Full wwPDB X-ray Structure Validation Report

Feb 1, 2016 – 10:25 PM GMT

PDB ID	:	4XQO
Title	:	Crystal structure of hemagglutinin from Jiangxi-Donghu (2013) H10N8 influenza virus in complex with 6'-SLN
Authors	:	Tzarum, N.; Zhang, H.; Zhu, X.; Wilson, I.A.
Deposited on	:	2015-01-19
Resolution	:	2.85 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

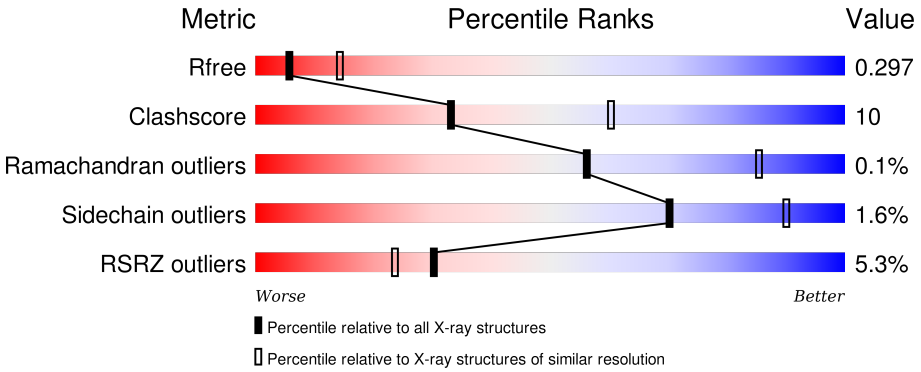
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.85 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.



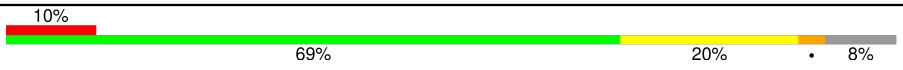
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2228 (2.90-2.82)
Clashscore	102246	2499 (2.90-2.82)
Ramachandran outliers	100387	2439 (2.90-2.82)
Sidechain outliers	100360	2442 (2.90-2.82)
RSRZ outliers	91569	2236 (2.90-2.82)

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 <=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	326	3% 79% 17%
1	C	326	3% 74% 22%
1	E	326	3% 74% 23%
2	B	181	8% 64% 29% 7%
2	D	181	8% 67% 24% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	181	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11538 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 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	0	0
			2424	1499	447	461	17			
1	C	315	Total	C	N	O	S	0	0	0
			2411	1491	445	458	17			
1	E	318	Total	C	N	O	S	0	0	0
			2437	1506	449	465	17			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	ALA	-	expression tag	UNP A0A059T4A1
A	9	ASP	-	expression tag	UNP A0A059T4A1
A	10	PRO	-	expression tag	UNP A0A059T4A1
C	8	ALA	-	expression tag	UNP A0A059T4A1
C	9	ASP	-	expression tag	UNP A0A059T4A1
C	10	PRO	-	expression tag	UNP A0A059T4A1
E	8	ALA	-	expression tag	UNP A0A059T4A1
E	9	ASP	-	expression tag	UNP A0A059T4A1
E	10	PRO	-	expression tag	UNP A0A059T4A1

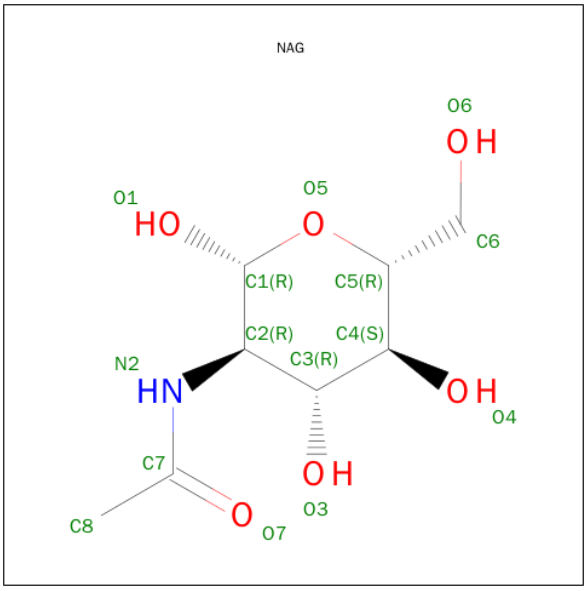
- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total	C	N	O	S	0	0	0
			1364	843	236	277	8			
2	D	168	Total	C	N	O	S	0	0	0
			1362	842	235	277	8			
2	F	166	Total	C	N	O	S	0	0	0
			1351	837	233	273	8			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	expression tag	UNP A0A059T4A1
B	176	GLY	-	expression tag	UNP A0A059T4A1
B	177	ARG	-	expression tag	UNP A0A059T4A1
B	178	LEU	-	expression tag	UNP A0A059T4A1
B	179	VAL	-	expression tag	UNP A0A059T4A1
B	180	PRO	-	expression tag	UNP A0A059T4A1
B	181	ARG	-	expression tag	UNP A0A059T4A1
D	175	SER	-	expression tag	UNP A0A059T4A1
D	176	GLY	-	expression tag	UNP A0A059T4A1
D	177	ARG	-	expression tag	UNP A0A059T4A1
D	178	LEU	-	expression tag	UNP A0A059T4A1
D	179	VAL	-	expression tag	UNP A0A059T4A1
D	180	PRO	-	expression tag	UNP A0A059T4A1
D	181	ARG	-	expression tag	UNP A0A059T4A1
F	175	SER	-	expression tag	UNP A0A059T4A1
F	176	GLY	-	expression tag	UNP A0A059T4A1
F	177	ARG	-	expression tag	UNP A0A059T4A1
F	178	LEU	-	expression tag	UNP A0A059T4A1
F	179	VAL	-	expression tag	UNP A0A059T4A1
F	180	PRO	-	expression tag	UNP A0A059T4A1
F	181	ARG	-	expression tag	UNP A0A059T4A1

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



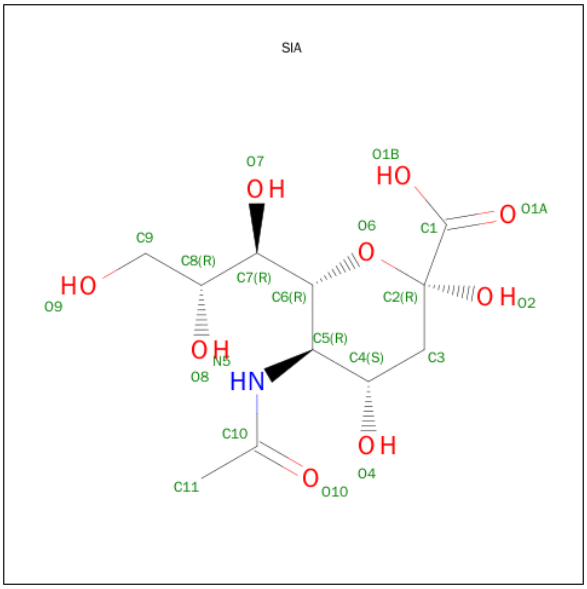
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

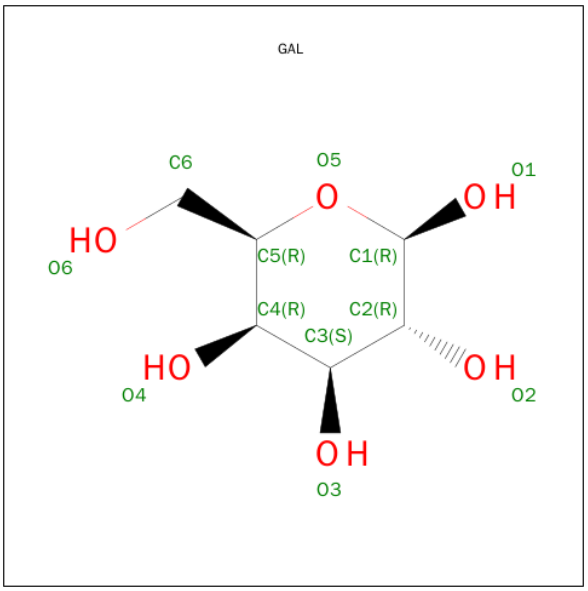
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 15 8 1 6	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is O-SIALIC ACID (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 20 11 1 8	0	0
4	C	1	Total C N O 20 11 1 8	0	0

- Molecule 5 is BETA-D-GALACTOSE (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 11 6 5	0	0
5	C	1	Total C O 12 6 6	0	0

- Molecule 6 is water.

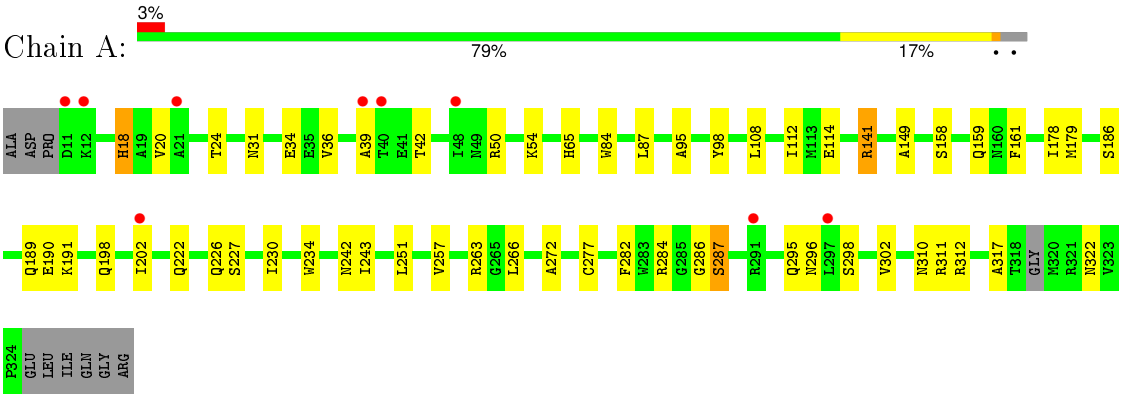
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	12	Total O 12 12	0	0
6	B	4	Total O 4 4	0	0
6	C	6	Total O 6 6	0	0
6	D	1	Total O 1 1	0	0
6	E	4	Total O 4 4	0	0

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 errors 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.

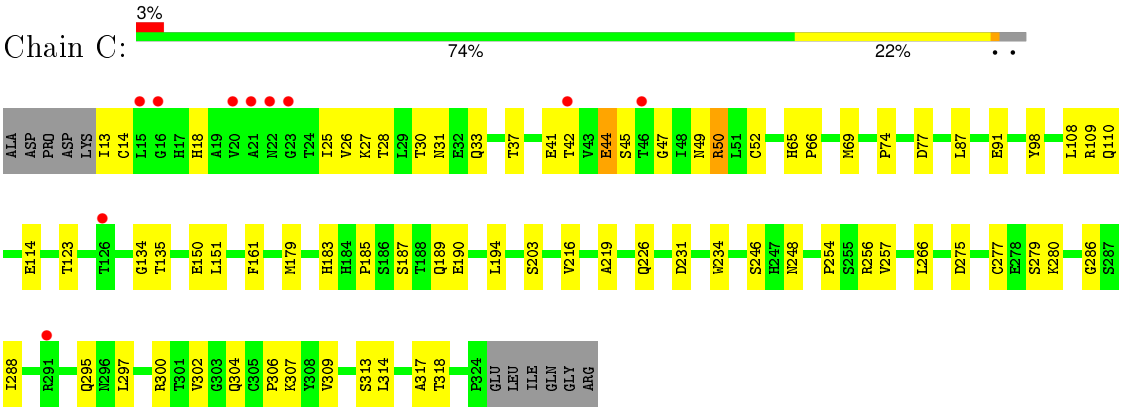
- Molecule 1: Hemagglutinin HA1 chain

Chain A:



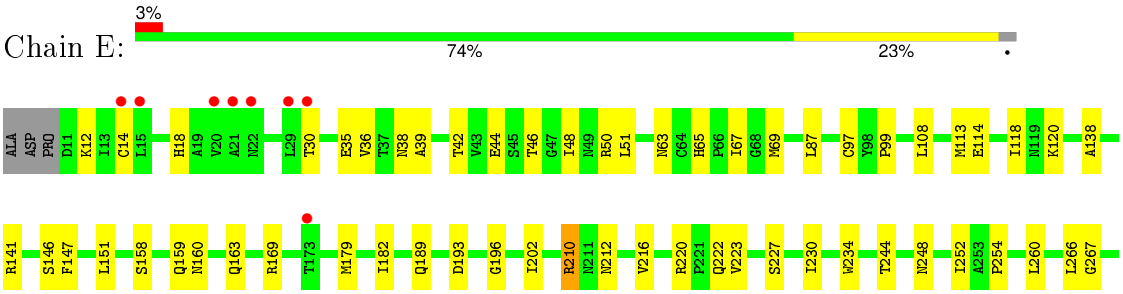
- Molecule 1: Hemagglutinin HA1 chain

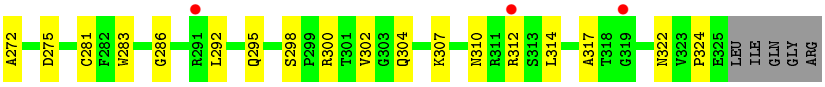
Chain C:



- Molecule 1: Hemagglutinin HA1 chain

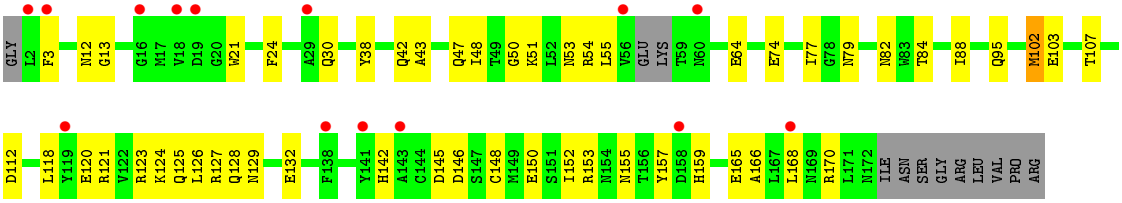
Chain E:





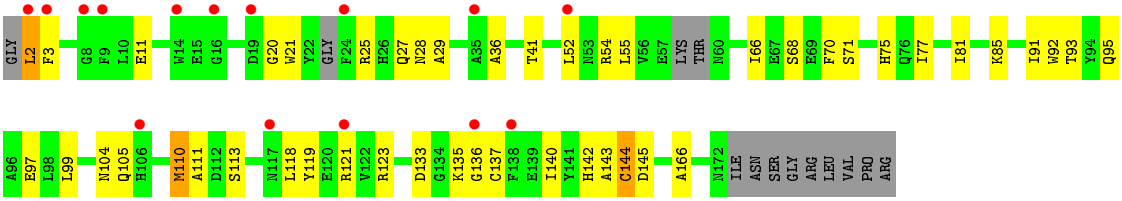
- Molecule 2: Hemagglutinin HA2 chain

Chain B:



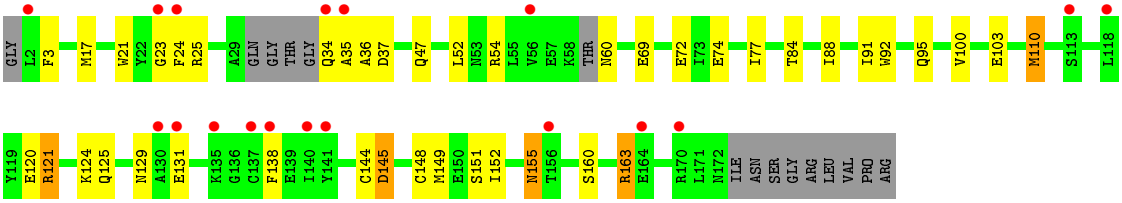
- Molecule 2: Hemagglutinin HA2 chain

Chain D:



- Molecule 2: Hemagglutinin HA2 chain

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	64.07Å 245.42Å 71.30Å 90.00° 112.46° 90.00°	Depositor
Resolution (Å)	46.17 – 2.85 46.17 – 2.81	Depositor EDS
% Data completeness (in resolution range)	88.3 (46.17-2.85) 80.8 (46.17-2.81)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 2.81Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.259 , 0.291 0.279 , 0.297	Depositor DCC
R_{free} test set	2105 reflections (5.63%)	DCC
Wilson B-factor (Å^2)	57.5	Xtriage
Anisotropy	0.542	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 29.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Outliers	1 of 42355 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	11538	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: SIA, GAL, NAG

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	0.24	0/2472	0.48	0/3348
1	C	0.24	0/2460	0.46	0/3334
1	E	0.23	0/2486	0.45	0/3368
2	B	0.26	0/1388	0.49	0/1873
2	D	0.26	0/1385	0.47	0/1867
2	F	0.27	0/1374	0.48	0/1851
All	All	0.25	0/11565	0.47	0/15641

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	F	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	155	ASN	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	2424	0	2375	36	0
1	C	2411	0	2363	63	0
1	E	2437	0	2386	52	0
2	B	1364	0	1259	43	0
2	D	1362	0	1254	42	0
2	F	1351	0	1249	32	0
3	A	57	0	52	1	0
3	C	14	0	13	0	0
3	E	28	0	25	1	0
4	A	20	0	17	2	0
4	C	20	0	17	3	0
5	A	11	0	9	0	0
5	C	12	0	10	1	0
6	A	12	0	0	3	0
6	B	4	0	0	1	0
6	C	6	0	0	0	0
6	D	1	0	0	0	0
6	E	4	0	0	1	0
All	All	11538	0	11029	232	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 10.

All (232) 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:C:44:GLU:OE1	1:C:45:SER:N	2.11	0.80
2:B:125:GLN:HE21	2:B:157:TYR:HB3	1.48	0.78
1:A:226:GLN:NE2	4:A:404:SIA:O1A	2.17	0.77
1:E:63:ASN:OD1	6:E:701:HOH:O	2.03	0.76
1:C:30:THR:HG22	2:D:105:GLN:HE22	1.51	0.74
1:A:226:GLN:NE2	4:A:404:SIA:O8	2.22	0.72
2:D:95:GLN:HE21	2:F:95:GLN:HE22	1.36	0.72
2:B:150:GLU:HG2	2:B:153:ARG:HD2	1.72	0.71
1:E:48:ILE:HD12	1:E:48:ILE:H	1.56	0.71
1:A:284:ARG:NH2	6:A:502:HOH:O	2.20	0.71
1:A:20:VAL:H	1:A:322:ASN:HD21	1.39	0.71
2:D:77:ILE:HD13	2:F:77:ILE:HD11	1.74	0.69
1:C:231:ASP:OD2	1:E:210:ARG:NH2	2.22	0.69
1:C:226:GLN:HE22	5:C:503:GAL:H62	1.60	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:51:LEU:HG	1:E:272:ALA:HB3	1.77	0.66
1:E:50:ARG:NH2	1:E:275:ASP:OD2	2.28	0.66
2:B:120:GLU:OE1	2:B:123:ARG:NH2	2.29	0.66
1:A:296:ASN:HD21	1:A:312:ARG:HA	1.61	0.65
1:A:295:GLN:NE2	1:A:298:SER:H	1.94	0.65
1:C:179:MET:HG2	1:C:234:TRP:HB3	1.79	0.64
2:B:132:GLU:OE1	2:D:123:ARG:NH1	2.31	0.63
2:F:145:ASP:N	2:F:145:ASP:OD2	2.23	0.63
2:F:121:ARG:HH11	2:F:155:ASN:HD21	1.47	0.63
1:C:42:THR:HG22	2:D:55:LEU:HD21	1.81	0.62
2:F:23:GLY:HA2	2:F:36:ALA:HA	1.82	0.61
2:B:150:GLU:HA	2:B:153:ARG:HB2	1.82	0.61
1:A:310:ASN:O	1:A:311:ARG:NH1	2.34	0.61
1:E:36:VAL:HG11	1:E:317:ALA:HB1	1.83	0.61
1:A:158:SER:HB2	1:A:159:GLN:HG2	1.82	0.61
2:B:148:CYS:O	2:B:152:ILE:HD12	2.00	0.61
1:C:18:HIS:CD2	2:D:21:TRP:HA	2.36	0.60
1:E:69:MET:HE1	1:E:87:LEU:HD21	1.82	0.60
1:E:169:ARG:NH1	3:E:601:NAG:O6	2.35	0.60
1:E:99:PRO:HB3	1:E:223:VAL:HB	1.84	0.59
2:D:142:HIS:O	2:D:142:HIS:ND1	2.34	0.59
2:F:54:ARG:NH2	2:F:103:GLU:OE2	2.34	0.59
1:A:295:GLN:HE21	1:A:298:SER:H	1.48	0.59
1:A:186:SER:OG	1:A:190:GLU:OE1	2.14	0.58
1:E:182:ILE:HD12	1:E:202:ILE:HD13	1.84	0.58
1:C:307:LYS:HG3	2:D:92:TRP:CE2	2.37	0.58
1:C:187:SER:OG	1:C:189:GLN:O	2.19	0.58
1:A:42:THR:HG22	2:B:55:LEU:HD21	1.85	0.58
2:D:142:HIS:HD2	2:D:166:ALA:HB2	1.68	0.58
1:C:203:SER:OG	1:C:246:SER:OG	2.22	0.58
1:C:18:HIS:CE1	1:C:37:THR:HG21	2.39	0.57
1:C:18:HIS:HD2	2:D:21:TRP:HA	1.68	0.57
1:A:202:ILE:HG12	1:A:251:LEU:HB2	1.87	0.57
1:C:110:GLN:O	1:C:114:GLU:HG3	2.05	0.57
1:C:13:ILE:HG22	2:D:140:ILE:HD11	1.87	0.56
2:B:30:GLN:OE1	2:B:30:GLN:N	2.38	0.56
2:D:21:TRP:H	2:D:41:THR:HG21	1.70	0.56
1:C:50:ARG:NH1	1:C:275:ASP:OD2	2.38	0.56
1:E:266:LEU:HD11	1:E:302:VAL:HG12	1.88	0.56
2:D:119:TYR:HE2	2:D:136:GLY:HA2	1.71	0.55
1:E:158:SER:H	1:E:159:GLN:HE21	1.54	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:161:PHE:O	1:A:198:GLN:NE2	2.36	0.55
1:C:288:ILE:HG12	1:C:297:LEU:HD12	1.89	0.55
2:B:123:ARG:HG2	2:B:123:ARG:HH11	1.71	0.55
1:E:67:ILE:HD12	1:E:108:LEU:HD23	1.89	0.55
1:C:41:GLU:HG3	1:C:42:THR:H	1.71	0.55
1:A:18:HIS:CD2	2:B:21:TRP:HA	2.42	0.55
1:E:179:MET:HG2	1:E:234:TRP:HB3	1.89	0.55
1:C:50:ARG:HG3	1:C:50:ARG:HH21	1.71	0.54
2:F:24:PHE:HE1	2:F:37:ASP:HB2	1.73	0.54
1:C:295:GLN:HG2	1:C:306:PRO:HG2	1.89	0.54
2:D:143:ALA:O	2:D:144:CYS:HB2	2.08	0.54
1:E:141:ARG:NH1	1:E:147:PHE:O	2.31	0.54
1:E:272:ALA:HB2	1:E:286:GLY:H	1.71	0.54
2:B:127:ARG:HD3	2:B:159:HIS:CD2	2.43	0.54
1:C:52:CYS:HB2	1:C:279:SER:HB2	1.90	0.54
1:A:189:GLN:NE2	6:A:504:HOH:O	2.23	0.53
1:C:108:LEU:HD22	1:C:234:TRP:CD1	2.44	0.53
1:A:50:ARG:O	1:A:287:SER:OG	2.25	0.53
2:D:145:ASP:OD1	2:D:145:ASP:N	2.40	0.53
1:E:35:GLU:HG2	1:E:322:ASN:HB3	1.90	0.53
1:C:41:GLU:HG3	1:C:42:THR:N	2.24	0.53
2:D:21:TRP:H	2:D:41:THR:CG2	2.21	0.53
2:D:54:ARG:HH12	2:D:99:LEU:HD11	1.74	0.53
1:E:283:TRP:CE3	1:E:298:SER:HB2	2.45	0.52
2:B:24:PHE:CD2	2:B:153:ARG:HG2	2.45	0.52
1:C:280:LYS:HB2	1:C:304:GLN:HB2	1.91	0.52
1:E:295:GLN:NE2	1:E:298:SER:H	2.08	0.52
1:C:91:GLU:HA	2:D:70:PHE:CG	2.44	0.52
1:E:113:MET:HB3	1:E:267:GLY:HA3	1.90	0.52
1:E:158:SER:O	1:E:159:GLN:HG2	2.10	0.51
2:B:118:LEU:HD11	2:B:121:ARG:HH21	1.75	0.51
1:C:309:VAL:HG23	2:D:93:THR:HA	1.92	0.51
2:F:17:MET:HE1	2:F:23:GLY:HA3	1.92	0.51
1:A:263:ARG:O	6:A:501:HOH:O	2.19	0.51
2:F:47:GLN:HB2	2:F:110:MET:HE1	1.93	0.51
2:B:102:MET:HG3	2:B:103:GLU:N	2.26	0.51
2:F:24:PHE:N	2:F:35:ALA:O	2.36	0.51
1:E:222:GLN:HG2	1:E:227:SER:OG	2.11	0.50
1:C:26:VAL:HG21	1:C:317:ALA:HB2	1.94	0.50
1:A:272:ALA:HB2	1:A:286:GLY:H	1.75	0.50
1:E:38:ASN:OD1	1:E:39:ALA:N	2.45	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:314:LEU:HD22	2:F:100:VAL:HG21	1.92	0.50
1:C:14:CYS:N	2:D:25:ARG:O	2.41	0.50
2:B:79:ASN:HA	2:B:82:ASN:HB2	1.95	0.49
1:E:44:GLU:OE2	1:E:46:THR:N	2.32	0.49
1:A:141:ARG:NH2	1:A:149:ALA:HB2	2.27	0.49
2:B:95:GLN:HE21	2:F:95:GLN:NE2	2.11	0.49
1:C:216:VAL:HG11	1:E:212:ASN:HB2	1.94	0.49
2:F:120:GLU:OE1	2:F:124:LYS:NZ	2.36	0.49
1:C:313:SER:O	1:C:314:LEU:HD23	2.13	0.49
2:B:165:GLU:O	2:B:168:LEU:HG	2.13	0.49
1:C:18:HIS:HE1	1:C:37:THR:HG21	1.77	0.48
1:E:304:GLN:NE2	2:F:60:ASN:O	2.44	0.48
1:A:282:PHE:CD2	1:A:287:SER:HB3	2.49	0.48
2:D:2:LEU:HD23	2:F:110:MET:HA	1.95	0.48
2:F:129:ASN:HD21	2:F:163:ARG:HA	1.78	0.48
2:D:11:GLU:HG2	2:D:135:LYS:HE3	1.96	0.48
2:B:129:ASN:HB3	2:B:142:HIS:HD2	1.77	0.48
1:C:49:ASN:HA	1:C:286:GLY:HA3	1.96	0.48
2:B:145:ASP:OD1	2:B:146:ASP:N	2.47	0.48
1:C:300:ARG:HG2	2:D:85:LYS:NZ	2.28	0.48
2:B:125:GLN:NE2	2:B:155:ASN:HA	2.29	0.48
1:E:160:ASN:HA	1:E:196:GLY:HA3	1.95	0.48
2:B:74:GLU:HB3	2:B:77:ILE:HG22	1.94	0.48
1:A:54:LYS:HB3	1:A:277:CYS:O	2.14	0.48
1:C:123:THR:HG22	1:C:257:VAL:HG23	1.96	0.48
1:C:28:THR:N	1:C:31:ASN:O	2.46	0.48
2:B:12:ASN:OD1	2:B:13:GLY:N	2.47	0.48
2:B:126:LEU:HD21	2:B:152:ILE:HG12	1.95	0.47
2:D:119:TYR:CE2	2:D:136:GLY:HA2	2.48	0.47
1:E:230:ILE:HD13	1:E:252:ILE:HG13	1.96	0.47
1:A:24:THR:HG21	1:A:39:ALA:HB3	1.96	0.47
1:E:281:CYS:HB2	1:E:304:GLN:O	2.15	0.47
1:A:179:MET:HG2	1:A:234:TRP:HB3	1.95	0.47
1:E:69:MET:O	1:E:120:LYS:NZ	2.44	0.47
1:C:98:TYR:CE2	1:C:226:GLN:HG2	2.50	0.47
1:C:49:ASN:O	1:C:50:ARG:HG2	2.15	0.47
2:D:95:GLN:HE21	2:F:95:GLN:NE2	2.09	0.46
2:B:64:GLU:OE2	6:B:201:HOH:O	2.21	0.46
1:E:14:CYS:N	2:F:25:ARG:O	2.36	0.46
1:C:108:LEU:HD13	1:C:234:TRP:CD2	2.50	0.46
1:E:12:LYS:HB2	2:F:138:PHE:O	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:128:GLN:OE1	2:B:170:ARG:NH1	2.43	0.46
1:E:151:LEU:HD23	1:E:254:PRO:HA	1.98	0.46
1:C:219:ALA:HB3	1:E:244:THR:HG21	1.98	0.46
2:D:3:PHE:CE2	2:D:113:SER:HB2	2.50	0.46
1:C:135:THR:O	4:C:502:SIA:H4	2.15	0.46
1:C:50:ARG:NH2	1:C:50:ARG:HG3	2.29	0.46
2:B:38:TYR:CZ	2:B:42:GLN:HG3	2.50	0.46
1:A:31:ASN:ND2	1:A:34:GLU:OE2	2.48	0.46
1:C:109:ARG:HH11	1:C:109:ARG:HG2	1.81	0.46
2:F:52:LEU:HD23	2:F:52:LEU:HA	1.80	0.46
1:E:18:HIS:ND1	2:F:21:TRP:HA	2.30	0.46
1:E:163:GLN:NE2	1:E:248:ASN:OD1	2.48	0.46
1:A:178:ILE:HG13	1:A:257:VAL:HG12	1.98	0.46
1:A:178:ILE:HG21	1:A:243:ILE:HD13	1.96	0.46
1:C:65:HIS:CG	1:C:66:PRO:HD2	2.51	0.46
2:D:118:LEU:HA	2:D:121:ARG:HG2	1.98	0.46
1:E:42:THR:HA	1:E:292:LEU:HD22	1.98	0.46
1:C:47:GLY:H	1:C:297:LEU:HD11	1.81	0.45
1:C:266:LEU:HD11	1:C:302:VAL:HG13	1.97	0.45
1:E:300:ARG:NE	2:F:69:GLU:OE2	2.30	0.45
2:D:28:ASN:OD1	2:D:29:ALA:N	2.49	0.45
1:E:65:HIS:CE1	1:E:67:ILE:HG12	2.52	0.45
2:B:51:LYS:HZ2	2:B:107:THR:HG1	1.55	0.45
2:D:75:HIS:NE2	1:E:114:GLU:OE1	2.47	0.45
1:C:74:PRO:HA	1:C:77:ASP:OD1	2.17	0.45
1:A:191:LYS:NZ	1:A:198:GLN:O	2.42	0.45
2:D:68:SER:OG	2:D:71:SER:OG	2.26	0.45
1:E:65:HIS:ND1	1:E:67:ILE:HG12	2.32	0.45
2:B:166:ALA:O	2:B:170:ARG:N	2.49	0.45
1:C:69:MET:HE1	1:C:87:LEU:HD21	1.99	0.45
1:E:48:ILE:HD12	1:E:48:ILE:N	2.30	0.44
2:D:66:ILE:HD11	2:D:85:LYS:HD3	1.99	0.44
2:D:28:ASN:CG	2:D:29:ALA:H	2.20	0.44
1:C:27:LYS:O	2:D:104:ASN:ND2	2.32	0.44
2:D:2:LEU:HD22	2:F:3:PHE:HE2	1.83	0.44
1:A:98:TYR:CE1	1:A:230:ILE:HG13	2.52	0.44
1:A:20:VAL:N	1:A:322:ASN:HD21	2.10	0.44
1:C:25:ILE:HG21	1:C:33:GLN:OE1	2.18	0.44
2:B:48:ILE:HD11	2:B:107:THR:HG23	2.00	0.43
2:D:20:GLY:HA3	2:D:36:ALA:HB1	1.99	0.43
2:F:160:SER:HA	2:F:163:ARG:HE	1.83	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:77:ILE:O	2:D:81:ILE:HG13	2.19	0.43
1:E:307:LYS:HG3	2:F:92:TRP:CE2	2.54	0.43
2:D:133:ASP:OD2	2:D:137:CYS:HB2	2.19	0.43
1:E:118:ILE:HG12	1:E:260:LEU:HD23	2.00	0.43
2:D:91:ILE:HD13	2:F:91:ILE:HG21	2.01	0.43
2:B:50:GLY:HA3	1:E:30:THR:O	2.19	0.42
2:B:142:HIS:HB3	2:B:166:ALA:HB2	2.02	0.42
1:C:109:ARG:NH1	1:C:109:ARG:HG2	2.34	0.42
2:B:84:THR:O	2:B:88:ILE:HG12	2.18	0.42
1:C:91:GLU:HA	2:D:70:PHE:CD2	2.54	0.42
2:B:121:ARG:HA	2:B:124:LYS:HE2	2.01	0.42
1:C:150:GLU:OE1	1:C:256:ARG:NE	2.51	0.42
1:C:151:LEU:HD23	1:C:254:PRO:HA	2.01	0.42
2:F:149:MET:HA	2:F:152:ILE:HG13	1.99	0.42
1:A:84:TRP:NE1	1:A:87:LEU:HB2	2.34	0.42
1:A:65:HIS:HB3	1:A:95:ALA:HB2	2.01	0.42
1:C:314:LEU:HD21	2:D:97:GLU:OE2	2.19	0.42
2:F:25:ARG:NE	2:F:34:GLN:OE1	2.53	0.42
2:B:123:ARG:HG2	2:B:123:ARG:NH1	2.34	0.42
2:B:123:ARG:HD3	2:B:132:GLU:OE2	2.20	0.42
1:C:194:LEU:HD11	4:C:502:SIA:O7	2.20	0.42
1:A:266:LEU:HD11	1:A:302:VAL:HG12	2.02	0.42
1:C:52:CYS:HB3	1:C:277:CYS:O	2.20	0.42
1:C:49:ASN:C	1:C:50:ARG:HG2	2.40	0.42
1:E:48:ILE:H	1:E:48:ILE:CD1	2.30	0.42
1:C:318:THR:HG22	2:D:52:LEU:HD21	2.01	0.42
2:B:54:ARG:HH11	2:B:54:ARG:HG3	1.84	0.41
3:A:402:NAG:H61	3:A:403:NAG:N2	2.35	0.41
1:E:314:LEU:HD23	1:E:314:LEU:HA	1.90	0.41
1:A:222:GLN:HG2	1:A:227:SER:HB2	2.02	0.41
2:B:51:LYS:NZ	2:B:103:GLU:O	2.54	0.41
1:C:183:HIS:O	1:C:185:PRO:HD3	2.19	0.41
1:A:114:GLU:HB3	1:A:263:ARG:NH1	2.36	0.41
1:E:216:VAL:O	1:E:220:ARG:NH2	2.53	0.41
2:B:54:ARG:NH1	2:B:54:ARG:HG3	2.36	0.41
2:B:53:ASN:HD22	2:B:53:ASN:HA	1.71	0.41
1:C:108:LEU:HD12	1:C:108:LEU:HA	1.89	0.41
2:F:125:GLN:HE22	2:F:155:ASN:HA	1.85	0.41
1:C:50:ARG:N	1:C:286:GLY:HA2	2.36	0.41
2:B:129:ASN:HB3	2:B:142:HIS:CD2	2.56	0.41
1:E:146:SER:OG	1:E:147:PHE:N	2.52	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:189:GLN:HG2	1:E:193:ASP:OD2	2.21	0.41
1:C:134:GLY:HA3	4:C:502:SIA:H113	2.02	0.40
2:B:3:PHE:HB2	2:B:112:ASP:CG	2.41	0.40
2:D:110:MET:HG3	2:D:111:ALA:N	2.36	0.40
1:E:97:CYS:HB2	1:E:138:ALA:O	2.21	0.40
1:A:108:LEU:O	1:A:112:ILE:HG13	2.21	0.40
1:A:36:VAL:HG11	1:A:317:ALA:HB1	2.04	0.40
1:C:189:GLN:HG3	1:C:190:GLU:N	2.36	0.40
2:F:148:CYS:O	2:F:151:SER:HB2	2.20	0.40
1:C:161:PHE:HB3	1:C:248:ASN:O	2.21	0.40
2:F:84:THR:O	2:F:88:ILE:HG12	2.20	0.40
2:B:128:GLN:NE2	2:F:131:GLU:OE2	2.55	0.40
2:B:43:ALA:O	2:B:47:GLN:HG3	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	312/326 (96%)	302 (97%)	10 (3%)	0	100	100
1	C	313/326 (96%)	295 (94%)	18 (6%)	0	100	100
1	E	316/326 (97%)	305 (96%)	10 (3%)	1 (0%)	46	76
2	B	165/181 (91%)	161 (98%)	4 (2%)	0	100	100
2	D	162/181 (90%)	149 (92%)	12 (7%)	1 (1%)	30	63
2	F	160/181 (88%)	157 (98%)	3 (2%)	0	100	100
All	All	1428/1521 (94%)	1369 (96%)	57 (4%)	2 (0%)	56	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	144	CYS
1	E	324	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	A	268/275 (98%)	264 (98%)	4 (2%)	72 91
1	C	266/275 (97%)	264 (99%)	2 (1%)	86 96
1	E	269/275 (98%)	266 (99%)	3 (1%)	80 94
2	B	144/154 (94%)	143 (99%)	1 (1%)	88 96
2	D	144/154 (94%)	141 (98%)	3 (2%)	61 86
2	F	143/154 (93%)	136 (95%)	7 (5%)	31 63
All	All	1234/1287 (96%)	1214 (98%)	20 (2%)	70 90

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	141	ARG
1	A	242	ASN
1	A	287	SER
2	B	102	MET
1	C	44	GLU
1	C	50	ARG
2	D	2	LEU
2	D	27	GLN
2	D	110	MET
1	E	210	ARG
1	E	310	ASN
1	E	312	ARG
2	F	72	GLU
2	F	74	GLU
2	F	110	MET
2	F	121	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	144	CYS
2	F	145	ASP
2	F	163	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (36) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	HIS
1	A	49	ASN
1	A	295	GLN
1	A	296	ASN
1	A	322	ASN
2	B	28	ASN
2	B	53	ASN
2	B	95	GLN
2	B	125	GLN
2	B	129	ASN
2	B	154	ASN
1	C	18	HIS
1	C	226	GLN
2	D	12	ASN
2	D	53	ASN
2	D	60	ASN
2	D	76	GLN
2	D	95	GLN
2	D	105	GLN
2	D	117	ASN
2	D	154	ASN
1	E	22	ASN
1	E	159	GLN
1	E	198	GLN
1	E	226	GLN
1	E	295	GLN
2	F	28	ASN
2	F	42	GLN
2	F	60	ASN
2	F	117	ASN
2	F	125	GLN
2	F	129	ASN
2	F	155	ASN
2	F	161	GLN
2	F	169	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	172	ASN

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

11 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	0.63	1 (7%)	15,19,21	0.51	0
3	NAG	A	402	1,3	14,14,15	0.73	1 (7%)	15,19,21	0.77	0
3	NAG	A	403	3	14,14,15	0.43	0	15,19,21	0.95	1 (6%)
4	SIA	A	404	5	16,20,21	0.35	0	18,28,31	1.08	3 (16%)
5	GAL	A	405	3,4	11,11,12	0.45	0	14,15,17	1.01	1 (7%)
3	NAG	A	406	5	15,15,15	0.45	0	17,21,21	1.36	3 (17%)
3	NAG	C	501	1	14,14,15	0.76	1 (7%)	15,19,21	0.77	0
4	SIA	C	502	5	16,20,21	0.30	0	18,28,31	1.04	1 (5%)
5	GAL	C	503	4	12,12,12	0.97	1 (8%)	17,17,17	0.95	1 (5%)
3	NAG	E	601	1,3	14,14,15	0.72	1 (7%)	15,19,21	0.76	0
3	NAG	E	602	3	14,14,15	0.32	0	15,19,21	0.72	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
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1
3	NAG	A	402	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	403	3	-	0/6/23/26	0/1/1/1
4	SIA	A	404	5	-	0/14/34/38	0/1/1/1
5	GAL	A	405	3,4	-	0/2/19/22	0/1/1/1
3	NAG	A	406	5	-	0/6/26/26	0/1/1/1
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
4	SIA	C	502	5	-	0/14/34/38	0/1/1/1
5	GAL	C	503	4	-	0/2/22/22	0/1/1/1
3	NAG	E	601	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	602	3	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	503	GAL	O1-C1	-3.21	1.27	1.39
3	E	601	NAG	O5-C1	-2.37	1.39	1.43
3	C	501	NAG	O5-C1	-2.34	1.39	1.43
3	A	402	NAG	O5-C1	-2.32	1.39	1.43
3	A	401	NAG	O5-C1	-2.24	1.40	1.43

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	502	SIA	C4-C5-N5	-2.89	104.13	110.41
3	A	403	NAG	C2-N2-C7	-2.78	119.46	123.04
4	A	404	SIA	C4-C5-N5	-2.39	105.22	110.41
3	A	406	NAG	C3-C2-N2	-2.38	105.73	110.66
5	C	503	GAL	O5-C1-C2	-2.36	106.03	109.80
5	A	405	GAL	O5-C1-C2	-2.20	107.29	110.86
4	A	404	SIA	C3-C4-C5	-2.09	109.15	111.47
4	A	404	SIA	C6-C5-N5	-2.01	107.56	111.07
3	A	406	NAG	C3-C4-C5	2.85	115.16	110.20
3	A	406	NAG	C4-C3-C2	3.48	115.25	110.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	402	NAG	1	0
3	A	403	NAG	1	0
4	A	404	SIA	2	0
4	C	502	SIA	3	0
5	C	503	GAL	1	0
3	E	601	NAG	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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^{th} 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	316/326 (96%)	0.13	9 (2%) 56 51	39, 53, 82, 99	0
1	C	315/326 (96%)	0.26	10 (3%) 51 44	40, 59, 87, 101	0
1	E	318/326 (97%)	0.22	11 (3%) 48 40	37, 60, 98, 113	0
2	B	169/181 (93%)	0.62	14 (8%) 14 9	45, 81, 100, 108	0
2	D	168/181 (92%)	0.70	15 (8%) 12 7	53, 83, 103, 112	0
2	F	166/181 (91%)	0.88	18 (10%) 8 4	43, 88, 108, 117	0
All	All	1452/1521 (95%)	0.39	77 (5%) 30 24	37, 65, 100, 117	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	23	GLY	7.1
1	E	14	CYS	5.0
2	D	8	GLY	4.5
2	F	170	ARG	4.3
1	C	16	GLY	4.2
2	F	118	LEU	4.2
1	C	42	THR	4.1
2	B	56	VAL	4.1
1	E	15	LEU	4.0
2	F	56	VAL	4.0
1	A	21	ALA	3.9
2	F	164	GLU	3.9
2	F	137	CYS	3.7
2	F	135	LYS	3.7
2	B	19	ASP	3.7
2	D	35	ALA	3.7
1	C	15	LEU	3.6
2	F	35	ALA	3.6
2	D	2	LEU	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	141	TYR	3.5
2	B	141	TYR	3.5
1	A	40	THR	3.5
1	E	291	ARG	3.4
2	B	18	VAL	3.2
2	D	106	HIS	3.2
2	D	138	PHE	3.2
1	E	319	GLY	3.2
1	E	22	ASN	3.2
2	D	136	GLY	3.1
1	C	21	ALA	3.0
1	C	46	THR	3.0
2	B	2	LEU	3.0
2	F	130	ALA	3.0
1	C	291	ARG	3.0
1	E	312	ARG	3.0
2	D	117	ASN	3.0
1	C	20	VAL	2.9
1	E	173	THR	2.9
2	B	29	ALA	2.9
1	E	21	ALA	2.8
1	A	48	ILE	2.7
1	C	22	ASN	2.7
2	F	140	ILE	2.6
2	F	2	LEU	2.6
2	F	34	GLN	2.6
1	C	126	THR	2.6
2	B	60	ASN	2.6
1	A	39	ALA	2.6
1	A	291	ARG	2.5
1	A	202	ILE	2.5
2	F	156	THR	2.4
1	A	11	ASP	2.4
2	B	158	ASP	2.4
1	A	297	LEU	2.4
1	C	23	GLY	2.3
2	D	24	PHE	2.3
2	F	113	SER	2.3
2	B	143	ALA	2.3
1	A	12	LYS	2.3
2	D	19	ASP	2.2
2	D	16	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	138	PHE	2.2
2	B	16	GLY	2.2
2	F	131	GLU	2.2
2	D	121	ARG	2.2
2	B	168	LEU	2.1
2	D	14	TRP	2.1
2	B	3	PHE	2.1
2	D	9	PHE	2.1
2	F	24	PHE	2.1
2	D	3	PHE	2.0
1	E	29	LEU	2.0
1	E	20	VAL	2.0
2	F	138	PHE	2.0
1	E	30	THR	2.0
2	D	52	LEU	2.0
2	B	119	TYR	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, 95^{th} 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
4	SIA	A	404	20/21	0.81	0.23	1.38	54,72,82,83	0
3	NAG	C	501	14/15	0.81	0.18	0.41	61,80,94,96	0
3	NAG	E	601	14/15	0.87	0.15	0.11	74,86,102,105	0
3	NAG	A	402	14/15	0.82	0.18	-0.20	64,83,92,101	0
4	SIA	C	502	20/21	0.91	0.18	-0.85	50,63,79,81	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	NAG	A	403	14/15	0.65	0.31	-	104,113,122,125	0
3	NAG	A	401	14/15	0.66	0.28	-	86,98,107,112	0
3	NAG	E	602	14/15	0.56	0.38	-	123,128,143,144	0
5	GAL	C	503	12/12	0.82	0.17	-	68,92,104,105	0
5	GAL	A	405	11/12	0.77	0.25	-	71,82,94,97	0
3	NAG	A	406	15/15	0.72	0.43	-	88,104,109,117	0

6.5 Other polymers

There are no such residues in this entry.