



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:55 PM GMT

PDB ID : 4JU0
Title : Crystal structure of 2009 pandemic influenza virus hemagglutinin mutant D225E complexed with human receptor analogue LSTc
Authors : Zhang, W.; Shi, Y.; Qi, J.; Gao, F.; Li, Q.; Fan, Z.; Yan, J.; Gao, G.F.
Deposited on : 2013-03-24
Resolution : 2.91 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

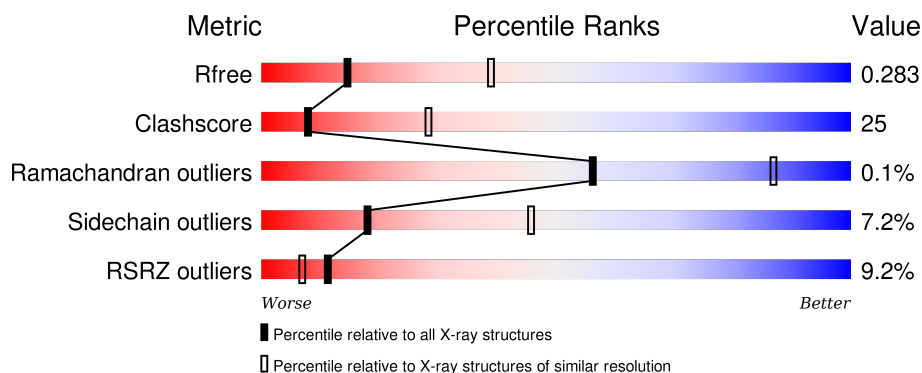
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.91 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	322	<div> <div>7%</div> <div>64%</div> <div>31%</div> <div>5%</div> </div>
1	C	322	<div> <div>6%</div> <div>59%</div> <div>35%</div> <div>•</div> </div>
1	E	322	<div> <div>5%</div> <div>62%</div> <div>33%</div> <div>5%</div> </div>
1	G	322	<div> <div>6%</div> <div>63%</div> <div>33%</div> <div>•</div> </div>
1	I	322	<div> <div>7%</div> <div>60%</div> <div>35%</div> <div>•</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	322	<div><div></div><div>7%</div><div>61%</div><div>34%</div><div></div><div>.</div></div>
2	B	164	<div><div></div><div>12%</div><div>66%</div><div>31%</div><div></div><div>..</div></div>
2	D	164	<div><div></div><div>10%</div><div>62%</div><div>36%</div><div></div><div>.</div></div>
2	F	164	<div><div></div><div>18%</div><div>56%</div><div>38%</div><div></div><div>..</div></div>
2	H	164	<div><div></div><div>4%</div><div>65%</div><div>32%</div><div></div><div>..</div></div>
2	J	164	<div><div></div><div>27%</div><div>49%</div><div>46%</div><div></div><div>.</div></div>
2	L	164	<div><div></div><div>19%</div><div>61%</div><div>33%</div><div></div><div>..</div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 23507 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	322	Total	C	N	O	S	0	0	0
			2513	1590	434	478	11			
1	C	321	Total	C	N	O	S	0	0	0
			2510	1587	433	479	11			
1	E	321	Total	C	N	O	S	0	0	0
			2510	1587	433	479	11			
1	G	322	Total	C	N	O	S	0	0	0
			2513	1590	434	478	11			
1	I	321	Total	C	N	O	S	0	0	0
			2510	1587	433	479	11			
1	K	321	Total	C	N	O	S	0	0	0
			2510	1587	433	479	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	228	GLU	ASP	ENGINEERED MUTATION	UNP C3W5S1
C	228	GLU	ASP	ENGINEERED MUTATION	UNP C3W5S1
E	228	GLU	ASP	ENGINEERED MUTATION	UNP C3W5S1
G	228	GLU	ASP	ENGINEERED MUTATION	UNP C3W5S1
I	228	GLU	ASP	ENGINEERED MUTATION	UNP C3W5S1
K	228	GLU	ASP	ENGINEERED MUTATION	UNP C3W5S1

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	162	Total	C	N	O	S	0	0	0
			1305	822	220	257	6			
2	D	164	Total	C	N	O	S	0	0	0
			1315	826	221	262	6			
2	F	161	Total	C	N	O	S	0	0	0
			1302	821	219	256	6			

Continued on next page...

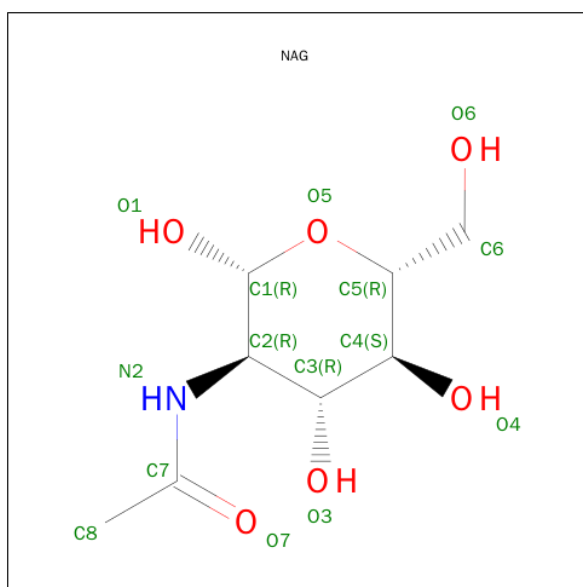
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	162	Total	C	N	O	S	0	0	0
			1305	822	220	257	6			
2	J	164	Total	C	N	O	S	0	0	0
			1315	826	221	262	6			
2	L	161	Total	C	N	O	S	0	0	0
			1302	821	219	256	6			

- Molecule 3 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	4	Total	C	N	O	0	0
			56	31	2	23		
3	C	4	Total	C	N	O	0	0
			56	31	2	23		
3	E	4	Total	C	N	O	0	0
			56	31	2	23		
3	G	4	Total	C	N	O	0	0
			56	31	2	23		
3	I	4	Total	C	N	O	0	0
			56	31	2	23		

- 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	C	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

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

- | Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---|----|---------|---------|
| 5 | E | 2 | Total | C | N | O | 0 | 0 |
| | | | 28 | 16 | 2 | 10 | | |

-
- Chemical structure of SIA (Sialic acid) is shown, highlighting the stereochemistry of the chiral centers. The structure includes a carboxylic acid group (C1, O1A, O1B), a hydroxyl group (C2, O2), a hydroxyl group (C3, O4), a hydroxyl group (C4, O5), a hydroxyl group (C5, O6), a hydroxyl group (C6, O7), a hydroxyl group (C7, O8), and a hydroxyl group (C8, O9). The stereochemistry is indicated by wedged and dashed bonds.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	K	1	Total	C	N	O	0	0
			20	11	1	8		

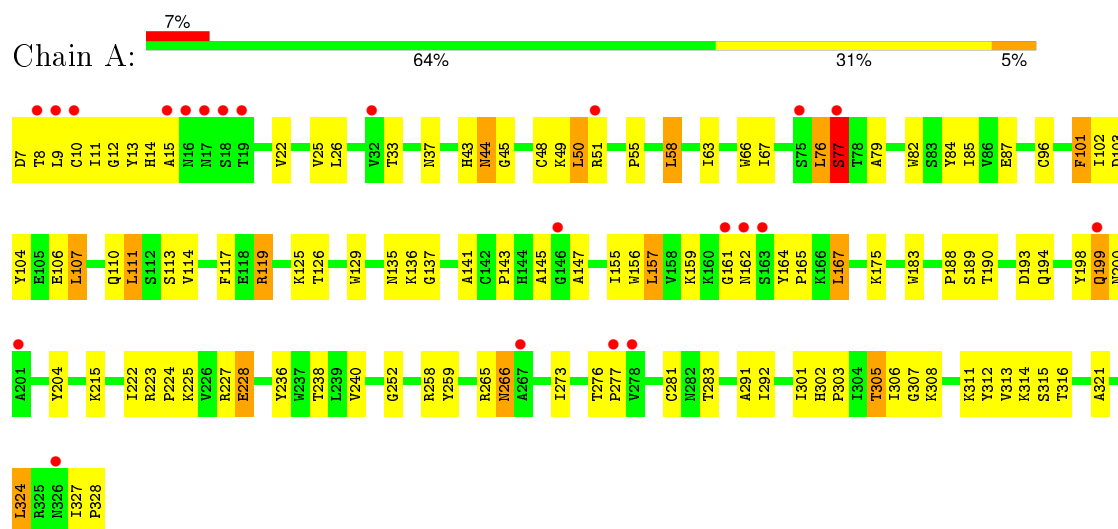
- 

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	22	Total 22	O 22	0	0
7	B	6	Total 6	O 6	0	0
7	C	20	Total 20	O 20	0	0
7	D	5	Total 5	O 5	0	0
7	E	18	Total 18	O 18	0	0
7	F	11	Total 11	O 11	0	0
7	G	16	Total 16	O 16	0	0
7	H	6	Total 6	O 6	0	0
7	I	25	Total 25	O 25	0	0
7	J	22	Total 22	O 22	0	0
7	K	26	Total 26	O 26	0	0
7	L	8	Total 8	O 8	0	0

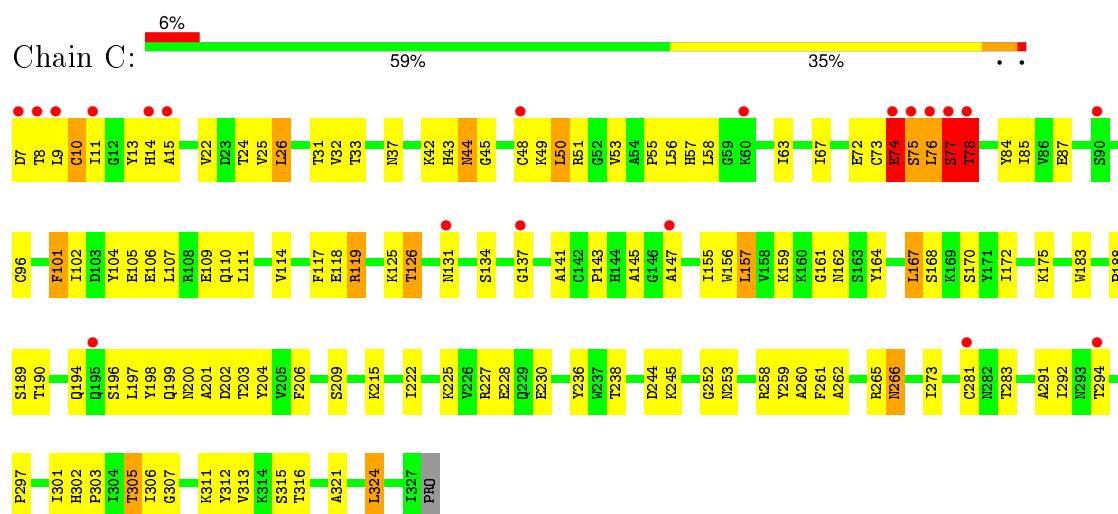
3 Residue-property plots [i](#)

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 ($\text{RSRZ} > 2$). 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

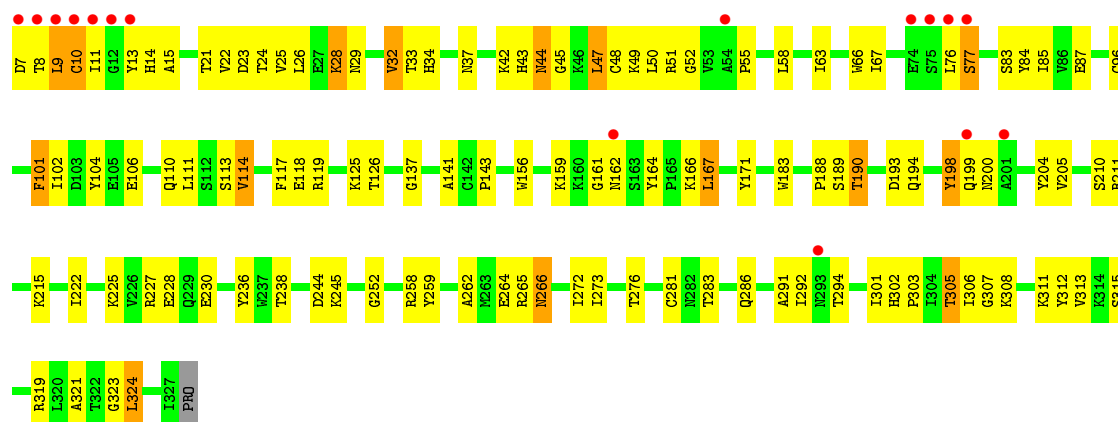


• Molecule 1: Hemagglutinin

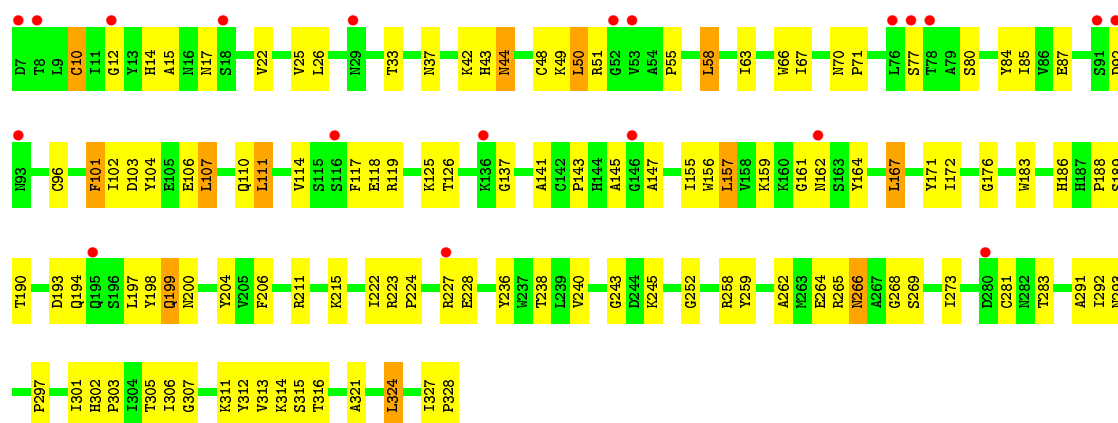


• Molecule 1: Hemagglutinin

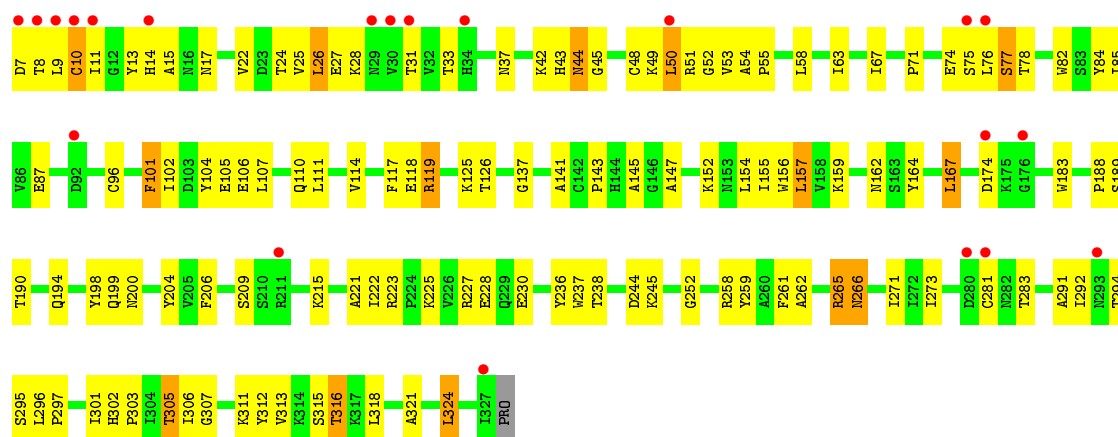




• Molecule 1: Hemagglutinin

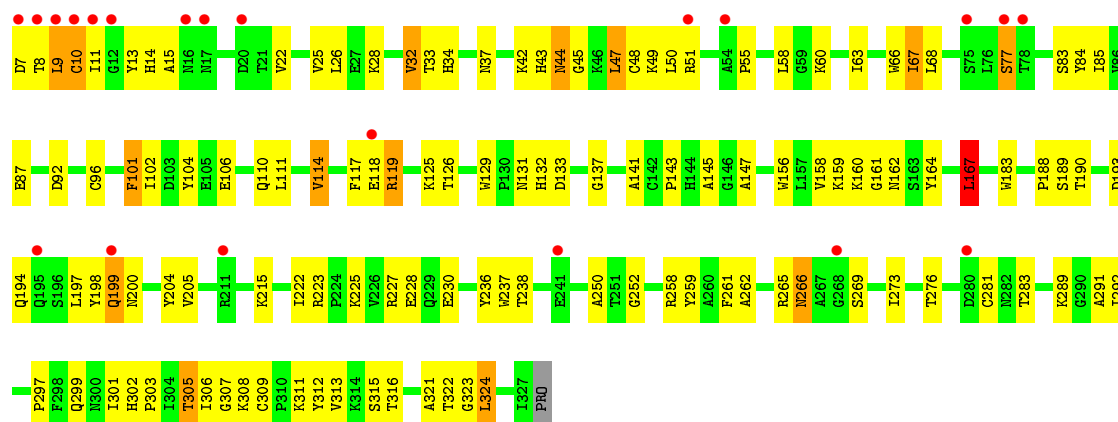


• Molecule 1: Hemagglutinin

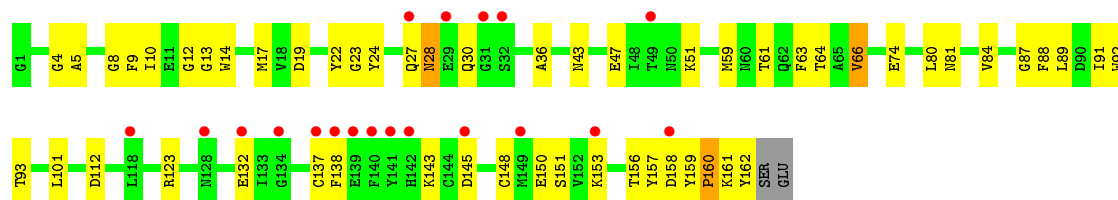


• Molecule 1: Hemagglutinin

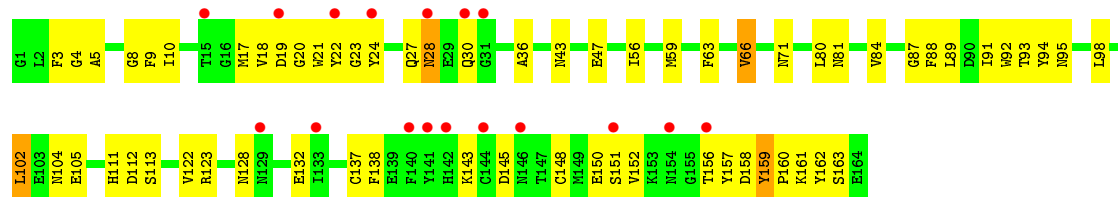




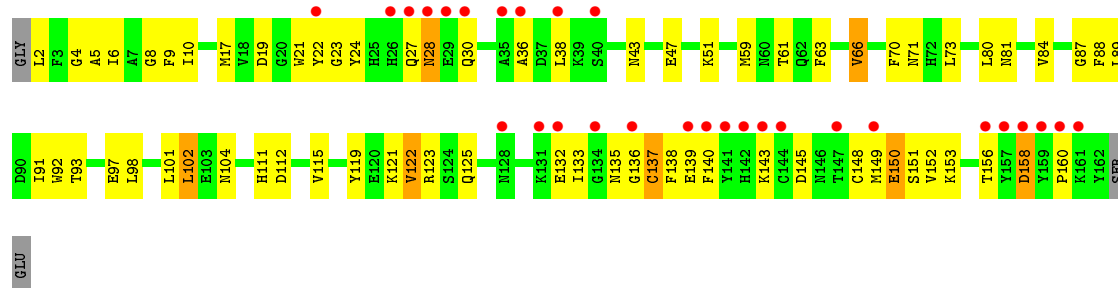
• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin

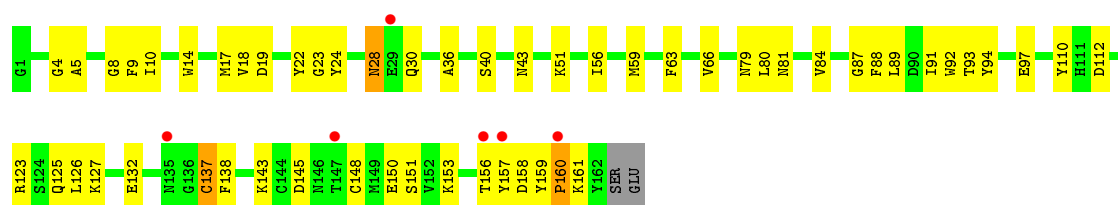


• Molecule 2: Hemagglutinin

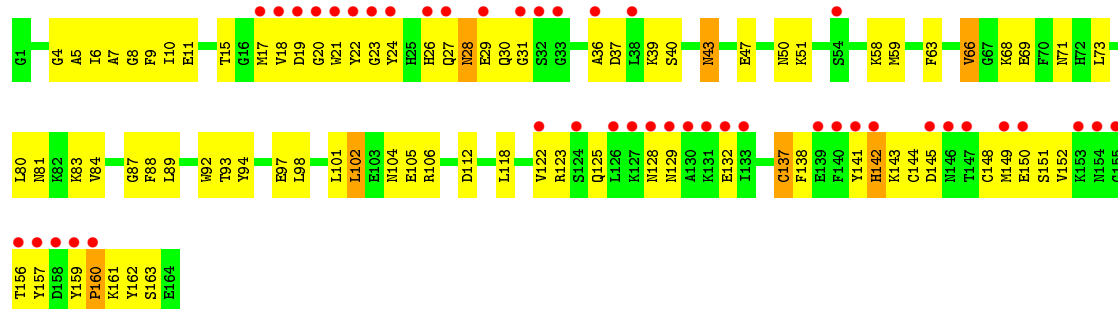


• Molecule 2: Hemagglutinin

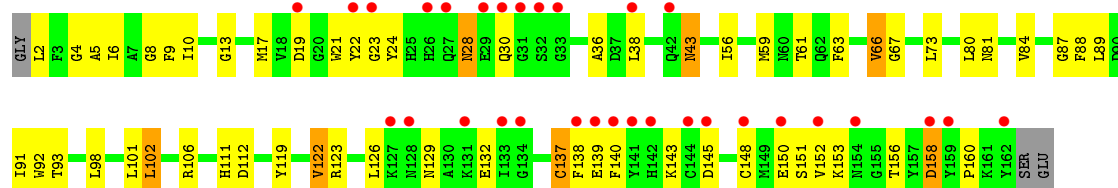




• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	66.87Å 116.81Å 116.50Å 62.06° 77.76° 81.54°	Depositor
Resolution (Å)	40.22 – 2.91 40.23 – 2.91	Depositor EDS
% Data completeness (in resolution range)	88.2 (40.22-2.91) 82.1 (40.23-2.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.59 (at 2.90Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.240 , 0.283 0.243 , 0.283	Depositor DCC
R_{free} test set	3175 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	54.6	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 53.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 62950 reflections	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	23507	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% 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: SIA, GAL, 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	0.25	0/2577	0.71	8/3501 (0.2%)
1	C	0.29	1/2573 (0.0%)	0.68	9/3494 (0.3%)
1	E	0.24	0/2573	0.60	8/3494 (0.2%)
1	G	0.26	0/2577	0.70	6/3501 (0.2%)
1	I	0.25	0/2573	0.60	6/3494 (0.2%)
1	K	0.24	0/2573	0.65	8/3494 (0.2%)
2	B	0.24	0/1333	0.54	1/1797 (0.1%)
2	D	0.26	0/1343	0.48	1/1811 (0.1%)
2	F	0.24	0/1330	0.49	0/1794
2	H	0.24	0/1333	0.50	1/1797 (0.1%)
2	J	0.24	0/1343	0.47	1/1811 (0.1%)
2	L	0.24	0/1330	0.49	0/1794
All	All	0.25	1/23458 (0.0%)	0.61	49/31782 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	E	0	1
1	K	0	1
2	J	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	73	CYS	CB-SG	-5.76	1.72	1.81

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	119	ARG	NE-CZ-NH1	-16.32	112.14	120.30
1	A	119	ARG	NE-CZ-NH1	-16.20	112.20	120.30
1	A	119	ARG	NE-CZ-NH2	16.09	128.34	120.30
1	G	119	ARG	NE-CZ-NH2	15.81	128.21	120.30
1	K	199	GLN	CB-CA-C	-14.58	81.24	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	77	SER	Peptide
2	J	141	TYR	Peptide
1	K	77	SER	Peptide

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	2513	0	2457	118	0
1	C	2510	0	2452	160	0
1	E	2510	0	2453	130	0
1	G	2513	0	2455	117	0
1	I	2510	0	2452	166	0
1	K	2510	0	2452	150	0
2	B	1305	0	1228	54	0
2	D	1315	0	1227	70	0
2	F	1302	0	1226	84	0
2	H	1305	0	1228	59	0
2	J	1315	0	1227	155	0
2	L	1302	0	1226	81	0
3	A	56	0	47	3	0
3	C	56	0	47	4	0
3	E	56	0	47	3	0
3	G	56	0	47	5	0
3	I	56	0	47	2	0
4	C	28	0	26	5	0
4	G	14	0	13	0	0
4	I	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	K	28	0	26	9	0
5	E	28	0	25	1	0
6	K	20	0	17	1	0
7	A	22	0	0	32	0
7	B	6	0	0	3	0
7	C	20	0	0	40	0
7	D	5	0	0	14	0
7	E	18	0	0	32	0
7	F	11	0	0	23	0
7	G	16	0	0	15	0
7	H	6	0	0	12	0
7	I	25	0	0	50	0
7	J	22	0	0	54	0
7	K	26	0	0	51	0
7	L	8	0	0	17	0
All	All	23507	0	22438	1137	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 25.

The worst 5 of 1137 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:56:LEU:CD1	1:C:76:LEU:HD12	1.48	1.42
1:C:56:LEU:HD12	1:C:76:LEU:CD1	1.49	1.40
1:C:8:THR:HB	7:D:204:HOH:O	1.18	1.31
1:C:110:GLN:HB2	7:C:714:HOH:O	1.20	1.29
1:G:245:LYS:HA	7:G:516:HOH:O	1.14	1.28

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	320/322 (99%)	296 (92%)	24 (8%)	0	100	100
1	C	319/322 (99%)	297 (93%)	22 (7%)	0	100	100
1	E	319/322 (99%)	297 (93%)	22 (7%)	0	100	100
1	G	320/322 (99%)	298 (93%)	22 (7%)	0	100	100
1	I	319/322 (99%)	298 (93%)	21 (7%)	0	100	100
1	K	319/322 (99%)	297 (93%)	22 (7%)	0	100	100
2	B	160/164 (98%)	149 (93%)	10 (6%)	1 (1%)	30	67
2	D	162/164 (99%)	150 (93%)	11 (7%)	1 (1%)	30	67
2	F	159/164 (97%)	148 (93%)	11 (7%)	0	100	100
2	H	160/164 (98%)	149 (93%)	10 (6%)	1 (1%)	30	67
2	J	162/164 (99%)	149 (92%)	12 (7%)	1 (1%)	30	67
2	L	159/164 (97%)	147 (92%)	12 (8%)	0	100	100
All	All	2878/2916 (99%)	2675 (93%)	199 (7%)	4 (0%)	56	87

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	160	PRO
2	H	160	PRO
2	J	160	PRO
2	D	159	TYR

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	282/283 (100%)	259 (92%)	23 (8%)	14	39
1	C	282/283 (100%)	260 (92%)	22 (8%)	16	41
1	E	282/283 (100%)	262 (93%)	20 (7%)	18	47
1	G	282/283 (100%)	263 (93%)	19 (7%)	20	50
1	I	282/283 (100%)	261 (93%)	21 (7%)	17	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	282/283 (100%)	263 (93%)	19 (7%)	20	50
2	B	139/141 (99%)	132 (95%)	7 (5%)	30	65
2	D	140/141 (99%)	131 (94%)	9 (6%)	22	53
2	F	139/141 (99%)	126 (91%)	13 (9%)	11	32
2	H	139/141 (99%)	132 (95%)	7 (5%)	30	65
2	J	140/141 (99%)	131 (94%)	9 (6%)	22	53
2	L	139/141 (99%)	126 (91%)	13 (9%)	11	32
All	All	2528/2544 (99%)	2346 (93%)	182 (7%)	18	46

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	22	TYR
1	G	107	LEU
1	K	324	LEU
2	F	38	LEU
2	F	150	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	30	GLN
1	G	195	GLN
1	K	266	ASN
2	F	125	GLN
1	G	44	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

22 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	SIA	A	801	3	16,20,21	0.32	0	18,28,31	2.19	1 (5%)
3	GAL	A	802	3	11,11,12	0.64	0	14,15,17	1.04	1 (7%)
3	NAG	A	803	3	14,14,15	0.62	0	15,19,21	0.63	0
3	GAL	A	804	3	11,11,12	0.62	0	14,15,17	0.66	0
3	SIA	C	603	3	16,20,21	0.37	0	18,28,31	2.20	1 (5%)
3	GAL	C	604	3	11,11,12	0.61	0	14,15,17	0.85	0
3	NAG	C	605	3	14,14,15	0.56	0	15,19,21	0.83	0
3	GAL	C	606	3	11,11,12	0.62	0	14,15,17	0.59	0
5	NAG	E	601	1,5	14,14,15	0.56	0	15,19,21	0.60	0
5	NAG	E	602	5	14,14,15	0.45	0	15,19,21	0.77	0
3	SIA	E	603	3	16,20,21	0.25	0	18,28,31	2.41	2 (11%)
3	GAL	E	604	3	11,11,12	0.69	0	14,15,17	0.87	0
3	NAG	E	605	3	14,14,15	0.61	0	15,19,21	0.65	0
3	GAL	E	606	3	11,11,12	0.60	0	14,15,17	0.63	0
3	SIA	G	402	3	16,20,21	0.39	0	18,28,31	2.15	1 (5%)
3	GAL	G	403	3	11,11,12	0.70	0	14,15,17	1.10	1 (7%)
3	NAG	G	404	3	14,14,15	0.56	0	15,19,21	0.80	0
3	GAL	G	405	3	11,11,12	0.70	0	14,15,17	1.30	2 (14%)
3	SIA	I	602	3	16,20,21	0.39	0	18,28,31	2.03	2 (11%)
3	GAL	I	603	3	11,11,12	0.66	0	14,15,17	1.25	2 (14%)
3	NAG	I	604	3	14,14,15	0.54	0	15,19,21	0.58	0
3	GAL	I	605	3	11,11,12	0.66	0	14,15,17	1.50	3 (21%)

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	SIA	A	801	3	-	0/14/34/38	0/1/1/1
3	GAL	A	802	3	-	0/2/19/22	0/1/1/1
3	NAG	A	803	3	-	0/6/23/26	0/1/1/1
3	GAL	A	804	3	-	0/2/19/22	0/1/1/1
3	SIA	C	603	3	-	0/14/34/38	0/1/1/1
3	GAL	C	604	3	-	0/2/19/22	0/1/1/1
3	NAG	C	605	3	-	0/6/23/26	0/1/1/1
3	GAL	C	606	3	-	0/2/19/22	0/1/1/1
5	NAG	E	601	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	602	5	-	0/6/23/26	0/1/1/1
3	SIA	E	603	3	-	0/14/34/38	0/1/1/1
3	GAL	E	604	3	-	0/2/19/22	0/1/1/1
3	NAG	E	605	3	-	0/6/23/26	0/1/1/1
3	GAL	E	606	3	-	0/2/19/22	0/1/1/1
3	SIA	G	402	3	-	0/14/34/38	0/1/1/1
3	GAL	G	403	3	-	0/2/19/22	0/1/1/1
3	NAG	G	404	3	-	0/6/23/26	0/1/1/1
3	GAL	G	405	3	-	0/2/19/22	0/1/1/1
3	SIA	I	602	3	-	0/14/34/38	0/1/1/1
3	GAL	I	603	3	-	0/2/19/22	0/1/1/1
3	NAG	I	604	3	-	0/6/23/26	0/1/1/1
3	GAL	I	605	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	603	SIA	O6-C2-C3	-9.77	91.08	109.86
3	C	603	SIA	O6-C2-C3	-8.89	92.76	109.86
3	A	801	SIA	O6-C2-C3	-8.88	92.80	109.86
3	G	402	SIA	O6-C2-C3	-8.71	93.11	109.86
3	I	602	SIA	O6-C2-C3	-7.81	94.85	109.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

11 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	801	SIA	2	0
3	A	803	NAG	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	603	SIA	2	0
3	C	605	NAG	2	0
5	E	602	NAG	1	0
3	E	603	SIA	3	0
3	E	605	NAG	1	0
3	G	402	SIA	5	0
3	G	404	NAG	2	0
3	I	602	SIA	1	0
3	I	603	GAL	1	0

5.6 Ligand geometry

7 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	C	601	1	14,14,15	0.76	0	15,19,21	1.18	1 (6%)
4	NAG	C	602	1	14,14,15	0.50	0	15,19,21	0.87	1 (6%)
4	NAG	G	401	1	14,14,15	0.45	0	15,19,21	1.13	1 (6%)
4	NAG	I	601	1	14,14,15	0.57	0	15,19,21	0.88	1 (6%)
4	NAG	K	601	1	14,14,15	0.45	0	15,19,21	1.94	4 (26%)
4	NAG	K	602	1	14,14,15	0.64	0	15,19,21	0.66	0
6	SIA	K	603	-	16,20,21	0.27	0	18,28,31	1.29	3 (16%)

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	C	601	1	-	0/6/23/26	0/1/1/1
4	NAG	C	602	1	-	0/6/23/26	0/1/1/1
4	NAG	G	401	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	I	601	1	-	0/6/23/26	0/1/1/1
4	NAG	K	601	1	-	0/6/23/26	0/1/1/1
4	NAG	K	602	1	-	0/6/23/26	0/1/1/1
6	SIA	K	603	-	-	0/14/34/38	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	K	601	NAG	C6-C5-C4	-3.01	105.58	113.02
6	K	603	SIA	C3-C4-C5	-2.82	108.33	111.47
4	K	601	NAG	C2-N2-C7	-2.52	119.80	123.04
4	K	601	NAG	C4-C3-C2	-2.32	107.62	111.23
4	C	602	NAG	C1-O5-C5	2.18	115.02	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	602	NAG	5	0
4	K	601	NAG	5	0
4	K	602	NAG	4	0
6	K	603	SIA	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	322/322 (100%)	0.39	22 (6%) 20 14	30, 55, 103, 206	0
1	C	321/322 (99%)	0.19	20 (6%) 24 17	29, 55, 102, 249	0
1	E	321/322 (99%)	0.17	16 (4%) 32 26	24, 56, 102, 220	0
1	G	322/322 (100%)	0.23	20 (6%) 24 17	29, 58, 98, 209	0
1	I	321/322 (99%)	0.43	21 (6%) 22 16	32, 64, 124, 236	0
1	K	321/322 (99%)	0.23	21 (6%) 22 16	26, 53, 96, 165	0
2	B	162/164 (98%)	0.57	19 (11%) 6 4	32, 72, 129, 175	0
2	D	164/164 (100%)	0.52	17 (10%) 8 5	36, 74, 130, 191	0
2	F	161/164 (98%)	0.84	29 (18%) 2 1	22, 74, 155, 210	0
2	H	162/164 (98%)	0.26	6 (3%) 45 38	28, 59, 101, 138	0
2	J	164/164 (100%)	1.41	44 (26%) 1 0	29, 81, 203, 324	0
2	L	161/164 (98%)	0.79	31 (19%) 2 1	30, 63, 163, 201	0
All	All	2902/2916 (99%)	0.43	266 (9%) 11 7	22, 60, 129, 324	0

The worst 5 of 266 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	159	TYR	12.3
2	J	158	ASP	11.4
1	I	9	LEU	9.0
2	J	20	GLY	8.9
2	J	129	ASN	8.5

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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
5	NAG	E	601	14/15	0.89	0.18	0.59	52,64,82,92	0
3	SIA	E	603	20/21	0.91	0.20	0.17	47,59,73,76	0
3	SIA	G	402	20/21	0.93	0.18	-0.52	44,68,95,96	0
3	SIA	I	602	20/21	0.95	0.16	-0.55	38,55,67,73	0
3	SIA	A	801	20/21	0.94	0.14	-0.79	34,46,60,66	0
3	SIA	C	603	20/21	0.94	0.14	-1.22	41,62,76,87	0
3	GAL	C	604	11/12	0.91	0.15	-1.45	53,69,77,79	0
3	GAL	G	403	11/12	0.87	0.19	-1.71	45,65,85,103	0
5	NAG	E	602	14/15	0.90	0.34	-	63,93,116,119	0
3	GAL	I	605	11/12	0.76	0.24	-	62,112,129,131	0
3	GAL	E	606	11/12	0.72	0.41	-	122,137,149,151	0
3	GAL	A	802	11/12	0.91	0.14	-	42,52,72,77	0
3	NAG	C	605	14/15	0.92	0.30	-	60,71,84,92	0
3	GAL	A	804	11/12	0.86	0.27	-	84,94,133,136	0
3	NAG	E	605	14/15	0.82	0.30	-	77,97,113,115	0
3	NAG	I	604	14/15	0.79	0.20	-	62,83,91,98	0
3	GAL	I	603	11/12	0.90	0.19	-	78,96,102,114	0
3	NAG	G	404	14/15	0.89	0.26	-	72,89,109,111	0
3	NAG	A	803	14/15	0.89	0.19	-	51,70,78,85	0
3	GAL	E	604	11/12	0.93	0.13	-	53,71,89,92	0
3	GAL	C	606	11/12	0.85	0.29	-	57,74,100,104	0
3	GAL	G	405	11/12	0.60	0.33	-	99,134,164,181	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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	NAG	K	602	14/15	0.75	0.32	1.82	76,105,115,115	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	NAG	C	602	14/15	0.92	0.18	0.33	25,57,74,79	0
6	SIA	K	603	20/21	0.95	0.16	-0.58	42,52,69,76	0
4	NAG	C	601	14/15	0.78	0.21	-0.67	66,93,106,114	0
4	NAG	K	601	14/15	0.79	0.30	-	119,128,136,137	0
4	NAG	I	601	14/15	0.69	0.33	-	101,130,147,150	0
4	NAG	G	401	14/15	0.81	0.37	-	78,108,129,132	0

6.5 Other polymers [i](#)

There are no such residues in this entry.