.49
2:H:17:VAL:HG12	2:H:18:GLU:HG2	1.94	0.49
2:H:32[A]:MET:HG2	2:H:141:TRP:CZ3	2.47	0.49
2:H:73[A]:ARG:O	2:H:73[A]:ARG:HD3	2.13	0.49
3:I:125[A]:THR:HG22	8:I:447:HOH:O	2.12	0.49
3:I:179:ILE:CD1	3:I:327:GLU:HB2	2.37	0.49
3:I:319:ASN:C	3:I:319:ASN:HD22	2.16	0.48
3:I:249:ILE:HD12	3:I:249:ILE:N	2.28	0.48
3:I:59:MET:CG	3:I:130[A]:MET:SD	2.99	0.48
2:H:65:LEU:C	2:H:65:LEU:HD23	2.32	0.48
2:H:61:GLU:HB3	2:H:85[A]:LEU:O	2.14	0.48
2:H:103:ILE:HG21	2:H:234[A]:LEU:HD13	1.95	0.48
2:H:185[A]:LYS:CG	2:H:186[A]:PRO:HA	2.42	0.48
2:H:34:PHE:CE1	2:H:67[B]:ARG:HD2	2.48	0.47
3:I:257:GLN:NE2	3:I:261:ASN:ND2	2.61	0.47
2:H:172:THR:HG21	2:H:176:ILE:HD11	1.97	0.47
2:H:67[B]:ARG:NH1	2:H:82[B]:ILE:CD1	2.77	0.47
3:I:175:ASN:N	3:I:175:ASN:HD22	2.13	0.47
2:H:185[A]:LYS:HB3	2:H:186(B)[A]:GLU:CB	2.44	0.47
2:H:60(B):PRO:HB2	2:H:60(C):PRO:HD3	1.97	0.47
2:H:67[A]:ARG:HG2	2:H:82[A]:ILE:HG12	1.96	0.47
3:I:107:GLN:O	3:I:186:PHE:HA	2.15	0.47
2:H:35:ARG:O	2:H:38:GLN:HA	2.15	0.47
3:I:178:VAL:CG2	3:I:323:ILE:HD11	2.45	0.47
2:H:173:ARG:HH21	3:I:350:ILE:HG23	1.80	0.46
2:H:174[B]:ILE:HD11	3:I:351:PHE:HB3	1.97	0.46
3:I:95:LEU:HD23	3:I:95:LEU:C	2.36	0.46
2:H:38:GLN:O	2:H:38:GLN:HG3	2.16	0.46
2:H:68:ILE:CD1	2:H:112[A]:VAL:HG21	2.44	0.46
3:I:186:PHE:HB3	3:I:334:VAL:HG22	1.96	0.45
3:I:298:VAL:HG12	8:I:559:HOH:O	2.15	0.45
3:I:60:LEU:O	3:I:60:LEU:HD23	2.15	0.45
3:I:290:GLU:HB2	3:I:333:VAL:HG22	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:29:TRP:CG	2:H:121:VAL:HB	2.52	0.44
2:H:35:ARG:HH11	2:H:37:PRO:HD2	1.82	0.44
3:I:234:ARG:CZ	3:I:363:LEU:HD13	2.48	0.44
2:H:32[B]:MET:SD	2:H:34:PHE:CZ	3.11	0.44
2:H:35:ARG:HD3	2:H:37:PRO:O	2.18	0.44
3:I:226[A]:LEU:C	3:I:226[A]:LEU:HD13	2.38	0.44
2:H:75[A]:ARG:NH2	8:H:1016:HOH:O	2.50	0.43
3:I:144:ARG:HD3	3:I:144:ARG:HA	1.81	0.43
2:H:67[B]:ARG:HH11	2:H:82[B]:ILE:CD1	2.31	0.43
2:H:88[A]:ILE:HD12	2:H:88[A]:ILE:C	2.38	0.43
3:I:117:ASP:HA	3:I:141:THR:O	2.18	0.43
2:H:109:LYS:HB3	8:H:1018:HOH:O	2.19	0.43
1:L:1(G):PHE:HD1	2:H:242:ILE:HD13	1.83	0.43
3:I:194:PHE:O	3:I:338:GLU:HB3	2.19	0.43
3:I:319:ASN:HD22	3:I:320:HIS:N	2.17	0.43
2:H:129:ALA:HA	2:H:210:MET:HE1	2.01	0.42
3:I:309:THR:C	3:I:324:GLN:NE2	2.73	0.42
3:I:165:ILE:HD13	3:I:183:TYR:HB2	2.01	0.42
2:H:66:VAL:HG22	2:H:83[B]:SER:O	2.19	0.42
3:I:192:THR:OG1	3:I:344:ALA:HB1	2.20	0.42
3:I:202:GLN:HG3	8:I:455:HOH:O	2.21	0.41
2:H:59:LEU:HD22	2:H:64:LEU:HD11	2.02	0.41
3:I:192:THR:OG1	3:I:344:ALA:HB2	2.20	0.41
3:I:375:ASP:HB3	3:I:376:ASN:H	1.56	0.41
2:H:178:ASP:O	2:H:233:ARG:HD2	2.20	0.41
3:I:299:LEU:N	3:I:300:PRO:CD	2.84	0.41
2:H:32[A]:MET:CG	2:H:141:TRP:CZ3	3.02	0.41
2:H:235:LYS:NZ	8:H:951:HOH:O	2.54	0.41
2:H:75[B]:ARG:HG2	2:H:75[B]:ARG:NH1	2.34	0.41
2:H:186(D)[B]:LYS:NZ	8:H:1029:HOH:O	2.54	0.41
3:I:257:GLN:HE22	3:I:261:ASN:HD21	1.69	0.41
3:I:290:GLU:CB	3:I:333:VAL:HG22	2.51	0.41
3:I:51:VAL:O	3:I:55[B]:MET:HG3	2.21	0.41
2:H:240[B]:LYS:O	2:H:244[B]:GLN:HG3	2.21	0.41
3:I:114:LEU:HD13	3:I:130[A]:MET:CE	2.51	0.41
3:I:95:LEU:HD23	3:I:95:LEU:O	2.20	0.40
2:H:93[A]:ARG:NH2	5:B:1:SGN:N2	2.68	0.40
3:I:57:LEU:HB2	3:I:76:LEU:HD21	2.03	0.40
2:H:211:GLY:HA2	2:H:229:THR:O	2.21	0.40
2:H:174[B]:ILE:HD13	3:I:351:PHE:HB3	2.00	0.40
3:I:192:THR:OG1	3:I:219:ARG:NH1	2.54	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:257:GLN:NE2	3:I:261:ASN:HD21	2.19	0.40
3:I:353:PHE:CD1	7:I:390:GOL:H12	2.57	0.40

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%)	14	3
All	All	687/703 (98%)	655 (95%)	28 (4%)	4 (1%)	25	8

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 [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	L	42/43 (98%)	42 (100%)	0	100	100
2	H	241/224 (108%)	237 (98%)	4 (2%)	60	38
3	I	296/347 (85%)	289 (98%)	7 (2%)	49	24
All	All	579/614 (94%)	568 (98%)	11 (2%)	55	34

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	35	ARG
2	H	65	LEU
2	H	94	TYR
2	H	182	CYS
3	I	78	LEU
3	I	192	THR
3	I	195	ASN
3	I	196	HIS
3	I	203	ASP
3	I	275	PHE
3	I	319	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	38	GLN
3	I	97	GLN
3	I	152	GLN
3	I	175	ASN
3	I	195	ASN
3	I	196	HIS
3	I	257	GLN
3	I	261	ASN
3	I	294	GLN
3	I	319	ASN
3	I	324	GLN
3	I	361	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

4 monosaccharides 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	NAG	A	1	2,4	14,14,15	0.52	0	17,19,21	0.69	0
4	FUC	A	2	4	10,10,11	0.48	0	14,14,16	0.35	0
5	SGN	B	1	5	15,16,20	3.74	3 (20%)	19,24,31	1.27	1 (5%)
5	IDS	B	2	5	12,15,17	1.66	2 (16%)	12,22,26	0.85	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	NAG	A	1	2,4	-	4/6/23/26	0/1/1/1
4	FUC	A	2	4	-	-	0/1/1/1
5	SGN	B	1	5	-	3/7/27/31	0/1/1/1
5	IDS	B	2	5	-	0/5/22/29	1/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1	SGN	O2S-S1	9.98	1.53	1.42
5	B	1	SGN	O1S-S1	9.89	1.53	1.42
5	B	2	IDS	O2-S	-4.50	1.44	1.57
5	B	1	SGN	S1-N2	2.95	1.63	1.59
5	B	2	IDS	O2-C2	-2.18	1.44	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	B	1	SGN	O2S-S1-O1S	-4.60	109.29	120.16

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	1	SGN	C1-C2-N2-S1
5	B	1	SGN	C3-C2-N2-S1
4	A	1	NAG	C8-C7-N2-C2
4	A	1	NAG	O7-C7-N2-C2
4	A	1	NAG	C1-C2-N2-C7
5	B	1	SGN	C2-N2-S1-O2S
4	A	1	NAG	C3-C2-N2-C7

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	2	IDS	C1-C2-C3-C4-C5-O5

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	2	IDS	3	0
5	B	1	SGN	4	0
4	A	1	NAG	1	0

## 5.6 Ligand geometry [i](#)

8 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	$\# Z  > 2$
7	GOL	H	3	-	5,5,5	0.31	0	5,5,5	0.28	0
7	GOL	H	4	-	5,5,5	0.32	0	5,5,5	0.24	0
7	GOL	I	391	-	5,5,5	0.32	0	5,5,5	0.28	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	SO4	H	781	-	4,4,4	0.26	0	6,6,6	0.06	0
6	SO4	I	389	-	4,4,4	0.26	0	6,6,6	0.05	0
6	SO4	I	388	-	4,4,4	0.26	0	6,6,6	0.05	0
7	GOL	H	782	-	5,5,5	0.32	0	5,5,5	0.15	0
7	GOL	I	390	-	5,5,5	0.31	0	5,5,5	0.28	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
7	GOL	I	391	-	-	0/4/4/4	-
7	GOL	H	782	-	-	0/4/4/4	-
7	GOL	H	3	-	-	0/4/4/4	-
7	GOL	I	390	-	-	0/4/4/4	-
7	GOL	H	4	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	H	4	GOL	1	0
7	I	391	GOL	1	0
7	I	390	GOL	6	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	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 28	9, 17, 35, 53	23 (9%)
3	I	355/395 (89%)	1.08	62 (17%) 1 1	14, 37, 59, 69	4 (1%)
All	All	652/703 (92%)	0.62	77 (11%) 4 3	9, 28, 54, 69	29 (4%)

All (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
3	I	104	ASP	6.4
3	I	83	SER	6.1
3	I	196	HIS	5.7
3	I	344	ALA	5.6
3	I	100	ASN	5.3
3	I	84	SER	5.2
3	I	347	THR	5.1
3	I	341	THR	5.0
2	H	36(A)	SER	5.0
3	I	101	GLN	4.5
3	I	78	LEU	4.4
3	I	105	GLY	4.1
3	I	195	ASN	4.0
1	L	14(L)	ASP	3.7
3	I	107	GLN	3.7
3	I	77	GLY	3.7
3	I	171	ASN	3.7
3	I	348	GLY	3.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	I	371[A]	MET	3.6
2	H	45[A]	SER	3.6
3	I	242	GLY	3.5
3	I	200	GLN	3.4
3	I	376	ASN	3.3
3	I	39	SER	3.3
3	I	342	ARG	3.2
3	I	322	ASN	3.2
2	H	147	THR	3.1
3	I	42	PRO	3.1
3	I	257	GLN	3.0
2	H	212	ILE	2.9
3	I	172	LEU	2.8
3	I	88	LEU	2.8
3	I	205	TYR	2.7
2	H	121	VAL	2.7
3	I	230	ASN	2.7
3	I	40	ALA	2.7
3	I	28	ASP	2.6
2	H	149[A]	LYS	2.6
2	H	200	VAL	2.6
3	I	311	HIS	2.6
2	H	44	ALA	2.5
3	I	147	ALA	2.5
3	I	99	LEU	2.5
3	I	170	LYS	2.5
3	I	211	VAL	2.5
2	H	199	PHE	2.5
3	I	199	THR	2.5
3	I	241	GLN	2.5
3	I	197	LYS	2.5
3	I	70	MET	2.4
3	I	44	GLN	2.4
3	I	210	THR	2.4
3	I	167	ASP	2.4
3	I	189	LYS	2.4
3	I	274	MET	2.3
3	I	166	VAL	2.3
3	I	94	GLN	2.3
1	L	1(R)	SER	2.2
3	I	73	LEU	2.2
3	I	273	LYS	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	I	173	ASP	2.2
3	I	86	LYS	2.1
3	I	358	LEU	2.1
3	I	67	SER	2.1
2	H	47	ILE	2.1
2	H	228	TYR	2.1
1	L	1(Q)	GLU	2.1
2	H	46	LEU	2.1
3	I	114	LEU	2.1
3	I	91	GLY	2.0
3	I	66	SER	2.0
3	I	97	GLN	2.0

## 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](#)

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	SGN	B	1	16/20	0.31	0.62	65,72,76,76	0
5	IDS	B	2	15/17	0.55	0.42	71,72,76,76	0
4	NAG	A	1	14/15	0.67	0.23	50,54,59,63	0
4	FUC	A	2	10/11	0.84	0.48	66,68,68,69	0

## 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
7	GOL	I	391	6/6	0.66	0.17	60,60,61,61	0
7	GOL	H	4	6/6	0.76	0.21	55,55,55,55	2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	SO4	I	389	5/5	0.82	0.26	76,76,76,76	2
7	GOL	H	3	6/6	0.84	0.18	48,49,49,50	3
6	SO4	H	781	5/5	0.87	0.14	77,77,77,77	3
7	GOL	I	390	6/6	0.87	0.12	44,45,45,46	2
7	GOL	H	782	6/6	0.91	0.12	20,24,25,27	3
6	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.