Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 06:54 PM GMT

PDB ID	:	4N62
Title	:	Crystal structure of hemagglutinin from an H7N9 influenza virus in complex with a sulfated receptor analog
Authors	:	Xu, R.; Wilson, I.A.
Deposited on	:	2013-10-11
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

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

with specific help available everywhere you see the (i) symbol.

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	Percentile Ranks	Value
Rfree		0.270
Clashscore		6
Ramachandran outliers		1.4%
Sidechain outliers		4.4%
RSRZ outliers		18.3%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
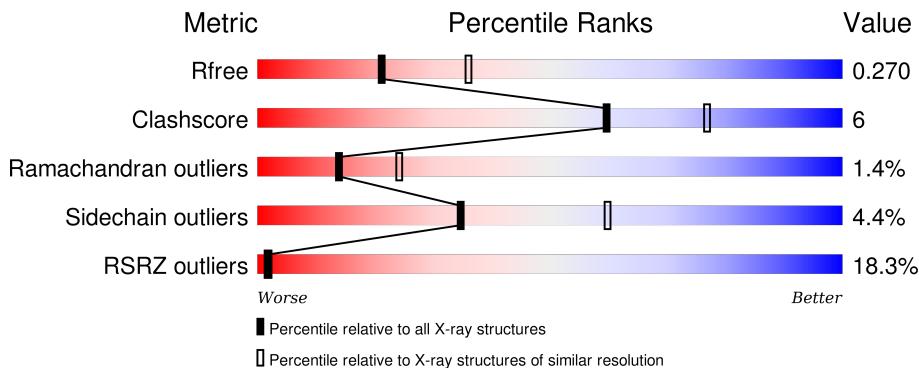
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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. 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	321	2% 81% 17% ••
1	C	321	30% 78% 19% ••
2	B	183	% 84% 7% • 8%
2	D	183	39% 83% 8% • 8%
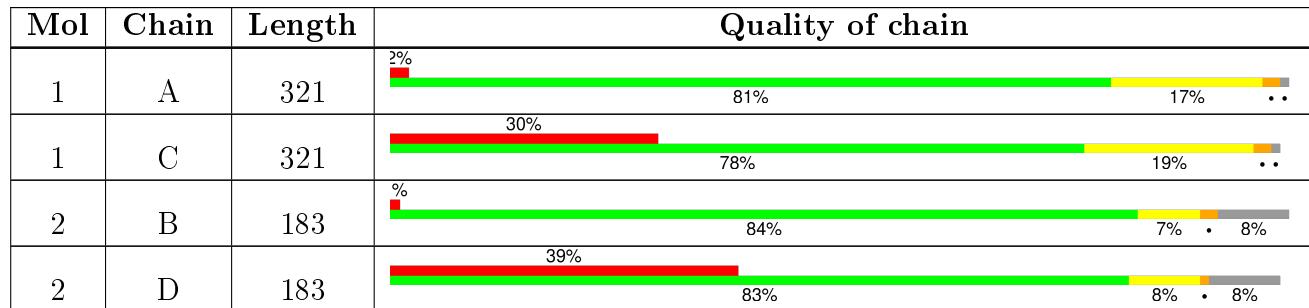
The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NGS	A	407	-	-	-	X
5	NGS	C	403	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7839 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 2427	C 1509	N 438	O 465	S 15	0	0	0
1	C	318	Total 2427	C 1509	N 438	O 465	S 15	0	0	0

- Molecule 2 is a protein called Hemagglutinin HA2.

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

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 a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	3	Total 39	C 22	N 2	O 15	0	0

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
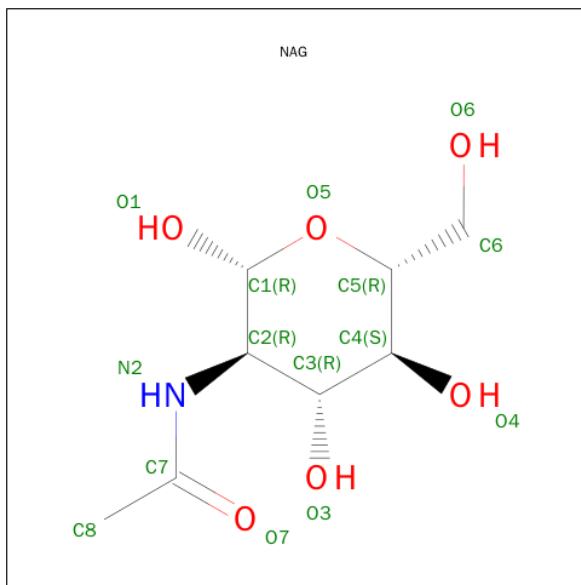
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 14	C 8	N 1	O 5	0	0
4	B	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	3	Total 50	C 25	N 2	O 22	S 1	0	0
5	C	3	Total 50	C 25	N 2	O 22	S 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	41	Total 41	O 41	0	0
6	B	25	Total 25	O 25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	6	Total O 6 6		0	0

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 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
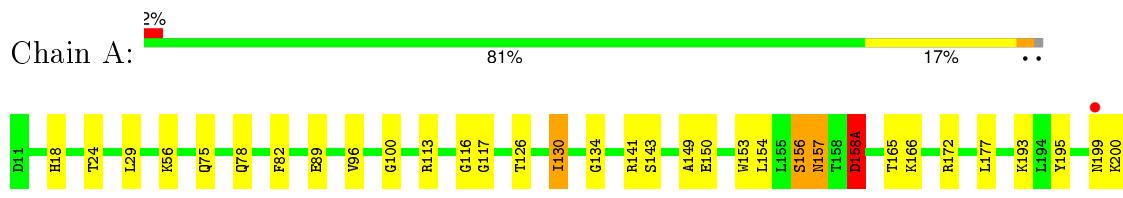
Chain A:

- Molecule 1: Hemagglutinin HA1

Chain C:

- Molecule 2: Hemagglutinin HA2

Chain B:

- Molecule 2: Hemagglutinin HA2

Chain D:
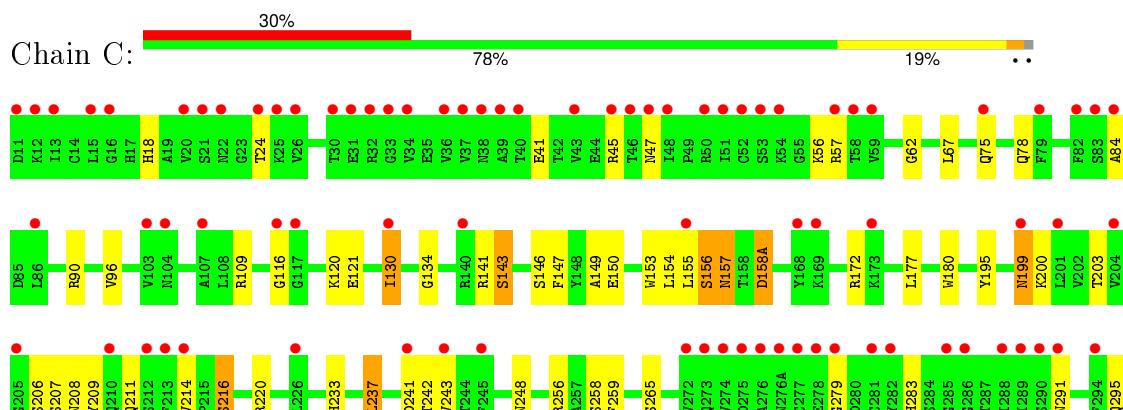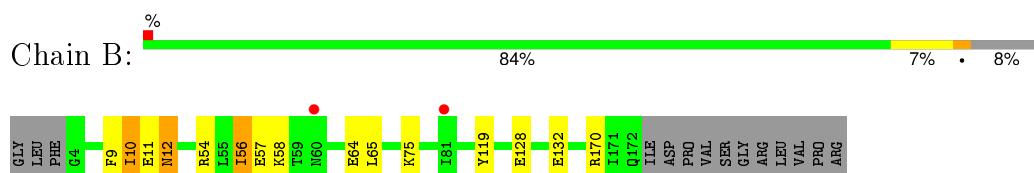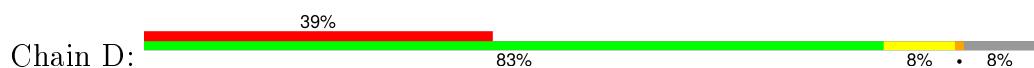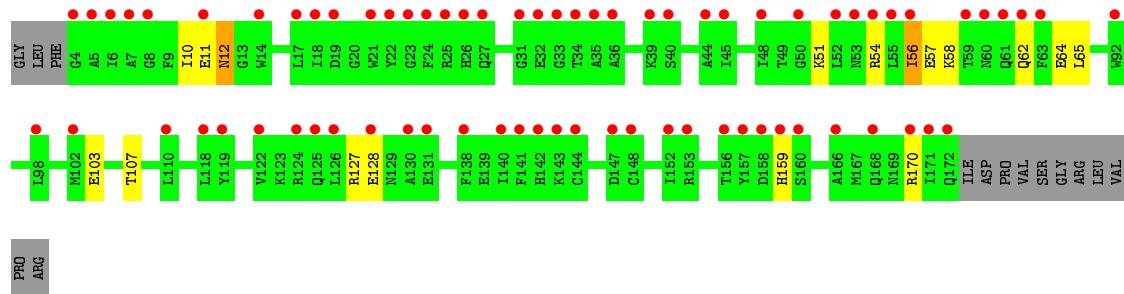
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	153.81Å 153.81Å 153.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.64 – 2.50 48.64 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.64-2.50) 99.7 (48.64-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.11 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.220 , 0.265 0.227 , 0.270	Depositor DCC
R_{free} test set	2121 reflections (5.33%)	DCC
Wilson B-factor (Å^2)	52.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 60.6	EDS
Estimated twinning fraction	0.031 for l,-k,h	Xtriage
L-test for twinning [2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Outliers	0 of 41990 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7839	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.91% 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 (i)

5.1 Standard geometry (i)

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

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	0.29	0/2474	0.48	0/3345
1	C	0.26	0/2474	0.43	0/3345
2	B	0.29	0/1396	0.43	0/1881
2	D	0.23	0/1396	0.39	0/1881
All	All	0.27	0/7740	0.44	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	2387	32	0
1	C	2427	0	2389	41	0
2	B	1373	0	1273	13	0
2	D	1373	0	1274	13	0
3	A	39	0	34	0	0
4	A	14	0	13	0	0
4	B	14	0	13	1	0
5	A	50	0	40	0	0
5	C	50	0	40	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	41	0	0	2	0
6	B	25	0	0	1	0
6	C	6	0	0	0	0
All	All	7839	0	7463	86	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 6.

All (86) 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	1.99	0.94
1:C:56:LYS:NZ	1:C:279:GLY:O	2.03	0.90
1:A:56:LYS:NZ	1:A:279:GLY:O	2.09	0.85
1:A:326:ILE:HG12	2:B:12:ASN:HB3	1.63	0.81
1:C:130:ILE:HD11	1:C:154:LEU:HB3	1.72	0.72
1:A:158(A):ASP:OD2	1:A:193:LYS:NZ	2.23	0.71
2:B:128:GLU:HG3	2:B:170:ARG:HH12	1.55	0.70
2:B:128:GLU:HG3	2:B:170:ARG:NH1	2.12	0.64
1:A:150:GLU:OE1	1:A:256:ARG:HD3	1.98	0.63
1:C:195:TYR:O	1:C:200:LYS:NZ	2.26	0.63
1:A:126:THR:O	1:A:166:LYS:NZ	2.20	0.62
1:C:206:SER:OG	1:C:207:SER:N	2.34	0.60
1:C:206:SER:HB3	1:C:209:TYR:HB3	1.83	0.60
1:A:325:GLU:HA	2:B:12:ASN:HB2	1.83	0.60
1:A:206:SER:HB3	1:A:209:TYR:HB3	1.85	0.59
1:C:150:GLU:OE1	1:C:256:ARG:HD3	2.01	0.59
2:D:128:GLU:HG3	2:D:170:ARG:HH12	1.68	0.58
1:A:206:SER:OG	1:A:207:SER:N	2.37	0.58
1:C:155:LEU:HD21	5:C:401:SIA:H111	1.87	0.57
2:B:12:ASN:ND2	6:B:314:HOH:O	2.37	0.57
1:A:291:ASN:OD1	2:B:58:LYS:NZ	2.27	0.57
1:A:291:ASN:HA	2:B:58:LYS:NZ	2.20	0.56
2:D:51:LYS:NZ	2:D:103:GLU:O	2.38	0.56
1:C:326:ILE:HG12	2:D:12:ASN:HB3	1.87	0.56
1:A:172:ARG:HD3	1:A:259:PHE:CZ	2.41	0.56
1:C:47:ASN:OD1	1:C:283:HIS:NE2	2.36	0.55
1:A:134:GLY:HA3	1:A:153:TRP:HB3	1.89	0.55
1:A:158(A):ASP:N	6:A:517:HOH:O	2.39	0.55
1:C:291:ASN:HB3	2:D:56:ILE:HG13	1.89	0.54
1:C:121:GLU:OE2	1:C:172:ARG:NH1	2.38	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:130:ILE:HD11	1:A:154:LEU:HB3	1.88	0.54
1:C:116:GLY:HA2	1:C:265:SER:HB3	1.90	0.53
1:A:237:LEU:HG	1:A:243:VAL:HG13	1.89	0.53
2:D:128:GLU:O	2:D:170:ARG:NH1	2.34	0.53
1:C:199:ASN:N	1:C:199:ASN:OD1	2.42	0.53
2:D:51:LYS:HZ2	2:D:107:THR:HG1	1.54	0.52
1:C:45:ARG:HD2	1:C:312:ARG:HB2	1.92	0.52
1:A:177:LEU:HB3	1:A:258:SER:HB2	1.92	0.51
1:A:141:ARG:HH12	1:A:149:ALA:HB2	1.77	0.50
2:D:128:GLU:HG3	2:D:170:ARG:NH1	2.25	0.50
1:A:116:GLY:HA2	1:A:265:SER:HB3	1.93	0.50
1:C:146:SER:OG	1:C:147:PHE:N	2.44	0.50
2:B:75:LYS:HD3	4:B:201:NAG:H81	1.94	0.50
1:C:200:LYS:HA	1:C:248:ASN:HD21	1.77	0.49
1:C:295:GLN:HG2	1:C:306:PRO:HG2	1.94	0.49
1:C:172:ARG:HD3	1:C:259:PHE:CZ	2.48	0.48
1:C:134:GLY:HA3	1:C:153:TRP:HB3	1.95	0.48
1:A:291:ASN:HA	2:B:58:LYS:HZ2	1.78	0.48
1:C:303:GLY:HA2	2:D:62:GLN:HG3	1.96	0.48
1:C:237:LEU:HG	1:C:243:VAL:HG13	1.94	0.48
1:C:156:SER:O	1:C:157:ASN:HB2	2.12	0.48
1:C:216:SER:O	1:C:220:ARG:NH2	2.32	0.47
1:A:291:ASN:HB3	2:B:56:ILE:HG13	1.97	0.47
1:A:156:SER:O	1:A:157:ASN:HB2	2.14	0.47
1:C:291:ASN:HA	2:D:58:LYS:NZ	2.30	0.47
1:C:296:ASN:ND2	1:C:309:VAL:O	2.46	0.46
1:C:141:ARG:O	1:C:143:SER:OG	2.31	0.46
1:C:177:LEU:HB3	1:C:258:SER:HB2	1.98	0.45
1:C:130:ILE:HG13	1:C:155:LEU:O	2.17	0.45
1:A:82:PHE:CZ	1:A:117:GLY:HA2	2.51	0.45
1:A:216:SER:O	1:A:220:ARG:NH2	2.34	0.45
2:B:119:TYR:OH	2:B:132:GLU:OE2	2.23	0.44
1:A:237:LEU:HD22	1:A:241:ASP:HB3	2.00	0.44
1:A:292:LEU:O	1:A:306:PRO:HB3	2.18	0.44
1:C:141:ARG:HH12	1:C:149:ALA:HB2	1.83	0.44
1:C:237:LEU:HA	1:C:237:LEU:HD23	1.86	0.44
1:A:325:GLU:HG3	2:B:12:ASN:HD22	1.84	0.43
1:A:195:TYR:O	1:A:200:LYS:NZ	2.32	0.43
1:A:100:GLY:HA3	1:A:230:ILE:O	2.20	0.42
1:C:67:LEU:HD21	1:C:109:ARG:HG2	2.01	0.42
1:C:62:GLY:HA2	1:C:90:ARG:HG3	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:127:ARG:HG3	2:D:159:HIS:CG	2.55	0.42
1:A:89:GLU:OE2	1:A:113:ARG:NE	2.44	0.41
1:C:41:GLU:OE1	1:C:313:SER:OG	2.21	0.41
1:C:146:SER:HG	1:C:147:PHE:H	1.68	0.41
1:A:213:PHE:CE1	1:A:233:HIS:CD2	3.08	0.41
1:A:229:ARG:NE	6:A:531:HOH:O	2.28	0.41
1:A:165:THR:HA	1:A:245:PHE:O	2.21	0.41
2:D:65:LEU:HD12	2:D:65:LEU:HA	1.86	0.41
1:C:200:LYS:O	1:C:214:VAL:HG13	2.21	0.41
1:C:57:ARG:HG2	1:C:84:ALA:HA	2.02	0.41
2:B:9:PHE:O	2:B:10:ILE:HB	2.21	0.41
1:C:241:ASP:OD1	1:C:242:THR:N	2.50	0.40
1:C:180:TRP:CE2	1:C:233:HIS:HB2	2.56	0.40
1:C:120:LYS:NZ	1:C:150:GLU:OE2	2.40	0.40
1:C:325:GLU:HA	2:D:12:ASN:HB2	2.03	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	316/321 (98%)	294 (93%)	18 (6%)	4 (1%)	15	26
1	C	316/321 (98%)	294 (93%)	18 (6%)	4 (1%)	15	26
2	B	167/183 (91%)	156 (93%)	8 (5%)	3 (2%)	11	18
2	D	167/183 (91%)	156 (93%)	8 (5%)	3 (2%)	11	18
All	All	966/1008 (96%)	900 (93%)	52 (5%)	14 (1%)	14	24

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	157	ASN
2	B	12	ASN
1	C	157	ASN
2	D	12	ASN
1	A	208	ASN
1	C	208	ASN
1	A	143	SER
1	A	158(A)	ASP
2	B	10	ILE
2	B	57	GLU
1	C	143	SER
1	C	158(A)	ASP
2	D	57	GLU
2	D	10	ILE

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	267/269 (99%)	253 (95%)	14 (5%)	29	51
1	C	267/269 (99%)	254 (95%)	13 (5%)	31	55
2	B	145/157 (92%)	140 (97%)	5 (3%)	44	72
2	D	145/157 (92%)	141 (97%)	4 (3%)	51	78
All	All	824/852 (97%)	788 (96%)	36 (4%)	35	60

All (36) 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	156	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	158(A)	ASP
1	A	199	ASN
1	A	203	THR
1	A	211	GLN
1	A	216	SER
1	A	237	LEU
2	B	11	GLU
2	B	54	ARG
2	B	56	ILE
2	B	64	GLU
2	B	65	LEU
1	C	18	HIS
1	C	24	THR
1	C	75	GLN
1	C	78	GLN
1	C	96	VAL
1	C	130	ILE
1	C	156	SER
1	C	158(A)	ASP
1	C	199	ASN
1	C	203	THR
1	C	211	GLN
1	C	216	SER
1	C	237	LEU
2	D	11	GLU
2	D	54	ARG
2	D	56	ILE
2	D	64	GLU

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)

9 carbohydrates 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	0.52	0	15,19,21	0.40	0
3	NAG	A	402	3	14,14,15	0.21	0	15,19,21	0.58	0
3	BMA	A	403	3	11,11,12	0.84	0	14,15,17	0.90	0
5	SIA	A	405	5	16,20,21	0.33	0	18,28,31	1.15	2 (11%)
5	GAL	A	406	5	11,11,12	1.31	1 (9%)	14,15,17	1.28	2 (14%)
5	NGS	A	407	5	19,19,19	1.58	4 (21%)	22,28,28	1.72	5 (22%)
5	SIA	C	401	5	16,20,21	0.30	0	18,28,31	1.07	2 (11%)
5	GAL	C	402	5	11,11,12	1.04	2 (18%)	14,15,17	1.23	1 (7%)
5	NGS	C	403	5	19,19,19	1.56	4 (21%)	22,28,28	1.60	5 (22%)

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
3	BMA	A	403	3	-	0/2/19/22	0/1/1/1
5	SIA	A	405	5	-	0/14/34/38	0/1/1/1
5	GAL	A	406	5	-	0/2/19/22	0/1/1/1
5	NGS	A	407	5	-	0/10/30/30	0/1/1/1
5	SIA	C	401	5	-	0/14/34/38	0/1/1/1
5	GAL	C	402	5	-	0/2/19/22	0/1/1/1
5	NGS	C	403	5	-	0/10/30/30	0/1/1/1

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	407	NGS	C1-C2	-2.82	1.49	1.53

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	403	NGS	C1-C2	-2.76	1.49	1.53
5	A	407	NGS	O5-C1	-2.33	1.38	1.43
5	C	403	NGS	O5-C1	-2.18	1.38	1.43
5	C	402	GAL	C1-C2	2.13	1.57	1.52
5	C	402	GAL	C2-C3	2.21	1.55	1.52
5	C	403	NGS	CH3-C	2.58	1.55	1.50
5	A	407	NGS	CH3-C	2.60	1.55	1.50
5	C	403	NGS	C-N	3.14	1.46	1.34
5	A	407	NGS	C-N	3.20	1.46	1.34
5	A	406	GAL	C2-C3	3.27	1.57	1.52

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	401	SIA	C3-C4-C5	-3.14	107.97	111.47
5	A	405	SIA	C3-C4-C5	-2.76	108.40	111.47
5	A	407	NGS	O9-S-O7	-2.60	101.14	112.46
5	C	403	NGS	O9-S-O7	-2.54	101.38	112.46
5	A	406	GAL	O3-C3-C4	-2.32	105.11	110.34
5	A	406	GAL	C1-O5-C5	2.12	114.94	112.25
5	C	403	NGS	CH3-C-N	2.16	120.24	116.11
5	C	401	SIA	O6-C6-C5	2.36	112.34	108.48
5	A	407	NGS	CH3-C-N	2.36	120.62	116.11
5	A	405	SIA	O6-C6-C5	2.41	112.43	108.48
5	C	402	GAL	C1-O5-C5	2.81	115.81	112.25
5	A	407	NGS	C3-C4-C5	2.96	115.35	110.20
5	C	403	NGS	C3-C4-C5	3.00	115.43	110.20
5	C	403	NGS	C4-C3-C2	3.14	114.78	110.43
5	C	403	NGS	O5-C5-C4	3.27	115.82	109.68
5	A	407	NGS	O5-C5-C4	3.31	115.89	109.68
5	A	407	NGS	C4-C3-C2	3.98	115.95	110.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	401	SIA	1	0

5.6 Ligand geometry (i)

2 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	#$|Z| > 2$
4	NAG	A	404	1	14,14,15	0.24	0	15,19,21	0.44	0
4	NAG	B	201	2	14,14,15	0.48	0	15,19,21	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
4	NAG	A	404	1	-	0/6/23/26	0/1/1/1
4	NAG	B	201	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	201	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 (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	318/321 (99%)	0.38	8 (2%)	61	65	50, 69, 98, 142	0
1	C	318/321 (99%)	1.57	96 (30%)	1	0	76, 119, 176, 213	0
2	B	169/183 (92%)	0.39	2 (1%)	81	83	46, 62, 96, 141	0
2	D	169/183 (92%)	2.21	72 (42%)	0	0	92, 160, 188, 232	0
All	All	974/1008 (96%)	1.09	178 (18%)	2	2	46, 93, 176, 232	0

All (178) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	319	GLY	13.5
1	C	327	PRO	12.1
2	D	7	ALA	11.3
2	D	4	GLY	9.0
1	C	58	THR	8.6
2	D	33	GLY	8.1
1	C	289	ILE	7.6
2	D	36	ALA	6.5
1	C	294	PHE	6.5
1	C	288	ILE	6.2
1	C	51	ILE	6.2
2	D	168	GLN	6.0
2	D	140	ILE	5.9
2	D	171	ILE	5.6
1	C	276(A)	ASN	5.6
1	C	323	VAL	5.6
2	D	34	THR	5.3
2	D	143	LYS	5.3
2	D	5	ALA	5.3
2	D	60	ASN	5.3
2	D	8	GLY	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	86	LEU	4.9
2	D	48	ILE	4.9
1	A	327	PRO	4.9
1	C	13	ILE	4.9
1	C	272	VAL	4.9
2	D	172	GLN	4.8
2	D	166	ALA	4.8
2	D	131	GLU	4.8
2	D	35	ALA	4.8
2	D	23	GLY	4.7
2	D	63	PHE	4.7
2	D	32	GLU	4.6
1	C	140	ARG	4.6
2	D	141	PHE	4.5
1	C	277	CYS	4.5
1	C	279	GLY	4.5
2	D	6	ILE	4.4
2	D	59	THR	4.4
1	C	33	GLY	4.3
1	C	38	ASN	4.2
1	C	16	GLY	4.2
2	D	144	CYS	4.1
1	C	54	LYS	4.1
1	C	282	TYR	4.1
1	C	12	LYS	4.0
2	D	18	ILE	4.0
2	D	24	PHE	4.0
2	D	52	LEU	4.0
2	D	138	PHE	3.9
2	D	122	VAL	3.9
1	C	173	LYS	3.9
2	D	44	ALA	3.9
1	C	45	ARG	3.8
2	D	157	TYR	3.8
1	A	214	VAL	3.7
1	C	317	ALA	3.7
1	C	281	CYS	3.6
2	D	27	GLN	3.6
1	C	47	ASN	3.6
2	D	170	ARG	3.6
2	B	60	ASN	3.6
2	D	152	ILE	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	205	GLY	3.5
1	C	52	CYS	3.5
1	C	326	ILE	3.4
1	C	46	THR	3.4
1	C	305	CYS	3.4
2	D	126	LEU	3.4
2	D	56	ILE	3.4
1	C	20	VAL	3.3
2	D	55	LEU	3.3
1	C	39	ALA	3.3
1	C	82	PHE	3.3
1	C	34	VAL	3.2
1	C	104	ASN	3.2
1	C	48	ILE	3.2
2	D	128	GLU	3.2
2	D	130	ALA	3.2
1	C	40	THR	3.2
1	C	273	GLN	3.2
1	C	84	ALA	3.2
2	D	54	ARG	3.1
1	C	37	VAL	3.1
1	A	212	SER	3.1
2	D	14	TRP	3.1
1	C	325	GLU	3.1
2	D	19	ASP	3.1
1	C	79	PHE	3.0
2	D	21	TRP	3.0
1	C	11	ASP	3.0
1	C	15	LEU	3.0
1	C	212	SER	3.0
1	C	278	GLU	3.0
1	C	324	PRO	3.0
1	C	117	GLY	3.0
1	C	199	ASN	3.0
2	D	119	TYR	3.0
2	D	142	HIS	3.0
1	C	116	GLY	3.0
1	C	155	LEU	2.9
2	D	124	ARG	2.9
1	C	31	GLU	2.9
1	C	22	ASN	2.9
1	C	275	ASP	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	11	GLU	2.9
1	C	43	VAL	2.9
2	D	98	LEU	2.9
1	C	318	THR	2.9
1	A	199	ASN	2.8
1	C	291	ASN	2.8
1	C	36	VAL	2.8
1	C	59	VAL	2.8
1	C	30	THR	2.8
2	D	102	MET	2.8
2	D	147	ASP	2.7
2	D	31	GLY	2.7
2	D	62	GLN	2.7
2	D	26	HIS	2.7
1	A	231	ASP	2.7
2	D	45	ILE	2.7
1	C	285	GLY	2.7
2	D	25	ARG	2.6
1	C	290	SER	2.6
1	C	57	ARG	2.6
1	C	214	VAL	2.6
1	A	326	ILE	2.5
1	C	276	ALA	2.5
1	C	130	ILE	2.5
1	C	168	TYR	2.5
1	C	32	ARG	2.5
2	D	156	THR	2.5
1	C	309	VAL	2.5
1	C	21	SER	2.4
1	C	274	VAL	2.4
2	D	17	LEU	2.4
2	D	110	LEU	2.4
1	C	314	LEU	2.4
1	C	103	VAL	2.4
1	C	26	VAL	2.4
2	D	40	SER	2.3
2	D	160	SER	2.3
1	C	297	ILE	2.3
1	C	169	LYS	2.3
2	D	158	ASP	2.3
2	D	153	ARG	2.3
1	C	320	MET	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	226	LEU	2.2
1	A	274	VAL	2.2
1	C	53	SER	2.2
1	C	286	GLY	2.2
1	C	316	LEU	2.2
2	D	22	TYR	2.2
1	C	315	LEU	2.2
1	C	75	GLN	2.2
2	D	148	CYS	2.1
1	C	107	ALA	2.1
1	C	243	VAL	2.1
1	C	25	LYS	2.1
2	D	53	ASN	2.1
1	C	83	SER	2.1
1	C	204	VAL	2.1
1	C	245	PHE	2.1
1	A	210	GLN	2.1
1	C	50	ARG	2.1
1	C	201	LEU	2.1
2	D	118	LEU	2.1
1	C	210	GLN	2.0
2	D	125	GLN	2.0
2	D	50	GLY	2.0
2	D	92	TRP	2.0
1	C	24	THR	2.0
2	D	39	LYS	2.0
2	D	159	HIS	2.0
1	C	213	PHE	2.0
2	D	61	GLN	2.0
1	C	241	ASP	2.0
2	B	81	ILE	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)

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.

Full wwPDB X-ray Structure Validation Report

The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	NGS	A	407	19/19	0.72	0.39	6.04	112,158,172,173	0
5	NGS	C	403	19/19	0.61	0.47	3.97	115,206,210,211	0
5	SIA	C	401	20/21	0.84	0.24	0.43	95,119,143,146	0
5	SIA	A	405	20/21	0.90	0.18	0.08	61,80,101,104	0
3	BMA	A	403	11/12	0.41	0.36	-	175,180,182,182	0
5	GAL	C	402	11/12	0.75	0.37	-	151,198,216,220	0
5	GAL	A	406	11/12	0.81	0.23	-	114,175,180,180	0
3	NAG	A	402	14/15	0.76	0.40	-	130,147,156,164	0
3	NAG	A	401	14/15	0.90	0.20	-	97,115,124,138	0

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
4	NAG	B	201	14/15	0.83	0.21	-0.24	80,92,96,99	0
4	NAG	A	404	14/15	0.82	0.18	-	127,140,146,149	0

6.5 Other polymers (i)

There are no such residues in this entry.