



Full wwPDB X-ray Structure Validation Report

Apr 24, 2014 – 01:18 AM EDT

PDB ID : 4LXV
Title : Crystal Structure of the Hemagglutinin from a H1N1pdm A/WASHINGTON/5/2011 virus
Authors : Yang, H.; Chang, J.C.; Guo, Z.; Carney, P.J.; Shore, D.A.; Donis, R.O.; Cox, N.J.; Villanueva, J.M.; Klimov, A.I.; Stevens, J.
Deposited on : 2013-07-30
Resolution : 3.00 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

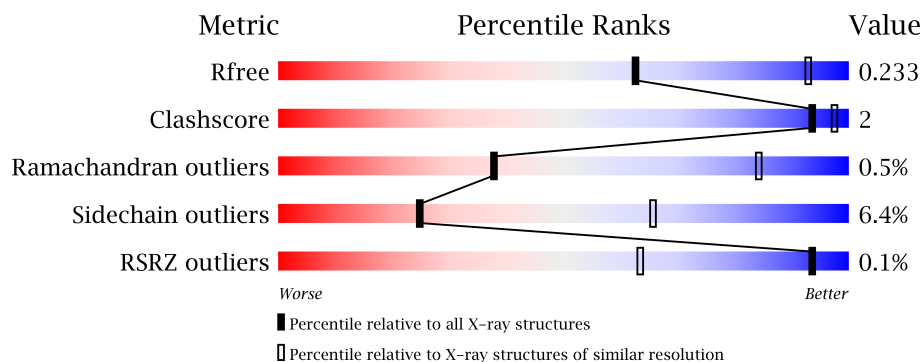
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.16 November 2013
Xtriage (Phenix) : dev-1439
EDS : stable22978
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22978

1 Overall quality at a glance

The reported resolution of this entry is 3.00 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}	66092	1216 (3.00-3.00)
Clashscore	79885	1594 (3.00-3.00)
Ramachandran outliers	78287	1537 (3.00-3.00)
Sidechain outliers	78261	1540 (3.00-3.00)
RSRZ outliers	66119	1217 (3.00-3.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	332	
1	C	332	
1	E	332	
1	G	332	
1	I	332	
1	K	332	
2	B	182	
2	D	182	
2	F	182	
2	H	182	
2	J	182	
2	L	182	

The following table lists non-polymeric compounds that are outliers for geometric or electron-

density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	NAG	A	403	-	X
4	NAG	E	403	-	X
4	NAG	K	401	-	X
4	NAG	K	402	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 23740 atoms, of which 0 are hydrogen and 0 are deuterium.

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 (with D amino acids) called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	0	0
			2523	1595	435	482	11			
1	C	323	Total	C	N	O	S	0	0	0
			2523	1595	435	482	11			
1	E	323	Total	C	N	O	S	0	0	0
			2523	1595	435	482	11			
1	G	323	Total	C	N	O	S	0	0	0
			2523	1595	435	482	11			
1	I	323	Total	C	N	O	S	0	0	0
			2523	1595	435	482	11			
1	K	323	Total	C	N	O	S	0	0	0
			2523	1595	435	482	11			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
A	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
A	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
A	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
A	0	SER	-	EXPRESSION TAG	UNP J7MFR5
C	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
C	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
C	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
C	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
C	0	SER	-	EXPRESSION TAG	UNP J7MFR5
E	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
E	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
E	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
E	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
E	0	SER	-	EXPRESSION TAG	UNP J7MFR5
G	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
G	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
G	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
G	0	SER	-	EXPRESSION TAG	UNP J7MFR5
I	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
I	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
I	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
I	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
I	0	SER	-	EXPRESSION TAG	UNP J7MFR5
K	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
K	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
K	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
K	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
K	0	SER	-	EXPRESSION TAG	UNP J7MFR5

- Molecule 2 is a protein (with D amino acids) called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	171	Total	C	N	O	S	0	0	0
			1387	871	236	274	6			
2	D	171	Total	C	N	O	S	0	0	0
			1387	871	236	274	6			
2	F	171	Total	C	N	O	S	0	0	0
			1387	871	236	274	6			
2	H	171	Total	C	N	O	S	0	0	0
			1387	871	236	274	6			
2	J	171	Total	C	N	O	S	0	0	0
			1387	871	236	274	6			
2	L	171	Total	C	N	O	S	0	0	0
			1387	871	236	274	6			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	EXPRESSION TAG	UNP J7MFR5
B	176	GLY	-	EXPRESSION TAG	UNP J7MFR5
B	177	ARG	-	EXPRESSION TAG	UNP J7MFR5
B	178	LEU	-	EXPRESSION TAG	UNP J7MFR5
B	179	VAL	-	EXPRESSION TAG	UNP J7MFR5
B	180	PRO	-	EXPRESSION TAG	UNP J7MFR5
B	181	ARG	-	EXPRESSION TAG	UNP J7MFR5
B	182	GLY	-	EXPRESSION TAG	UNP J7MFR5
D	175	SER	-	EXPRESSION TAG	UNP J7MFR5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	176	GLY	-	EXPRESSION TAG	UNP J7MFR5
D	177	ARG	-	EXPRESSION TAG	UNP J7MFR5
D	178	LEU	-	EXPRESSION TAG	UNP J7MFR5
D	179	VAL	-	EXPRESSION TAG	UNP J7MFR5
D	180	PRO	-	EXPRESSION TAG	UNP J7MFR5
D	181	ARG	-	EXPRESSION TAG	UNP J7MFR5
D	182	GLY	-	EXPRESSION TAG	UNP J7MFR5
F	175	SER	-	EXPRESSION TAG	UNP J7MFR5
F	176	GLY	-	EXPRESSION TAG	UNP J7MFR5
F	177	ARG	-	EXPRESSION TAG	UNP J7MFR5
F	178	LEU	-	EXPRESSION TAG	UNP J7MFR5
F	179	VAL	-	EXPRESSION TAG	UNP J7MFR5
F	180	PRO	-	EXPRESSION TAG	UNP J7MFR5
F	181	ARG	-	EXPRESSION TAG	UNP J7MFR5
F	182	GLY	-	EXPRESSION TAG	UNP J7MFR5
H	175	SER	-	EXPRESSION TAG	UNP J7MFR5
H	176	GLY	-	EXPRESSION TAG	UNP J7MFR5
H	177	ARG	-	EXPRESSION TAG	UNP J7MFR5
H	178	LEU	-	EXPRESSION TAG	UNP J7MFR5
H	179	VAL	-	EXPRESSION TAG	UNP J7MFR5
H	180	PRO	-	EXPRESSION TAG	UNP J7MFR5
H	181	ARG	-	EXPRESSION TAG	UNP J7MFR5
H	182	GLY	-	EXPRESSION TAG	UNP J7MFR5
J	175	SER	-	EXPRESSION TAG	UNP J7MFR5
J	176	GLY	-	EXPRESSION TAG	UNP J7MFR5
J	177	ARG	-	EXPRESSION TAG	UNP J7MFR5
J	178	LEU	-	EXPRESSION TAG	UNP J7MFR5
J	179	VAL	-	EXPRESSION TAG	UNP J7MFR5
J	180	PRO	-	EXPRESSION TAG	UNP J7MFR5
J	181	ARG	-	EXPRESSION TAG	UNP J7MFR5
J	182	GLY	-	EXPRESSION TAG	UNP J7MFR5
L	175	SER	-	EXPRESSION TAG	UNP J7MFR5
L	176	GLY	-	EXPRESSION TAG	UNP J7MFR5
L	177	ARG	-	EXPRESSION TAG	UNP J7MFR5
L	178	LEU	-	EXPRESSION TAG	UNP J7MFR5
L	179	VAL	-	EXPRESSION TAG	UNP J7MFR5
L	180	PRO	-	EXPRESSION TAG	UNP J7MFR5
L	181	ARG	-	EXPRESSION TAG	UNP J7MFR5
L	182	GLY	-	EXPRESSION TAG	UNP J7MFR5

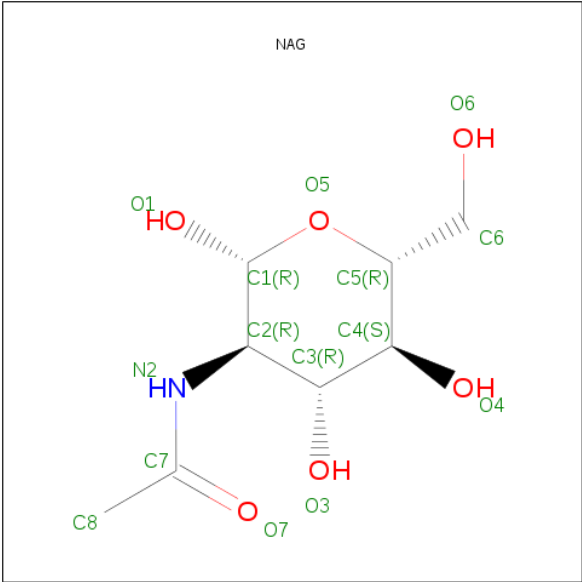
- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	C	2	Total	C	N	O	0	0
			28	16	2	10		
3	E	2	Total	C	N	O	0	0
			28	16	2	10		
3	G	2	Total	C	N	O	0	0
			28	16	2	10		
3	I	2	Total	C	N	O	0	0
			28	16	2	10		

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
A	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
A	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
A	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
A	0	SER	-	EXPRESSION TAG	UNP J7MFR5
C	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
C	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
C	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
C	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
C	0	SER	-	EXPRESSION TAG	UNP J7MFR5
E	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
E	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
E	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
E	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
E	0	SER	-	EXPRESSION TAG	UNP J7MFR5
G	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
G	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
G	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
G	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
G	0	SER	-	EXPRESSION TAG	UNP J7MFR5
I	-4	ALA	-	EXPRESSION TAG	UNP J7MFR5
I	-3	ASP	-	EXPRESSION TAG	UNP J7MFR5
I	-2	LEU	-	EXPRESSION TAG	UNP J7MFR5
I	-1	GLY	-	EXPRESSION TAG	UNP J7MFR5
I	0	SER	-	EXPRESSION TAG	UNP J7MFR5

- 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	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	I	1	Total	C	N	O	0	0
			14	8	1	5		
4	K	1	Total	C	N	O	0	0
			14	8	1	5		
4	K	1	Total	C	N	O	0	0
			14	8	1	5		

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

Chain A: 



• Molecule 1: Hemagglutinin

Chain C: 



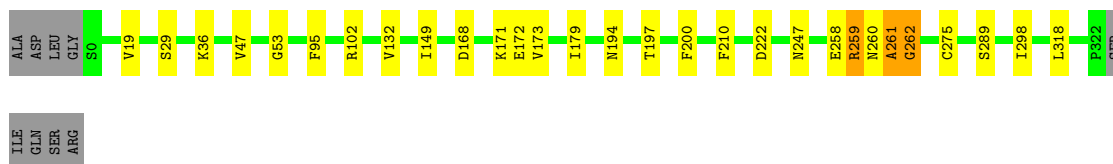
• Molecule 1: Hemagglutinin

Chain E: 



• Molecule 1: Hemagglutinin

Chain G: 



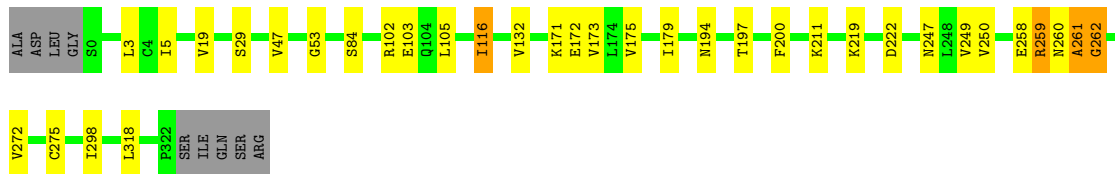
• Molecule 1: Hemagglutinin

Chain I: 



- Molecule 1: Hemagglutinin

Chain K:



- Molecule 2: Hemagglutinin

Chain B:



- Molecule 2: Hemagglutinin

Chain D:



- Molecule 2: Hemagglutinin

Chain F:



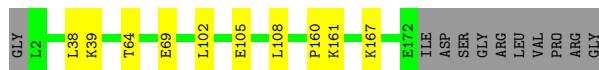
- Molecule 2: Hemagglutinin

Chain H:



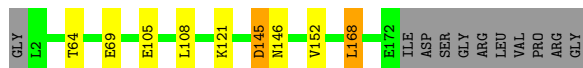
- Molecule 2: Hemagglutinin

Chain J:



- Molecule 2: Hemagglutinin

Chain L:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	73.24Å 226.00Å 271.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.78 – 3.00 49.73 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.2 (49.78-3.00) 98.2 (49.73-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.99 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.205 , 0.230 0.208 , 0.233	Depositor DCC
R_{free} test set	4492 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	59.6	Xtriage
Anisotropy	0.337	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 17.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 89589 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	23740	wwPDB-VP
Average B, all atoms (Å ²)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.59	0/2586	0.76	1/3515 (0.0%)
1	C	0.61	0/2586	0.79	3/3515 (0.1%)
1	E	0.61	0/2586	0.79	3/3515 (0.1%)
1	G	0.58	0/2586	0.76	2/3515 (0.1%)
1	I	0.60	0/2586	0.77	2/3515 (0.1%)
1	K	0.58	0/2586	0.78	3/3515 (0.1%)
2	B	0.53	0/1415	0.67	0/1906
2	D	0.54	0/1415	0.70	0/1906
2	F	0.54	0/1415	0.74	4/1906 (0.2%)
2	H	0.56	0/1415	0.68	0/1906
2	J	0.59	0/1415	0.71	0/1906
2	L	0.57	0/1415	0.75	4/1906 (0.2%)
All	All	0.58	0/24006	0.75	22/32526 (0.1%)

There are no bond length outliers.

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	145	ASP	CB-CG-OD2	-7.78	111.30	118.30
2	L	145	ASP	CB-CG-OD2	-7.66	111.41	118.30
2	L	145	ASP	N-CA-CB	-7.57	96.97	110.60
2	F	145	ASP	N-CA-CB	-7.31	97.44	110.60
1	E	275	CYS	N-CA-CB	-7.21	97.61	110.60
1	G	149	ILE	CG1-CB-CG2	-7.11	95.77	111.40
1	A	149	ILE	CG1-CB-CG2	-6.94	96.13	111.40
1	C	288	THR	CB-CA-C	-6.76	93.36	111.60
1	K	275	CYS	CA-CB-SG	6.65	125.97	114.00
1	I	288	THR	CB-CA-C	-6.62	93.73	111.60
1	C	1	ASP	CB-CG-OD2	6.46	124.11	118.30
1	I	1	ASP	CB-CG-OD2	6.42	124.08	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	275	CYS	N-CA-CB	-6.38	99.11	110.60
2	L	145	ASP	CB-CG-OD1	6.12	123.81	118.30
2	F	168	LEU	CA-CB-CG	5.99	129.07	115.30
2	L	168	LEU	CA-CB-CG	5.94	128.96	115.30
2	F	145	ASP	CB-CG-OD1	5.90	123.61	118.30
1	E	275	CYS	CA-CB-SG	5.86	124.55	114.00
1	E	168	ASP	CB-CG-OD2	5.16	122.94	118.30
1	C	116	ILE	CB-CA-C	-5.14	101.32	111.60
1	G	168	ASP	CB-CG-OD2	5.10	122.89	118.30
1	K	116	ILE	CB-CA-C	-5.08	101.44	111.60

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2523	0	0	6	0
1	C	2523	0	0	5	0
1	E	2523	0	0	6	0
1	G	2523	0	0	6	0
1	I	2523	0	0	5	0
1	K	2523	0	0	6	0
2	B	1387	0	0	2	0
2	D	1387	0	0	1	0
2	F	1387	0	0	2	0
2	H	1387	0	0	1	0
2	J	1387	0	0	2	0
2	L	1387	0	0	2	0
3	A	28	0	25	0	0
3	C	28	0	25	0	0
3	E	28	0	25	0	0
3	G	28	0	25	0	0
3	I	28	0	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	28	0	26	0	0
4	C	14	0	13	0	0
4	E	28	0	26	1	0
4	G	28	0	26	0	0
4	I	14	0	13	0	0
4	K	28	0	26	0	0
All	All	23740	0	255	38	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (38) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:200:PHE:CD2	1:C:200:PHE:C	2.76	0.57
1:G:102:ARG:NE	2:H:69:GLU:OE1	2.38	0.56
1:E:200:PHE:C	1:E:200:PHE:CD1	2.78	0.56
1:K:200:PHE:C	1:K:200:PHE:CD1	2.81	0.53
1:C:261:ALA:O	1:C:262:GLY:C	2.46	0.53
1:I:200:PHE:CD2	1:I:200:PHE:C	2.81	0.53
1:A:102:ARG:NE	2:B:69:GLU:OE1	2.44	0.51
1:E:72:THR:O	1:K:219:LYS:NZ	2.44	0.50
2:F:145:ASP:OD1	2:F:146:ASN:N	2.44	0.50
1:G:261:ALA:O	1:G:262:GLY:C	2.49	0.50
1:K:261:ALA:O	1:K:262:GLY:C	2.50	0.50
1:A:200:PHE:C	1:A:200:PHE:CD2	2.85	0.50
1:A:261:ALA:O	1:A:262:GLY:C	2.49	0.50
1:I:261:ALA:O	1:I:262:GLY:C	2.51	0.49
1:C:102:ARG:NE	2:D:69:GLU:OE1	2.46	0.49
4:E:404:NAG:C1	4:E:404:NAG:C8	2.90	0.49
1:I:102:ARG:NE	2:J:69:GLU:OE1	2.45	0.49
1:G:200:PHE:C	1:G:200:PHE:CD2	2.89	0.47
1:E:261:ALA:O	1:E:262:GLY:C	2.54	0.46
2:L:145:ASP:OD1	2:L:146:ASN:N	2.49	0.46
1:E:102:ARG:NE	2:F:69:GLU:OE1	2.49	0.45
1:C:179:ILE:CD1	1:C:210:PHE:CG	3.00	0.44
1:A:179:ILE:CD1	1:A:210:PHE:CG	3.01	0.44
1:E:197:THR:OG1	1:E:247:ASN:OD1	2.35	0.44
1:I:197:THR:OG1	1:I:247:ASN:OD1	2.35	0.44
1:A:172:GLU:OE2	1:A:259:ARG:NE	2.51	0.44
1:A:197:THR:OG1	1:A:247:ASN:OD1	2.37	0.43
1:G:197:THR:OG1	1:G:247:ASN:OD1	2.36	0.43
1:G:179:ILE:CD1	1:G:210:PHE:CG	3.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:J:160:PRO:O	2:J:161:LYS:C	2.56	0.43
1:K:172:GLU:OE2	1:K:259:ARG:NE	2.52	0.43
1:G:172:GLU:OE2	1:G:259:ARG:NE	2.52	0.43
1:K:197:THR:OG1	1:K:247:ASN:OD1	2.36	0.42
1:C:197:THR:OG1	1:C:247:ASN:OD1	2.37	0.42
1:I:172:GLU:OE2	1:I:259:ARG:NE	2.53	0.42
1:K:102:ARG:NE	2:L:69:GLU:OE1	2.52	0.42
1:E:172:GLU:OE2	1:E:259:ARG:NE	2.53	0.41
2:B:160:PRO:O	2:B:161:LYS:C	2.60	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

There are no protein chains in this entry.

5.3.2 Protein sidechains ⓘ

There are no protein chains in this entry.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

10 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	12,14,15	0.90	1 (8%)	15,19,21	1.70	4 (26%)
3	NAG	A	402	3	12,14,15	0.97	1 (8%)	15,19,21	1.74	3 (20%)
3	NAG	C	401	1,3	12,14,15	1.11	1 (8%)	15,19,21	1.68	4 (26%)
3	NAG	C	402	3	12,14,15	0.96	1 (8%)	15,19,21	1.65	5 (33%)
3	NAG	E	401	1,3	12,14,15	0.80	1 (8%)	15,19,21	1.61	1 (6%)
3	NAG	E	402	3	12,14,15	0.82	0	15,19,21	1.72	3 (20%)
3	NAG	G	401	1,3	12,14,15	0.70	0	15,19,21	1.90	2 (13%)
3	NAG	G	402	3	12,14,15	0.68	0	15,19,21	2.10	4 (26%)
3	NAG	I	401	1,3	12,14,15	0.68	0	15,19,21	1.15	1 (6%)
3	NAG	I	402	3	12,14,15	0.88	0	15,19,21	1.17	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	402	3	-	0/6/23/26	0/1/1/1
3	NAG	C	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	402	3	-	0/6/23/26	0/1/1/1
3	NAG	E	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	402	3	-	0/6/23/26	0/1/1/1
3	NAG	G	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	402	3	-	0/6/23/26	0/1/1/1
3	NAG	I	401	1,3	-	0/6/23/26	0/1/1/1
3	NAG	I	402	3	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	401	NAG	O5-C5	-3.16	1.40	1.45
3	C	402	NAG	O5-C5	-2.49	1.41	1.45
3	A	402	NAG	C3-C2	2.45	1.56	1.52
3	E	401	NAG	O5-C5	-2.28	1.41	1.45
3	A	401	NAG	O5-C5	-2.10	1.42	1.45

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	401	NAG	O5-C5-C6	5.05	112.28	106.98
3	G	402	NAG	O5-C5-C6	5.00	112.23	106.98
3	E	402	NAG	O5-C5-C6	4.67	111.88	106.98
3	E	401	NAG	O5-C5-C6	4.63	111.84	106.98
3	C	402	NAG	O5-C5-C4	-4.18	105.34	110.65
3	A	402	NAG	O5-C5-C6	4.09	111.28	106.98
3	A	402	NAG	C4-C3-C2	3.50	118.09	110.74
3	G	402	NAG	C3-C4-C5	-3.49	103.91	110.17
3	G	402	NAG	O4-C4-C5	3.48	118.47	109.25
3	A	401	NAG	O7-C7-N2	3.39	128.80	121.90
3	C	401	NAG	O5-C5-C6	-3.37	103.45	106.98
3	C	401	NAG	C3-C2-N2	-3.33	106.63	111.62
3	G	401	NAG	C2-N2-C7	-3.04	119.72	123.39
3	I	401	NAG	O5-C5-C6	3.03	110.16	106.98
3	A	401	NAG	O7-C7-C8	-2.69	116.89	122.04
3	A	401	NAG	C3-C4-C5	-2.68	105.36	110.17
3	E	402	NAG	C4-C3-C2	2.65	116.31	110.74
3	E	402	NAG	O3-C3-C4	-2.61	104.52	110.36
3	A	402	NAG	O5-C5-C4	-2.60	107.35	110.65
3	G	402	NAG	O3-C3-C2	2.44	114.14	109.16
3	A	401	NAG	O5-C5-C6	2.42	109.52	106.98
3	C	401	NAG	C3-C4-C5	-2.36	105.94	110.17
3	C	402	NAG	O5-C5-C6	2.27	109.36	106.98
3	C	401	NAG	O5-C5-C4	-2.19	107.88	110.65
3	C	402	NAG	O7-C7-C8	-2.17	117.88	122.04
3	C	402	NAG	C3-C4-C5	-2.04	106.50	110.17
3	C	402	NAG	O4-C4-C5	2.01	114.59	109.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry ⓘ

10 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	403	1	12,14,15	0.57	0	15,19,21	0.88	1 (6%)
4	NAG	A	404	1	12,14,15	0.68	0	15,19,21	1.15	1 (6%)
4	NAG	C	403	1	12,14,15	1.04	0	15,19,21	1.69	4 (26%)
4	NAG	E	403	-	12,14,15	1.03	0	15,19,21	1.75	1 (6%)
4	NAG	E	404	1	12,14,15	0.94	2 (16%)	15,19,21	1.25	2 (13%)
4	NAG	G	403	1	12,14,15	0.58	0	15,19,21	0.89	1 (6%)
4	NAG	G	404	1	12,14,15	0.76	0	15,19,21	3.02	2 (13%)
4	NAG	I	403	1	12,14,15	0.82	0	15,19,21	2.32	3 (20%)
4	NAG	K	401	1	12,14,15	0.66	0	15,19,21	1.98	7 (46%)
4	NAG	K	402	1	12,14,15	0.68	0	15,19,21	2.42	7 (46%)

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	403	1	1/1/5/7	0/6/23/26	0/1/1/1
4	NAG	A	404	1	-	0/6/23/26	0/1/1/1
4	NAG	C	403	1	-	0/6/23/26	0/1/1/1
4	NAG	E	403	-	-	0/6/23/26	0/1/1/1
4	NAG	E	404	1	1/1/5/7	0/6/23/26	0/1/1/1
4	NAG	G	403	1	-	0/6/23/26	0/1/1/1
4	NAG	G	404	1	-	0/6/23/26	0/1/1/1
4	NAG	I	403	1	-	0/6/23/26	0/1/1/1
4	NAG	K	401	1	-	0/6/23/26	0/1/1/1
4	NAG	K	402	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	404	NAG	O5-C5	-2.26	1.42	1.45
4	E	404	NAG	C2-N2	-2.07	1.44	1.46

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	404	NAG	O5-C5-C6	10.50	118.00	106.98
4	I	403	NAG	O5-C5-C6	7.87	115.24	106.98
4	K	402	NAG	O5-C5-C6	6.03	113.31	106.98
4	E	403	NAG	O5-C5-C6	5.08	112.31	106.98
4	C	403	NAG	O5-C5-C6	4.38	111.58	106.98
4	K	401	NAG	O3-C3-C4	3.81	118.86	110.36
4	K	402	NAG	C3-C2-N2	-3.37	106.57	111.62
4	K	401	NAG	O5-C5-C4	3.32	114.86	110.65
4	K	402	NAG	O5-C5-C4	3.11	114.60	110.65
4	I	403	NAG	C4-C3-C2	3.08	117.22	110.74
4	E	404	NAG	C2-N2-C7	-3.04	119.72	123.39
4	K	402	NAG	C3-C4-C5	2.95	115.47	110.17
4	K	401	NAG	O4-C4-C3	2.91	116.85	110.36
4	C	403	NAG	C4-C3-C2	2.70	116.41	110.74
4	A	404	NAG	C3-C2-N2	-2.69	107.59	111.62
4	K	402	NAG	C4-C3-C2	2.56	116.12	110.74
4	K	401	NAG	C4-C3-C2	-2.42	105.65	110.74
4	A	403	NAG	O5-C5-C6	2.38	109.48	106.98
4	G	403	NAG	O5-C5-C6	2.37	109.47	106.98
4	K	401	NAG	C3-C4-C5	-2.31	106.02	110.17
4	E	404	NAG	C3-C2-N2	-2.27	108.22	111.62
4	K	402	NAG	O7-C7-C8	-2.26	117.71	122.04
4	C	403	NAG	O6-C6-C5	2.22	119.11	111.37
4	K	402	NAG	C8-C7-N2	2.19	120.27	116.12
4	C	403	NAG	O4-C4-C5	2.18	115.01	109.25
4	K	401	NAG	O7-C7-N2	2.13	126.24	121.90
4	G	404	NAG	C2-N2-C7	2.10	125.94	123.39
4	K	401	NAG	O7-C7-C8	-2.08	118.05	122.04
4	I	403	NAG	O5-C5-C4	-2.06	108.04	110.65

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	E	404	NAG	C1
4	A	403	NAG	C1

There are no torsion outliers.

There are no ring outliers.

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, 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	323/332 (97%)	-0.25	2 (0%) 86 32	34, 59, 90, 153	0
1	C	323/332 (97%)	-0.25	0 100 100	38, 57, 89, 138	0
1	E	323/332 (97%)	-0.28	0 100 100	37, 61, 89, 107	0
1	G	323/332 (97%)	-0.30	0 100 100	41, 63, 96, 138	0
1	I	323/332 (97%)	-0.25	0 100 100	38, 62, 92, 133	0
1	K	323/332 (97%)	-0.26	0 100 100	43, 65, 95, 140	0
2	B	171/182 (93%)	-0.20	1 (0%) 86 32	39, 71, 105, 136	0
2	D	171/182 (93%)	0.01	1 (0%) 86 32	44, 78, 124, 159	0
2	F	171/182 (93%)	-0.19	0 100 100	31, 77, 128, 152	0
2	H	171/182 (93%)	-0.13	0 100 100	44, 71, 109, 133	0
2	J	171/182 (93%)	-0.32	0 100 100	42, 62, 95, 136	0
2	L	171/182 (93%)	-0.24	0 100 100	43, 67, 112, 134	0
All	All	2964/3084 (96%)	-0.24	4 (0%) 93 63	31, 64, 107, 159	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	172	GLU	4.0
1	A	71	SER	3.3
2	B	172	GLU	2.8
1	A	70	LEU	2.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	NAG	G	402	14/15	0.24	7.22	68,93,106,107	0
3	NAG	I	401	14/15	0.33	2.05	56,61,63,68	0
3	NAG	A	401	14/15	0.25	1.05	56,65,72,75	0
3	NAG	G	401	14/15	0.15	0.77	62,69,84,90	0
3	NAG	C	401	14/15	0.19	0.13	54,65,70,80	0
3	NAG	E	401	14/15	0.14	-1.06	35,44,50,53	0
3	NAG	C	402	14/15	0.36	-	82,92,105,109	0
3	NAG	E	402	14/15	0.18	-	57,67,80,82	0
3	NAG	A	402	14/15	0.34	-	80,93,104,106	0
3	NAG	I	402	14/15	0.35	-	64,75,81,83	0

6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	NAG	A	403	14/15	0.33	23.33	48,92,100,101	0
4	NAG	E	403	14/15	0.24	5.70	90,116,128,134	0
4	NAG	K	402	14/15	0.26	3.08	83,100,110,116	0
4	NAG	K	401	14/15	0.31	2.21	63,79,82,85	0
4	NAG	C	403	14/15	0.29	1.77	67,74,76,78	0
4	NAG	I	403	14/15	0.45	1.39	92,111,118,125	0
4	NAG	E	404	14/15	0.22	1.37	106,126,139,140	0
4	NAG	A	404	14/15	0.31	0.42	92,115,124,127	0
4	NAG	G	404	14/15	0.26	0.14	84,93,98,102	0
4	NAG	G	403	14/15	0.28	-	58,98,103,104	0

6.5 Other polymers ⓘ

There are no such residues in this entry.