



wwPDB X-ray Structure Validation Summary Report

Nov 13, 2014 – 08:23 PM EST

PDB ID : 4D69
Title : SOYBEAN AGGLUTININ FROM GLYCINE MAX IN COMPLEX WITH THE ANTIGEN Tn
Authors : Madariaga, D.; Martinez-Saez, N.; Somovilla, V.J.; Coelho, H.; Valero-Gonzalez, J.; Castro-Lopez, J.; Asensio, J.L.; Jimenez-Barbero, J.; Busto, J.H.; Avenoza, A.; Marcelo, F.; Hurtado-Guerrero, R.; Corzana, F.; Pereg-rina, J.M.
Deposited on : 2014-11-10
Resolution : 2.70 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

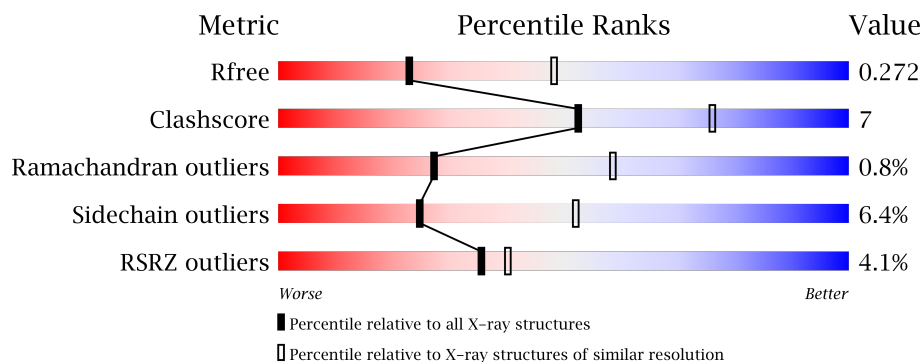
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24195
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.1.3
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24195

1 Overall quality at a glance

The reported resolution of this entry is 2.70 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	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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	253	
1	B	253	
1	C	253	
1	D	253	
1	E	253	
1	F	253	
1	G	253	
1	H	253	
1	I	253	
1	J	253	
1	K	253	
1	L	253	
2	O	5	
2	P	5	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	Q	5	<div><div></div></div>
2	R	5	<div><div></div></div>
2	S	5	<div><div></div></div>
2	T	5	<div><div></div></div>
2	U	5	<div><div></div></div>
2	V	5	<div><div></div></div>
2	W	5	<div><div></div></div>
2	X	5	<div><div></div></div>
2	Y	5	<div><div></div></div>
2	Z	5	<div><div></div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 22512 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 called LECTIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	234	Total	C	N	O	S	0	0	0
			1809	1157	301	350	1			
1	B	232	Total	C	N	O	S	0	0	0
			1793	1148	297	347	1			
1	C	231	Total	C	N	O	S	0	0	0
			1790	1147	298	344	1			
1	D	234	Total	C	N	O	S	0	0	0
			1809	1157	301	350	1			
1	E	231	Total	C	N	O	S	0	0	0
			1790	1147	298	344	1			
1	F	234	Total	C	N	O	S	0	0	0
			1809	1157	301	350	1			
1	G	233	Total	C	N	O	S	0	0	0
			1803	1154	300	348	1			
1	H	232	Total	C	N	O	S	0	1	0
			1805	1156	300	348	1			
1	I	234	Total	C	N	O	S	0	0	0
			1809	1157	301	350	1			
1	J	232	Total	C	N	O	S	0	0	0
			1797	1151	299	346	1			
1	K	232	Total	C	N	O	S	0	1	0
			1805	1156	300	348	1			
1	L	232	Total	C	N	O	S	0	0	0
			1799	1152	299	347	1			

- Molecule 2 is a protein called SHORT ANTIGEN PEPTIDE.

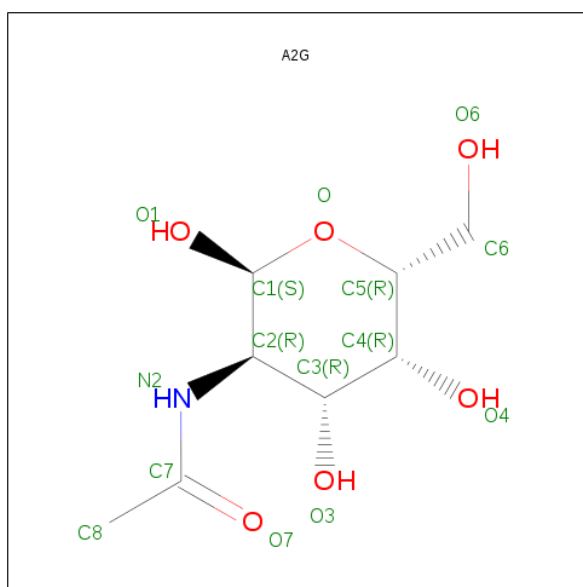
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	O	4	Total	C	N	O	0	0	0
			33	19	7	7			
2	P	2	Total	C	N	O	0	0	0
			15	8	2	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	Q	2	Total	C	N	O	0	0	0
			15	8	2	5			
2	R	3	Total	C	N	O	0	0	0
			26	14	6	6			
2	S	1	Total	C	N	O	0	0	0
			7	4	1	2			
2	T	2	Total	C	N	O	0	0	0
			15	8	2	5			
2	U	3	Total	C	N	O	0	0	0
			22	13	3	6			
2	V	3	Total	C	N	O	0	0	0
			26	14	6	6			
2	W	2	Total	C	N	O	0	0	0
			15	8	2	5			
2	X	4	Total	C	N	O	0	0	0
			33	19	7	7			
2	Y	3	Total	C	N	O	0	0	0
			26	14	6	6			
2	Z	2	Total	C	N	O	0	0	0
			15	8	2	5			

- Molecule 3 is SUGAR (N-ACETYL-2-DEOXY-2-AMINO-GALACTOSE) (three-letter code: A2G) (formula: $C_8H_{15}NO_6$).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1	Total	C	N	O	0	0
			14	8	1	5		
3	F	1	Total	C	N	O	0	0
			14	8	1	5		
3	G	1	Total	C	N	O	0	0
			14	8	1	5		
3	H	1	Total	C	N	O	0	0
			14	8	1	5		
3	I	1	Total	C	N	O	0	0
			14	8	1	5		
3	J	1	Total	C	N	O	0	0
			14	8	1	5		
3	K	1	Total	C	N	O	0	0
			14	8	1	5		
3	L	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	2	Total	Mn	0	0
			2	2		
4	J	2	Total	Mn	0	0
			2	2		
4	D	2	Total	Mn	0	0
			2	2		
4	K	2	Total	Mn	0	0
			2	2		
4	E	2	Total	Mn	0	0
			2	2		
4	H	2	Total	Mn	0	0
			2	2		
4	B	2	Total	Mn	0	0
			2	2		
4	I	2	Total	Mn	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	2	Total 2	Mn 2	0	0
4	A	2	Total 2	Mn 2	0	0
4	L	2	Total 2	Mn 2	0	0
4	F	2	Total 2	Mn 2	0	0

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total 28	C 16	N 2	O 10	0	0
5	B	2	Total 28	C 16	N 2	O 10	0	0
5	C	2	Total 28	C 16	N 2	O 10	0	0
5	D	2	Total 28	C 16	N 2	O 10	0	0
5	E	2	Total 28	C 16	N 2	O 10	0	0
5	F	2	Total 28	C 16	N 2	O 10	0	0
5	G	2	Total 28	C 16	N 2	O 10	0	0
5	H	2	Total 28	C 16	N 2	O 10	0	0
5	I	2	Total 28	C 16	N 2	O 10	0	0
5	J	2	Total 28	C 16	N 2	O 10	0	0
5	K	2	Total 28	C 16	N 2	O 10	0	0
5	L	2	Total 28	C 16	N 2	O 10	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	11	Total 11	O 11	0	0

Continued on next page...

Continued from previous page...

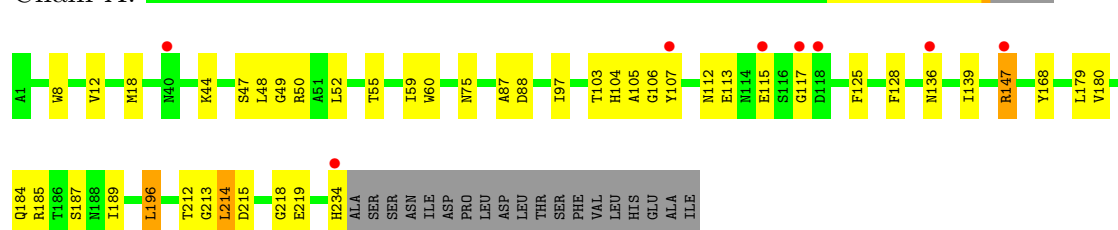
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	7	Total 7	O 7	0	0
6	C	9	Total 9	O 9	0	0
6	D	16	Total 16	O 16	0	0
6	E	11	Total 11	O 11	0	0
6	F	10	Total 10	O 10	0	0
6	G	7	Total 7	O 7	0	0
6	H	11	Total 11	O 11	0	0
6	I	12	Total 12	O 12	0	0
6	J	5	Total 5	O 5	0	0
6	K	7	Total 7	O 7	0	0
6	L	12	Total 12	O 12	0	0

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.

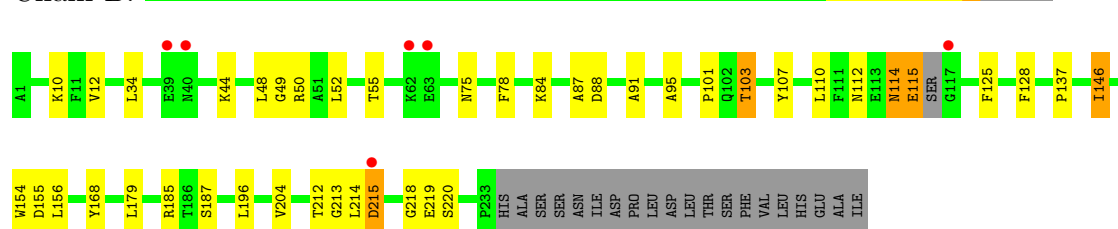
• Molecule 1: LECTIN

Chain A:



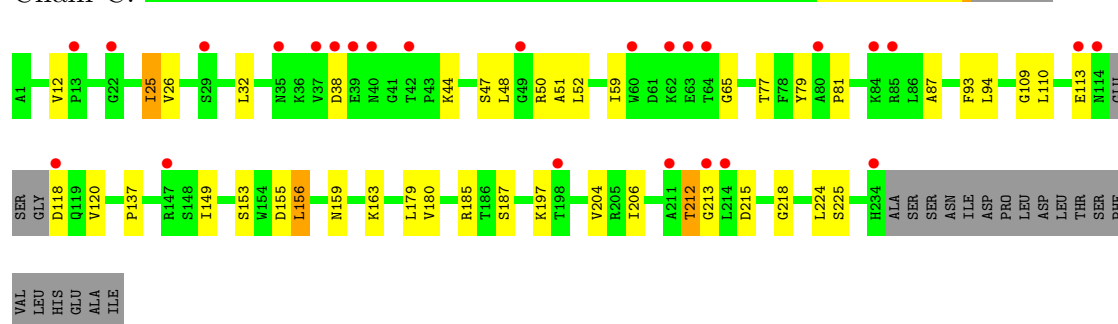
• Molecule 1: LECTIN

Chain B:



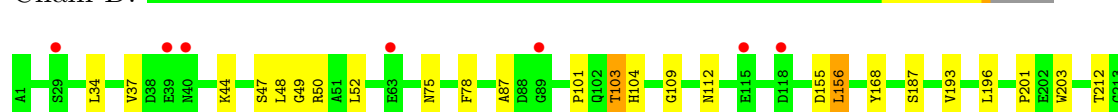
• Molecule 1: LECTIN

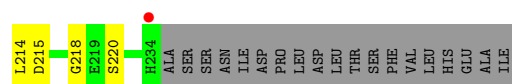
Chain C:



• Molecule 1: LECTIN

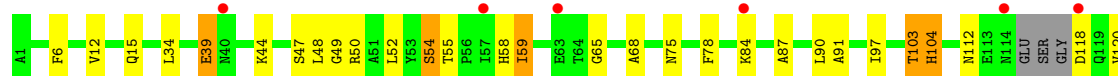
Chain D:





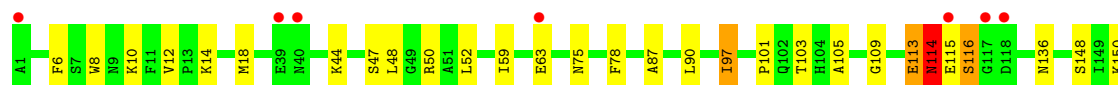
• Molecule 1: LECTIN

Chain E:



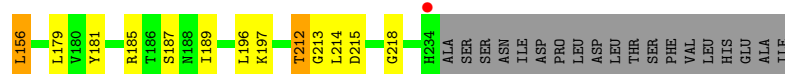
• Molecule 1: LECTIN

Chain F:



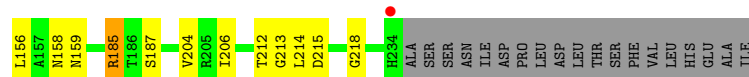
• Molecule 1: LECTIN

Chain G:



• Molecule 1: LECTIN

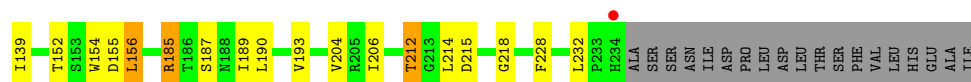
Chain H:



• Molecule 1: LECTIN

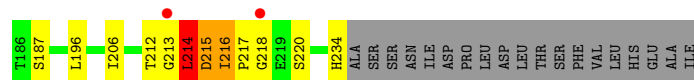
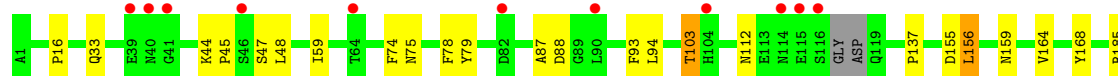
Chain I:





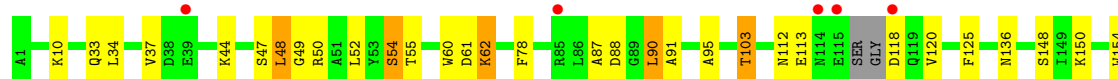
- Molecule 1: LECTIN

Chain J:



- Molecule 1: LECTIN

Chain K:



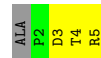
- Molecule 1: LECTIN

Chain L:



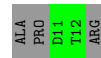
- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain O:



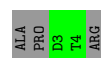
- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain P:



- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain Q:



- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain R: 

ALA	PRO	D3	T4	R5
-----	-----	----	----	----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain S: 

ALA	PRO	ASP	T4	ARG
-----	-----	-----	----	-----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain T: 

ALA	PRO	D3	T4	ARG
-----	-----	----	----	-----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain U: 

ALA	P2	D3	T4	ARG
-----	----	----	----	-----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain V: 

ALA	PRO	D3	T4	R5
-----	-----	----	----	----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain W: 

ALA	PRO	D3	T4	ARG
-----	-----	----	----	-----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain X: 

ALA	P2	D3	T4	R5
-----	----	----	----	----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain Y: 

ALA	PRO	D3	T4	R5
-----	-----	----	----	----

- Molecule 2: SHORT ANTIGEN PEPTIDE

Chain Z: 

ALA
PRO
D3
T4
ARG

4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	114.23Å 114.23Å 202.89Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	202.89 – 2.70 19.96 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (202.89-2.70) 100.0 (19.96-2.70)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 2.71Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.230 , 0.274 0.233 , 0.272	Depositor DCC
R_{free} test set	2232 reflections (2.83%)	DCC
Wilson B-factor (Å ²)	48.9	Xtriage
Anisotropy	0.197	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 23.5	EDS
Estimated twinning fraction	0.009 for -h,-k,l 0.037 for h,-h-k,-l 0.024 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 81145 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	22512	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.22% 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: MN, NAG, A2G

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	0.61	0/1858	0.77	1/2540 (0.0%)
1	B	0.57	0/1840	0.78	2/2514 (0.1%)
1	C	0.62	0/1838	0.77	0/2512
1	D	0.61	0/1858	0.77	0/2540
1	E	0.75	1/1838 (0.1%)	0.83	3/2512 (0.1%)
1	F	0.63	0/1858	0.80	1/2540 (0.0%)
1	G	0.59	0/1851	0.75	1/2529 (0.0%)
1	H	0.77	2/1856 (0.1%)	0.86	4/2536 (0.2%)
1	I	0.61	0/1858	0.77	2/2540 (0.1%)
1	J	0.56	0/1845	0.76	1/2521 (0.0%)
1	K	0.77	1/1856 (0.1%)	0.79	3/2536 (0.1%)
1	L	0.61	0/1847	0.83	2/2524 (0.1%)
2	O	0.73	0/33	0.83	0/43
2	P	0.78	0/14	0.74	0/18
2	Q	0.95	0/14	0.70	0/18
2	R	0.70	0/25	0.86	0/32
2	S	0.92	0/6	0.55	0/7
2	T	1.02	0/14	0.76	0/18
2	U	0.71	0/22	0.70	0/29
2	V	0.79	0/25	0.80	0/32
2	W	0.87	0/14	1.21	0/18
2	X	0.52	0/33	1.12	0/43
2	Y	0.77	0/25	1.08	0/32
2	Z	0.70	0/14	1.11	0/18
All	All	0.65	4/22442 (0.0%)	0.79	20/30652 (0.1%)

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	H	0	2
1	K	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	54	SER	C-N	-19.55	0.89	1.34
1	E	54	SER	C-N	-16.63	0.95	1.34
1	H	54	SER	C-N	-16.57	0.95	1.34
1	H	55	THR	C-N	-13.48	1.08	1.34

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	135	PRO	N-CA-C	13.28	146.63	112.10
1	L	136	ASN	N-CA-CB	-12.82	87.53	110.60
1	H	55	THR	O-C-N	-10.88	100.42	121.10
1	E	54	SER	C-N-CA	10.39	147.67	121.70
1	B	115	GLU	CB-CA-C	-9.74	90.92	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	54	SER	Mainchain,Peptide
1	K	54	SER	Peptide

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	1809	0	1767	31	0
1	B	1793	0	1754	25	0
1	C	1790	0	1752	31	0
1	D	1809	0	1767	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1790	0	1751	25	1
1	F	1809	0	1767	39	0
1	G	1803	0	1761	19	1
1	H	1805	0	1765	18	0
1	I	1809	0	1767	30	0
1	J	1797	0	1759	29	0
1	K	1805	0	1765	29	0
1	L	1799	0	1758	20	0
2	O	33	0	30	2	0
2	P	15	0	9	0	0
2	Q	15	0	9	0	0
2	R	26	0	22	1	0
2	S	7	0	5	0	0
2	T	15	0	9	0	0
2	U	22	0	17	0	0
2	V	26	0	22	1	0
2	W	15	0	9	0	0
2	X	33	0	30	10	0
2	Y	26	0	22	1	0
2	Z	15	0	9	0	0
3	A	14	0	13	2	0
3	B	14	0	13	2	0
3	C	14	0	13	0	0
3	D	14	0	13	0	0
3	E	14	0	13	0	0
3	F	14	0	13	5	0
3	G	14	0	13	0	0
3	H	14	0	13	0	0
3	I	14	0	13	0	0
3	J	14	0	13	9	0
3	K	14	0	13	0	0
3	L	14	0	13	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
4	E	2	0	0	0	0
4	F	2	0	0	0	0
4	G	2	0	0	0	0
4	H	2	0	0	0	0
4	I	2	0	0	0	0
4	J	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	K	2	0	0	0	0
4	L	2	0	0	0	0
5	A	28	0	25	3	0
5	B	28	0	25	1	0
5	C	28	0	25	2	0
5	D	28	0	25	1	0
5	E	28	0	25	1	0
5	F	28	0	25	1	0
5	G	28	0	25	1	0
5	H	28	0	25	1	0
5	I	28	0	25	1	0
5	J	28	0	25	1	0
5	K	28	0	25	1	0
5	L	28	0	25	1	0
6	A	11	0	0	0	0
6	B	7	0	0	1	0
6	C	9	0	0	0	0
6	D	16	0	0	0	0
6	E	11	0	0	0	0
6	F	10	0	0	0	0
6	G	7	0	0	0	0
6	H	11	0	0	1	0
6	I	12	0	0	1	0
6	J	5	0	0	1	0
6	K	7	0	0	1	0
6	L	12	0	0	1	0
All	All	22512	0	21782	313	1

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 7.

The worst 5 of 313 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:114:ASN:OD1	1:F:116:SER:N	1.74	1.20
1:F:87:ALA:HB2	1:F:218:GLY:HA3	1.25	1.16
1:J:87:ALA:HB2	1:J:218:GLY:HA3	1.21	1.11
1:J:216:ILE:HG12	3:J:1235:A2G:C6	1.83	1.09
1:A:87:ALA:HB2	1:A:218:GLY:HA3	1.35	1.08

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:55:THR:OG1	1:G:55:THR:OG1[3.655]	1.63	0.57

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	232/253 (92%)	217 (94%)	13 (6%)	2 (1%)	25	55
1	B	228/253 (90%)	217 (95%)	11 (5%)	0	100	100
1	C	227/253 (90%)	216 (95%)	10 (4%)	1 (0%)	43	76
1	D	232/253 (92%)	217 (94%)	14 (6%)	1 (0%)	43	76
1	E	227/253 (90%)	215 (95%)	10 (4%)	2 (1%)	25	55
1	F	232/253 (92%)	218 (94%)	10 (4%)	4 (2%)	14	33
1	G	229/253 (90%)	213 (93%)	15 (7%)	1 (0%)	43	76
1	H	229/253 (90%)	216 (94%)	12 (5%)	1 (0%)	43	76
1	I	232/253 (92%)	216 (93%)	14 (6%)	2 (1%)	25	55
1	J	228/253 (90%)	215 (94%)	10 (4%)	3 (1%)	18	43
1	K	229/253 (90%)	218 (95%)	10 (4%)	1 (0%)	43	76
1	L	228/253 (90%)	214 (94%)	12 (5%)	2 (1%)	25	55
2	O	2/5 (40%)	2 (100%)	0	0	100	100
2	R	1/5 (20%)	0	1 (100%)	0	100	100
2	U	1/5 (20%)	0	0	1 (100%)	0	0
2	V	1/5 (20%)	1 (100%)	0	0	100	100
2	X	2/5 (40%)	2 (100%)	0	0	100	100
2	Y	1/5 (20%)	0	1 (100%)	0	100	100
All	All	2761/3066 (90%)	2597 (94%)	143 (5%)	21 (1%)	27	58

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	114	ASN
1	J	214	LEU
1	A	106	GLY
1	F	116	SER
1	G	156	LEU

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	202/219 (92%)	190 (94%)	12 (6%)	28	56
1	B	200/219 (91%)	187 (94%)	13 (6%)	24	51
1	C	200/219 (91%)	190 (95%)	10 (5%)	34	66
1	D	202/219 (92%)	193 (96%)	9 (4%)	38	70
1	E	200/219 (91%)	185 (92%)	15 (8%)	19	43
1	F	202/219 (92%)	186 (92%)	16 (8%)	18	39
1	G	201/219 (92%)	185 (92%)	16 (8%)	17	37
1	H	202/219 (92%)	191 (95%)	11 (5%)	31	61
1	I	202/219 (92%)	189 (94%)	13 (6%)	25	52
1	J	201/219 (92%)	192 (96%)	9 (4%)	38	70
1	K	202/219 (92%)	186 (92%)	16 (8%)	18	39
1	L	201/219 (92%)	187 (93%)	14 (7%)	21	47
2	O	4/4 (100%)	4 (100%)	0	100	100
2	P	2/4 (50%)	2 (100%)	0	100	100
2	Q	2/4 (50%)	2 (100%)	0	100	100
2	R	3/4 (75%)	2 (67%)	1 (33%)	0	0
2	S	1/4 (25%)	1 (100%)	0	100	100
2	T	2/4 (50%)	2 (100%)	0	100	100
2	U	3/4 (75%)	3 (100%)	0	100	100
2	V	3/4 (75%)	3 (100%)	0	100	100
2	W	2/4 (50%)	2 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	X	4/4 (100%)	3 (75%)	1 (25%)	1	2
2	Y	3/4 (75%)	3 (100%)	0	100	100
2	Z	2/4 (50%)	2 (100%)	0	100	100
All	All	2446/2676 (91%)	2290 (94%)	156 (6%)	25	52

5 of 156 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	185	ARG
1	G	196	LEU
1	L	94	LEU
1	F	212	THR
1	G	52	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	15	GLN
1	H	102	GLN
1	L	102	GLN
1	G	102	GLN
1	G	159	ASN

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 ⓘ

24 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	$\# Z > 2$
5	NAG	A	1238	1,5	12,14,15	0.74	1 (8%)	15,19,21	1.78	4 (26%)
5	NAG	A	1239	5	12,14,15	0.74	1 (8%)	15,19,21	1.03	1 (6%)
5	NAG	B	1237	1,5	12,14,15	0.91	0	15,19,21	1.88	5 (33%)
5	NAG	B	1238	5	12,14,15	0.62	0	15,19,21	1.35	1 (6%)
5	NAG	C	1238	1,5	12,14,15	0.66	0	15,19,21	2.00	6 (40%)
5	NAG	C	1239	5	12,14,15	0.87	1 (8%)	15,19,21	1.76	4 (26%)
5	NAG	D	1238	1,5	12,14,15	1.10	2 (16%)	15,19,21	2.31	9 (60%)
5	NAG	D	1239	5	12,14,15	0.87	0	15,19,21	2.61	6 (40%)
5	NAG	E	1238	1,5	12,14,15	0.81	0	15,19,21	1.66	5 (33%)
5	NAG	E	1239	5	12,14,15	0.81	1 (8%)	15,19,21	1.24	1 (6%)
5	NAG	F	1235	1,5	12,14,15	0.90	1 (8%)	15,19,21	2.39	8 (53%)
5	NAG	F	1236	5	12,14,15	0.73	0	15,19,21	2.51	5 (33%)
5	NAG	G	1238	1,5	12,14,15	0.63	0	15,19,21	2.00	4 (26%)
5	NAG	G	1239	5	12,14,15	0.55	0	15,19,21	1.25	2 (13%)
5	NAG	H	1238	1,5	12,14,15	0.76	0	15,19,21	1.78	4 (26%)
5	NAG	H	1239	5	12,14,15	1.12	1 (8%)	15,19,21	1.94	5 (33%)
5	NAG	I	1238	1,5	12,14,15	0.63	0	15,19,21	1.98	6 (40%)
5	NAG	I	1239	5	12,14,15	0.76	1 (8%)	15,19,21	1.47	4 (26%)
5	NAG	J	1238	1,5	12,14,15	0.98	1 (8%)	15,19,21	2.16	5 (33%)
5	NAG	J	1239	5	12,14,15	0.57	0	15,19,21	1.98	2 (13%)
5	NAG	K	1238	1,5	12,14,15	0.78	0	15,19,21	1.96	5 (33%)
5	NAG	K	1239	5	12,14,15	0.72	0	15,19,21	1.89	4 (26%)
5	NAG	L	1238	1,5	12,14,15	1.06	1 (8%)	15,19,21	2.21	4 (26%)
5	NAG	L	1239	5	12,14,15	0.84	0	15,19,21	1.79	2 (13%)

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	A	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1239	5	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	B	1237	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	1238	5	-	0/6/23/26	0/1/1/1
5	NAG	C	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	C	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	D	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	D	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	E	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	F	1235	1,5	-	0/6/23/26	0/1/1/1
5	NAG	F	1236	5	-	0/6/23/26	0/1/1/1
5	NAG	G	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	G	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	H	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	H	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	I	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	I	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	J	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	J	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	K	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	K	1239	5	-	0/6/23/26	0/1/1/1
5	NAG	L	1238	1,5	-	0/6/23/26	0/1/1/1
5	NAG	L	1239	5	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	1239	NAG	O5-C5	-2.55	1.41	1.45
5	D	1238	NAG	C4-C5	2.52	1.58	1.53
5	H	1239	NAG	C2-N2	2.36	1.49	1.46
5	D	1238	NAG	C2-N2	-2.24	1.44	1.46
5	I	1239	NAG	O5-C5	-2.17	1.42	1.45

The worst 5 of 102 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	1236	NAG	O5-C5-C4	7.71	120.44	110.65
5	D	1239	NAG	O5-C5-C4	7.07	119.63	110.65
5	J	1239	NAG	O5-C5-C4	6.47	118.86	110.65
5	L	1238	NAG	C8-C7-N2	5.32	126.18	116.12
5	H	1239	NAG	O5-C5-C4	4.82	116.77	110.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry

Of 36 ligands modelled in this entry, 24 are monoatomic - leaving 12 for Mogul analysis.

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	A2G	A	1235	2	12,14,15	0.63	0	15,19,21	1.53	3 (20%)
3	A2G	B	1234	2	12,14,15	0.52	0	15,19,21	1.62	3 (20%)
3	A2G	C	1235	2	12,14,15	0.81	0	15,19,21	1.75	3 (20%)
3	A2G	D	1235	2	12,14,15	0.44	0	15,19,21	1.34	2 (13%)
3	A2G	E	1235	2	12,14,15	0.64	0	15,19,21	1.02	1 (6%)
3	A2G	F	1239	2	12,14,15	0.57	0	15,19,21	1.90	5 (33%)
3	A2G	G	1235	2	12,14,15	0.60	0	15,19,21	1.32	1 (6%)
3	A2G	H	1235	2	12,14,15	0.53	0	15,19,21	1.55	3 (20%)
3	A2G	I	1235	2	12,14,15	0.47	0	15,19,21	1.30	2 (13%)
3	A2G	J	1235	2	12,14,15	0.63	0	15,19,21	1.04	0
3	A2G	K	1235	2	12,14,15	0.89	0	15,19,21	1.82	3 (20%)
3	A2G	L	1235	2	12,14,15	0.58	0	15,19,21	0.76	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	A2G	A	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	B	1234	2	-	0/6/23/26	0/1/1/1
3	A2G	C	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	D	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	E	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	F	1239	2	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	A2G	G	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	H	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	I	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	J	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	K	1235	2	-	0/6/23/26	0/1/1/1
3	A2G	L	1235	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	K	1235	A2G	C3-C2-N2	-4.83	104.37	111.62
3	B	1234	A2G	O-C5-C6	4.62	111.83	106.98
3	C	1235	A2G	C2-N2-C7	4.23	128.51	123.39
3	F	1239	A2G	O-C5-C6	3.66	110.82	106.98
3	D	1235	A2G	C3-C2-N2	-3.56	106.28	111.62

There are no chirality outliers.

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	234/253 (92%)	0.19	8 (3%) 43 48	43, 56, 98, 127	0
1	B	232/253 (91%)	0.07	6 (2%) 53 59	44, 56, 83, 111	0
1	C	231/253 (91%)	0.64	26 (11%) 6 6	47, 61, 85, 118	0
1	D	234/253 (92%)	0.06	8 (3%) 43 48	44, 49, 63, 87	0
1	E	231/253 (91%)	0.05	7 (3%) 48 54	44, 49, 65, 83	0
1	F	234/253 (92%)	0.01	8 (3%) 43 48	44, 49, 64, 107	0
1	G	233/253 (92%)	0.08	10 (4%) 34 38	43, 50, 76, 116	0
1	H	232/253 (91%)	0.01	6 (2%) 53 59	43, 50, 66, 116	0
1	I	234/253 (92%)	0.01	4 (1%) 67 73	43, 49, 61, 85	0
1	J	232/253 (91%)	0.26	13 (5%) 24 25	42, 66, 108, 141	0
1	K	232/253 (91%)	-0.05	6 (2%) 53 59	44, 49, 70, 113	0
1	L	232/253 (91%)	0.09	9 (3%) 37 42	44, 51, 75, 125	0
2	O	4/5 (80%)	0.45	0 100 100	47, 50, 73, 77	0
2	P	2/5 (40%)	0.93	0 100 100	77, 77, 77, 86	0
2	Q	2/5 (40%)	1.09	0 100 100	72, 72, 72, 81	0
2	R	3/5 (60%)	1.01	0 100 100	76, 76, 76, 82	0
2	S	1/5 (20%)	0.07	0 100 100	71, 71, 71, 71	0
2	T	2/5 (40%)	0.54	0 100 100	63, 63, 63, 63	0
2	U	3/5 (60%)	0.67	0 100 100	61, 61, 62, 82	0
2	V	3/5 (60%)	1.52	1 (33%) 1 1	71, 71, 73, 98	0
2	W	2/5 (40%)	1.59	1 (50%) 0 0	70, 70, 70, 70	0
2	X	4/5 (80%)	1.98	2 (50%) 0 0	65, 71, 73, 91	0
2	Y	3/5 (60%)	0.94	0 100 100	50, 50, 55, 75	0
2	Z	2/5 (40%)	1.28	0 100 100	61, 61, 61, 64	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å ²)	Q<0.9
All	All	2822/3096 (91%)	0.13	115 (4%)	35	40	42, 51, 83, 141	0

The worst 5 of 115 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	40	ASN	5.6
1	C	234	HIS	5.4
1	G	117	GLY	5.0
1	L	115	GLU	4.6
1	B	40	ASN	4.3

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 ⓘ

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
5	NAG	K	1239	14/15	0.31	3.37	45,65,74,75	0
5	NAG	I	1239	14/15	0.26	3.13	58,65,75,76	0
5	NAG	H	1239	14/15	0.36	3.09	41,50,58,59	0
5	NAG	J	1239	14/15	0.30	2.77	50,59,71,73	0
5	NAG	L	1239	14/15	0.28	2.11	47,53,64,66	0
5	NAG	B	1238	14/15	0.31	1.55	50,58,63,67	0
5	NAG	F	1236	14/15	0.27	1.47	44,48,55,60	0
5	NAG	H	1238	14/15	0.23	1.40	23,35,39,41	0
5	NAG	C	1238	14/15	0.29	1.28	44,60,77,79	0
5	NAG	E	1239	14/15	0.26	1.08	42,47,54,67	0
5	NAG	G	1239	14/15	0.25	0.93	49,55,64,65	0
5	NAG	D	1239	14/15	0.25	0.61	44,58,64,75	0
5	NAG	J	1238	14/15	0.21	0.46	32,44,50,50	0
5	NAG	C	1239	14/15	0.29	0.32	61,69,90,91	0
5	NAG	F	1235	14/15	0.19	0.31	23,33,36,40	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	NAG	K	1238	14/15	0.18	0.19	31,50,63,71	0
5	NAG	D	1238	14/15	0.17	-0.06	24,39,51,52	0
5	NAG	L	1238	14/15	0.15	-0.07	23,33,43,50	0
5	NAG	A	1239	14/15	0.22	-0.32	52,57,59,59	0
5	NAG	A	1238	14/15	0.15	-0.34	29,39,46,48	0
5	NAG	G	1238	14/15	0.14	-0.40	28,41,44,46	0
5	NAG	E	1238	14/15	0.16	-0.45	21,32,46,50	0
5	NAG	I	1238	14/15	0.14	-0.46	29,41,52,53	0
5	NAG	B	1237	14/15	0.15	-0.55	29,45,50,50	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(\AA^2)	Q<0.9
3	A2G	C	1235	14/15	0.22	-0.58	59,74,79,81	0
3	A2G	J	1235	14/15	0.19	-0.65	57,72,78,79	0
3	A2G	A	1235	14/15	0.21	-0.73	53,69,74,74	0
4	MN	A	1237	1/1	0.18	-0.80	86,86,86,86	0
3	A2G	B	1234	14/15	0.17	-1.05	47,52,59,60	0
4	MN	E	1237	1/1	0.13	-1.12	60,60,60,60	0
3	A2G	F	1239	14/15	0.11	-1.17	34,39,42,44	0
3	A2G	K	1235	14/15	0.11	-1.39	33,40,50,57	0
3	A2G	I	1235	14/15	0.12	-1.43	26,29,32,32	0
4	MN	L	1237	1/1	0.10	-1.46	75,75,75,75	0
4	MN	B	1235	1/1	0.08	-1.53	55,55,55,55	0
3	A2G	G	1235	14/15	0.14	-1.55	48,58,61,63	0
3	A2G	L	1235	14/15	0.11	-1.58	40,46,50,52	0
3	A2G	H	1235	14/15	0.16	-1.61	41,44,49,52	0
3	A2G	D	1235	14/15	0.14	-1.63	33,39,42,46	0
4	MN	H	1236	1/1	0.08	-1.64	41,41,41,41	0
4	MN	J	1236	1/1	0.14	-1.69	63,63,63,63	0
4	MN	D	1237	1/1	0.07	-1.79	55,55,55,55	0
4	MN	B	1236	1/1	0.11	-1.86	60,60,60,60	0
4	MN	A	1236	1/1	0.08	-1.88	61,61,61,61	0
4	MN	I	1236	1/1	0.07	-1.91	43,43,43,43	0
4	MN	G	1237	1/1	0.11	-1.93	66,66,66,66	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	MN	E	1236	1/1	0.09	-2.19	44,44,44,44	0
4	MN	H	1237	1/1	0.06	-2.69	55,55,55,55	0
4	MN	F	1238	1/1	0.08	-2.71	67,67,67,67	0
4	MN	J	1237	1/1	0.07	-2.76	76,76,76,76	0
4	MN	G	1236	1/1	0.09	-2.82	59,59,59,59	0
4	MN	L	1236	1/1	0.04	-2.85	53,53,53,53	0
4	MN	C	1237	1/1	0.08	-2.92	65,65,65,65	0
3	A2G	E	1235	14/15	0.10	-2.93	34,40,45,50	0
4	MN	I	1237	1/1	0.04	-3.07	46,46,46,46	0
4	MN	D	1236	1/1	0.02	-3.34	37,37,37,37	0
4	MN	K	1237	1/1	0.06	-3.56	65,65,65,65	0
4	MN	F	1237	1/1	0.03	-3.83	49,49,49,49	0
4	MN	K	1236	1/1	0.05	-3.86	43,43,43,43	0
4	MN	C	1236	1/1	0.06	-4.11	65,65,65,65	0

6.5 Other polymers ⓘ

There are no such residues in this entry.