



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 9, 2017 – 11:07 am GMT

PDB ID : 5T6N
Title : Crystal structure of the A/Hong Kong/1/1968 (H3N2) influenza virus hemagglutinin in complex with the antiviral drug arbidol
Authors : Kadam, R.U.; Wilson, I.A.
Deposited on : 2016-09-01
Resolution : 2.54 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

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : recalc28906
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28906

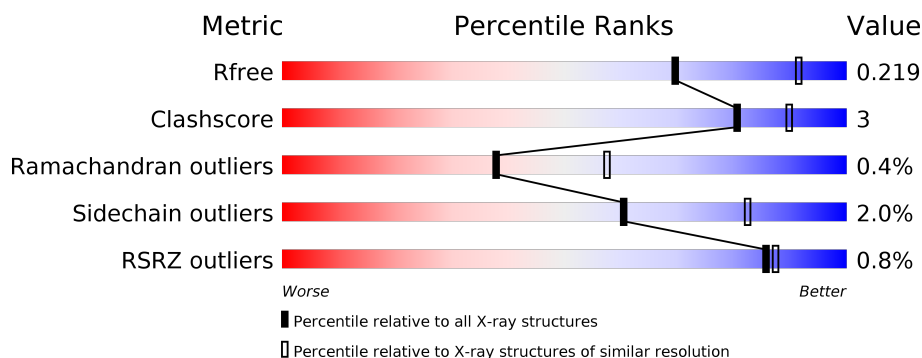
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.54 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}	100719	4993 (2.58-2.50)
Clashscore	112137	5755 (2.58-2.50)
Ramachandran outliers	110173	5652 (2.58-2.50)
Sidechain outliers	110143	5654 (2.58-2.50)
RSRZ outliers	101464	5026 (2.58-2.50)

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	A	323	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>8%</div> <div>..</div> </div> </div>
1	C	323	<div> <div>%</div> <div> <div></div> <div>93%</div> <div>5%</div> <div>..</div> </div> </div>
1	E	323	<div> <div></div> <div> <div></div> <div>90%</div> <div>7%</div> <div>..</div> </div> </div>
2	B	174	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>11%</div> <div>..</div> </div> </div>
2	D	174	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>9%</div> <div>.</div> </div> </div>
2	F	174	<div> <div></div> <div> <div></div> <div>89%</div> <div>7%</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
3	NAG	A	401	-	-	-	X
3	NAG	A	402	-	-	-	X
3	NAG	A	408	-	-	-	X
3	NAG	C	403	-	-	-	X
3	NAG	C	408	-	-	-	X
3	NAG	F	201	-	-	-	X
6	75U	D	201	-	-	-	X
7	SO4	C	410	-	-	-	X
8	GOL	C	411	-	-	-	X
8	GOL	E	410	-	-	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 12992 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 Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	318	Total	C	N	O	S	0	8	0
			2519	1578	443	484	14			
1	C	317	Total	C	N	O	S	0	8	0
			2506	1567	443	482	14			
1	E	318	Total	C	N	O	S	0	9	0
			2533	1587	447	485	14			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q91MA7
A	8	ASP	-	expression tag	UNP Q91MA7
A	9	PRO	-	expression tag	UNP Q91MA7
A	10	GLY	-	expression tag	UNP Q91MA7
C	7	ALA	-	expression tag	UNP Q91MA7
C	8	ASP	-	expression tag	UNP Q91MA7
C	9	PRO	-	expression tag	UNP Q91MA7
C	10	GLY	-	expression tag	UNP Q91MA7
E	7	ALA	-	expression tag	UNP Q91MA7
E	8	ASP	-	expression tag	UNP Q91MA7
E	9	PRO	-	expression tag	UNP Q91MA7
E	10	GLY	-	expression tag	UNP Q91MA7

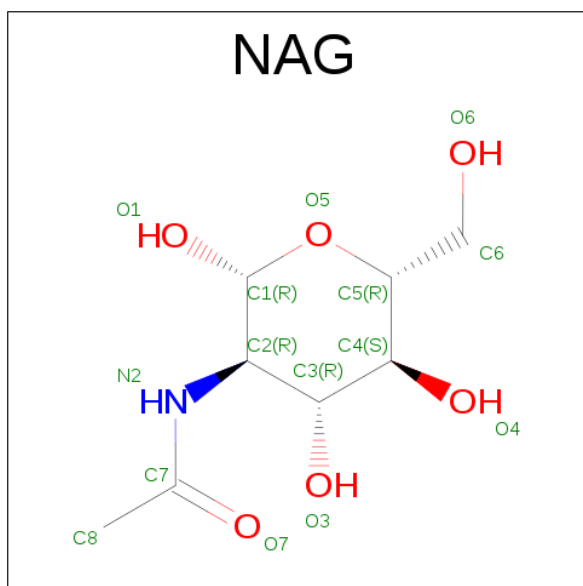
- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	5	0
			1429	889	248	285	7			
2	D	171	Total	C	N	O	S	0	5	0
			1419	883	246	283	7			
2	F	171	Total	C	N	O	S	0	7	0
			1435	893	248	287	7			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	123	GLY	ARG	conflict	UNP P03436
D	123	GLY	ARG	conflict	UNP P03436
F	123	GLY	ARG	conflict	UNP P03436

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



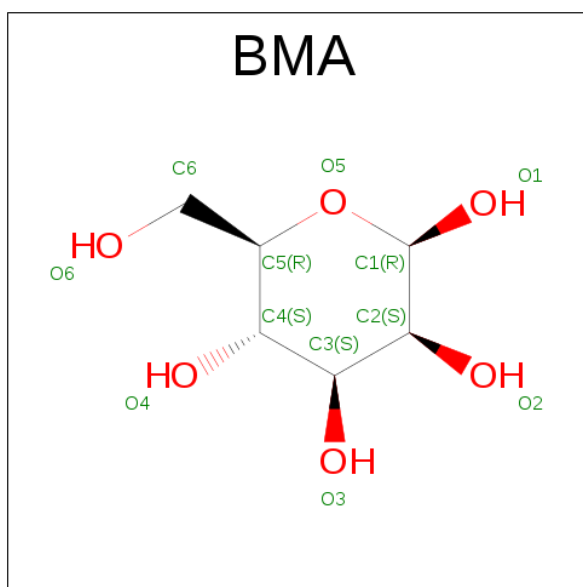
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1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

Continued on next page...

Continued from previous page...

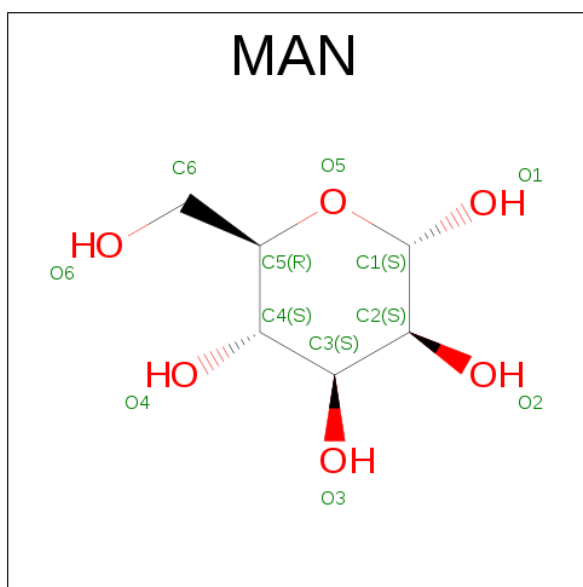
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1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

- Molecule 4 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C₆H₁₂O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	C	1	Total	C	O	0	0
			11	6	5		
4	E	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C₆H₁₂O₆).



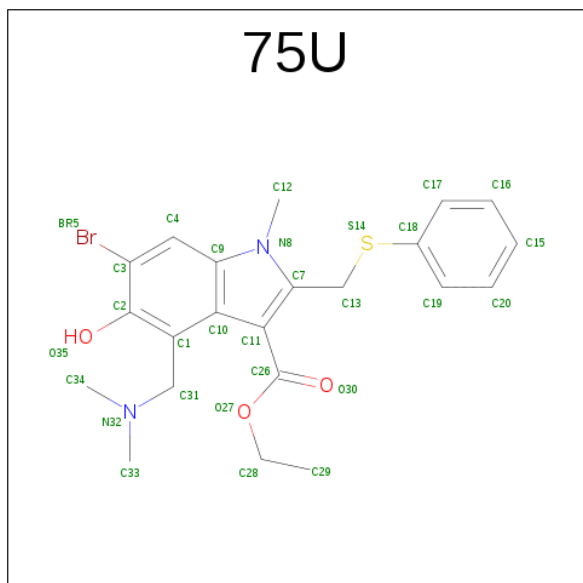
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		

Continued on next page...

Continued from previous page...

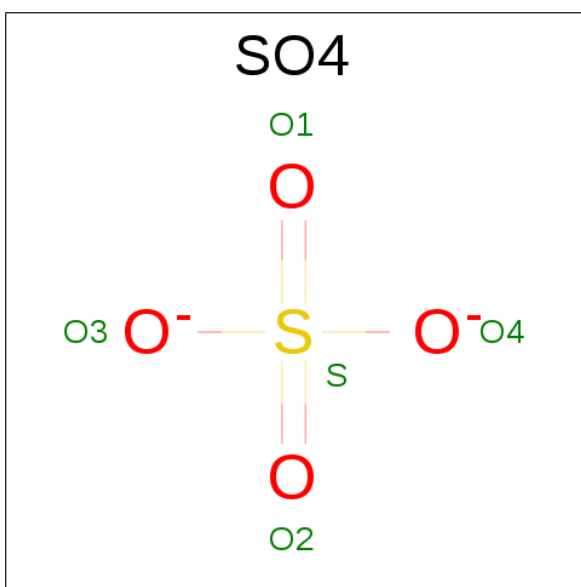
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			10	6	4		
5	E	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is ethyl 6-bromo-4-[(dimethylamino)methyl]-5-hydroxy-1-methyl-2-[(phenylsulfonyl)methyl]-1H-indole-3-carboxylate (three-letter code: 75U) (formula: C₂₂H₂₅BrN₂O₃S).



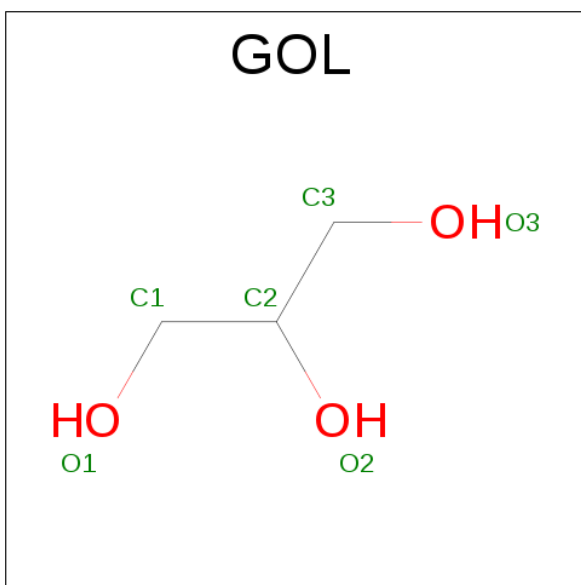
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	B	1	Total	Br	C	N	O	S	0	0
			29	1	22	2	3	1		
6	D	1	Total	Br	C	N	O	S	0	0
			29	1	22	2	3	1		
6	D	1	Total	Br	C	N	O	S	0	0
			29	1	22	2	3	1		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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	C	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			6	3	3		
8	E	1	Total	C	O	0	0
			6	3	3		

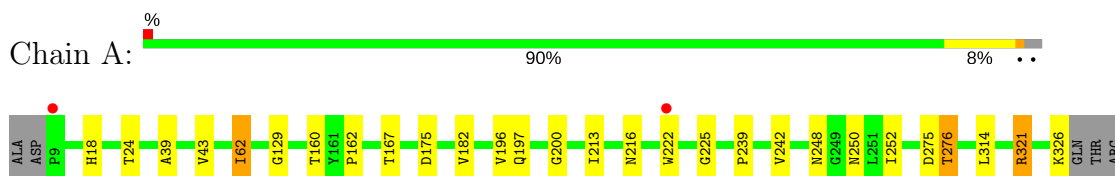
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	107	Total	O	0	0
			107	107		
9	B	99	Total	O	0	0
			99	99		
9	C	122	Total	O	0	0
			122	122		
9	D	103	Total	O	0	0
			103	103		
9	E	112	Total	O	0	0
			112	112		
9	F	74	Total	O	0	0
			74	74		

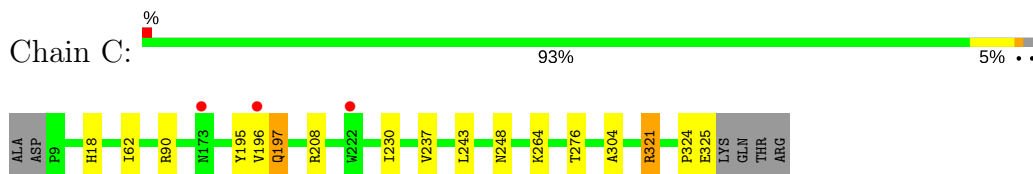
3 Residue-property plots [i](#)

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.

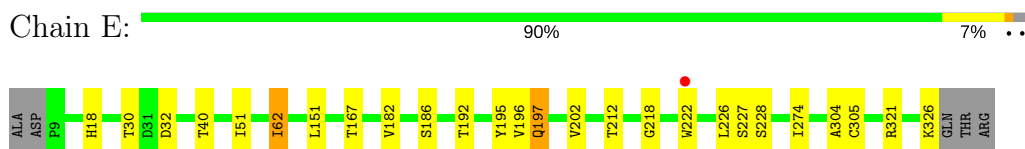
• Molecule 1: Hemagglutinin HA1



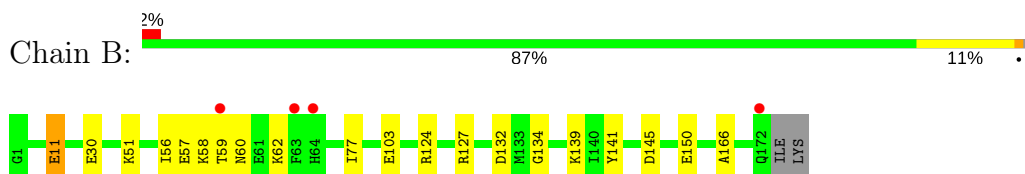
• Molecule 1: Hemagglutinin HA1



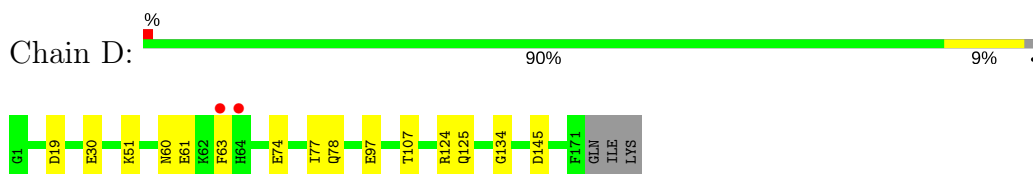
• Molecule 1: Hemagglutinin HA1




• Molecule 2: Hemagglutinin HA2



• Molecule 2: Hemagglutinin HA2



• Molecule 2: Hemagglutinin HA2

Chain F:  89% 7% ...



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	105.94Å 151.81Å 349.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.22 – 2.54 48.22 – 2.54	Depositor EDS
% Data completeness (in resolution range)	98.3 (48.22-2.54) 98.3 (48.22-2.54)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 2.54Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.180 , 0.220 0.176 , 0.219	Depositor DCC
R_{free} test set	4464 reflections (4.89%)	DCC
Wilson B-factor (Å ²)	33.7	Xtriage
Anisotropy	0.494	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 49.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	12992	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, BMA, NAG, 75U, SO4, 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	0.63	1/2578 (0.0%)	0.68	0/3512
1	C	0.62	0/2563	0.68	0/3491
1	E	0.63	0/2592	0.68	0/3530
2	B	0.65	0/1453	0.75	1/1953 (0.1%)
2	D	0.68	0/1443	0.72	0/1940
2	F	0.71	1/1459 (0.1%)	0.77	1/1962 (0.1%)
All	All	0.65	2/12088 (0.0%)	0.70	2/16388 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	0	1
2	F	0	2
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	103	GLU	CG-CD	-5.16	1.44	1.51
1	A	252	ILE	C-N	-5.14	1.22	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	90	ASP	CB-CG-OD1	5.75	123.48	118.30
2	B	132	ASP	CB-CG-OD1	5.52	123.27	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	60	ASN	Peptide
2	F	60	ASN	Peptide
2	F	62	LYS	Peptide

5.2 Too-close contacts

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	2519	0	2457	18	0
1	C	2506	0	2451	10	0
1	E	2533	0	2475	18	0
2	B	1429	0	1348	13	0
2	D	1419	0	1338	14	0
2	F	1435	0	1351	13	0
3	A	98	0	88	4	0
3	B	14	0	13	0	0
3	C	112	0	99	0	0
3	D	28	0	25	0	0
3	E	84	0	75	2	0
3	F	14	0	13	0	0
4	A	11	0	9	2	0
4	C	11	0	10	0	0
4	E	11	0	8	4	0
5	A	11	0	10	0	0
5	E	21	0	17	2	0
6	B	29	0	0	0	0
6	D	58	0	0	1	0
7	B	10	0	0	2	0
7	C	5	0	0	0	0
7	D	5	0	0	0	0
8	C	6	0	8	1	0
8	E	6	0	8	0	0
9	A	107	0	0	1	0
9	B	99	0	0	0	0
9	C	122	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	D	103	0	0	1	0
9	E	112	0	0	1	0
9	F	74	0	0	2	0
All	All	12992	0	11803	74	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 3.

All (74) 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)
1:A:326:LYS:HD2	2:B:11:GLU:HG2	1.65	0.77
2:B:77:ILE:HD12	2:D:77[A]:ILE:HD11	1.69	0.74
2:D:51:LYS:HG3	1:E:30[A]:THR:HG22	1.73	0.69
4:E:405:BMA:H2	5:E:406:MAN:H5	1.77	0.66
1:E:196:VAL:HG22	1:E:197:GLN:OE1	1.99	0.63
1:A:222:TRP:CZ2	1:A:225:GLY:HA2	2.35	0.60
1:A:160:THR:HA	1:A:196:VAL:HG21	1.84	0.60
2:B:134:GLY:HA2	2:F:124:ARG:HD3	1.85	0.59
1:E:304:ALA:HA	2:F:61:GLU:HA	1.82	0.59
1:C:195:TYR:O	1:C:197:GLN:N	2.36	0.58
2:D:19:ASP:OD1	2:D:19:ASP:N	2.37	0.57
2:F:59:THR:HG22	2:F:60:ASN:H	1.70	0.56
2:B:124:ARG:HD3	2:D:134:GLY:HA2	1.86	0.56
2:D:124:ARG:HD3	2:F:134:GLY:HA2	1.89	0.55
2:B:56:ILE:O	2:B:56:ILE:HG22	2.07	0.55
1:A:216:ASN:CG	1:E:212:THR:HG21	2.28	0.52
1:E:305:CYS:O	2:F:60:ASN:HA	2.09	0.51
2:D:51:LYS:CG	1:E:30[A]:THR:HG22	2.40	0.51
1:A:222:TRP:CH2	4:E:405:BMA:H5	2.46	0.51
2:D:74:GLU:OE1	2:D:78:GLN:NE2	2.44	0.50
1:E:151:LEU:O	9:E:501:HOH:O	2.20	0.50
1:E:51:ILE:HB	1:E:274[B]:ILE:HD13	1.93	0.50
1:A:275:ASP:OD1	1:A:276:THR:N	2.44	0.50
3:A:405:NAG:O3	4:A:406:BMA:O5	2.19	0.50
2:D:97:GLU:HG2	6:D:201:75U:C4	2.42	0.49
9:D:316:HOH:O	1:E:30[A]:THR:HG21	2.11	0.49
1:A:24:THR:HG21	1:A:39:ALA:HB3	1.94	0.49
1:C:324:PRO:O	1:C:325:GLU:HG3	2.13	0.49
2:F:139:LYS:HD3	2:F:141:TYR:CZ	2.49	0.48
2:B:30:GLU:OE2	2:B:145:ASP:HB2	2.13	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:200:GLY:HA3	1:A:250:ASN:OD1	2.15	0.47
1:A:222:TRP:CZ2	4:E:405:BMA:H5	2.50	0.47
2:B:51:LYS:HD3	2:B:103:GLU:OE1	2.15	0.47
1:A:167:THR:HB	3:A:404:NAG:H62	1.98	0.46
1:C:264:LYS:HD3	2:D:63:PHE:CE2	2.50	0.46
1:C:321[A]:ARG:NH2	9:C:505:HOH:O	2.40	0.46
3:A:404:NAG:O4	3:A:405:NAG:O7	2.33	0.45
2:B:141:TYR:O	2:B:166:ALA:HA	2.16	0.45
1:E:222:TRP:HA	1:E:226:LEU:O	2.16	0.45
2:F:59:THR:O	2:F:60:ASN:HB2	2.16	0.45
2:B:77:ILE:HD12	2:D:77[A]:ILE:CD1	2.45	0.44
1:C:237:VAL:HG21	1:C:243:LEU:HB2	1.99	0.44
2:F:51:LYS:HB2	2:F:51:LYS:HE2	1.77	0.44
3:A:405:NAG:HO3	4:A:406:BMA:C1	2.28	0.44
2:B:77:ILE:HD13	2:F:77:ILE:HD13	2.00	0.44
2:D:51:LYS:HE2	2:D:107:THR:OG1	2.17	0.43
1:C:230:ILE:HG21	1:C:230:ILE:HD13	1.78	0.43
1:A:321[B]:ARG:HB2	9:A:575:HOH:O	2.18	0.43
1:E:186:SER:HA	1:E:218:GLY:O	2.19	0.42
1:E:192:THR:HA	1:E:195:TYR:O	2.19	0.42
1:E:62:ILE:HG13	1:E:62:ILE:H	1.50	0.42
1:A:197:GLN:NE2	1:A:248:ASN:O	2.52	0.42
1:C:90:ARG:HD2	8:C:411:GOL:H11	2.00	0.42
1:A:43:VAL:HG23	1:A:314:LEU:HB2	2.02	0.42
1:E:326:LYS:HD3	2:F:11:GLU:HG3	2.00	0.42
1:C:304:ALA:HA	2:D:61:GLU:HA	2.00	0.42
2:F:124:ARG:HD2	9:F:312:HOH:O	2.19	0.42
2:F:57:GLU:HB3	2:F:58:LYS:H	1.46	0.42
2:B:127:ARG:NH2	7:B:204:SO4:O2	2.36	0.42
1:C:197:GLN:NE2	1:C:248:ASN:O	2.53	0.42
1:E:167:THR:HB	3:E:403:NAG:H62	2.01	0.42
4:E:405:BMA:C2	5:E:406:MAN:H5	2.48	0.41
1:E:222:TRP:HD1	1:E:227:SER:HG	1.67	0.41
2:F:153:ARG:NH1	9:F:304:HOH:O	2.49	0.41
2:B:139:LYS:HD2	7:B:203:SO4:O3	2.20	0.41
1:C:264:LYS:HB3	2:D:63:PHE:CG	2.55	0.41
1:A:129:GLY:HA3	1:A:162:PRO:HG2	2.01	0.41
1:A:182:VAL:HG21	1:A:213:ILE:HB	2.03	0.41
1:A:216:ASN:CB	1:E:212:THR:HG21	2.51	0.41
1:E:182:VAL:HG22	1:E:202:VAL:HG21	2.03	0.41
2:B:57:GLU:O	2:B:58:LYS:C	2.59	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:30:GLU:OE2	2:D:145:ASP:HB2	2.20	0.40
1:A:222:TRP:HB2	3:E:404:NAG:H2	2.04	0.40
1:A:175:ASP:OD1	1:A:239:PRO:HD3	2.22	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	A	324/323 (100%)	313 (97%)	10 (3%)	1 (0%)	44	64
1	C	323/323 (100%)	311 (96%)	10 (3%)	2 (1%)	28	46
1	E	325/323 (101%)	314 (97%)	10 (3%)	1 (0%)	44	64
2	B	175/174 (101%)	166 (95%)	9 (5%)	0	100	100
2	D	174/174 (100%)	164 (94%)	10 (6%)	0	100	100
2	F	176/174 (101%)	164 (93%)	10 (6%)	2 (1%)	17	28
All	All	1497/1491 (100%)	1432 (96%)	59 (4%)	6 (0%)	38	57

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	62	ILE
2	F	57	GLU
2	F	60	ASN
1	A	62	ILE
1	C	62	ILE
1	C	196	VAL

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	A	287/283 (101%)	281 (98%)	6 (2%)	59	81
1	C	286/283 (101%)	280 (98%)	6 (2%)	59	81
1	E	288/283 (102%)	281 (98%)	7 (2%)	54	78
2	B	151/148 (102%)	146 (97%)	5 (3%)	43	68
2	D	150/148 (101%)	149 (99%)	1 (1%)	87	95
2	F	152/148 (103%)	149 (98%)	3 (2%)	60	82
All	All	1314/1293 (102%)	1286 (98%)	28 (2%)	60	81

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	62	ILE
1	A	242	VAL
1	A	276	THR
1	A	321[A]	ARG
1	A	321[B]	ARG
2	B	11	GLU
2	B	59	THR
2	B	60	ASN
2	B	62	LYS
2	B	150	GLU
1	C	18	HIS
1	C	197	GLN
1	C	208	ARG
1	C	276	THR
1	C	321[A]	ARG
1	C	321[B]	ARG
2	D	125	GLN
1	E	18	HIS
1	E	32	ASP
1	E	40	THR
1	E	197	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	228	SER
1	E	321[A]	ARG
1	E	321[B]	ARG
2	F	51	LYS
2	F	63	PHE
2	F	77	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

40 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	$\# Z > 2$
3	NAG	A	401	1	14,14,15	0.32	0	15,19,21	0.54	0
3	NAG	A	402	1	14,14,15	0.15	0	15,19,21	0.69	1 (6%)
3	NAG	A	403	1	14,14,15	0.45	0	15,19,21	0.86	1 (6%)
3	NAG	A	404	1,3	14,14,15	0.50	0	15,19,21	0.69	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NAG	A	405	3,4	14,14,15	1.15	1 (7%)	15,19,21	1.64	2 (13%)
4	BMA	A	406	3,5	11,11,12	0.44	0	13,15,17	2.27	5 (38%)
5	MAN	A	407	4	11,11,12	0.79	0	13,15,17	1.12	2 (15%)
3	NAG	A	408	1,3	14,14,15	0.37	0	15,19,21	0.55	0
3	NAG	A	409	3	14,14,15	0.17	0	15,19,21	0.67	0
6	75U	B	201	-	29,31,31	3.19	8 (27%)	33,44,44	2.42	11 (33%)
3	NAG	B	202	2	14,14,15	0.33	0	15,19,21	0.47	0
7	SO4	B	203	-	4,4,4	0.18	0	6,6,6	0.14	0
7	SO4	B	204	-	4,4,4	0.16	0	6,6,6	0.20	0
3	NAG	C	401	1,3	14,14,15	0.42	0	15,19,21	0.59	0
3	NAG	C	402	3	14,14,15	0.22	0	15,19,21	0.53	0
3	NAG	C	403	1,3	14,14,15	0.28	0	15,19,21	0.74	0
3	NAG	C	404	3	14,14,15	0.30	0	15,19,21	0.45	0
3	NAG	C	405	1,3	14,14,15	0.54	0	15,19,21	0.44	0
3	NAG	C	406	3,4	14,14,15	0.42	0	15,19,21	0.43	0
4	BMA	C	407	3	11,11,12	0.91	1 (9%)	13,15,17	1.34	3 (23%)
3	NAG	C	408	1,3	14,14,15	0.42	0	15,19,21	0.52	0
3	NAG	C	409	3	14,14,15	0.34	0	15,19,21	0.44	0
7	SO4	C	410	-	4,4,4	0.10	0	6,6,6	0.25	0
8	GOL	C	411	-	5,5,5	0.34	0	5,5,5	0.32	0
6	75U	D	201	-	29,31,31	3.62	8 (27%)	33,44,44	2.19	9 (27%)
6	75U	D	202	-	29,31,31	3.59	9 (31%)	33,44,44	2.13	10 (30%)
3	NAG	D	203	3,2	14,14,15	0.43	0	15,19,21	0.55	0
3	NAG	D	204	3	14,14,15	0.22	0	15,19,21	0.40	0
7	SO4	D	205	-	4,4,4	0.13	0	6,6,6	0.07	0
3	NAG	E	401	1	14,14,15	0.29	0	15,19,21	0.81	1 (6%)
3	NAG	E	402	1	14,14,15	0.61	0	15,19,21	0.78	1 (6%)
3	NAG	E	403	1,3	14,14,15	0.25	0	15,19,21	0.61	0
3	NAG	E	404	3,4	14,14,15	0.31	0	15,19,21	0.55	0
4	BMA	E	405	3,5	11,11,12	0.93	1 (9%)	13,15,17	1.69	2 (15%)
5	MAN	E	406	4	9,10,12	0.67	0	13,14,17	1.12	1 (7%)
5	MAN	E	407	4	11,11,12	0.74	0	13,15,17	1.00	1 (7%)
3	NAG	E	408	1,3	14,14,15	0.55	0	15,19,21	0.51	0
3	NAG	E	409	3	14,14,15	0.26	0	15,19,21	0.57	0
8	GOL	E	410	-	5,5,5	0.36	0	5,5,5	0.23	0
3	NAG	F	201	2	14,14,15	0.42	0	15,19,21	0.70	1 (6%)

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. '-' means

no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1
3	NAG	A	402	1	-	0/6/23/26	0/1/1/1
3	NAG	A	403	1	-	0/6/23/26	0/1/1/1
3	NAG	A	404	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	405	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	406	3,5	-	0/2/19/22	0/1/1/1
5	MAN	A	407	4	-	0/2/19/22	0/1/1/1
3	NAG	A	408	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	409	3	-	0/6/23/26	0/1/1/1
6	75U	B	201	-	-	0/16/16/16	0/3/3/3
3	NAG	B	202	2	-	0/6/23/26	0/1/1/1
7	SO4	B	203	-	-	0/0/0/0	0/0/0/0
7	SO4	B	204	-	-	0/0/0/0	0/0/0/0
3	NAG	C	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	402	3	-	0/6/23/26	0/1/1/1
3	NAG	C	403	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	404	3	-	0/6/23/26	0/1/1/1
3	NAG	C	405	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	406	3,4	-	0/6/23/26	0/1/1/1
4	BMA	C	407	3	-	0/2/19/22	0/1/1/1
3	NAG	C	408	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	409	3	-	0/6/23/26	0/1/1/1
7	SO4	C	410	-	-	0/0/0/0	0/0/0/0
8	GOL	C	411	-	-	0/4/4/4	0/0/0/0
6	75U	D	201	-	-	0/16/16/16	0/3/3/3
6	75U	D	202	-	-	0/16/16/16	0/3/3/3
3	NAG	D	203	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	204	3	-	0/6/23/26	0/1/1/1
7	SO4	D	205	-	-	0/0/0/0	0/0/0/0
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
3	NAG	E	402	1	-	0/6/23/26	0/1/1/1
3	NAG	E	403	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	404	3,4	-	0/6/23/26	0/1/1/1
4	BMA	E	405	3,5	-	0/2/19/22	0/1/1/1
5	MAN	E	406	4	-	0/0/17/22	0/1/1/1
5	MAN	E	407	4	-	0/2/19/22	0/1/1/1
3	NAG	E	408	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	409	3	-	0/6/23/26	0/1/1/1
8	GOL	E	410	-	-	0/4/4/4	0/0/0/0
3	NAG	F	201	2	-	0/6/23/26	0/1/1/1

All (28) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	202	75U	C31-C1	-15.16	1.41	1.51
6	D	201	75U	C31-C1	-15.09	1.41	1.51
6	B	201	75U	C31-C1	-12.67	1.42	1.51
6	D	201	75U	C11-C26	-8.02	1.35	1.49
6	B	201	75U	C11-C26	-7.65	1.36	1.49
6	D	202	75U	C11-C26	-7.63	1.36	1.49
6	D	202	75U	C10-C9	-4.97	1.36	1.42
6	D	201	75U	C10-C9	-4.83	1.36	1.42
6	B	201	75U	C10-C9	-4.77	1.36	1.42
6	D	201	75U	C1-C10	-3.89	1.37	1.44
6	B	201	75U	C4-C9	-3.68	1.33	1.40
6	D	202	75U	C1-C10	-3.67	1.37	1.44
6	D	201	75U	C4-C9	-3.62	1.33	1.40
6	D	202	75U	C4-C9	-3.38	1.33	1.40
6	B	201	75U	C1-C10	-3.21	1.38	1.44
6	D	201	75U	C11-C10	-2.98	1.37	1.41
6	D	202	75U	C11-C7	-2.49	1.35	1.40
6	B	201	75U	C11-C10	-2.41	1.37	1.41
6	D	201	75U	C11-C7	-2.37	1.35	1.40
6	B	201	75U	C11-C7	-2.24	1.35	1.40
6	D	202	75U	C11-C10	-2.20	1.38	1.41
4	E	405	BMA	C2-C3	2.11	1.55	1.52
4	C	407	BMA	C1-C2	2.54	1.58	1.52
6	D	202	75U	C18-S14	2.84	1.82	1.76
6	D	202	75U	C13-C7	2.95	1.54	1.49
6	B	201	75U	C13-C7	3.06	1.54	1.49
6	D	201	75U	C13-C7	3.09	1.54	1.49
3	A	405	NAG	O5-C1	4.07	1.50	1.43

All (51) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	202	75U	C4-C3-C2	-7.09	115.30	121.58
6	D	201	75U	C4-C3-C2	-6.41	115.90	121.58
6	B	201	75U	C4-C3-C2	-5.97	116.28	121.58
6	B	201	75U	C28-O27-C26	-4.34	106.89	116.45
6	B	201	75U	C31-C1-C2	-4.27	113.99	120.44
6	D	202	75U	C31-C1-C2	-4.20	114.10	120.44
6	D	201	75U	C31-C1-C2	-4.04	114.34	120.44
6	D	201	75U	C28-O27-C26	-3.83	108.02	116.45
4	A	406	BMA	C3-C4-C5	-3.42	104.20	110.22
6	B	201	75U	O27-C26-O30	-2.63	118.25	123.64
5	E	406	MAN	O5-C1-C2	-2.60	106.71	110.79

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	201	75U	C34-N32-C33	-2.53	102.95	109.72
5	E	407	MAN	O2-C2-C3	-2.33	105.59	110.17
6	D	202	75U	O27-C26-O30	-2.28	118.98	123.64
6	D	201	75U	O27-C26-O30	-2.27	118.99	123.64
5	A	407	MAN	O2-C2-C3	-2.27	105.72	110.17
4	A	406	BMA	C2-C3-C4	-2.07	107.27	110.88
6	D	202	75U	BR5-C3-C4	2.05	122.34	119.07
4	A	406	BMA	O5-C1-C2	2.09	114.06	110.79
6	D	201	75U	C4-C9-C10	2.09	124.04	120.90
3	F	201	NAG	C1-O5-C5	2.14	115.12	112.17
3	A	402	NAG	C1-O5-C5	2.20	115.20	112.17
5	A	407	MAN	C1-O5-C5	2.25	115.27	112.17
4	C	407	BMA	C1-O5-C5	2.27	115.30	112.17
4	C	407	BMA	O5-C1-C2	2.28	114.37	110.79
6	D	202	75U	C12-N8-C7	2.29	126.33	124.09
4	C	407	BMA	C1-C2-C3	2.38	112.67	109.65
6	D	202	75U	C4-C9-C10	2.38	124.47	120.90
6	B	201	75U	C4-C9-C10	2.47	124.59	120.90
6	D	202	75U	O27-C26-C11	2.53	116.71	112.39
3	A	405	NAG	C2-N2-C7	2.57	126.69	122.94
3	A	403	NAG	C1-O5-C5	2.57	115.71	112.17
3	E	402	NAG	C1-O5-C5	2.57	115.71	112.17
3	E	401	NAG	C1-O5-C5	2.69	115.87	112.17
4	E	405	BMA	C1-O5-C5	2.80	116.03	112.17
6	B	201	75U	C1-C31-N32	2.98	117.78	112.45
6	D	202	75U	BR5-C3-C2	3.06	122.36	118.82
6	D	201	75U	BR5-C3-C2	3.15	122.46	118.82
6	B	201	75U	C12-N8-C7	3.18	127.21	124.09
6	D	201	75U	C1-C31-N32	3.20	118.18	112.45
6	D	202	75U	C1-C31-N32	3.23	118.22	112.45
6	D	201	75U	O27-C26-C11	3.57	118.49	112.39
4	A	406	BMA	C1-C2-C3	3.57	114.18	109.65
6	B	201	75U	O27-C26-C11	3.96	119.17	112.39
6	D	202	75U	C31-C1-C10	4.27	126.79	120.92
4	E	405	BMA	O3-C3-C2	4.47	118.16	110.02
6	D	201	75U	C31-C1-C10	4.48	127.08	120.92
4	A	406	BMA	C1-O5-C5	4.81	118.79	112.17
6	B	201	75U	C31-C1-C10	5.05	127.86	120.92
6	B	201	75U	BR5-C3-C2	5.21	124.84	118.82
3	A	405	NAG	C1-O5-C5	5.30	119.47	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

11 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	404	NAG	2	0
3	A	405	NAG	3	0
4	A	406	BMA	2	0
7	B	203	SO4	1	0
7	B	204	SO4	1	0
8	C	411	GOL	1	0
6	D	201	75U	1	0
3	E	403	NAG	1	0
3	E	404	NAG	1	0
4	E	405	BMA	4	0
5	E	406	MAN	2	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	A	318/323 (98%)	-0.42	2 (0%) 89 90	15, 31, 55, 82	0
1	C	317/323 (98%)	-0.41	3 (0%) 84 86	15, 32, 57, 81	0
1	E	318/323 (98%)	-0.50	1 (0%) 93 94	14, 30, 52, 79	0
2	B	172/174 (98%)	-0.31	4 (2%) 61 64	15, 27, 63, 99	0
2	D	171/174 (98%)	-0.38	2 (1%) 79 81	13, 27, 54, 81	0
2	F	171/174 (98%)	-0.33	0 100 100	12, 27, 55, 87	0
All	All	1467/1491 (98%)	-0.41	12 (0%) 86 87	12, 30, 55, 99	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	222	TRP	3.6
2	D	64	HIS	3.1
1	C	173	ASN	2.7
2	B	63	PHE	2.7
1	C	196	VAL	2.6
1	A	222	TRP	2.4
2	B	59	THR	2.4
1	E	222	TRP	2.3
2	B	64	HIS	2.1
2	B	172	GLN	2.1
2	D	63	PHE	2.1
1	A	9	PRO	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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
8	GOL	C	411	6/6	0.92	0.44	18.29	44,50,67,84	0
8	GOL	E	410	6/6	0.87	0.56	11.31	52,54,59,69	0
3	NAG	F	201	14/15	0.88	0.31	6.76	47,62,71,72	0
3	NAG	A	401	14/15	0.73	0.28	6.15	59,82,87,88	0
3	NAG	C	408	14/15	0.93	0.19	5.60	36,47,59,61	0
3	NAG	A	402	14/15	0.86	0.17	3.05	42,59,70,70	0
3	NAG	A	408	14/15	0.94	0.14	2.65	36,47,56,71	0
7	SO4	C	410	5/5	0.89	0.29	2.45	72,79,100,117	0
6	75U	D	201	29/29	0.92	0.20	2.34	31,41,59,166	0
3	NAG	C	403	14/15	0.88	0.21	2.27	33,64,76,93	0
3	NAG	A	405	14/15	0.74	0.32	1.87	58,97,112,115	0
3	NAG	E	401	14/15	0.90	0.16	1.85	43,52,63,64	0
3	NAG	C	401	14/15	0.84	0.19	1.68	61,71,98,102	0
6	75U	D	202	29/29	0.92	0.19	1.59	27,47,65,157	0
3	NAG	E	402	14/15	0.93	0.19	1.52	43,54,60,61	0
6	75U	B	201	29/29	0.89	0.20	1.17	30,53,70,152	0
3	NAG	E	404	14/15	0.94	0.22	1.08	41,68,79,88	0
3	NAG	A	403	14/15	0.91	0.16	0.60	41,56,63,65	0
3	NAG	C	406	14/15	0.94	0.17	-0.10	37,59,80,80	0
3	NAG	C	405	14/15	0.95	0.14	-0.28	45,54,64,64	0
7	SO4	D	205	5/5	0.97	0.12	-0.46	56,56,61,62	0
7	SO4	B	204	5/5	0.99	0.10	-1.02	47,53,62,62	0
7	SO4	B	203	5/5	0.97	0.10	-1.03	51,56,68,86	0
3	NAG	C	404	14/15	0.76	0.35	-	74,99,116,116	0
3	NAG	D	204	14/15	0.88	0.44	-	99,118,136,147	0
3	NAG	E	408	14/15	0.95	0.15	-	36,48,61,68	0
3	NAG	D	203	14/15	0.88	0.32	-	50,66,91,108	0
4	BMA	E	405	11/12	0.88	0.34	-	96,109,125,139	0
3	NAG	A	409	14/15	0.83	0.28	-	66,87,103,105	0
4	BMA	A	406	11/12	0.82	0.32	-	110,119,128,141	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	C	409	14/15	0.89	0.29	-	62,80,89,95	0
3	NAG	E	409	14/15	0.81	0.25	-	60,80,90,92	0
3	NAG	C	402	14/15	0.74	0.31	-	52,103,126,132	0
4	BMA	C	407	11/12	0.79	0.21	-	70,83,86,86	0
3	NAG	E	403	14/15	0.95	0.19	-	43,56,67,69	0
5	MAN	E	407	11/12	0.51	0.56	-	146,153,158,159	0
5	MAN	A	407	11/12	0.46	0.44	-	145,158,164,170	0
3	NAG	B	202	14/15	0.91	0.25	-	48,56,61,70	0
3	NAG	A	404	14/15	0.90	0.24	-	61,72,76,80	0
5	MAN	E	406	10/12	0.64	0.33	-	112,130,139,143	0

6.5 Other polymers [i](#)

There are no such residues in this entry.