



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:28 AM GMT

PDB ID : 3LZG
Title : Crystal structure of a 2009 H1N1 influenza virus hemagglutinin
Authors : Xu, R.; Wilson, I.A.
Deposited on : 2010-03-01
Resolution : 2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

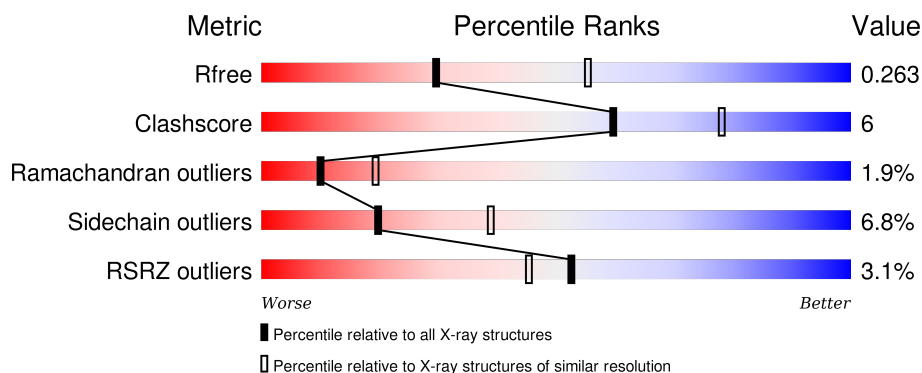
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	329	<div> <div>76%</div> <div>21%</div> <div>..</div> </div>
1	C	329	<div> <div>3%</div> <div>79%</div> <div>17%</div> <div>..</div> </div>
1	E	329	<div> <div>2%</div> <div>73%</div> <div>22%</div> <div>..</div> </div>
1	G	329	<div> <div>75%</div> <div>22%</div> <div>..</div> </div>
1	I	329	<div> <div>3%</div> <div>77%</div> <div>20%</div> <div>..</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	329	<div><div></div><div>3%</div><div>77%</div><div>21%</div><div>..</div></div>
2	B	177	<div><div></div><div>%</div><div>84%</div><div>15%</div><div>.</div></div>
2	D	177	<div><div></div><div>9%</div><div>79%</div><div>16%</div><div>..</div></div>
2	F	177	<div><div></div><div>7%</div><div>81%</div><div>15%</div><div>..</div></div>
2	H	177	<div><div></div><div>2%</div><div>79%</div><div>15%</div><div>..</div></div>
2	J	177	<div><div></div><div>9%</div><div>81%</div><div>15%</div><div>..</div></div>
2	L	177	<div><div></div><div></div><div>76%</div><div>19%</div><div>..</div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 24137 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 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	0	0
			2525	1597	435	482	11			
1	C	323	Total	C	N	O	S	0	0	0
			2525	1597	435	482	11			
1	E	323	Total	C	N	O	S	0	0	0
			2525	1597	435	482	11			
1	G	323	Total	C	N	O	S	0	0	0
			2525	1597	435	482	11			
1	I	323	Total	C	N	O	S	0	0	0
			2525	1597	435	482	11			
1	K	325	Total	C	N	O	S	0	0	0
			2536	1604	437	484	11			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
A	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
C	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
C	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
E	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
E	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
G	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
G	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
I	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
I	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
K	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
K	10	GLY	-	EXPRESSION TAG	UNP C3W5S1

- Molecule 2 is a protein called Hemagglutinin, HA2 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	175	Total	C	N	O	S	0	0	0
			1411	884	238	283	6			
2	D	171	Total	C	N	O	S	0	0	0
			1380	866	234	274	6			
2	F	171	Total	C	N	O	S	0	0	0
			1380	866	234	274	6			
2	H	171	Total	C	N	O	S	0	0	0
			1380	866	234	274	6			
2	J	172	Total	C	N	O	S	0	0	0
			1389	871	235	277	6			
2	L	172	Total	C	N	O	S	0	0	0
			1389	871	235	277	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	EXPRESSION TAG	UNP C3W5S1
B	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
B	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
D	175	SER	-	EXPRESSION TAG	UNP C3W5S1
D	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
D	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
F	175	SER	-	EXPRESSION TAG	UNP C3W5S1
F	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
F	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
H	175	SER	-	EXPRESSION TAG	UNP C3W5S1
H	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
H	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
J	175	SER	-	EXPRESSION TAG	UNP C3W5S1
J	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
J	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
L	175	SER	-	EXPRESSION TAG	UNP C3W5S1
L	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
L	177	ARG	-	EXPRESSION TAG	UNP C3W5S1

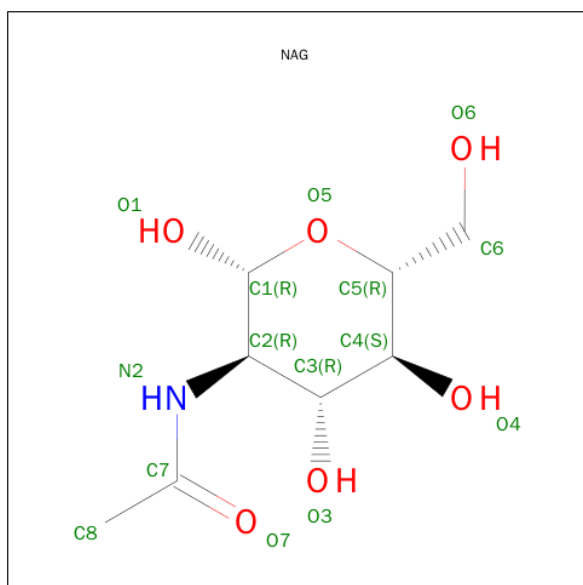
- Molecule 3 is a polymer of unknown type called SUGAR (3-MER).

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

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		
4	C	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	K	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	I	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	69	Total	O	0	0
			69	69		
7	B	32	Total	O	0	0
			32	32		

Continued on next page...

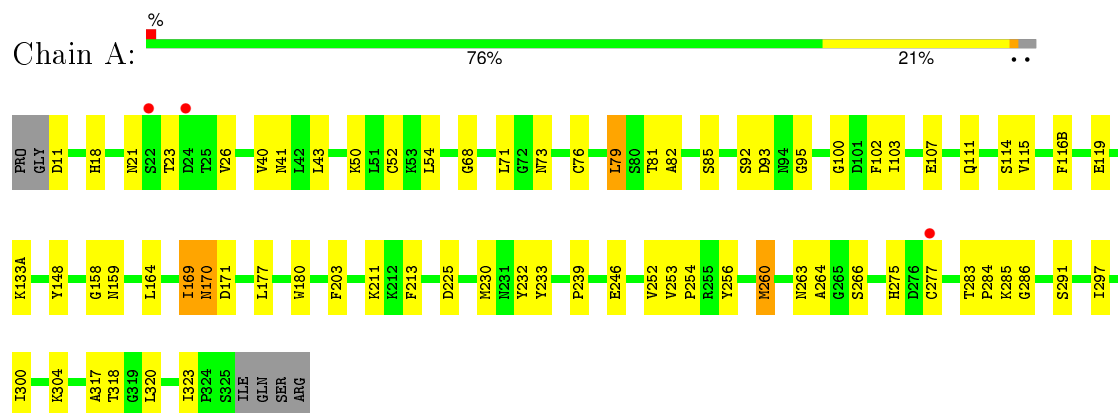
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	54	Total 54	O 54	0	0
7	D	23	Total 23	O 23	0	0
7	E	60	Total 60	O 60	0	0
7	F	23	Total 23	O 23	0	0
7	G	59	Total 59	O 59	0	0
7	H	24	Total 24	O 24	0	0
7	I	39	Total 39	O 39	0	0
7	J	20	Total 20	O 20	0	0
7	K	34	Total 34	O 34	0	0
7	L	26	Total 26	O 26	0	0

3 Residue-property plots [i](#)

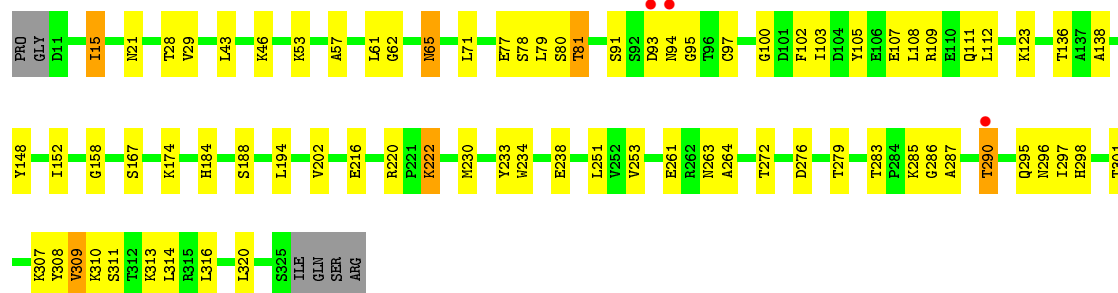
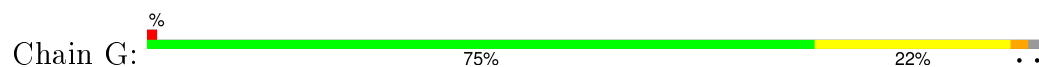
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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, HA1 SUBUNIT

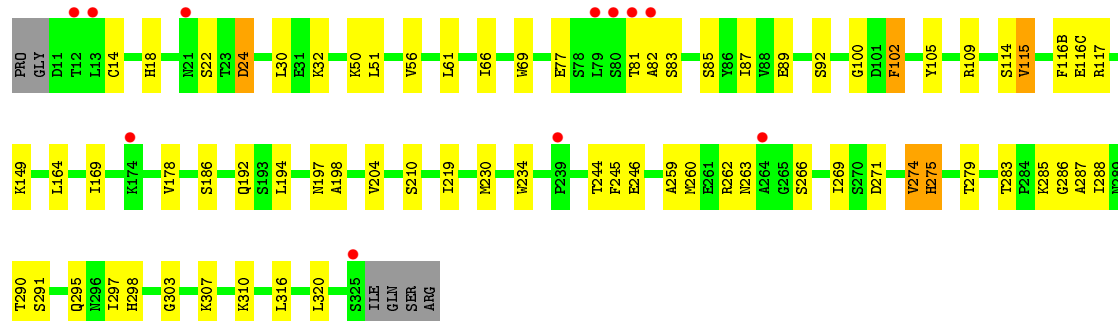
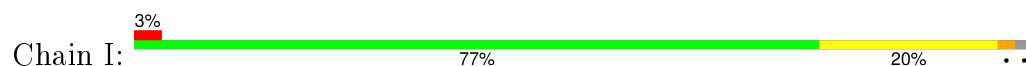




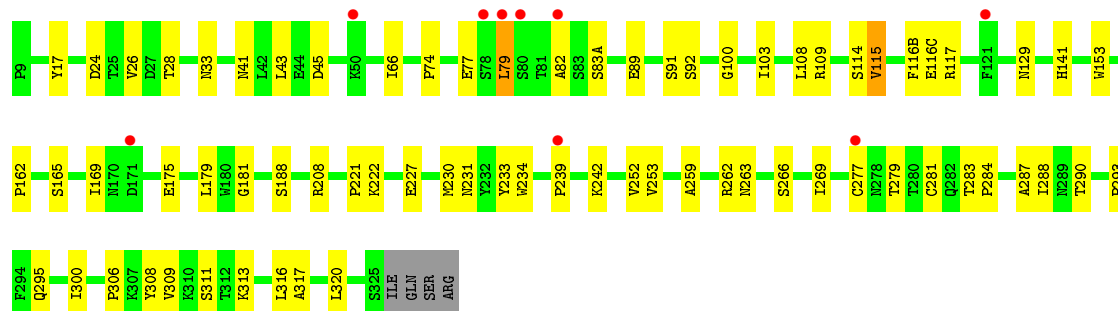
• Molecule 1: Hemagglutinin, HA1 SUBUNIT



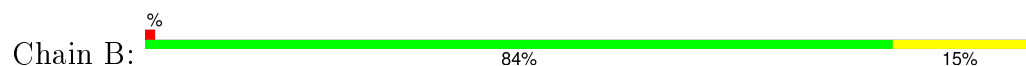
• Molecule 1: Hemagglutinin, HA1 SUBUNIT



• Molecule 1: Hemagglutinin, HA1 SUBUNIT

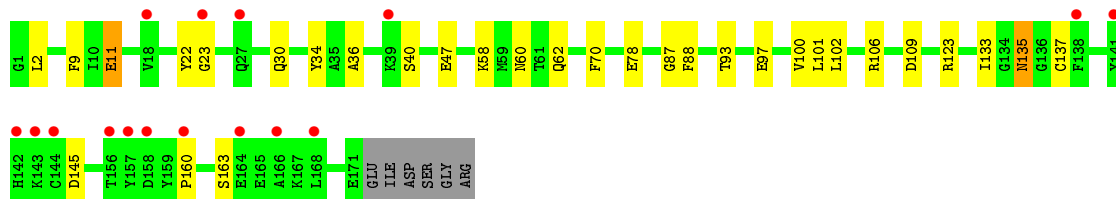
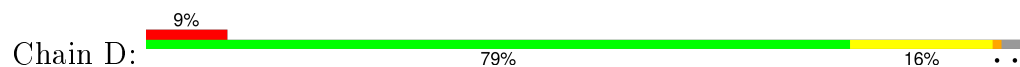


• Molecule 2: Hemagglutinin, HA2 SUBUNIT

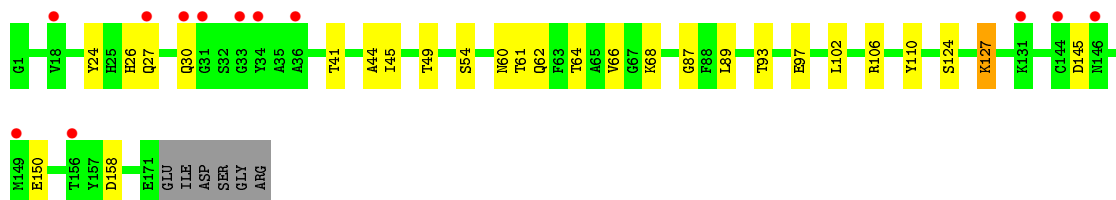
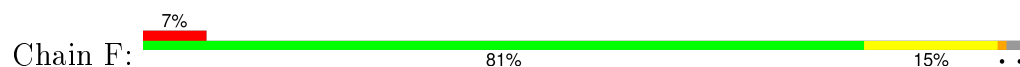




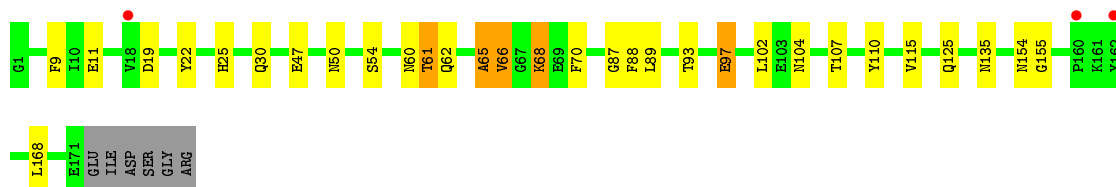
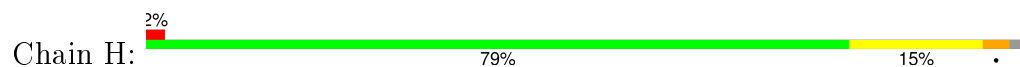
• Molecule 2: Hemagglutinin, HA2 SUBUNIT



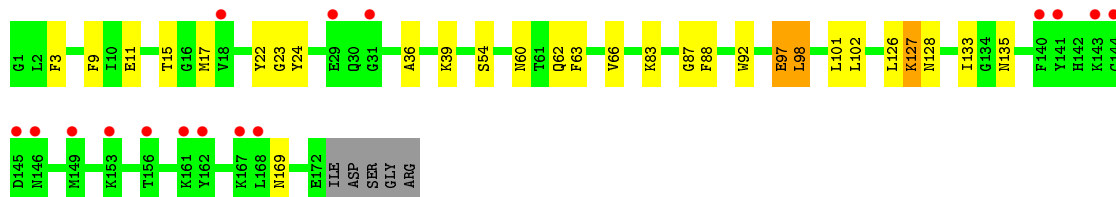
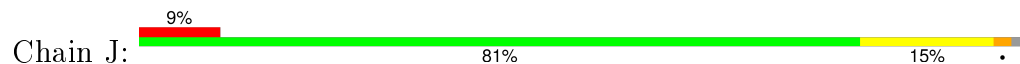
• Molecule 2: Hemagglutinin, HA2 SUBUNIT



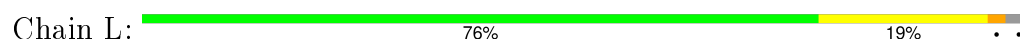
• Molecule 2: Hemagglutinin, HA2 SUBUNIT

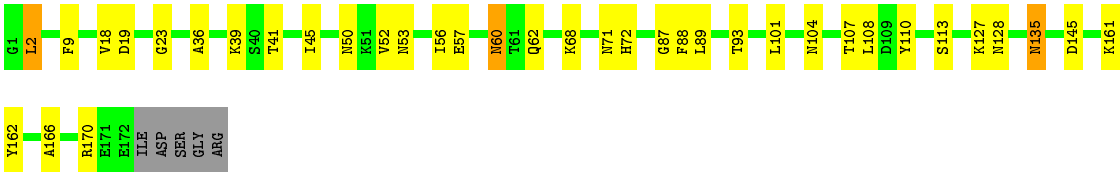


• Molecule 2: Hemagglutinin, HA2 SUBUNIT



• Molecule 2: Hemagglutinin, HA2 SUBUNIT





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	67.68Å 118.31Å 120.07Å 117.37° 92.67° 100.84°	Depositor
Resolution (Å)	45.00 – 2.60 43.52 – 2.60	Depositor EDS
% Data completeness (in resolution range)	91.7 (45.00-2.60) 75.7 (43.52-2.60)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.181 , 0.252 0.188 , 0.263	Depositor DCC
R_{free} test set	4548 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	50.3	Xtriage
Anisotropy	0.277	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 53.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 90481 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	24137	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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, 0.375 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: BMA, NAG, 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	0.55	0/2589	0.84	0/3519
1	C	0.52	0/2589	0.83	1/3519 (0.0%)
1	E	0.51	0/2589	0.81	0/3519
1	G	0.54	0/2589	0.86	0/3519
1	I	0.49	0/2589	0.80	0/3519
1	K	0.48	0/2601	0.76	0/3535
2	B	0.50	0/1439	0.75	0/1939
2	D	0.50	0/1408	0.72	0/1897
2	F	0.50	0/1408	0.73	0/1897
2	H	0.50	0/1408	0.77	0/1897
2	J	0.48	0/1417	0.73	0/1909
2	L	0.51	0/1417	0.78	0/1909
All	All	0.51	0/24043	0.79	1/32578 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	289	ASN	C-N-CA	5.09	134.43	121.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2525	0	2474	30	0
1	C	2525	0	2475	33	0
1	E	2525	0	2475	45	0
1	G	2525	0	2476	30	0
1	I	2525	0	2475	36	0
1	K	2536	0	2486	34	0
2	B	1411	0	1329	14	0
2	D	1380	0	1303	16	0
2	F	1380	0	1303	12	0
2	H	1380	0	1303	21	0
2	J	1389	0	1309	13	0
2	L	1389	0	1309	19	0
3	A	39	0	34	1	0
4	A	28	0	25	0	0
4	C	28	0	25	2	0
5	E	14	0	13	1	0
5	K	14	0	13	0	0
6	I	61	0	52	2	0
7	A	69	0	0	0	0
7	B	32	0	0	0	0
7	C	54	0	0	0	0
7	D	23	0	0	0	0
7	E	60	0	0	0	0
7	F	23	0	0	0	0
7	G	59	0	0	0	0
7	H	24	0	0	2	0
7	I	39	0	0	1	0
7	J	20	0	0	0	0
7	K	34	0	0	1	0
7	L	26	0	0	0	0
All	All	24137	0	22879	264	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.

The worst 5 of 264 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:E:279:THR:HG21	1:E:287:ALA:HB1	1.28	1.11
1:E:290:THR:HG21	1:E:304:LYS:O	1.75	0.87
1:A:283:THR:HB	1:A:286:GLY:O	1.75	0.87
1:C:283:THR:HB	1:C:286:GLY:O	1.78	0.83

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:114:SER:HB2	1:I:266:SER:HB2	1.68	0.76

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	321/329 (98%)	295 (92%)	18 (6%)	8 (2%)	7	12
1	C	321/329 (98%)	294 (92%)	20 (6%)	7 (2%)	8	15
1	E	321/329 (98%)	292 (91%)	24 (8%)	5 (2%)	12	24
1	G	321/329 (98%)	294 (92%)	18 (6%)	9 (3%)	6	10
1	I	321/329 (98%)	293 (91%)	21 (6%)	7 (2%)	8	15
1	K	323/329 (98%)	293 (91%)	26 (8%)	4 (1%)	16	33
2	B	173/177 (98%)	160 (92%)	12 (7%)	1 (1%)	30	56
2	D	169/177 (96%)	149 (88%)	17 (10%)	3 (2%)	11	21
2	F	169/177 (96%)	151 (89%)	15 (9%)	3 (2%)	11	21
2	H	169/177 (96%)	153 (90%)	11 (6%)	5 (3%)	5	8
2	J	170/177 (96%)	161 (95%)	5 (3%)	4 (2%)	7	13
2	L	170/177 (96%)	158 (93%)	11 (6%)	1 (1%)	30	56
All	All	2948/3036 (97%)	2693 (91%)	198 (7%)	57 (2%)	10	19

5 of 57 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	ALA
1	A	92	SER
1	A	170	ASN
1	A	264	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	77	GLU

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	284/289 (98%)	268 (94%)	16 (6%)	26	50
1	C	284/289 (98%)	268 (94%)	16 (6%)	26	50
1	E	284/289 (98%)	267 (94%)	17 (6%)	24	47
1	G	284/289 (98%)	253 (89%)	31 (11%)	8	14
1	I	284/289 (98%)	267 (94%)	17 (6%)	24	47
1	K	285/289 (99%)	269 (94%)	16 (6%)	26	50
2	B	151/152 (99%)	141 (93%)	10 (7%)	21	40
2	D	147/152 (97%)	138 (94%)	9 (6%)	23	46
2	F	147/152 (97%)	136 (92%)	11 (8%)	17	33
2	H	147/152 (97%)	140 (95%)	7 (5%)	31	58
2	J	148/152 (97%)	136 (92%)	12 (8%)	15	28
2	L	148/152 (97%)	134 (90%)	14 (10%)	11	20
All	All	2593/2646 (98%)	2417 (93%)	176 (7%)	20	39

5 of 176 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	15	ILE
1	G	261	GLU
2	L	2	LEU
1	G	29	VAL
1	G	136	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	J	62	GLN

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 ⓘ

12 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	330	1,3	14,14,15	1.70	2 (14%)	15,19,21	2.51	4 (26%)
3	NAG	A	331	3	14,14,15	1.53	1 (7%)	15,19,21	1.39	3 (20%)
3	BMA	A	332	3	11,11,12	1.82	2 (18%)	14,15,17	1.98	4 (28%)
4	NAG	A	333	1,4	14,14,15	1.83	5 (35%)	15,19,21	2.48	4 (26%)
4	NAG	A	334	4	14,14,15	2.20	3 (21%)	15,19,21	1.57	3 (20%)
4	NAG	C	330	1,4	14,14,15	1.02	1 (7%)	15,19,21	1.35	4 (26%)
4	NAG	C	331	4	14,14,15	1.85	4 (28%)	15,19,21	2.98	5 (33%)
6	NAG	I	501	1,6	14,14,15	1.81	5 (35%)	15,19,21	1.29	2 (13%)
6	NAG	I	502	6	14,14,15	1.18	1 (7%)	15,19,21	2.70	3 (20%)
6	BMA	I	503	6	11,11,12	1.60	3 (27%)	14,15,17	3.57	6 (42%)
6	MAN	I	504	6	11,11,12	2.55	7 (63%)	14,15,17	2.67	7 (50%)
6	MAN	I	505	6	11,11,12	3.51	5 (45%)	14,15,17	2.66	5 (35%)

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	330	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	331	3	-	0/6/23/26	0/1/1/1
3	BMA	A	332	3	-	0/2/19/22	0/1/1/1
4	NAG	A	333	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	334	4	-	0/6/23/26	0/1/1/1
4	NAG	C	330	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	331	4	-	0/6/23/26	0/1/1/1
6	NAG	I	501	1,6	-	0/6/23/26	0/1/1/1
6	NAG	I	502	6	-	0/6/23/26	0/1/1/1
6	BMA	I	503	6	-	0/2/19/22	0/1/1/1
6	MAN	I	504	6	-	0/2/19/22	0/1/1/1
6	MAN	I	505	6	-	0/2/19/22	0/1/1/1

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	I	504	MAN	O3-C3	2.01	1.47	1.43
6	I	501	NAG	O4-C4	2.05	1.47	1.43
6	I	501	NAG	C4-C3	2.16	1.58	1.52
6	I	502	NAG	C1-C2	2.24	1.55	1.52
4	C	330	NAG	C1-C2	2.25	1.55	1.52

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	331	NAG	C4-C3-C2	-6.37	101.33	111.23
3	A	330	NAG	O4-C4-C3	-3.01	103.57	110.34
6	I	503	BMA	C1-C2-C3	-2.98	106.01	109.54
3	A	332	BMA	C3-C4-C5	-2.94	105.06	110.20
6	I	504	MAN	O6-C6-C5	-2.79	102.11	111.33

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	330	NAG	1	0
3	A	331	NAG	1	0
4	C	330	NAG	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	I	502	NAG	1	0
6	I	503	BMA	1	0
6	I	505	MAN	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$
5	NAG	E	330	1	14,14,15	1.74	3 (21%)	15,19,21	1.60	4 (26%)
5	NAG	K	601	1	14,14,15	1.52	4 (28%)	15,19,21	2.44	6 (40%)

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
5	NAG	E	330	1	-	0/6/23/26	0/1/1/1
5	NAG	K	601	1	-	0/6/23/26	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	K	601	NAG	O5-C1	2.16	1.47	1.43
5	K	601	NAG	C4-C5	2.28	1.57	1.53
5	K	601	NAG	O5-C5	2.33	1.48	1.43
5	E	330	NAG	O4-C4	2.70	1.49	1.43
5	E	330	NAG	C4-C5	2.89	1.59	1.53

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
5	K	601	NAG	C4-C3-C2	-3.55	105.70	111.23
5	K	601	NAG	C3-C4-C5	-2.47	105.89	110.20
5	E	330	NAG	O5-C5-C6	-2.12	102.75	107.35
5	E	330	NAG	C4-C3-C2	2.07	114.45	111.23
5	K	601	NAG	O3-C3-C4	2.16	115.20	110.34

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	E	330	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	323/329 (98%)	-0.30	3 (0%) 85 83	32, 50, 74, 95	0
1	C	323/329 (98%)	-0.09	11 (3%) 49 41	33, 61, 92, 122	0
1	E	323/329 (98%)	-0.17	6 (1%) 70 64	35, 58, 84, 118	0
1	G	323/329 (98%)	-0.24	3 (0%) 85 83	35, 55, 77, 100	0
1	I	323/329 (98%)	-0.08	11 (3%) 49 41	44, 64, 91, 108	0
1	K	325/329 (98%)	0.07	9 (2%) 56 49	52, 74, 99, 133	0
2	B	175/177 (98%)	-0.08	1 (0%) 90 88	37, 63, 85, 100	0
2	D	171/177 (96%)	0.43	16 (9%) 11 6	39, 73, 126, 148	0
2	F	171/177 (96%)	0.38	12 (7%) 19 13	38, 74, 119, 144	0
2	H	171/177 (96%)	0.08	3 (1%) 71 66	46, 65, 88, 98	0
2	J	172/177 (97%)	0.40	16 (9%) 11 7	45, 72, 109, 120	0
2	L	172/177 (97%)	-0.11	0 100 100	47, 63, 83, 103	0
All	All	2972/3036 (97%)	-0.02	91 (3%) 52 45	32, 62, 100, 148	0

The worst 5 of 91 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	81	THR	6.8
1	G	94	ASN	5.9
1	K	79	LEU	5.8
1	K	82	ALA	5.4
2	D	144	CYS	5.3

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 [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	A	333	14/15	0.88	0.29	1.55	84,88,92,93	0
4	NAG	C	330	14/15	0.92	0.15	-0.46	83,84,87,88	0
3	NAG	A	330	14/15	0.95	0.12	-0.70	58,62,66,68	0
6	NAG	I	501	14/15	0.95	0.12	-2.02	69,72,77,78	0
6	BMA	I	503	11/12	0.80	0.15	-	79,82,84,85	0
3	BMA	A	332	11/12	0.79	0.16	-	85,88,92,93	0
6	MAN	I	504	11/12	0.71	0.16	-	81,83,88,89	0
4	NAG	C	331	14/15	0.83	0.25	-	87,89,96,96	0
3	NAG	A	331	14/15	0.89	0.14	-	74,79,82,83	0
6	MAN	I	505	11/12	0.54	0.26	-	84,86,88,91	0
6	NAG	I	502	14/15	0.95	0.11	-	71,74,78,81	0
4	NAG	A	334	14/15	0.79	0.36	-	94,98,102,103	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
5	NAG	K	601	14/15	0.91	0.17	-0.16	133,137,144,144	0
5	NAG	E	330	14/15	0.93	0.16	-0.37	91,95,101,101	0

6.5 Other polymers [i](#)

There are no such residues in this entry.