Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 12:56 AM BST

PDB ID	:	5XEB
Title	:	Structure of the envelope glycoprotein of Dhori virus
Authors	:	Peng, R.; Shi, Y.; Qi, J.; Gao, G.F.
Deposited on	:	2017-04-03
Resolution	:	2.50 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
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:

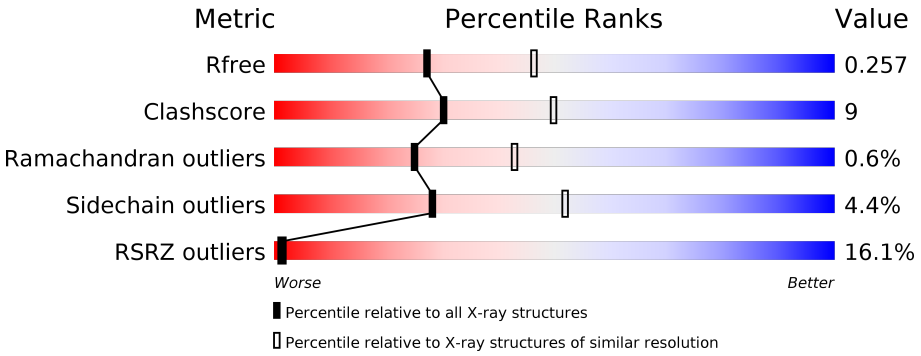
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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 <=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	474	7% 67% 14% • 18%
1	B	474	4% 67% 14% • 18%
1	C	474	28% 60% 19% • 18%

2 Entry composition

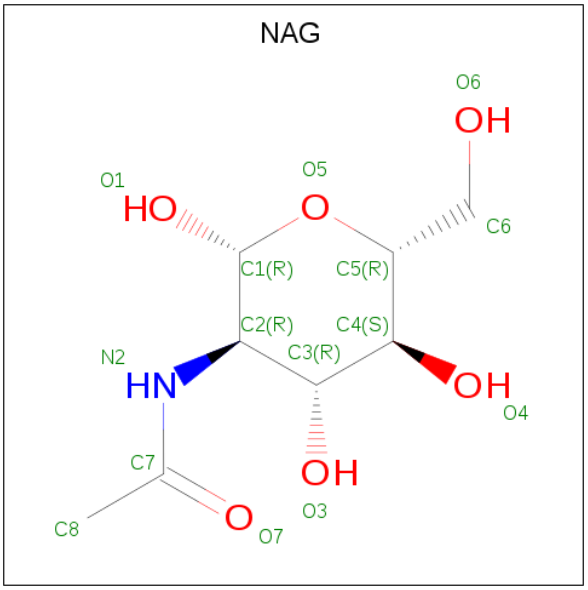
There are 3 unique types of molecules in this entry. The entry contains 9601 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	0	0
			3114	1982	532	580	20			
1	B	390	Total	C	N	O	S	0	0	0
			3114	1982	532	580	20			
1	C	389	Total	C	N	O	S	0	1	0
			3116	1983	533	581	19			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 3 is water.

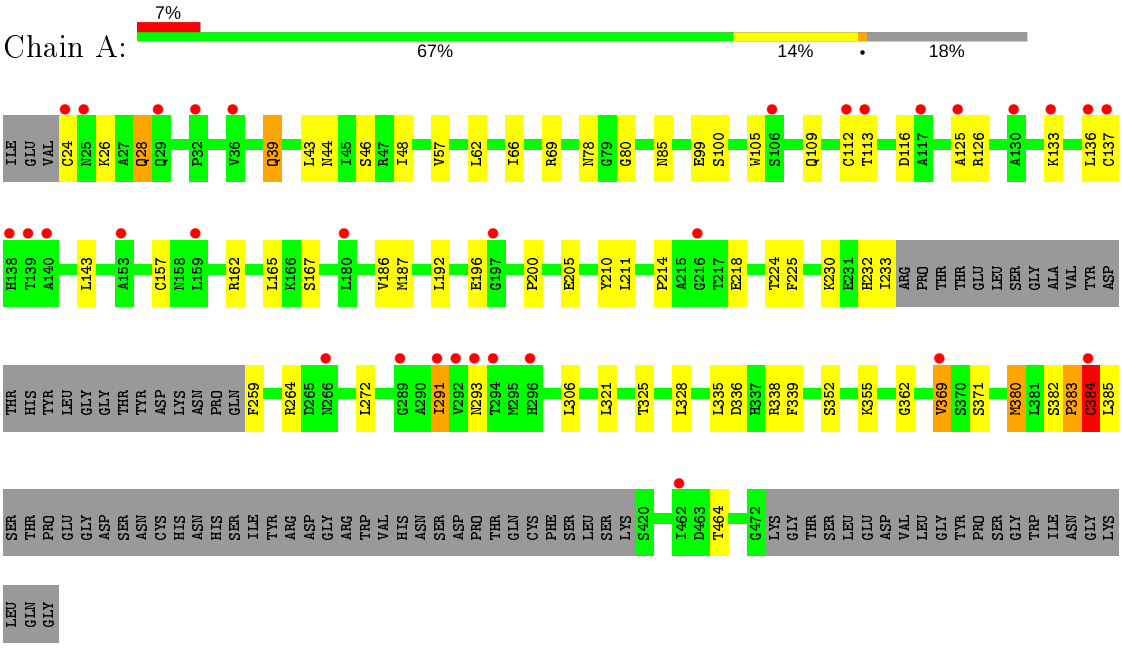
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	105	Total	O	0	0
			105	105		
3	B	72	Total	O	0	0
			72	72		
3	C	24	Total	O	0	0
			24	24		

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.

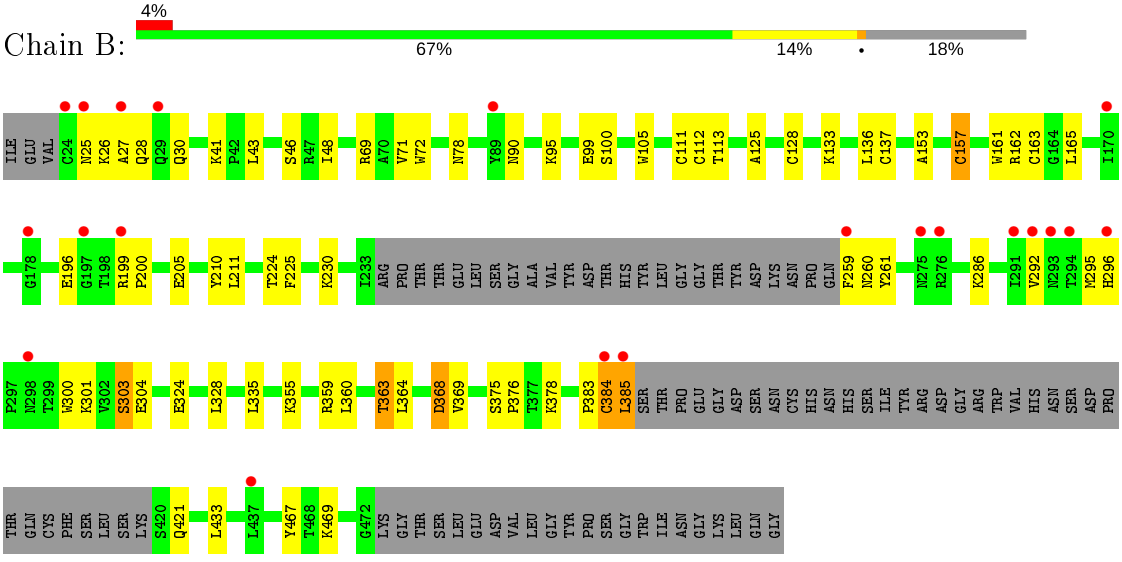
- Molecule 1: Envelope glycoprotein

Chain A:



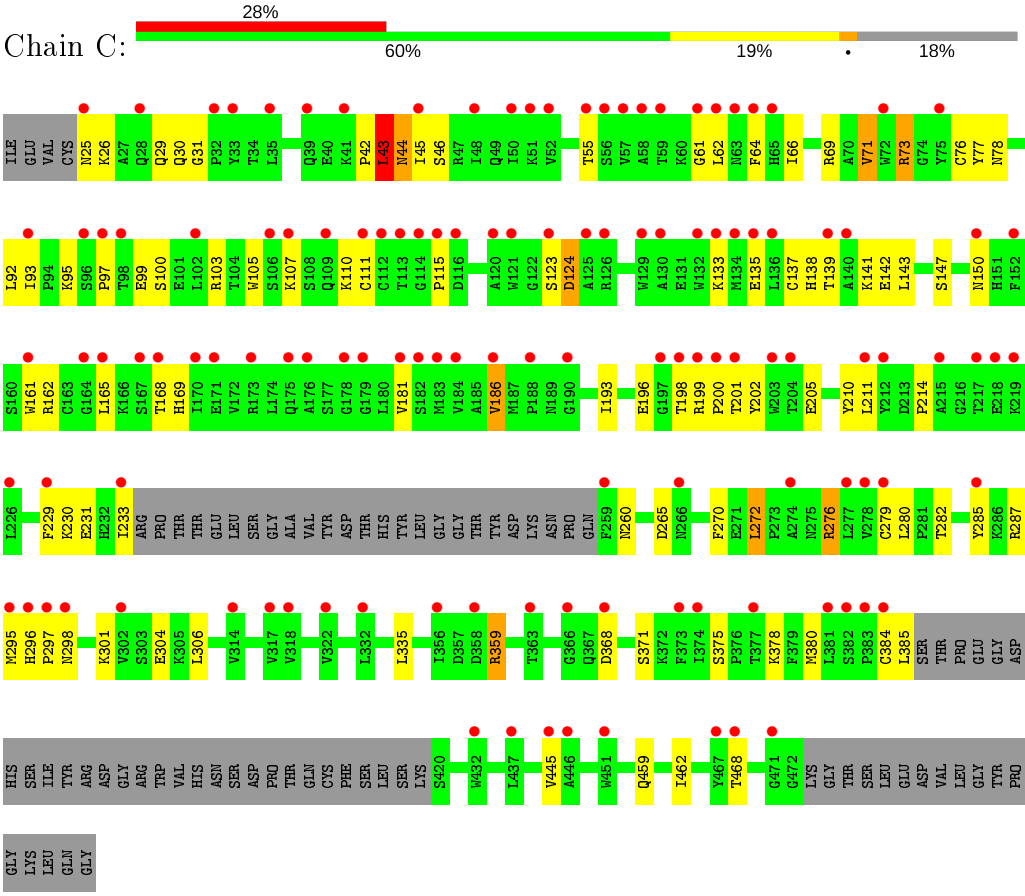
- Molecule 1: Envelope glycoprotein

Chain B:



• Molecule 1: Envelope glycoprotein

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	106.73Å 106.73Å 134.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.94 – 2.50 34.94 – 2.50	Depositor EDS
% Data completeness (in resolution range)	72.3 (34.94-2.50) 72.3 (34.94-2.50)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.49 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155)	Depositor
R, R_{free}	0.231 , 0.257 0.229 , 0.257	Depositor DCC
R_{free} test set	2170 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	36.7	Xtriage
Anisotropy	0.239	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 24.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l 0.407 for h,-h-k,-l 0.000 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9601	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.02 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.2255e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: 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	0.26	0/3192	0.46	0/4330
1	B	0.26	0/3192	0.46	0/4330
1	C	0.27	0/3194	0.51	1/4333 (0.0%)
All	All	0.26	0/9578	0.48	1/12993 (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
1	C	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	279	CYS	C-N-CA	8.90	143.94	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	43	LEU	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	3114	0	3036	46	1
1	B	3114	0	3037	50	0
1	C	3116	0	3036	77	0
2	A	28	0	26	1	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
3	A	105	0	0	23	1
3	B	72	0	0	15	0
3	C	24	0	0	5	0
All	All	9601	0	9161	173	1

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 9.

All (173) 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:B:128:CYS:SG	3:B:725:HOH:O	2.29	0.91
1:B:199:ARG:HD2	1:B:200:PRO:HD2	1.61	0.82
1:B:384:CYS:SG	1:B:385:LEU:N	2.58	0.73
1:C:111:CYS:SG	1:C:141:LYS:NZ	2.54	0.73
1:C:45:ILE:HD11	1:C:229:PHE:HB3	1.70	0.73
1:A:336:ASP:OD1	3:A:701:HOH:O	2.08	0.71
1:A:26:LYS:HE3	1:A:28:GLN:H	1.56	0.70
1:A:44:ASN:HD22	2:A:602:NAG:C7	2.05	0.69
1:B:300:TRP:O	3:B:701:HOH:O	2.11	0.68
1:C:97:PRO:O	1:C:202:TYR:OH	2.14	0.66
1:A:464:THR:OG1	3:A:702:HOH:O	2.13	0.66
1:C:280:LEU:HD22	1:C:285:TYR:HE2	1.60	0.65
1:A:80:GLY:N	3:A:703:HOH:O	2.30	0.65
1:B:375:SER:HB3	1:B:378:LYS:HB2	1.77	0.65
1:B:95:LYS:HB2	3:B:704:HOH:O	1.97	0.64
1:C:287:ARG:NH1	3:C:704:HOH:O	2.29	0.64
1:B:78:ASN:OD1	3:B:702:HOH:O	2.15	0.64
1:C:138:HIS:HB3	1:C:141:LYS:HE3	1.79	0.64
1:A:78:ASN:OD1	3:A:703:HOH:O	2.16	0.63
1:C:115:PRO:O	1:C:161:TRP:NE1	2.29	0.63
1:C:110:LYS:HE3	1:C:137:CYS:HB2	1.82	0.62
1:C:64:PHE:CE2	1:C:66:ILE:HD11	2.35	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:230:LYS:O	3:B:703:HOH:O	2.16	0.62
1:C:359:ARG:NH2	1:C:368:ASP:OD1	2.33	0.61
1:A:352:SER:O	3:A:704:HOH:O	2.16	0.60
1:B:95:LYS:O	3:B:704:HOH:O	2.17	0.60
1:A:369:VAL:HA	1:A:382:SER:O	2.02	0.59
1:C:270:PHE:HE1	1:C:272:LEU:HD23	1.67	0.59
1:A:39:GLN:OE1	3:A:705:HOH:O	2.17	0.59
1:A:24:CYS:N	3:A:733:HOH:O	2.35	0.59
1:A:232:HIS:HA	1:A:259:PHE:HA	1.85	0.58
1:B:112:CYS:HA	1:B:137:CYS:HA	1.86	0.58
1:B:368:ASP:HB3	1:B:385:LEU:HA	1.86	0.58
1:B:69:ARG:HA	1:B:162:ARG:O	2.05	0.57
1:C:78:ASN:HD22	1:C:150[B]:ASN:HA	1.70	0.56
1:C:265:ASP:OD2	1:C:285:TYR:OH	2.17	0.56
1:C:375:SER:HG	1:C:378:LYS:H	1.54	0.56
1:B:28:GLN:O	1:B:30:GLN:N	2.36	0.56
1:A:371:SER:HA	1:A:380:MET:O	2.06	0.55
1:A:113:THR:HG22	1:A:136:LEU:O	2.06	0.55
1:A:339:PHE:HB2	3:A:701:HOH:O	2.07	0.55
1:A:187:MET:SD	3:A:709:HOH:O	2.58	0.55
1:A:362:GLY:HA2	1:A:369:VAL:HG12	1.88	0.55
1:B:421:GLN:N	3:B:708:HOH:O	2.39	0.54
1:A:109:GLN:OE1	3:A:706:HOH:O	2.18	0.54
1:A:48:ILE:HG12	3:A:708:HOH:O	2.08	0.54
1:A:116:ASP:OD2	3:A:707:HOH:O	2.19	0.53
1:B:69:ARG:HG3	1:B:163:CYS:HB3	1.90	0.53
1:C:66:ILE:O	1:C:165:LEU:HA	2.08	0.53
1:A:167:SER:OG	3:A:706:HOH:O	2.19	0.53
1:C:186:VAL:O	3:C:701:HOH:O	2.19	0.52
1:C:76:CYS:SG	3:C:714:HOH:O	2.58	0.52
1:B:99:GLU:HG3	1:B:211:LEU:HD13	1.93	0.51
1:C:71:VAL:HG23	1:C:93:ILE:HB	1.92	0.51
1:A:230:LYS:NZ	3:A:741:HOH:O	2.44	0.51
1:C:147:SER:HB3	1:C:468:THR:HG21	1.93	0.51
1:A:66:ILE:HG12	1:A:210:TYR:HD1	1.76	0.50
1:A:99:GLU:HG3	1:A:211:LEU:HD13	1.94	0.50
1:B:199:ARG:HD2	1:B:200:PRO:CD	2.38	0.50
1:C:66:ILE:HD12	1:C:210:TYR:HD1	1.77	0.50
1:C:61:GLY:O	1:C:169:HIS:ND1	2.45	0.50
1:C:293:ASN:HB3	1:C:295:MET:O	2.12	0.50
1:B:48:ILE:HD13	1:B:230:LYS:HE2	1.94	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:205:GLU:OE1	1:C:210:TYR:OH	2.20	0.49
1:B:205:GLU:HG2	1:B:210:TYR:HE2	1.78	0.49
1:B:324:GLU:OE2	3:B:705:HOH:O	2.19	0.49
1:B:224:THR:OG1	1:B:225:PHE:N	2.46	0.49
1:C:91:ASP:OD2	1:C:93:ILE:HG13	2.12	0.49
1:C:200:PRO:HB3	1:C:211:LEU:HD11	1.94	0.49
1:B:25:ASN:OD1	1:B:26:LYS:N	2.46	0.48
1:C:45:ILE:HD13	1:C:231:GLU:HB2	1.94	0.48
1:C:55:THR:HG21	1:C:181:VAL:HG11	1.96	0.48
1:B:359:ARG:NH2	1:B:368:ASP:OD1	2.42	0.48
1:C:138:HIS:C	1:C:141:LYS:HE3	2.34	0.47
1:C:66:ILE:HD12	1:C:210:TYR:CD1	2.49	0.47
1:C:233:ILE:HG13	3:C:713:HOH:O	2.14	0.47
1:B:230:LYS:HA	1:B:260:ASN:O	2.14	0.47
1:C:282:THR:N	3:C:706:HOH:O	2.35	0.47
1:A:264:ARG:NH2	3:A:714:HOH:O	2.43	0.47
1:C:196:GLU:HB2	1:C:198:THR:HG22	1.95	0.47
1:C:105:TRP:CZ3	1:C:111:CYS:HB2	2.50	0.47
1:A:46:SER:O	3:A:708:HOH:O	2.20	0.47
1:C:186:VAL:HG12	1:C:445:VAL:HG11	1.96	0.47
1:C:45:ILE:HG13	1:C:46:SER:N	2.28	0.47
1:A:205:GLU:OE1	3:A:709:HOH:O	2.21	0.47
1:A:383:PRO:HB2	1:A:384:CYS:H	1.53	0.47
1:B:46:SER:OG	3:B:703:HOH:O	2.20	0.47
1:B:133:LYS:HA	1:B:133:LYS:HD3	1.76	0.47
1:B:113:THR:HG22	1:B:136:LEU:O	2.14	0.47
1:C:142:GLU:OE2	1:C:162:ARG:NE	2.47	0.47
1:C:186:VAL:HG11	1:C:445:VAL:HG21	1.97	0.47
1:B:295:MET:HG2	1:B:296:HIS:H	1.79	0.46
1:B:105:TRP:HB3	1:B:165:LEU:HD22	1.96	0.46
1:C:138:HIS:CE1	1:C:139:THR:HG22	2.51	0.46
1:B:125:ALA:HA	1:B:157:CYS:HB3	1.96	0.46
1:C:124:ASP:OD1	1:C:124:ASP:N	2.48	0.46
1:C:198:THR:HG23	1:C:199:ARG:N	2.30	0.46
1:C:25:ASN:OD1	1:C:26:LYS:N	2.49	0.46
1:C:71:VAL:O	1:C:92:LEU:HA	2.16	0.46
1:B:46:SER:N	3:B:703:HOH:O	2.48	0.46
1:C:138:HIS:CG	1:C:139:THR:H	2.34	0.46
1:A:291:ILE:HG13	1:A:291:ILE:H	1.55	0.46
1:B:360:LEU:O	1:B:363:THR:HG22	2.16	0.46
1:A:126:ARG:NH2	3:A:723:HOH:O	2.29	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:141:LYS:HB2	1:C:162:ARG:HA	1.98	0.45
1:C:45:ILE:HD12	1:C:230:LYS:O	2.17	0.45
1:A:26:LYS:HB3	1:A:28:GLN:HB3	1.97	0.45
1:A:85:ASN:ND2	3:A:736:HOH:O	2.41	0.45
1:A:133:LYS:HA	1:A:133:LYS:HD3	1.82	0.45
1:B:286:LYS:HE3	1:B:292:VAL:HG23	1.99	0.45
1:C:200:PRO:CB	1:C:211:LEU:HD11	2.47	0.45
1:B:368:ASP:CG	1:B:385:LEU:HD23	2.38	0.45
1:B:71:VAL:HG22	1:B:161:TRP:CZ3	2.52	0.45
1:C:384:CYS:SG	1:C:385:LEU:N	2.89	0.45
1:A:57:VAL:HA	1:A:218:GLU:O	2.17	0.44
1:C:69:ARG:HA	1:C:162:ARG:O	2.16	0.44
1:A:143:LEU:HD12	3:A:702:HOH:O	2.18	0.44
1:A:230:LYS:N	3:A:708:HOH:O	2.40	0.44
1:C:115:PRO:O	1:C:159:LEU:HD23	2.17	0.44
1:C:168:THR:OG1	1:C:169:HIS:N	2.50	0.44
1:C:138:HIS:ND1	1:C:139:THR:HG22	2.33	0.44
1:C:200:PRO:HA	1:C:214:PRO:HD3	2.00	0.44
1:B:467:TYR:CZ	1:B:469:LYS:HE3	2.53	0.43
1:A:205:GLU:HG2	1:A:210:TYR:HE2	1.83	0.43
1:C:99:GLU:OE2	1:C:211:LEU:HD13	2.18	0.43
1:C:69:ARG:HB3	1:C:95:LYS:O	2.18	0.43
1:C:371:SER:HA	1:C:380:MET:O	2.19	0.43
1:A:112:CYS:HA	1:A:137:CYS:HA	2.00	0.43
1:A:69:ARG:HA	1:A:162:ARG:O	2.18	0.43
1:A:321:LEU:O	1:A:325:THR:HG23	2.18	0.43
1:B:303:SER:OG	3:B:707:HOH:O	2.21	0.43
1:C:301:LYS:HD2	1:C:304:GLU:HG3	2.01	0.43
1:B:383:PRO:O	1:B:384:CYS:HB2	2.18	0.43
1:C:92:LEU:HD11	1:C:143:LEU:HD21	2.01	0.43
1:B:196:GLU:O	3:B:706:HOH:O	2.21	0.43
1:C:77:TYR:HB2	1:C:155:HIS:ND1	2.34	0.43
1:B:259:PHE:N	3:B:734:HOH:O	2.52	0.43
1:C:78:ASN:HA	1:C:86:THR:HG23	2.01	0.43
1:C:44:ASN:HB3	1:C:45:ILE:H	1.59	0.42
1:C:205:GLU:HB2	1:C:210:TYR:CE2	2.54	0.42
1:C:233:ILE:HG12	1:C:260:ASN:OD1	2.19	0.42
1:B:26:LYS:HD2	1:B:27:ALA:H	1.83	0.42
1:C:73:ARG:HB2	1:C:159:LEU:HD12	2.01	0.42
1:C:133:LYS:O	1:C:135:GLU:HG3	2.19	0.42
1:A:125:ALA:HA	1:A:157:CYS:HB3	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:224:THR:OG1	1:A:225:PHE:N	2.49	0.42
1:A:355:LYS:HB2	3:A:704:HOH:O	2.20	0.42
1:C:201:THR:O	1:C:211:LEU:HD12	2.20	0.42
1:C:295:MET:HG3	1:C:296:HIS:N	2.35	0.42
1:C:296:HIS:N	1:C:297:PRO:HD3	2.35	0.42
1:B:355:LYS:NZ	1:B:376:PRO:HA	2.35	0.42
1:C:105:TRP:HB3	1:C:165:LEU:HD22	2.02	0.41
1:A:105:TRP:HB3	1:A:165:LEU:HD22	2.03	0.41
1:B:230:LYS:HD3	1:B:261:TYR:CE1	2.55	0.41
1:C:295:MET:HG3	1:C:296:HIS:H	1.84	0.41
1:C:459:GLN:OE1	1:C:462:ILE:HD11	2.20	0.41
1:C:93:ILE:H	1:C:93:ILE:HD12	1.85	0.41
1:B:105:TRP:CD2	1:B:111:CYS:HA	2.56	0.41
1:C:276:ARG:NH2	1:C:298:ASN:OD1	2.54	0.41
1:A:200:PRO:HA	1:A:214:PRO:HD3	2.01	0.41
1:B:72:TRP:CH2	1:B:90:ASN:HB3	2.55	0.41
1:B:301:LYS:HB3	1:B:304:GLU:HG3	2.02	0.41
1:B:421:GLN:O	3:B:708:HOH:O	2.22	0.41
1:B:78:ASN:O	1:B:153:ALA:N	2.53	0.41
1:C:103:ARG:HG3	1:C:107:LYS:HZ3	1.86	0.41
1:C:306:LEU:HD23	1:C:445:VAL:HB	2.02	0.41
1:A:186:VAL:HG22	1:A:192:LEU:HD22	2.02	0.41
1:B:368:ASP:OD2	1:B:385:LEU:HD23	2.21	0.41
1:C:99:GLU:HG2	1:C:202:TYR:CG	2.56	0.41
1:C:43:LEU:HB3	1:C:44:ASN:H	1.69	0.41
1:B:369:VAL:O	3:B:710:HOH:O	2.22	0.40
1:A:232:HIS:O	3:A:710:HOH:O	2.21	0.40
1:C:29:GLN:O	1:C:31:GLY:N	2.53	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:338:ARG:NH1	3:A:701:HOH:O[2_555]	2.11	0.09

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	384/474 (81%)	363 (94%)	19 (5%)	2 (0%)	29	48
1	B	384/474 (81%)	365 (95%)	18 (5%)	1 (0%)	41	61
1	C	384/474 (81%)	353 (92%)	27 (7%)	4 (1%)	15	28
All	All	1152/1422 (81%)	1081 (94%)	64 (6%)	7 (1%)	25	43

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	383	PRO
1	B	384	CYS
1	C	43	LEU
1	C	30	GLN
1	C	44	ASN
1	A	384	CYS
1	C	42	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	344/417 (82%)	327 (95%)	17 (5%)	25	47
1	B	344/417 (82%)	332 (96%)	12 (4%)	36	62
1	C	344/417 (82%)	328 (95%)	16 (5%)	26	49
All	All	1032/1251 (82%)	987 (96%)	45 (4%)	28	52

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	GLN
1	A	39	GLN
1	A	43	LEU
1	A	62	LEU
1	A	100	SER
1	A	196	GLU
1	A	233	ILE
1	A	272	LEU
1	A	291	ILE
1	A	293	ASN
1	A	306	LEU
1	A	328	LEU
1	A	335	LEU
1	A	369	VAL
1	A	380	MET
1	A	384	CYS
1	A	385	LEU
1	B	41	LYS
1	B	43	LEU
1	B	100	SER
1	B	157	CYS
1	B	303	SER
1	B	328	LEU
1	B	335	LEU
1	B	363	THR
1	B	364	LEU
1	B	368	ASP
1	B	385	LEU
1	B	433	LEU
1	C	43	LEU
1	C	62	LEU
1	C	71	VAL
1	C	73	ARG
1	C	100	SER
1	C	123	SER
1	C	124	ASP
1	C	159	LEU
1	C	186	VAL
1	C	193	ILE
1	C	220	LYS
1	C	272	LEU
1	C	276	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	293	ASN
1	C	335	LEU
1	C	359	ARG

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	$\#\|Z\| > 2$
2	NAG	A	602	1	14,14,15	0.15	0	17,19,21	0.42	0
2	NAG	B	601	1	14,14,15	0.24	0	17,19,21	0.36	0
2	NAG	C	601	1	14,14,15	0.47	0	17,19,21	0.78	1 (5%)
2	NAG	A	601	1	14,14,15	0.30	0	17,19,21	0.40	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
2	NAG	A	602	1	-	0/6/23/26	0/1/1/1
2	NAG	B	601	1	-	0/6/23/26	0/1/1/1
2	NAG	C	601	1	-	2/6/23/26	0/1/1/1
2	NAG	A	601	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	601	NAG	C1-O5-C5	2.79	115.97	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	601	NAG	O5-C5-C6-O6
2	C	601	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	602	NAG	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

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	390/474 (82%)	0.32	32 (8%)	11	11	15, 43, 91, 125	0
1	B	390/474 (82%)	0.32	21 (5%)	25	27	16, 45, 94, 171	0
1	C	389/474 (82%)	1.80	135 (34%)	0	0	49, 93, 150, 193	0
All	All	1169/1422 (82%)	0.81	188 (16%)	1	1	15, 58, 128, 193	0

All (188) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	294	THR	14.9
1	A	384	CYS	9.1
1	C	114	GLY	8.4
1	C	190	GLY	8.3
1	C	294	THR	7.5
1	B	24	CYS	7.2
1	C	198	THR	7.2
1	C	107	LYS	7.2
1	B	292	VAL	7.1
1	C	170	ILE	6.9
1	C	226	LEU	6.7
1	C	197	GLY	6.3
1	C	366	GLY	6.2
1	C	223	SER	6.2
1	C	48	ILE	6.1
1	C	437	LEU	5.9
1	C	125	ALA	5.7
1	C	186	VAL	5.7
1	C	51	LYS	5.6
1	C	115	PRO	5.6
1	C	171	GLU	5.5
1	A	136	LEU	5.5
1	C	173	ARG	5.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	50	ILE	5.3
1	C	199	ARG	5.1
1	A	24	CYS	5.0
1	A	29	GLN	5.0
1	C	93	ILE	4.9
1	C	225	PHE	4.9
1	C	296	HIS	4.9
1	C	25	ASN	4.9
1	C	274	ALA	4.8
1	C	297	PRO	4.8
1	A	113	THR	4.8
1	C	165	LEU	4.8
1	C	113	THR	4.8
1	A	369	VAL	4.7
1	C	201	THR	4.6
1	C	98	THR	4.6
1	B	296	HIS	4.4
1	C	106	SER	4.4
1	B	275	ASN	4.4
1	C	102	LEU	4.3
1	C	183	MET	4.3
1	C	140	ALA	4.3
1	C	212	TYR	4.2
1	C	277	LEU	4.2
1	A	140	ALA	4.1
1	B	385	LEU	4.1
1	C	179	GLY	4.1
1	C	259	PHE	4.0
1	C	120	ALA	4.0
1	C	293	ASN	4.0
1	C	219	LYS	4.0
1	C	200	PRO	3.9
1	C	64	PHE	3.9
1	A	137	CYS	3.8
1	C	298	ASN	3.7
1	C	314	VAL	3.7
1	B	384	CYS	3.7
1	C	135	GLU	3.7
1	C	139	THR	3.7
1	C	130	ALA	3.7
1	A	130	ALA	3.6
1	A	293	ASN	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	218	GLU	3.6
1	A	36	VAL	3.5
1	A	112	CYS	3.5
1	C	134	MET	3.5
1	C	65	HIS	3.4
1	C	56	SER	3.4
1	C	381	LEU	3.4
1	A	289	GLY	3.4
1	C	154	TYR	3.3
1	C	302	VAL	3.3
1	C	152	PHE	3.3
1	B	89	TYR	3.3
1	B	25	ASN	3.2
1	C	292	VAL	3.2
1	C	52	VAL	3.2
1	C	318	VAL	3.2
1	C	266	ASN	3.2
1	C	184	VAL	3.1
1	C	374	ILE	3.1
1	C	377	THR	3.1
1	C	175	GLN	3.0
1	C	295	MET	3.0
1	C	215	ALA	3.0
1	C	233	ILE	3.0
1	C	129	TRP	3.0
1	A	133	LYS	3.0
1	A	139	THR	3.0
1	C	217	THR	3.0
1	C	167	SER	3.0
1	A	216	GLY	3.0
1	C	39	GLN	3.0
1	C	291	ILE	2.9
1	C	62	LEU	2.9
1	C	111	CYS	2.9
1	C	126	ARG	2.9
1	A	296	HIS	2.9
1	A	25	ASN	2.8
1	C	211	LEU	2.8
1	C	132	TRP	2.8
1	B	199	ARG	2.8
1	B	293	ASN	2.8
1	C	33	TYR	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	467	TYR	2.8
1	C	322	VAL	2.7
1	A	462	ILE	2.7
1	C	28	GLN	2.7
1	C	41	LYS	2.7
1	C	178	GLY	2.7
1	A	138	HIS	2.7
1	C	451	TRP	2.7
1	B	291	ILE	2.6
1	C	176	ALA	2.6
1	C	133	LYS	2.6
1	C	45	ILE	2.6
1	B	197	GLY	2.6
1	A	294	THR	2.6
1	C	97	PRO	2.5
1	C	109	GLN	2.5
1	C	382	SER	2.5
1	A	117	ALA	2.5
1	C	58	ALA	2.5
1	C	89	TYR	2.5
1	C	35	LEU	2.5
1	C	153	ALA	2.5
1	A	197	GLY	2.5
1	A	266	ASN	2.5
1	C	432	TRP	2.5
1	C	123	SER	2.4
1	A	291	ILE	2.4
1	C	203	TRP	2.4
1	C	332	LEU	2.4
1	C	229	PHE	2.4
1	A	153	ALA	2.4
1	C	116	ASP	2.4
1	C	150[A]	ASN	2.3
1	C	57	VAL	2.3
1	C	182	SER	2.3
1	C	164	GLY	2.3
1	B	170	ILE	2.3
1	B	437	LEU	2.3
1	A	125	ALA	2.3
1	C	363	THR	2.3
1	B	178	GLY	2.3
1	C	61	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	168	THR	2.3
1	C	96	SER	2.2
1	C	373	PHE	2.2
1	C	278	VAL	2.2
1	C	59	THR	2.2
1	A	292	VAL	2.2
1	C	356	ILE	2.2
1	B	29	GLN	2.2
1	B	298	ASN	2.2
1	C	32	PRO	2.2
1	C	188	PRO	2.2
1	C	471	GLY	2.2
1	C	181	VAL	2.2
1	C	445	VAL	2.2
1	C	55	THR	2.2
1	C	63	ASN	2.2
1	C	112	CYS	2.2
1	C	279	CYS	2.2
1	B	276	ARG	2.1
1	C	204	THR	2.1
1	A	180	LEU	2.1
1	C	72	TRP	2.1
1	C	468	THR	2.1
1	C	383	PRO	2.1
1	C	161	TRP	2.1
1	A	106	SER	2.1
1	A	32	PRO	2.1
1	C	285	TYR	2.1
1	C	136	LEU	2.1
1	C	446	ALA	2.1
1	C	317	VAL	2.1
1	B	259	PHE	2.1
1	C	368	ASP	2.0
1	C	75	TYR	2.0
1	C	121	TRP	2.0
1	C	384	CYS	2.0
1	B	27	ALA	2.0
1	A	159	LEU	2.0
1	C	358	ASP	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 monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{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	B-factors(Å^2)	Q<0.9
2	NAG	A	601	14/15	0.64	0.23	57,74,82,84	0
2	NAG	A	602	14/15	0.73	0.23	120,123,126,128	0
2	NAG	C	601	14/15	0.75	0.24	77,92,99,100	0
2	NAG	B	601	14/15	0.78	0.18	51,63,75,75	0

6.5 Other polymers (i)

There are no such residues in this entry.