



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 9, 2017 – 04:02 PM EST

PDB ID : 4N63
Title : Crystal structure of hemagglutinin from an H7N9 influenza virus in complex with an O-linked glycan receptor
Authors : Xu, R.; Wilson, I.A.
Deposited on : unknown
Resolution : 2.75 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

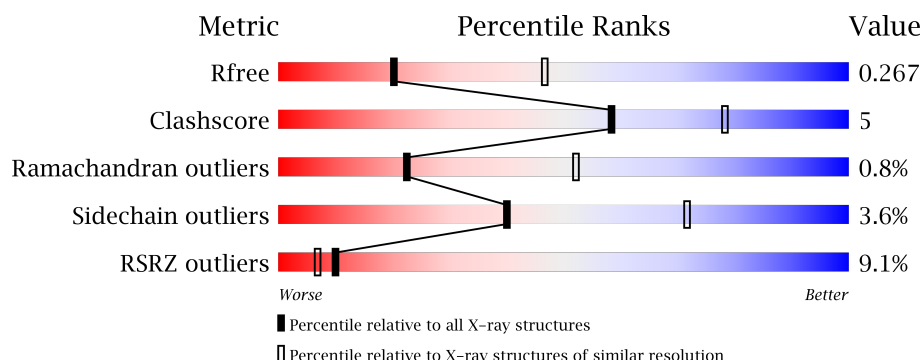
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3666 (2.80-2.72)
Clashscore	112137	4174 (2.80-2.72)
Ramachandran outliers	110173	4103 (2.80-2.72)
Sidechain outliers	110143	4106 (2.80-2.72)
RSRZ outliers	101464	3697 (2.80-2.72)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	321	<div> <div>13%</div> <div>80%</div> <div>18%</div> <div>..</div> </div>
1	C	321	<div> <div>13%</div> <div>82%</div> <div>16%</div> <div>..</div> </div>
2	B	183	<div> <div>82%</div> <div>10%</div> <div>8%</div> </div>
2	D	183	<div> <div>24%</div> <div>85%</div> <div>8%</div> <div>8%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	B	201	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7824 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	318	Total	C	N	O	S	0	0	0
			2427	1509	438	465	15			
1	C	318	Total	C	N	O	S	0	0	0
			2427	1509	438	465	15			

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total	C	N	O	S	0	0	0
			1373	845	241	280	7			
2	D	169	Total	C	N	O	S	0	0	0
			1373	845	241	280	7			

There are 14 discrepancies between the modelled and reference sequences:

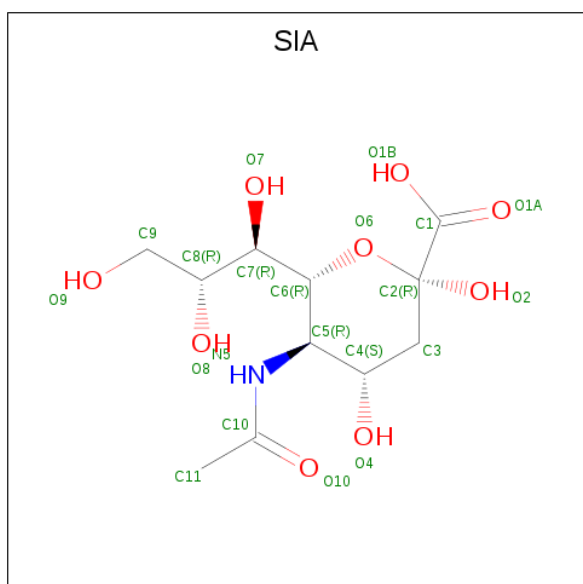
Chain	Residue	Modelled	Actual	Comment	Reference
B	177	SER	-	EXPRESSION TAG	UNP R4NN21
B	178	GLY	-	EXPRESSION TAG	UNP R4NN21
B	179	ARG	-	EXPRESSION TAG	UNP R4NN21
B	180	LEU	-	EXPRESSION TAG	UNP R4NN21
B	181	VAL	-	EXPRESSION TAG	UNP R4NN21
B	182	PRO	-	EXPRESSION TAG	UNP R4NN21
B	183	ARG	-	EXPRESSION TAG	UNP R4NN21
D	177	SER	-	EXPRESSION TAG	UNP R4NN21
D	178	GLY	-	EXPRESSION TAG	UNP R4NN21
D	179	ARG	-	EXPRESSION TAG	UNP R4NN21
D	180	LEU	-	EXPRESSION TAG	UNP R4NN21
D	181	VAL	-	EXPRESSION TAG	UNP R4NN21
D	182	PRO	-	EXPRESSION TAG	UNP R4NN21
D	183	ARG	-	EXPRESSION TAG	UNP R4NN21

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



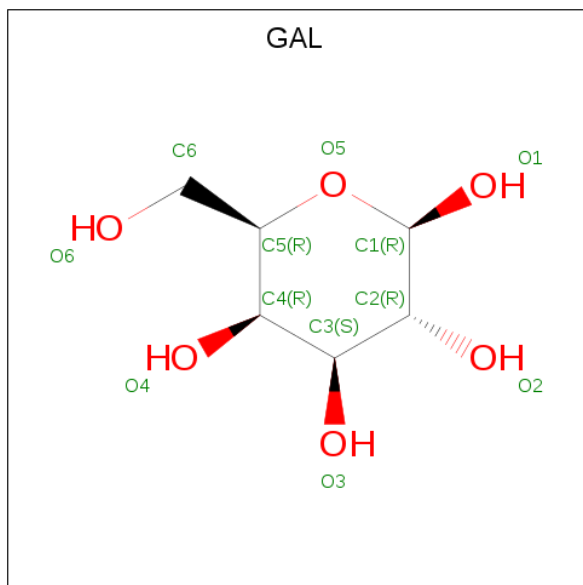
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		

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



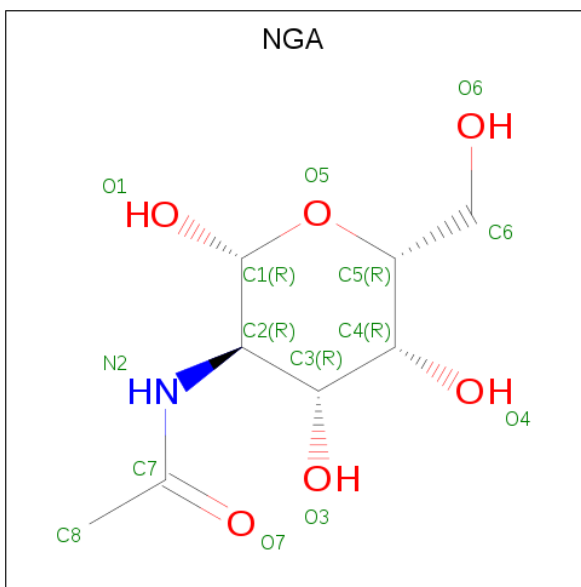
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			20	11	1	8		
4	C	1	Total	C	N	O	0	0
			20	11	1	8		

- 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	0	0
			11	6	5		
5	C	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is N-ACETYL-D-GALACTOSAMINE (three-letter code: NGA) (formula: $C_8H_{15}NO_6$).



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

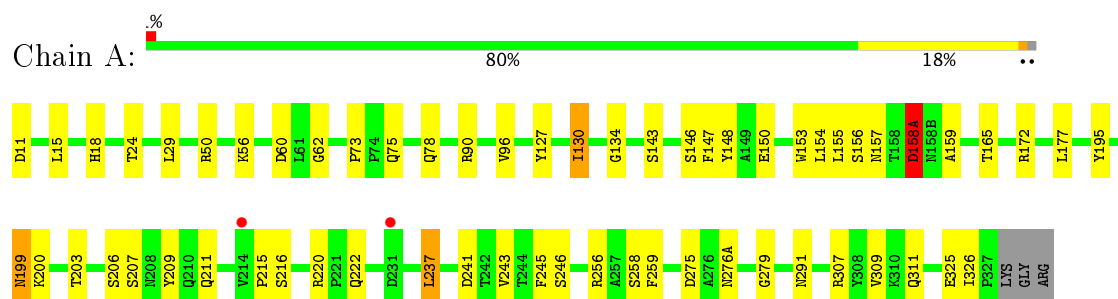
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	28	Total	O	0	0
			28	28		
7	B	24	Total	O	0	0
			24	24		
7	C	7	Total	O	0	0
			7	7		
7	D	3	Total	O	0	0
			3	3		

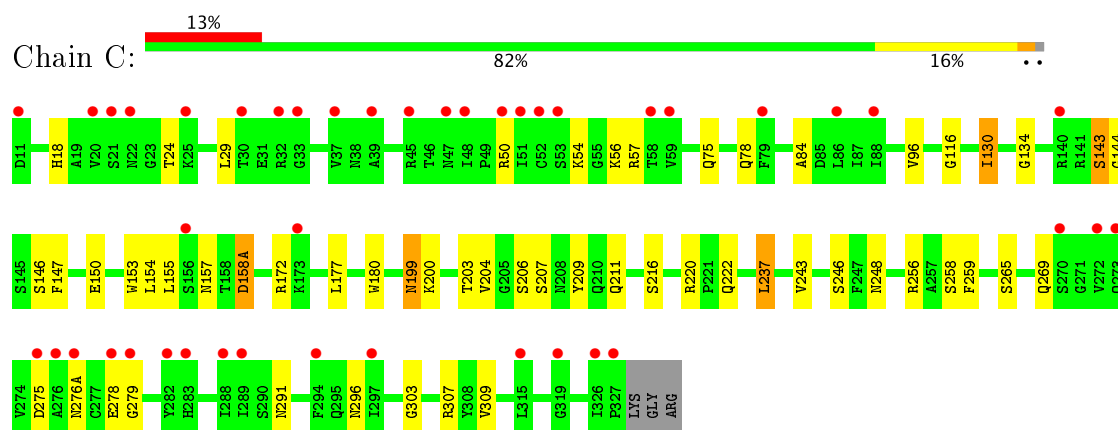
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

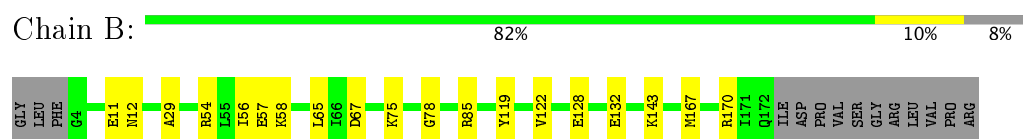
• Molecule 1: Hemagglutinin HA1



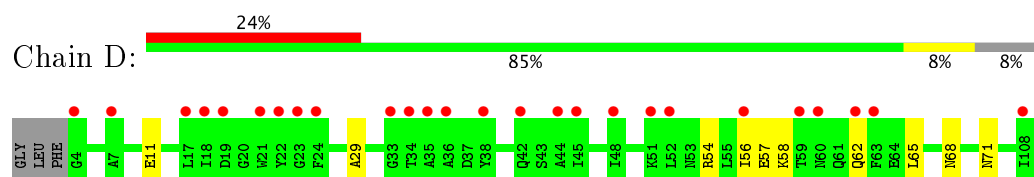
• Molecule 1: Hemagglutinin HA1

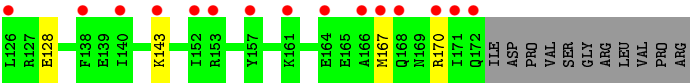


• Molecule 2: Hemagglutinin HA2



• Molecule 2: Hemagglutinin HA2





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	153.61Å 153.61Å 153.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.60 – 2.75 42.60 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.2 (42.60-2.75) 98.9 (42.60-2.75)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.8.2 _1309	Depositor
R, R_{free}	0.213 , 0.269 0.212 , 0.267	Depositor DCC
R_{free} test set	1580 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	64.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 54.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.035 for l,-k,h	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7824	wwPDB-VP
Average B, all atoms (Å ²)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.06% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SIA, GAL, NAG, NGA

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	0/2474	0.42	0/3345
1	C	0.23	0/2474	0.41	0/3345
2	B	0.24	0/1396	0.40	0/1881
2	D	0.23	0/1396	0.38	0/1881
All	All	0.23	0/7740	0.40	0/10452

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2427	0	2388	38	0
1	C	2427	0	2389	33	0
2	B	1373	0	1273	16	0
2	D	1373	0	1274	10	0
3	A	42	0	37	1	0
3	B	14	0	13	2	0
3	C	14	0	12	1	0
4	A	20	0	17	1	0
4	C	20	0	17	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	11	0	9	0	0
5	C	11	0	9	0	0
6	A	15	0	14	0	0
6	C	15	0	14	0	0
7	A	28	0	0	1	0
7	B	24	0	0	0	0
7	C	7	0	0	0	0
7	D	3	0	0	0	0
All	All	7824	0	7466	83	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 5.

All (83) 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:291:ASN:OD1	2:D:58:LYS:NZ	1.93	1.01
2:B:128:GLU:HG3	2:B:170:ARG:HH12	1.36	0.90
1:C:56:LYS:NZ	1:C:279:GLY:O	2.13	0.81
2:D:128:GLU:HG3	2:D:170:ARG:HH12	1.47	0.79
1:A:56:LYS:NZ	1:A:279:GLY:O	2.16	0.78
1:A:222:GLN:NE2	3:A:405:NAG:O7	2.17	0.77
2:B:128:GLU:HG3	2:B:170:ARG:NH1	2.05	0.71
2:D:128:GLU:HG3	2:D:170:ARG:NH1	2.05	0.71
1:C:155:LEU:HD21	4:C:401:SIA:H111	1.73	0.71
1:C:222:GLN:NE2	3:C:403:NAG:O7	2.24	0.69
1:C:203:THR:HG1	1:C:246:SER:HG	1.40	0.68
2:B:29:ALA:H	2:B:143:LYS:HZ1	1.42	0.67
2:D:29:ALA:H	2:D:143:LYS:HZ1	1.44	0.63
1:C:203:THR:OG1	1:C:246:SER:OG	2.17	0.62
1:A:177:LEU:HB3	1:A:258:SER:HB2	1.80	0.62
1:A:275:ASP:OD1	1:A:276(A):ASN:ND2	2.33	0.62
1:C:275:ASP:OD1	1:C:276(A):ASN:ND2	2.33	0.61
1:A:155:LEU:HD21	4:A:403:SIA:H111	1.83	0.60
1:A:146:SER:OG	1:A:147:PHE:N	2.34	0.59
1:A:291:ASN:OD1	2:B:58:LYS:NZ	2.35	0.58
2:B:75:LYS:HD3	3:B:201:NAG:H81	1.84	0.58
1:C:130:ILE:HD11	1:C:154:LEU:HB3	1.84	0.58
1:C:134:GLY:HA3	1:C:153:TRP:HB3	1.85	0.58
1:A:11:ASP:N	7:A:507:HOH:O	2.37	0.57
1:A:134:GLY:HA3	1:A:153:TRP:HB3	1.85	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:146:SER:OG	1:C:147:PHE:N	2.36	0.57
1:C:199:ASN:N	1:C:199:ASN:OD1	2.37	0.56
1:A:130:ILE:HD11	1:A:154:LEU:HB3	1.87	0.56
1:A:50:ARG:NE	1:A:275:ASP:OD2	2.38	0.56
2:B:119:TYR:OH	2:B:132:GLU:OE2	2.13	0.55
1:C:206:SER:HB3	1:C:209:TYR:HB3	1.88	0.55
1:C:206:SER:OG	1:C:207:SER:N	2.39	0.55
1:A:150:GLU:OE1	1:A:256:ARG:HD3	2.08	0.54
1:C:291:ASN:HA	2:D:58:LYS:NZ	2.23	0.53
1:A:206:SER:OG	1:A:207:SER:N	2.41	0.53
1:A:325:GLU:HG3	2:B:12:ASN:HD22	1.75	0.52
1:A:237:LEU:HG	1:A:243:VAL:HG13	1.92	0.52
1:C:216:SER:O	1:C:220:ARG:NH2	2.38	0.52
1:A:203:THR:OG1	1:A:246:SER:OG	2.22	0.51
1:A:325:GLU:HA	2:B:12:ASN:HB2	1.92	0.51
2:B:128:GLU:O	2:B:170:ARG:NH1	2.37	0.51
1:C:303:GLY:HA2	2:D:62:GLN:HG3	1.93	0.50
1:C:150:GLU:OE1	1:C:256:ARG:HD3	2.10	0.50
1:A:326:ILE:HG12	2:B:12:ASN:HB3	1.94	0.50
1:A:291:ASN:HA	2:B:58:LYS:HZ2	1.77	0.50
2:B:78:GLY:HA3	3:B:201:NAG:H82	1.95	0.48
1:A:237:LEU:HD22	1:A:241:ASP:HB3	1.96	0.48
1:C:180:TRP:CE2	1:C:204:VAL:HG21	2.49	0.48
1:A:172:ARG:HD3	1:A:259:PHE:CZ	2.49	0.48
1:C:237:LEU:HG	1:C:243:VAL:HG13	1.96	0.47
1:A:216:SER:O	1:A:220:ARG:NH2	2.37	0.47
1:C:296:ASN:ND2	1:C:309:VAL:O	2.42	0.47
1:C:291:ASN:HA	2:D:58:LYS:HZ2	1.79	0.47
1:A:206:SER:HB3	1:A:209:TYR:HB3	1.97	0.46
1:C:200:LYS:HA	1:C:248:ASN:HD21	1.79	0.46
1:C:54:LYS:HD3	1:C:278:GLU:HB2	1.98	0.46
1:A:199:ASN:OD1	1:A:199:ASN:N	2.49	0.45
1:A:291:ASN:HA	2:B:58:LYS:NZ	2.32	0.45
1:C:57:ARG:HG2	1:C:84:ALA:HA	1.98	0.45
1:A:60:ASP:OD1	1:A:90:ARG:NE	2.34	0.45
1:C:172:ARG:HD3	1:C:259:PHE:CZ	2.52	0.44
1:C:177:LEU:HB3	1:C:258:SER:HB2	2.00	0.44
1:A:195:TYR:O	1:A:200:LYS:NZ	2.28	0.43
1:C:307:ARG:HH11	1:C:307:ARG:HG2	1.83	0.43
1:C:116:GLY:HA2	1:C:265:SER:HB3	2.00	0.43
1:A:127:TYR:HB3	1:A:130:ILE:HG21	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:291:ASN:HB3	2:D:56:ILE:HG13	2.01	0.42
1:C:269:GLN:HE22	2:D:68:ASN:CG	2.23	0.42
1:A:130:ILE:HA	1:A:156:SER:O	2.20	0.42
1:A:309:VAL:HB	1:A:311:GLN:OE1	2.19	0.42
1:A:15:LEU:HD11	2:B:122:VAL:HG21	2.01	0.42
2:B:67:ASP:OD2	2:B:85:ARG:NH2	2.43	0.42
1:C:143:SER:OG	1:C:144:GLY:N	2.52	0.42
1:A:73:PRO:HG3	1:A:148:TYR:CE1	2.55	0.41
1:A:307:ARG:HG2	1:A:307:ARG:HH11	1.85	0.41
2:D:68:ASN:HB3	2:D:71:ASN:O	2.20	0.41
1:C:50:ARG:NE	1:C:275:ASP:OD2	2.53	0.41
1:A:158(A):ASP:C	1:A:159:ALA:H	2.23	0.41
1:A:291:ASN:HB3	2:B:56:ILE:HG13	2.01	0.41
1:A:62:GLY:O	1:A:90:ARG:HB2	2.20	0.41
1:A:165:THR:HA	1:A:245:PHE:O	2.21	0.41
1:C:237:LEU:HA	1:C:237:LEU:HD23	1.89	0.40
1:A:200:LYS:HB2	1:A:215:PRO:HG3	2.04	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	316/321 (98%)	293 (93%)	20 (6%)	3 (1%)	20	49
1	C	316/321 (98%)	293 (93%)	20 (6%)	3 (1%)	20	49
2	B	167/183 (91%)	158 (95%)	8 (5%)	1 (1%)	28	59
2	D	167/183 (91%)	157 (94%)	9 (5%)	1 (1%)	28	59
All	All	966/1008 (96%)	901 (93%)	57 (6%)	8 (1%)	22	52

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	157	ASN
2	B	57	GLU
1	C	157	ASN
1	A	158(A)	ASP
1	C	158(A)	ASP
2	D	57	GLU
1	A	143	SER
1	C	143	SER

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	267/269 (99%)	256 (96%)	11 (4%)	35	67
1	C	267/269 (99%)	256 (96%)	11 (4%)	35	67
2	B	145/157 (92%)	141 (97%)	4 (3%)	49	79
2	D	145/157 (92%)	141 (97%)	4 (3%)	49	79
All	All	824/852 (97%)	794 (96%)	30 (4%)	40	71

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	24	THR
1	A	29	LEU
1	A	75	GLN
1	A	78	GLN
1	A	96	VAL
1	A	130	ILE
1	A	158(A)	ASP
1	A	199	ASN
1	A	211	GLN
1	A	237	LEU
2	B	11	GLU
2	B	54	ARG
2	B	65	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	167	MET
1	C	18	HIS
1	C	24	THR
1	C	29	LEU
1	C	75	GLN
1	C	78	GLN
1	C	96	VAL
1	C	130	ILE
1	C	158(A)	ASP
1	C	199	ASN
1	C	211	GLN
1	C	237	LEU
2	D	11	GLU
2	D	54	ARG
2	D	65	LEU
2	D	167	MET

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

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	1,3	14,14,15	0.55	0	15,19,21	0.72	1 (6%)
3	NAG	A	402	3	14,14,15	0.22	0	15,19,21	0.49	0
4	SIA	A	403	5	17,20,21	0.53	0	19,28,31	1.13	2 (10%)
5	GAL	A	404	3,4	11,11,12	1.12	2 (18%)	13,15,17	1.17	1 (7%)
3	NAG	A	405	5,6	14,14,15	0.42	0	15,19,21	0.48	0
6	NGA	A	406	3	15,15,15	0.51	0	21,21,21	1.45	4 (19%)
3	NAG	B	201	2	14,14,15	0.33	0	15,19,21	0.45	0
4	SIA	C	401	5	17,20,21	0.53	0	19,28,31	1.16	1 (5%)
5	GAL	C	402	3,4	11,11,12	1.23	2 (18%)	13,15,17	1.19	2 (15%)
3	NAG	C	403	5,6	14,14,15	0.27	0	15,19,21	0.56	0
6	NGA	C	404	3	15,15,15	0.46	0	21,21,21	1.16	1 (4%)

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	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	402	3	-	0/6/23/26	0/1/1/1
4	SIA	A	403	5	-	0/14/34/38	0/1/1/1
5	GAL	A	404	3,4	-	0/2/19/22	0/1/1/1
3	NAG	A	405	5,6	-	0/6/23/26	0/1/1/1
6	NGA	A	406	3	-	0/6/26/26	0/1/1/1
3	NAG	B	201	2	-	0/6/23/26	0/1/1/1
4	SIA	C	401	5	-	0/14/34/38	0/1/1/1
5	GAL	C	402	3,4	-	0/2/19/22	0/1/1/1
3	NAG	C	403	5,6	-	0/6/23/26	0/1/1/1
6	NGA	C	404	3	-	0/6/26/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	402	GAL	C1-C2	2.08	1.57	1.52
5	A	404	GAL	C1-C2	2.12	1.57	1.52
5	A	404	GAL	C2-C3	2.43	1.55	1.52
5	C	402	GAL	C2-C3	2.83	1.56	1.52

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	401	SIA	C3-C4-C5	-4.09	106.51	111.46
4	A	403	SIA	C3-C4-C5	-3.80	106.87	111.46
6	A	406	NGA	C1-C2-N2	-3.60	106.56	110.73
3	A	401	NAG	O5-C1-C2	-2.15	108.48	111.47
6	A	406	NGA	C4-C3-C2	2.00	113.30	110.33
5	C	402	GAL	C1-O5-C5	2.04	114.98	112.17
6	A	406	NGA	O3-C3-C4	2.13	114.98	110.36
4	A	403	SIA	O6-C2-C3	2.34	114.03	109.82
6	A	406	NGA	O4-C4-C3	2.44	115.66	110.36
5	C	402	GAL	C1-C2-C3	2.64	113.00	109.65
5	A	404	GAL	C1-C2-C3	2.91	113.34	109.65
6	C	404	NGA	O5-C1-C2	3.00	112.53	109.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	403	SIA	1	0
3	A	405	NAG	1	0
3	B	201	NAG	2	0
4	C	401	SIA	1	0
3	C	403	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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	318/321 (99%)	0.00	2 (0%) 89 87	47, 69, 104, 172	0
1	C	318/321 (99%)	0.77	43 (13%) 3 2	70, 115, 173, 227	0
2	B	169/183 (92%)	0.03	0 100 100	42, 61, 106, 172	0
2	D	169/183 (92%)	1.44	44 (26%) 1 0	92, 153, 189, 252	0
All	All	974/1008 (96%)	0.51	89 (9%) 10 7	42, 91, 170, 252	0

All (89) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	60	ASN	8.4
2	D	168	GLN	8.3
1	C	319	GLY	8.0
2	D	36	ALA	7.7
1	C	327	PRO	7.5
2	D	172	GLN	7.2
2	D	7	ALA	6.2
2	D	18	ILE	6.1
2	D	4	GLY	5.9
1	C	51	ILE	5.4
2	D	35	ALA	5.3
2	D	171	ILE	4.9
1	C	289	ILE	4.8
1	C	58	THR	4.6
1	C	37	VAL	4.3
2	D	48	ILE	4.3
1	C	22	ASN	4.3
1	C	11	ASP	4.3
1	C	33	GLY	4.0
1	C	45	ARG	4.0
2	D	19	ASP	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	52	LEU	3.9
2	D	21	TRP	3.8
2	D	45	ILE	3.8
1	C	21	SER	3.7
2	D	170	ARG	3.6
1	C	279	GLY	3.6
2	D	38	TYR	3.6
1	C	294	PHE	3.5
1	C	20	VAL	3.5
2	D	167	MET	3.4
2	D	59	THR	3.4
2	D	17	LEU	3.4
1	C	288	ILE	3.4
1	C	276(A)	ASN	3.2
2	D	33	GLY	3.2
1	C	47	ASN	3.2
2	D	24	PHE	3.0
2	D	56	ILE	3.0
2	D	126	LEU	3.0
2	D	23	GLY	2.9
2	D	166	ALA	2.9
2	D	161	LYS	2.9
2	D	157	TYR	2.9
1	C	48	ILE	2.9
2	D	140	ILE	2.9
1	C	272	VAL	2.9
1	C	276	ALA	2.9
1	C	275	ASP	2.8
1	C	30	THR	2.8
1	C	156	SER	2.8
2	D	119	TYR	2.8
1	C	59	VAL	2.7
1	C	86	LEU	2.7
2	D	34	THR	2.7
1	C	50	ARG	2.7
2	D	118	LEU	2.6
1	C	25	LYS	2.6
2	D	51	LYS	2.6
1	C	297	ILE	2.5
2	D	138	PHE	2.5
1	C	173	LYS	2.5
2	D	44	ALA	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	140	ARG	2.4
1	C	79	PHE	2.4
2	D	22	TYR	2.4
1	C	32	ARG	2.4
2	D	42	GLN	2.3
2	D	152	ILE	2.3
2	D	153	ARG	2.3
1	C	315	LEU	2.3
2	D	125	GLN	2.3
2	D	62	GLN	2.3
2	D	63	PHE	2.3
1	C	273	GLN	2.2
1	A	214	VAL	2.2
1	C	53	SER	2.2
1	C	282	TYR	2.2
1	C	270	SER	2.2
2	D	164	GLU	2.1
1	C	39	ALA	2.1
1	C	278	GLU	2.1
1	C	88	ILE	2.1
1	A	231	ASP	2.1
1	C	326	ILE	2.1
2	D	108	ILE	2.1
2	D	143	LYS	2.0
1	C	283	HIS	2.0
1	C	52	CYS	2.0

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	B	201	14/15	0.78	0.24	2.25	86,108,112,113	0
4	SIA	C	401	20/21	0.93	0.21	0.43	64,99,124,134	0
4	SIA	A	403	20/21	0.95	0.17	-0.44	42,69,90,100	0
5	GAL	A	404	11/12	0.96	0.16	-	61,77,85,86	0
3	NAG	A	402	14/15	0.81	0.41	-	131,148,164,167	0
3	NAG	C	403	14/15	0.76	0.38	-	182,208,213,213	0
6	NGA	A	406	15/15	0.84	0.21	-	107,129,140,149	0
6	NGA	C	404	15/15	0.72	0.26	-	156,164,187,197	0
3	NAG	A	405	14/15	0.92	0.22	-	94,122,133,138	0
5	GAL	C	402	11/12	0.89	0.16	-	96,109,135,151	0
3	NAG	A	401	14/15	0.87	0.20	-	100,109,124,140	0

6.5 Other polymers [i](#)

There are no such residues in this entry.