Full wwPDB X-ray Structure Validation Report (i)

Mar 17, 2016 – 10:45 PM EDT

PDB ID	:	5E65
Title	:	The complex structure of Hemagglutinin-esterase-fusion mutant protein from the influenza D virus with receptor analog 9-O-Ac-3'SLN (Tr322)
Authors	:	Song, H.; Qi, J.; Shi, Y.; Gao, G.F.
Deposited on	:	2015-10-09
Resolution	:	2.20 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:

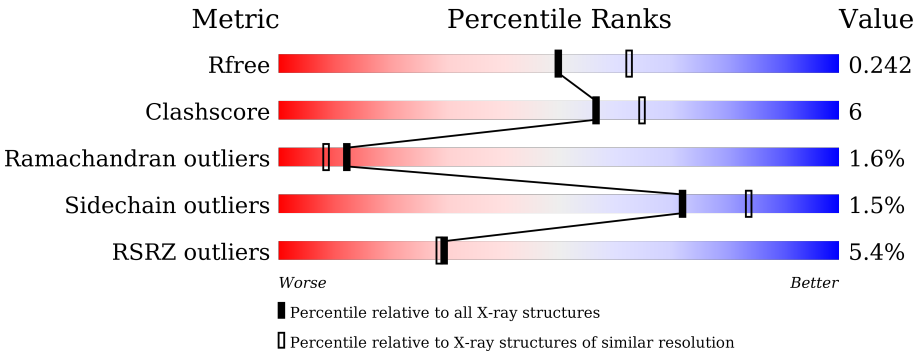
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027107
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0122
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027107

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 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	427	5% 90% 9% ·
1	C	427	% 88% 11% ·
2	B	166	11% 66% 21% · 10%
2	D	166	11% 63% 25% · · 10%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 9829 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-esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	0	0
			3240	2039	540	639	22			
1	C	424	Total	C	N	O	S	0	0	0
			3240	2039	540	639	22			

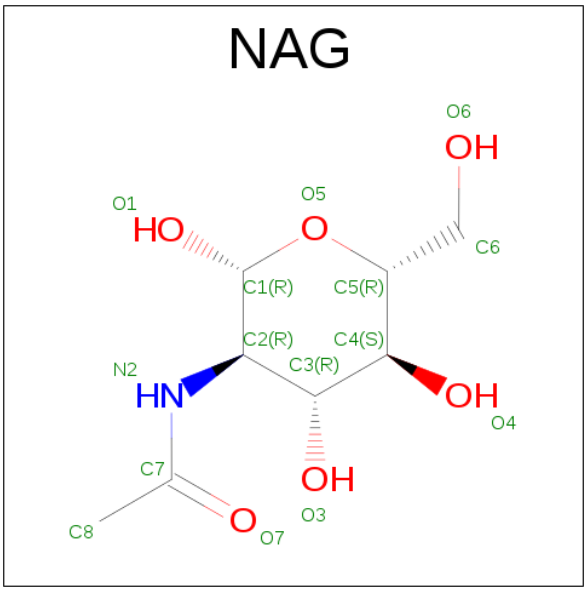
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	57	ALA	SER	engineered mutation	UNP K9LG83
A	356	ALA	ASP	engineered mutation	UNP K9LG83
A	359	ALA	HIS	engineered mutation	UNP K9LG83
C	57	ALA	SER	engineered mutation	UNP K9LG83
C	356	ALA	ASP	engineered mutation	UNP K9LG83
C	359	ALA	HIS	engineered mutation	UNP K9LG83

- Molecule 2 is a protein called Hemagglutinin-esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	149	Total	C	N	O	S	0	0	0
			1110	695	194	217	4			
2	D	149	Total	C	N	O	S	0	0	0
			1110	695	194	217	4			

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



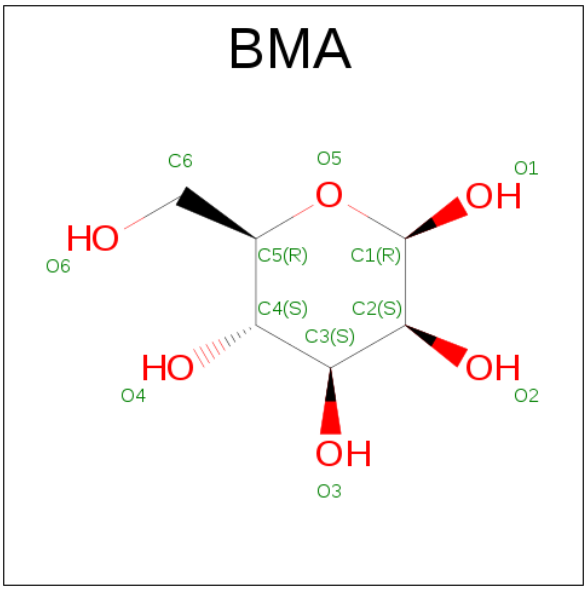
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0
3	A	1	Total C N O 14 8 1 5	0	0
3	B	1	Total C N O 14 8 1 5	0	0
3	B	1	Total C N O 14 8 1 5	0	0
3	B	1	Total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

Continued on next page...

Continued from previous page...

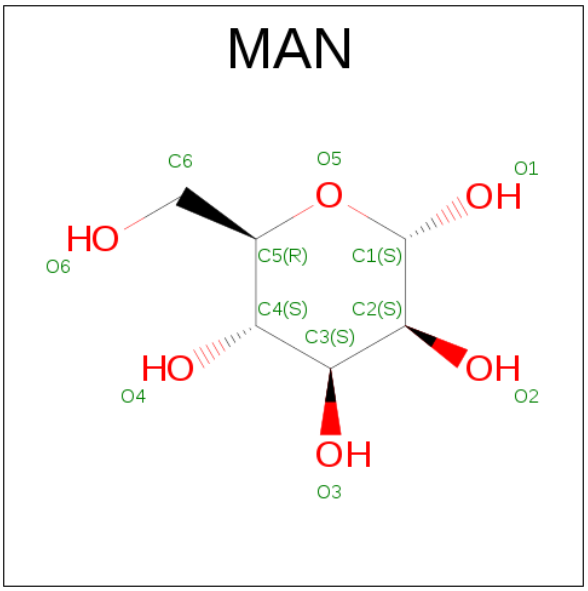
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is BETA-D-MANNOSE (three-letter code: BMA) (formula: $C_6H_{12}O_6$).



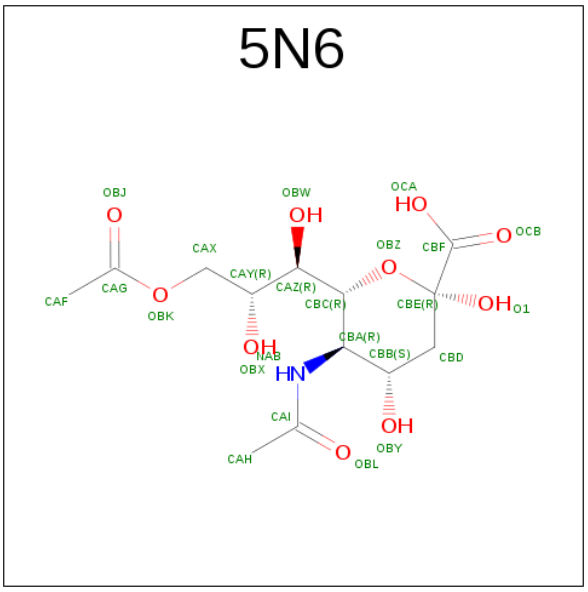
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 11 6 5	0	0
4	A	1	Total C O 11 6 5	0	0
4	C	1	Total C O 11 6 5	0	0

- Molecule 5 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: $C_6H_{12}O_6$).



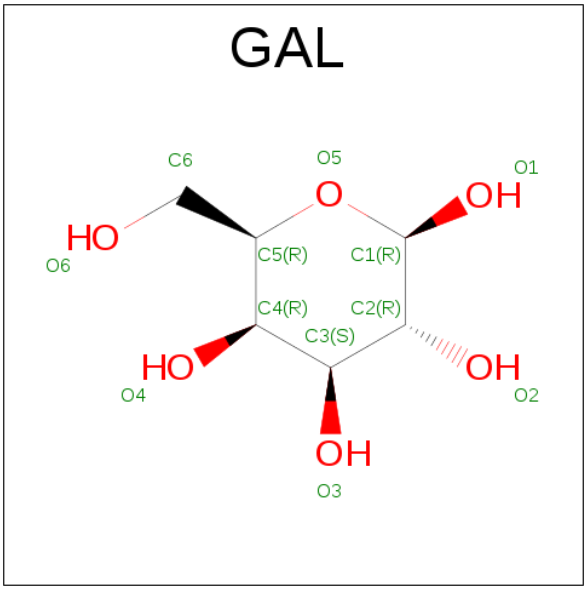
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 11 6 5	0	0
5	A	1	Total C O 11 6 5	0	0
5	A	1	Total C O 11 6 5	0	0
5	A	1	Total C O 11 6 5	0	0
5	A	1	Total C O 11 6 5	0	0
5	A	1	Total C O 11 6 5	0	0
5	C	1	Total C O 11 6 5	0	0
5	C	1	Total C O 11 6 5	0	0
5	C	1	Total C O 11 6 5	0	0

- Molecule 6 is (2 {R},4 {S},5 {R},6 {R})-5-acetamido-6-[(1 {R},2 {R})-3-acetyloxy-1,2-bis(oxidanyl)propyl]-2,4-bis(oxidanyl)oxane-2-carboxylic acid (three-letter code: 5N6) (formula: $C_{13}H_{21}NO_{10}$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			23	13	1	9		
6	C	1	Total	C	N	O	0	0
			23	13	1	9		

- Molecule 7 is BETA-D-GALACTOSE (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



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

- Molecule 8 is water.

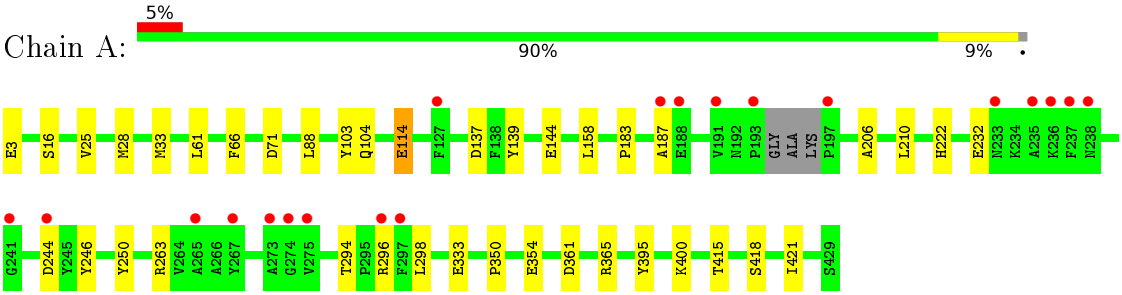
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	264	Total O 264 264	0	0
8	B	68	Total O 68 68	0	0
8	C	304	Total O 304 304	0	0
8	D	39	Total O 39 39	0	0

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.

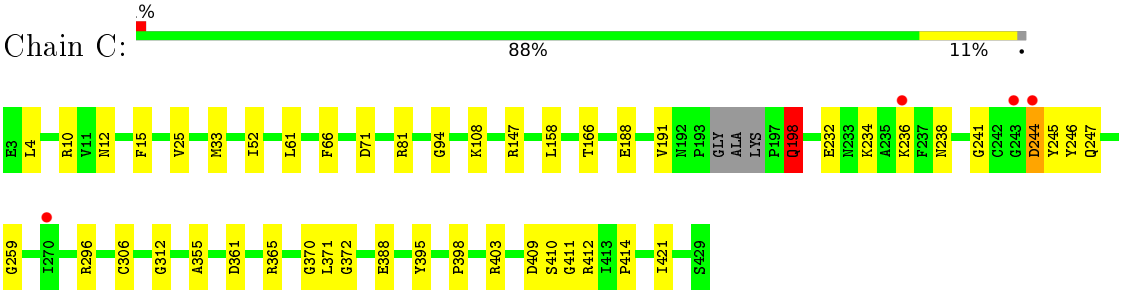
- Molecule 1: Hemagglutinin-esterase

Chain A:



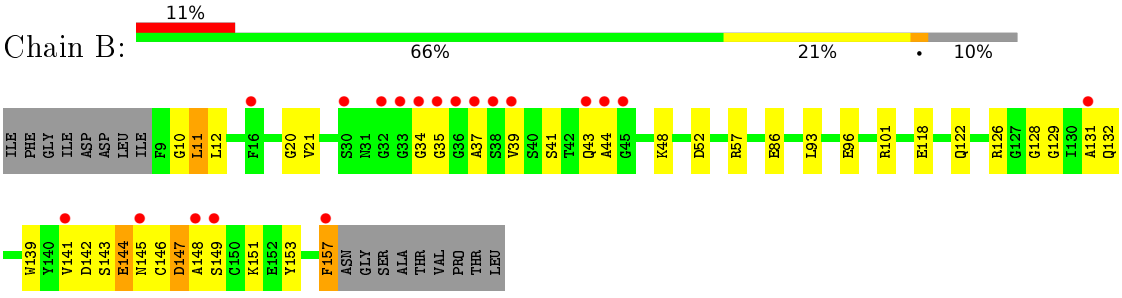
- Molecule 1: Hemagglutinin-esterase

Chain C:



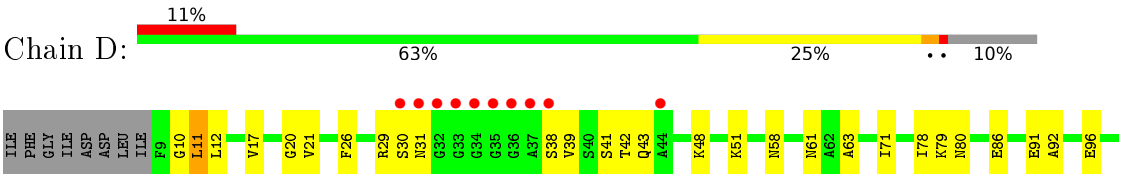
- Molecule 2: Hemagglutinin-esterase

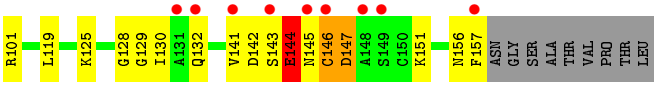
Chain B:



- Molecule 2: Hemagglutinin-esterase

Chain D:




R101
L119
K125
G128
G129
I130
A131
Q132
V141
D142
S143
E144
N145
G146
D147
A148
S149
C150
K151
N156
F157
ASN
GLY
SER
ALA
THR
VAL
PRO
THR
LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	164.90Å 164.90Å 164.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.60 – 2.20 47.60 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (47.60-2.20) 99.6 (47.60-2.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.30 (at 2.20Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.203 , 0.243 0.204 , 0.242	Depositor DCC
R_{free} test set	3795 reflections (5.04%)	DCC
Wilson B-factor (Å^2)	33.4	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 32.9	EDS
Estimated twinning fraction	0.021 for l,-k,h	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 75446 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9829	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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: GAL, BMA, NAG, 5N6, 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	0.28	0/3315	0.46	0/4494
1	C	0.29	0/3315	0.49	1/4494 (0.0%)
2	B	0.29	0/1127	0.46	0/1515
2	D	0.32	0/1127	0.49	0/1515
All	All	0.29	0/8884	0.47	1/12018 (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
2	D	0	1
All	All	0	2

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	198	GLN	CA-CB-CG	5.78	126.12	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	244	ASP	Peptide
2	D	144	GLU	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3240	0	3114	29	0
1	C	3240	0	3114	34	0
2	B	1110	0	1067	23	0
2	D	1110	0	1067	33	0
3	A	84	0	73	2	0
3	B	42	0	38	1	0
3	C	84	0	74	0	0
3	D	42	0	38	2	0
4	A	22	0	17	1	0
4	C	11	0	9	0	0
5	A	66	0	57	2	0
5	C	33	0	28	0	0
6	A	23	0	0	1	0
6	C	23	0	0	1	0
7	A	12	0	11	0	0
7	C	12	0	11	1	0
8	A	264	0	0	9	1
8	B	68	0	0	2	0
8	C	304	0	0	4	0
8	D	39	0	0	3	0
All	All	9829	0	8718	109	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 6.

All (109) 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)
6:A:715:5N6:OBL	6:A:715:5N6:OBW	1.89	0.90
3:A:708:NAG:O3	4:A:709:BMA:H2	1.81	0.81
2:D:63:ALA:O	8:D:801:HOH:O	2.02	0.76
1:A:415:THR:O	2:B:101:ARG:NH1	2.19	0.75
3:B:702:NAG:O4	8:B:801:HOH:O	2.07	0.73
1:A:33:MET:HE1	2:B:101:ARG:HB2	1.73	0.69
2:B:48:LYS:NZ	2:B:52:ASP:OD2	2.24	0.68

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:388:GLU:OE1	8:C:802:HOH:O	2.12	0.68
1:C:188:GLU:OE1	8:C:801:HOH:O	2.12	0.68
2:B:126:ARG:NE	2:B:153:TYR:O	2.27	0.66
2:B:96:GLU:OE2	8:B:802:HOH:O	2.14	0.66
1:A:16:SER:OG	8:A:802:HOH:O	2.14	0.65
1:A:144:GLU:OE1	8:A:803:HOH:O	2.14	0.64
1:A:333:GLU:OE1	8:A:804:HOH:O	2.15	0.63
2:D:142:ASP:O	2:D:144:GLU:N	2.33	0.61
2:D:141:VAL:HG23	2:D:145:ASN:HD21	1.65	0.61
2:B:142:ASP:O	2:B:144:GLU:N	2.25	0.61
1:C:371:LEU:N	1:C:372:GLY:HA3	2.16	0.60
2:D:30:SER:OG	2:D:31:ASN:N	2.31	0.60
2:D:41:SER:OG	2:D:42:THR:N	2.35	0.60
1:C:25:VAL:HG13	1:C:421:ILE:HG23	1.84	0.60
2:D:130:ILE:CG2	2:D:145:ASN:HD22	2.16	0.58
1:A:114:GLU:OE2	1:A:139:TYR:OH	2.11	0.58
5:A:712:MAN:O2	8:A:805:HOH:O	2.17	0.58
1:A:361:ASP:O	1:A:365:ARG:HG3	2.04	0.58
1:C:147:ARG:NH1	8:C:814:HOH:O	2.37	0.57
1:A:250:TYR:OH	8:A:801:HOH:O	1.89	0.57
1:A:350:PRO:HD3	1:C:410:SER:HB3	1.86	0.57
2:D:10:GLY:O	2:D:12:LEU:N	2.31	0.56
1:C:409:ASP:OD1	1:C:411:GLY:N	2.31	0.56
1:A:104:GLN:NE2	8:A:814:HOH:O	2.39	0.55
2:D:141:VAL:CG2	2:D:145:ASN:HD21	2.19	0.55
2:B:145:ASN:OD1	2:B:146:CYS:N	2.39	0.54
2:D:78:ILE:HG22	2:D:80:ASN:H	1.73	0.54
2:D:51:LYS:HD3	2:D:51:LYS:N	2.22	0.54
1:A:137:ASP:OD2	8:A:806:HOH:O	2.18	0.53
1:A:244:ASP:OD2	1:A:263:ARG:HG3	2.09	0.53
2:D:79:LYS:NZ	8:D:802:HOH:O	2.22	0.53
2:D:92:ALA:O	2:D:96:GLU:HG3	2.09	0.53
1:A:25:VAL:HG13	1:A:421:ILE:HG23	1.90	0.53
1:A:418:SER:HA	2:B:57:ARG:HD2	1.91	0.53
2:D:29:ARG:O	2:D:38:SER:OG	2.26	0.53
1:C:12:ASN:H	1:C:15:PHE:HD1	1.56	0.53
1:A:88:LEU:N	1:A:114:GLU:OE2	2.39	0.52
2:D:156:ASN:O	2:D:157:PHE:HB2	2.09	0.52
2:D:26:PHE:HD1	2:D:43:GLN:N	2.07	0.52
1:C:412:ARG:NH2	2:D:91:GLU:HA	2.24	0.52
1:C:94:GLY:HA2	8:C:906:HOH:O	2.09	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:28:MET:SD	3:A:713:NAG:H5	2.50	0.51
6:C:711:5N6:CBF	7:C:712:GAL:H4	2.40	0.51
2:D:146:CYS:SG	2:D:151:LYS:N	2.83	0.51
2:D:43:GLN:NE2	8:D:803:HOH:O	2.44	0.51
1:C:52:ILE:HD11	1:C:81:ARG:HG2	1.93	0.50
1:C:33:MET:HE2	1:C:414:PRO:HB2	1.93	0.50
1:A:88:LEU:HD23	1:A:103:TYR:CD2	2.48	0.48
1:C:71:ASP:HA	1:C:365:ARG:HG2	1.94	0.48
2:D:61:ASN:HD22	3:D:703:NAG:H61	1.77	0.48
1:A:395:TYR:CD2	2:B:86:GLU:HG2	2.48	0.48
2:B:10:GLY:O	2:B:12:LEU:N	2.43	0.48
1:C:234:LYS:NZ	1:C:238:ASN:HD21	2.11	0.48
1:A:246:TYR:H	1:A:294:THR:HG23	1.79	0.48
1:C:108:LYS:HD2	1:C:312:GLY:HA2	1.95	0.48
1:C:4:LEU:H	2:D:30:SER:HB3	1.78	0.48
1:C:10:ARG:HE	2:D:17:VAL:HG11	1.79	0.47
2:B:147:ASP:OD1	2:B:148:ALA:N	2.48	0.47
1:C:238:ASN:HA	1:C:241:GLY:O	2.14	0.47
1:C:395:TYR:HD2	2:D:86:GLU:HG2	1.80	0.46
2:D:26:PHE:HD1	2:D:43:GLN:H	1.63	0.46
1:C:395:TYR:CD2	2:D:86:GLU:HG2	2.50	0.46
2:B:44:ALA:HA	2:B:157:PHE:CE2	2.51	0.46
1:C:370:GLY:C	1:C:372:GLY:HA3	2.36	0.46
1:A:61:LEU:HA	1:A:66:PHE:CD1	2.51	0.46
1:C:198:GLN:HE21	1:C:198:GLN:HB3	1.27	0.46
2:B:146:CYS:SG	2:B:151:LYS:N	2.89	0.45
1:C:33:MET:HE1	2:D:101:ARG:HB2	1.97	0.45
1:A:3:GLU:N	2:B:141:VAL:O	2.50	0.45
2:B:142:ASP:C	2:B:144:GLU:H	2.13	0.45
2:B:128:GLY:HA2	2:B:129:GLY:HA2	1.61	0.45
1:C:234:LYS:HZ3	1:C:238:ASN:HD21	1.66	0.44
2:B:118:GLU:O	2:B:122:GLN:HG3	2.17	0.44
2:D:48:LYS:HD2	2:D:119:LEU:HD13	1.98	0.44
5:A:706:MAN:O6	8:A:807:HOH:O	2.21	0.44
1:C:244:ASP:HB3	1:C:246:TYR:CD2	2.53	0.43
1:A:206:ALA:HB2	1:A:298:LEU:HD21	2.01	0.43
1:C:403:ARG:HB3	2:D:71:ILE:HB	2.00	0.43
1:C:166:THR:OG1	1:C:188:GLU:OE2	2.32	0.43
1:C:191:VAL:HB	1:C:296:ARG:HD3	2.01	0.43
2:D:145:ASN:CG	2:D:146:CYS:HB3	2.38	0.43
1:A:71:ASP:HA	1:A:365:ARG:HG2	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:131:ALA:HB1	2:B:139:TRP:CD1	2.54	0.43
1:C:61:LEU:HA	1:C:66:PHE:CD1	2.54	0.43
1:C:398:PRO:HB2	2:D:78:ILE:O	2.19	0.43
2:B:145:ASN:OD1	2:B:146:CYS:HB2	2.19	0.43
1:C:33:MET:CE	2:D:101:ARG:HB2	2.49	0.42
2:D:128:GLY:HA2	2:D:129:GLY:HA2	1.70	0.42
1:C:232:GLU:OE1	1:C:236:LYS:NZ	2.31	0.42
1:C:247:GLN:O	1:C:259:GLY:HA3	2.20	0.42
1:A:187:ALA:HA	1:A:296:ARG:O	2.19	0.42
2:D:58:ASN:OD1	3:D:703:NAG:H2	2.20	0.42
1:C:361:ASP:O	1:C:365:ARG:HG3	2.20	0.41
1:A:183:PRO:HB3	8:A:834:HOH:O	2.20	0.41
1:A:210:LEU:HB3	1:A:222:HIS:CD2	2.56	0.41
1:A:244:ASP:OD2	1:A:263:ARG:N	2.54	0.41
1:A:354:GLU:OE2	1:A:365:ARG:NH2	2.47	0.41
2:B:34:GLY:HA2	2:B:35:GLY:HA2	1.68	0.41
2:D:125:LYS:HE2	2:D:125:LYS:HB3	1.76	0.41
2:B:147:ASP:OD1	2:B:149:SER:N	2.54	0.40
1:A:400:LYS:NZ	2:B:86:GLU:OE2	2.53	0.40
2:B:93:LEU:HA	2:B:93:LEU:HD12	1.90	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)
8:A:1061:HOH:O	8:A:1063:HOH:O[9_555]	2.04	0.16

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	420/427 (98%)	403 (96%)	17 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	420/427 (98%)	398 (95%)	20 (5%)	2 (0%)	34 35
2	B	147/166 (89%)	122 (83%)	15 (10%)	10 (7%)	1 0
2	D	147/166 (89%)	124 (84%)	17 (12%)	6 (4%)	3 1
All	All	1134/1186 (96%)	1047 (92%)	69 (6%)	18 (2%)	12 8

All (18) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	147	ASP
1	C	245	TYR
2	D	11	LEU
2	D	143	SER
2	B	11	LEU
2	B	39	VAL
2	B	143	SER
1	C	355	ALA
2	D	147	ASP
2	B	41	SER
2	B	37	ALA
2	B	43	GLN
2	D	21	VAL
2	B	21	VAL
2	B	132	GLN
2	D	132	GLN
2	D	20	GLY
2	B	20	GLY

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	357/358 (100%)	354 (99%)	3 (1%)	86 93
1	C	357/358 (100%)	354 (99%)	3 (1%)	86 93
2	B	109/123 (89%)	106 (97%)	3 (3%)	51 63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	D	109/123 (89%)	104 (95%)	5 (5%)	33 40
All	All	932/962 (97%)	918 (98%)	14 (2%)	72 84

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	114	GLU
1	A	158	LEU
1	A	232	GLU
2	B	11	LEU
2	B	144	GLU
2	B	157	PHE
1	C	158	LEU
1	C	198	GLN
1	C	306	CYS
2	D	11	LEU
2	D	39	VAL
2	D	144	GLU
2	D	146	CYS
2	D	147	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	104	GLN
2	B	73	HIS
1	C	198	GLN
2	D	61	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 (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

34 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	701	1,3	14,14,15	0.35	0	15,19,21	0.34	0
3	NAG	A	702	3,4	14,14,15	0.55	0	15,19,21	0.47	0
4	BMA	A	703	3,5	11,11,12	0.74	0	15,15,17	1.22	2 (13%)
5	MAN	A	704	5,4	11,11,12	0.89	1 (9%)	15,15,17	1.05	2 (13%)
5	MAN	A	705	5	11,11,12	0.85	0	15,15,17	1.12	1 (6%)
5	MAN	A	706	5	11,11,12	0.94	0	15,15,17	1.10	0
3	NAG	A	707	1,3	14,14,15	0.28	0	15,19,21	0.29	0
3	NAG	A	708	3,4	14,14,15	0.20	0	15,19,21	0.83	0
4	BMA	A	709	3,5	11,11,12	1.33	2 (18%)	15,15,17	1.47	2 (13%)
5	MAN	A	710	5,4	11,11,12	0.87	1 (9%)	15,15,17	1.98	4 (26%)
5	MAN	A	711	5	11,11,12	0.88	1 (9%)	15,15,17	0.94	1 (6%)
5	MAN	A	712	4	11,11,12	1.19	1 (9%)	15,15,17	0.88	0
3	NAG	A	713	1,3	14,14,15	0.29	0	15,19,21	0.41	0
3	NAG	A	714	3	14,14,15	0.33	0	15,19,21	0.49	0
6	5N6	A	715	7	20,23,24	1.91	7 (35%)	20,32,35	1.45	4 (20%)
7	GAL	A	716	6	12,12,12	1.14	1 (8%)	17,17,17	1.26	2 (11%)
3	NAG	B	701	3,2	14,14,15	0.47	0	15,19,21	0.40	0
3	NAG	B	702	3	14,14,15	0.21	0	15,19,21	0.60	0
3	NAG	B	703	2	14,14,15	0.26	0	15,19,21	0.52	0
3	NAG	C	701	1,3	14,14,15	0.38	0	15,19,21	0.32	0
3	NAG	C	702	3,4	14,14,15	0.66	1 (7%)	15,19,21	0.47	0
4	BMA	C	703	3,5	11,11,12	0.77	0	15,15,17	1.06	1 (6%)
5	MAN	C	704	5,4	11,11,12	0.91	0	15,15,17	1.29	1 (6%)
5	MAN	C	705	5	11,11,12	0.93	1 (9%)	15,15,17	0.94	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	MAN	C	706	5	11,11,12	0.81	0	15,15,17	0.89	0
3	NAG	C	707	1,3	14,14,15	0.31	0	15,19,21	0.53	0
3	NAG	C	708	3	14,14,15	0.20	0	15,19,21	0.58	0
3	NAG	C	709	1,3	14,14,15	0.39	0	15,19,21	0.56	0
3	NAG	C	710	3	14,14,15	0.33	0	15,19,21	0.53	0
6	5N6	C	711	7	20,23,24	2.25	11 (55%)	20,32,35	1.32	3 (15%)
7	GAL	C	712	6	12,12,12	1.43	3 (25%)	17,17,17	1.23	2 (11%)
3	NAG	D	701	3,2	14,14,15	0.41	0	15,19,21	0.48	0
3	NAG	D	702	3	14,14,15	0.30	0	15,19,21	0.58	0
3	NAG	D	703	2	14,14,15	0.50	0	15,19,21	1.17	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. 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	701	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	702	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	703	3,5	-	0/2/19/22	0/1/1/1
5	MAN	A	704	5,4	-	0/2/19/22	0/1/1/1
5	MAN	A	705	5	-	0/2/19/22	0/1/1/1
5	MAN	A	706	5	-	0/2/19/22	0/1/1/1
3	NAG	A	707	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	708	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	709	3,5	-	0/2/19/22	0/1/1/1
5	MAN	A	710	5,4	-	0/2/19/22	0/1/1/1
5	MAN	A	711	5	-	0/2/19/22	0/1/1/1
5	MAN	A	712	4	-	0/2/19/22	0/1/1/1
3	NAG	A	713	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	714	3	-	0/6/23/26	0/1/1/1
6	5N6	A	715	7	-	0/17/37/41	0/1/1/1
7	GAL	A	716	6	-	0/2/22/22	0/1/1/1
3	NAG	B	701	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	702	3	-	0/6/23/26	0/1/1/1
3	NAG	B	703	2	-	0/6/23/26	0/1/1/1
3	NAG	C	701	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	702	3,4	-	0/6/23/26	0/1/1/1
4	BMA	C	703	3,5	-	0/2/19/22	0/1/1/1
5	MAN	C	704	5,4	-	0/2/19/22	0/1/1/1
5	MAN	C	705	5	-	0/2/19/22	0/1/1/1
5	MAN	C	706	5	-	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	707	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	708	3	-	0/6/23/26	0/1/1/1
3	NAG	C	709	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	710	3	-	0/6/23/26	0/1/1/1
6	5N6	C	711	7	-	0/17/37/41	0/1/1/1
7	GAL	C	712	6	-	0/2/22/22	0/1/1/1
3	NAG	D	701	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	702	3	-	0/6/23/26	0/1/1/1
3	NAG	D	703	2	-	0/6/23/26	0/1/1/1

All (30) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	711	5N6	OBY-CBB	-4.74	1.32	1.43
6	C	711	5N6	CAZ-CBC	-3.66	1.48	1.52
5	A	712	MAN	O5-C1	-3.51	1.38	1.43
6	C	711	5N6	CBB-CBA	-3.50	1.49	1.53
6	A	715	5N6	OBK-CAX	-3.41	1.37	1.45
6	A	715	5N6	CBD-CBB	-3.23	1.47	1.52
6	A	715	5N6	CBD-CBE	-3.21	1.47	1.52
6	A	715	5N6	CBC-CBA	-2.99	1.47	1.53
6	C	711	5N6	CAY-CAZ	-2.95	1.47	1.53
6	A	715	5N6	OBY-CBB	-2.85	1.37	1.43
6	C	711	5N6	OBZ-CBC	-2.75	1.39	1.43
4	A	709	BMA	C2-C3	-2.70	1.48	1.52
6	C	711	5N6	CBD-CBB	-2.58	1.48	1.52
7	C	712	GAL	C4-C5	-2.55	1.47	1.53
6	A	715	5N6	OBX-CAY	-2.54	1.37	1.43
6	C	711	5N6	CBC-CBA	-2.51	1.48	1.53
7	C	712	GAL	O4-C4	-2.43	1.37	1.43
5	A	711	MAN	O5-C1	-2.35	1.39	1.43
7	C	712	GAL	O2-C2	-2.25	1.37	1.43
6	C	711	5N6	CAX-CAY	-2.24	1.48	1.51
5	A	704	MAN	O5-C1	-2.22	1.40	1.43
6	C	711	5N6	OBX-CAY	-2.22	1.38	1.43
6	C	711	5N6	OBK-CAX	-2.14	1.40	1.45
5	C	705	MAN	O5-C1	-2.11	1.40	1.43
6	C	711	5N6	OBZ-CBE	-2.07	1.38	1.43
3	C	702	NAG	O5-C1	-2.05	1.40	1.43
7	A	716	GAL	C3-C2	-2.04	1.47	1.52
6	A	715	5N6	CAX-CAY	-2.01	1.48	1.51
5	A	710	MAN	C1-C2	2.40	1.58	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	709	BMA	O5-C5	2.43	1.48	1.43

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	709	BMA	C2-C3-C4	-3.96	104.14	111.05
6	A	715	5N6	CAX-OBK-CAG	-3.81	107.56	117.12
6	C	711	5N6	CBC-CBA-NAB	-3.08	105.79	111.06
6	A	715	5N6	OBX-CAY-CAX	-2.88	103.82	110.09
5	A	711	MAN	O2-C2-C3	-2.55	105.05	110.19
4	A	709	BMA	O2-C2-C3	-2.45	105.25	110.19
5	A	710	MAN	O2-C2-C3	-2.36	105.43	110.19
7	C	712	GAL	O5-C5-C6	-2.33	100.36	106.38
6	A	715	5N6	CBB-CBA-NAB	-2.31	105.06	110.31
4	A	703	BMA	O2-C2-C3	-2.28	105.58	110.19
5	A	704	MAN	O2-C2-C3	-2.28	105.58	110.19
7	C	712	GAL	C1-C2-C3	-2.25	106.99	110.68
6	C	711	5N6	OBL-CAI-CAH	-2.18	118.06	122.07
7	A	716	GAL	O4-C4-C5	2.02	114.56	109.23
6	C	711	5N6	OBZ-CBC-CBA	2.05	111.84	108.48
6	A	715	5N6	CBD-CBB-CBA	2.08	113.80	111.47
7	A	716	GAL	C4-C3-C2	2.08	114.62	110.79
5	A	704	MAN	C1-O5-C5	2.26	115.47	112.14
5	A	710	MAN	C1-C2-C3	2.44	112.51	109.55
5	A	705	MAN	C1-O5-C5	2.56	115.91	112.14
4	A	703	BMA	C1-O5-C5	2.78	116.23	112.14
5	A	710	MAN	O5-C1-C2	2.79	115.35	110.89
4	C	703	BMA	C1-O5-C5	2.90	116.40	112.14
3	D	703	NAG	C1-O5-C5	3.39	117.12	112.14
5	C	704	MAN	C1-O5-C5	3.55	117.36	112.14
5	A	710	MAN	C1-O5-C5	5.50	120.23	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	706	MAN	1	0
3	A	708	NAG	1	0
4	A	709	BMA	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	712	MAN	1	0
3	A	713	NAG	1	0
6	A	715	5N6	1	0
3	B	702	NAG	1	0
6	C	711	5N6	1	0
7	C	712	GAL	1	0
3	D	703	NAG	2	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	424/427 (99%)	-0.18	20 (4%) 35 34	24, 35, 85, 131	0
1	C	424/427 (99%)	-0.44	4 (0%) 85 85	20, 33, 58, 89	0
2	B	149/166 (89%)	0.72	19 (12%) 5 4	22, 48, 128, 201	0
2	D	149/166 (89%)	0.67	19 (12%) 5 4	26, 49, 141, 173	0
All	All	1146/1186 (96%)	-0.05	62 (5%) 29 29	20, 37, 91, 201	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	34	GLY	23.4
2	B	32	GLY	18.8
2	D	33	GLY	15.3
2	B	37	ALA	14.2
2	D	32	GLY	12.1
2	D	34	GLY	7.9
2	D	37	ALA	7.2
2	B	36	GLY	6.6
2	D	35	GLY	6.4
2	D	31	ASN	6.1
2	B	33	GLY	6.0
1	A	191	VAL	6.0
1	A	193	PRO	5.6
2	B	44	ALA	5.3
2	D	38	SER	4.8
2	D	145	ASN	4.6
2	B	145	ASN	4.5
1	A	237	PHE	4.4
2	D	149	SER	4.4
2	B	45	GLY	4.4
2	D	30	SER	4.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	157	PHE	4.3
1	A	236	LYS	4.2
2	D	36	GLY	4.1
1	A	265	ALA	3.8
2	B	38	SER	3.7
2	D	143	SER	3.7
1	A	273	ALA	3.6
2	B	35	GLY	3.4
1	A	197	PRO	3.4
2	B	141	VAL	3.3
2	D	44	ALA	3.3
1	A	241	GLY	3.3
2	D	141	VAL	3.3
1	A	274	GLY	3.2
2	B	131	ALA	3.2
1	C	270	ILE	3.2
2	D	148	ALA	3.2
2	D	131	ALA	3.1
1	A	127	PHE	3.1
2	B	39	VAL	3.1
2	B	43	GLN	3.1
1	A	296	ARG	3.0
1	A	297	PHE	2.8
1	C	243	GLY	2.6
1	C	244	ASP	2.5
1	A	235	ALA	2.5
1	A	187	ALA	2.4
1	A	244	ASP	2.4
2	D	157	PHE	2.4
1	A	233	ASN	2.4
1	A	238	ASN	2.4
2	B	30	SER	2.3
2	B	16	PHE	2.3
1	A	275	VAL	2.3
1	A	267	TYR	2.3
2	D	132	GLN	2.3
2	D	146	CYS	2.2
1	C	236	LYS	2.2
2	B	148	ALA	2.1
1	A	188	GLU	2.0
2	B	149	SER	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
3	NAG	B	701	14/15	0.95	0.13	1.24	27,33,35,37	0
3	NAG	B	702	14/15	0.95	0.12	1.23	36,40,47,53	0
6	5N6	C	711	23/24	0.93	0.14	0.43	43,52,60,62	0
3	NAG	D	702	14/15	0.94	0.12	0.35	34,40,45,47	0
5	MAN	C	706	11/12	0.91	0.10	0.07	31,34,38,41	0
3	NAG	C	709	14/15	0.94	0.13	-0.01	43,48,63,65	0
5	MAN	A	706	11/12	0.93	0.10	-0.02	30,34,37,40	0
3	NAG	D	701	14/15	0.96	0.10	-0.22	27,30,34,35	0
3	NAG	A	713	14/15	0.89	0.13	-0.49	53,56,62,63	0
6	5N6	A	715	23/24	0.88	0.18	-0.51	71,81,85,87	0
3	NAG	C	702	14/15	0.95	0.10	-	24,29,35,37	0
3	NAG	A	702	14/15	0.94	0.12	-	28,35,38,40	0
3	NAG	D	703	14/15	0.60	0.44	-	71,79,82,83	0
3	NAG	C	707	14/15	0.96	0.10	-	23,30,34,40	0
3	NAG	C	710	14/15	0.89	0.13	-	45,50,60,72	0
3	NAG	A	707	14/15	0.95	0.09	-	27,33,36,36	0
3	NAG	A	708	14/15	0.93	0.12	-	29,35,40,44	0
3	NAG	A	701	14/15	0.97	0.08	-	33,36,44,49	0
5	MAN	A	705	11/12	0.87	0.10	-	31,36,39,47	0
5	MAN	C	704	11/12	0.97	0.07	-	23,30,35,40	0
5	MAN	A	712	11/12	0.96	0.10	-	36,39,44,47	0
3	NAG	A	714	14/15	0.85	0.16	-	56,63,76,85	0
3	NAG	C	701	14/15	0.95	0.09	-	30,34,36,39	0
4	BMA	A	703	11/12	0.94	0.10	-	29,32,35,38	0
7	GAL	C	712	12/12	0.78	0.25	-	68,88,102,113	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	MAN	A	704	11/12	0.98	0.10	-	26,30,33,38	0
3	NAG	B	703	14/15	0.81	0.26	-	51,59,65,66	0
4	BMA	A	709	11/12	0.83	0.23	-	39,43,52,56	0
5	MAN	A	710	11/12	0.94	0.09	-	38,41,47,49	0
4	BMA	C	703	11/12	0.94	0.09	-	29,33,36,41	0
7	GAL	A	716	12/12	0.69	0.30	-	96,109,121,131	0
3	NAG	C	708	14/15	0.86	0.16	-	39,47,57,57	0
5	MAN	C	705	11/12	0.93	0.10	-	36,40,45,52	0
5	MAN	A	711	11/12	0.93	0.19	-	43,47,54,57	0

6.5 Other polymers

There are no such residues in this entry.