



wwPDB X-ray Structure Validation Summary Report ⓘ

May 30, 2016 – 11:12 AM EDT

PDB ID : 5HUF
Title : The crystal structure of hemagglutinin from A/gyrfalcon/Washington/41088-6/2014 influenza virus
Authors : Yang, H.; Carney, P.J.; Guo, Z.; Chang, J.C.; Stevens, J.
Deposited on : 2016-01-27
Resolution : 2.81 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:

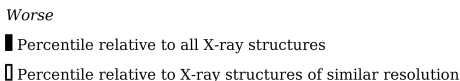
MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027674
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027674

i

X-RAY DIFFRACTION

A.

Metric	Percentile Range	Value
--------	------------------	-------



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R _{free}	91344	2676 (2.84-2.80)
Clashscore	102246	3124 (2.84-2.80)
Ramachandran outliers	100387	3072 (2.84-2.80)
Sidechain outliers	100360	3074 (2.84-2.80)
RSRZ outliers	91569	2690 (2.84-2.80)

Mol	Chain	Length	Quality of chain			
1	A	334		74%	19%	• • •
1	C	334		73%	20%	• • •
1	E	334		74%	20%	• • •
2	B	181		73%	19%	• • •
2	D	181		72%	19%	• • •
2	F	181		71%	19%	5% • •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	D	201	-	-	-	X
3	NAG	F	201	-	-	-	X
4	MAN	A	504	X	-	-	-
4	MAN	C	504	X	-	-	-
4	MAN	E	504	X	-	-	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12219 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	0	0
			2559	1618	445	481	15			
1	C	324	Total	C	N	O	S	0	0	0
			2559	1618	445	481	15			
1	E	324	Total	C	N	O	S	0	0	0
			2559	1618	445	481	15			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	expression tag	UNP A0A0C4X0C0
A	-3	ASP	-	expression tag	UNP A0A0C4X0C0
A	-2	LEU	-	expression tag	UNP A0A0C4X0C0
A	-1	GLY	-	expression tag	UNP A0A0C4X0C0
C	-4	ALA	-	expression tag	UNP A0A0C4X0C0
C	-3	ASP	-	expression tag	UNP A0A0C4X0C0
C	-2	LEU	-	expression tag	UNP A0A0C4X0C0
C	-1	GLY	-	expression tag	UNP A0A0C4X0C0
E	-4	ALA	-	expression tag	UNP A0A0C4X0C0
E	-3	ASP	-	expression tag	UNP A0A0C4X0C0
E	-2	LEU	-	expression tag	UNP A0A0C4X0C0
E	-1	GLY	-	expression tag	UNP A0A0C4X0C0

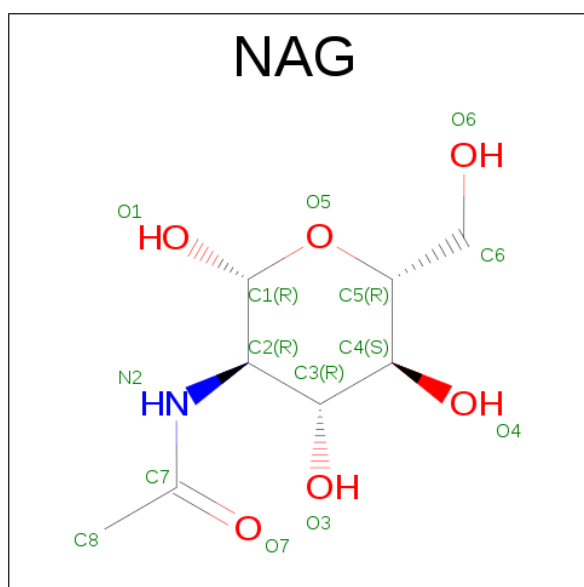
- Molecule 2 is a protein called hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total	C	N	O	S	0	0	0
			1419	884	244	283	8			
2	D	176	Total	C	N	O	S	0	0	0
			1419	884	244	283	8			
2	F	176	Total	C	N	O	S	0	0	0
			1419	884	244	283	8			

There are 21 discrepancies between the modelled and reference sequences:

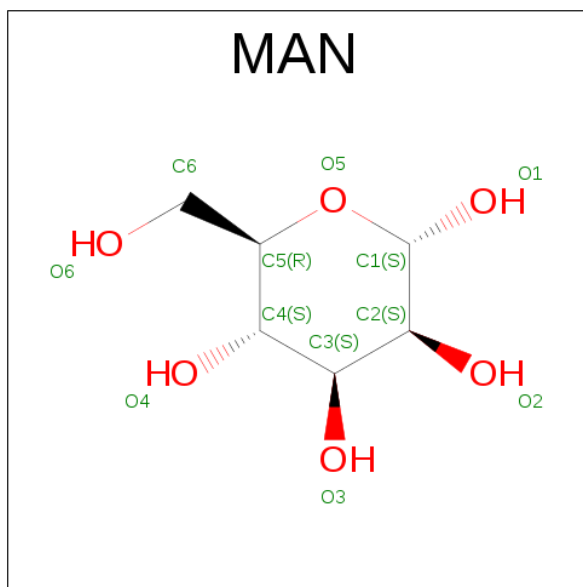
Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	expression tag	UNP A0A0C4X0C0
B	176	GLY	-	expression tag	UNP A0A0C4X0C0
B	177	ARG	-	expression tag	UNP A0A0C4X0C0
B	178	LEU	-	expression tag	UNP A0A0C4X0C0
B	179	VAL	-	expression tag	UNP A0A0C4X0C0
B	180	PRO	-	expression tag	UNP A0A0C4X0C0
B	181	ARG	-	expression tag	UNP A0A0C4X0C0
D	175	SER	-	expression tag	UNP A0A0C4X0C0
D	176	GLY	-	expression tag	UNP A0A0C4X0C0
D	177	ARG	-	expression tag	UNP A0A0C4X0C0
D	178	LEU	-	expression tag	UNP A0A0C4X0C0
D	179	VAL	-	expression tag	UNP A0A0C4X0C0
D	180	PRO	-	expression tag	UNP A0A0C4X0C0
D	181	ARG	-	expression tag	UNP A0A0C4X0C0
F	175	SER	-	expression tag	UNP A0A0C4X0C0
F	176	GLY	-	expression tag	UNP A0A0C4X0C0
F	177	ARG	-	expression tag	UNP A0A0C4X0C0
F	178	LEU	-	expression tag	UNP A0A0C4X0C0
F	179	VAL	-	expression tag	UNP A0A0C4X0C0
F	180	PRO	-	expression tag	UNP A0A0C4X0C0
F	181	ARG	-	expression tag	UNP A0A0C4X0C0

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (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		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
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	C	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	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	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	E	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	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	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C₆H₁₂O₆).

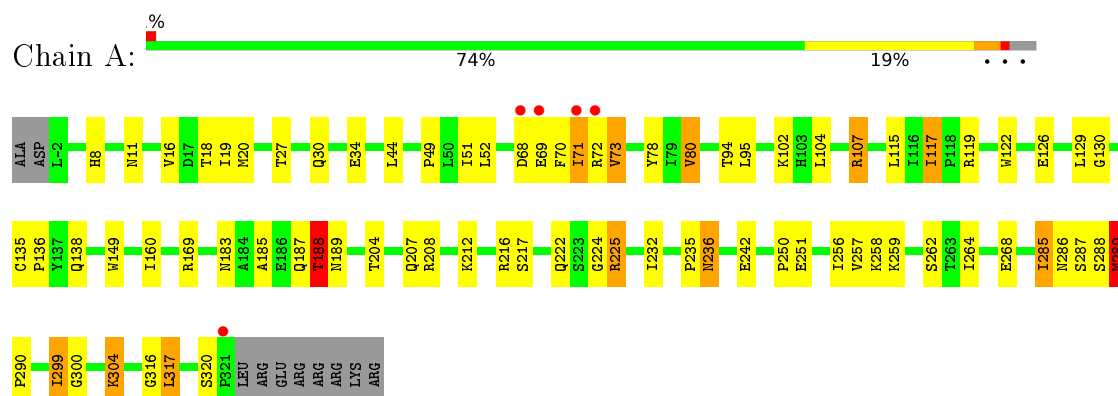


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	C	1	Total	C	O	0	0
			11	6	5		
4	E	1	Total	C	O	0	0
			11	6	5		

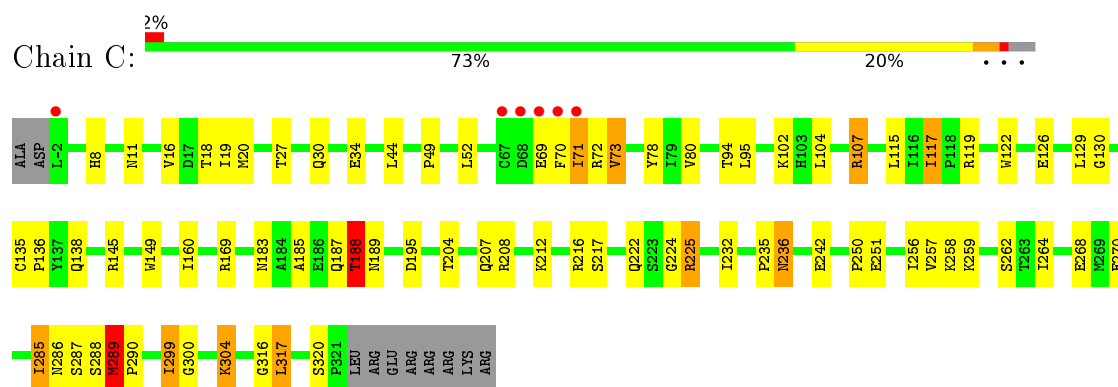
3 Residue-property plots [i](#)

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.

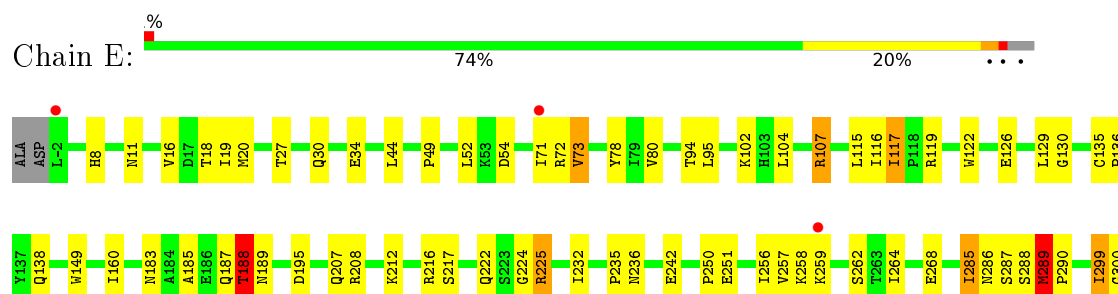
• Molecule 1: hemagglutinin HA1

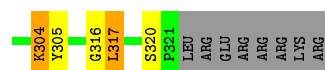


• Molecule 1: hemagglutinin HA1



• Molecule 1: hemagglutinin HA1





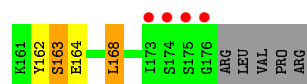
• Molecule 2: hemagglutinin HA2

Chain B: 73% 19% . . .



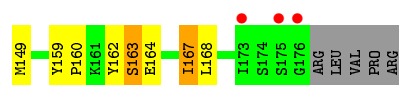
• Molecule 2: hemagglutinin HA2

Chain D: 3% 72% 19% . . .



• Molecule 2: hemagglutinin HA2

Chain F: 3% 71% 19% 5% . . .



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	125.33 Å 252.38 Å 70.62 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.66 – 2.81 49.66 – 2.81	Depositor EDS
% Data completeness (in resolution range)	96.0 (49.66-2.81) 96.1 (49.66-2.81)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.90 (at 2.81 Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.228 , 0.256 0.228 , 0.256	Depositor DCC
R_{free} test set	2722 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	57.2	Xtriage
Anisotropy	0.236	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 20.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12219	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% 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: 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	0.57	0/2622	0.81	1/3565 (0.0%)
1	C	0.61	0/2622	0.83	2/3565 (0.1%)
1	E	0.62	0/2622	0.83	1/3565 (0.0%)
2	B	0.70	1/1446 (0.1%)	0.83	1/1943 (0.1%)
2	D	0.63	1/1446 (0.1%)	0.82	2/1943 (0.1%)
2	F	0.63	0/1446	0.81	2/1943 (0.1%)
All	All	0.62	2/12204 (0.0%)	0.82	9/16524 (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
2	B	0	4
2	D	0	4
2	F	0	4
All	All	0	12

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	64	GLU	CG-CD	5.52	1.60	1.51
2	D	64	GLU	CG-CD	5.22	1.59	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	289	MET	CG-SD-CE	6.28	110.24	100.20
1	C	289	MET	CG-SD-CE	5.57	109.10	100.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	289	MET	CG-SD-CE	5.53	109.05	100.20
2	F	65	ALA	N-CA-C	5.36	125.48	111.00
2	B	65	ALA	N-CA-C	5.30	125.30	111.00

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	5	ALA	Peptide
2	B	61	THR	Peptide
2	B	64	GLU	Peptide
2	B	67	GLY	Peptide
2	D	5	ALA	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	2559	0	2512	37	0
1	C	2559	0	2512	33	0
1	E	2559	0	2512	37	0
2	B	1419	0	1329	21	0
2	D	1419	0	1329	24	0
2	F	1419	0	1329	27	0
3	A	56	0	50	0	0
3	B	28	0	25	0	0
3	C	56	0	50	0	0
3	D	28	0	25	1	0
3	E	56	0	50	0	0
3	F	28	0	25	0	0
4	A	11	0	10	0	0
4	C	11	0	10	0	0
4	E	11	0	10	0	0
All	All	12219	0	11778	167	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 7.

The worst 5 of 167 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:A:117:ILE:O	1:A:117:ILE:HD13	1.74	0.87
1:C:117:ILE:O	1:C:117:ILE:HD13	1.74	0.87
1:E:117:ILE:HD13	1:E:117:ILE:O	1.75	0.87
2:B:59:MET:O	2:B:61:THR:N	2.10	0.84
2:D:59:MET:O	2:D:61:THR:N	2.10	0.83

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	322/334 (96%)	287 (89%)	31 (10%)	4 (1%)	16	45
1	C	322/334 (96%)	289 (90%)	29 (9%)	4 (1%)	16	45
1	E	322/334 (96%)	287 (89%)	32 (10%)	3 (1%)	21	53
2	B	174/181 (96%)	161 (92%)	8 (5%)	5 (3%)	6	19
2	D	174/181 (96%)	160 (92%)	10 (6%)	4 (2%)	8	25
2	F	174/181 (96%)	159 (91%)	11 (6%)	4 (2%)	8	25
All	All	1488/1545 (96%)	1343 (90%)	121 (8%)	24 (2%)	12	36

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	71	ILE
2	B	60	ASN
1	C	71	ILE
2	D	60	ASN
1	E	71	ILE

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	288/297 (97%)	252 (88%)	36 (12%)	6	17
1	C	288/297 (97%)	250 (87%)	38 (13%)	5	14
1	E	288/297 (97%)	253 (88%)	35 (12%)	6	18
2	B	150/155 (97%)	130 (87%)	20 (13%)	5	14
2	D	150/155 (97%)	130 (87%)	20 (13%)	5	14
2	F	150/155 (97%)	127 (85%)	23 (15%)	3	10
All	All	1314/1356 (97%)	1142 (87%)	172 (13%)	5	14

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	217	SER
2	D	29	GLU
2	F	66	VAL
1	C	236	ASN
1	C	287	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	222	GLN
1	C	236	ASN
1	E	183	ASN
1	C	183	ASN
1	E	222	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

21 non-standard protein/DNA/RNA residues 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	501	1	14,14,15	1.01	1 (7%)	15,19,21	1.80	4 (26%)
3	NAG	A	502	1,3	14,14,15	0.64	0	15,19,21	1.31	1 (6%)
3	NAG	A	503	3,4	14,14,15	0.74	0	15,19,21	1.29	1 (6%)
4	MAN	A	504	3	11,11,12	0.74	0	15,15,17	1.58	3 (20%)
3	NAG	A	505	1	14,14,15	0.61	0	15,19,21	1.49	3 (20%)
3	NAG	B	201	3,2	14,14,15	1.00	1 (7%)	15,19,21	1.96	4 (26%)
3	NAG	B	202	3	14,14,15	0.71	0	15,19,21	1.63	2 (13%)
3	NAG	C	501	1	14,14,15	0.57	0	15,19,21	2.42	3 (20%)
3	NAG	C	502	1,3	14,14,15	0.94	1 (7%)	15,19,21	2.12	3 (20%)
3	NAG	C	503	3,4	14,14,15	0.77	0	15,19,21	1.15	0
4	MAN	C	504	3	11,11,12	0.67	0	15,15,17	2.81	5 (33%)
3	NAG	C	505	1	14,14,15	0.70	0	15,19,21	1.60	2 (13%)
3	NAG	D	201	3,2	14,14,15	0.65	0	15,19,21	1.20	3 (20%)
3	NAG	D	202	3	14,14,15	0.57	0	15,19,21	2.30	3 (20%)
3	NAG	E	501	1	14,14,15	0.88	1 (7%)	15,19,21	1.49	1 (6%)
3	NAG	E	502	1,3	14,14,15	0.60	0	15,19,21	2.52	6 (40%)
3	NAG	E	503	3,4	14,14,15	0.70	0	15,19,21	1.41	2 (13%)
4	MAN	E	504	3	11,11,12	0.90	1 (9%)	15,15,17	1.67	4 (26%)
3	NAG	E	505	1	14,14,15	0.78	0	15,19,21	1.26	1 (6%)
3	NAG	F	201	3,2	14,14,15	0.52	0	15,19,21	1.27	2 (13%)
3	NAG	F	202	3	14,14,15	0.68	0	15,19,21	1.59	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
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1
3	NAG	A	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	503	3,4	-	0/6/23/26	0/1/1/1
4	MAN	A	504	3	1/1/4/5	0/2/19/22	1/1/1/1
3	NAG	A	505	1	-	0/6/23/26	0/1/1/1
3	NAG	B	201	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	202	3	-	0/6/23/26	0/1/1/1
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
3	NAG	C	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	503	3,4	-	0/6/23/26	0/1/1/1
4	MAN	C	504	3	1/1/4/5	0/2/19/22	0/1/1/1
3	NAG	C	505	1	-	0/6/23/26	0/1/1/1
3	NAG	D	201	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	202	3	-	0/6/23/26	0/1/1/1
3	NAG	E	501	1	-	0/6/23/26	0/1/1/1
3	NAG	E	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	503	3,4	-	0/6/23/26	0/1/1/1
4	MAN	E	504	3	1/1/4/5	0/2/19/22	0/1/1/1
3	NAG	E	505	1	-	0/6/23/26	0/1/1/1
3	NAG	F	201	3,2	-	0/6/23/26	0/1/1/1
3	NAG	F	202	3	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	504	MAN	C2-C3	2.09	1.55	1.52
3	C	502	NAG	C1-C2	2.21	1.55	1.52
3	E	501	NAG	C1-C2	2.40	1.55	1.52
3	B	201	NAG	C1-C2	2.62	1.56	1.52
3	A	501	NAG	C1-C2	2.95	1.56	1.52

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	502	NAG	O4-C4-C3	-3.53	102.39	110.36
4	E	504	MAN	O5-C1-C2	-3.15	105.86	110.89
4	C	504	MAN	C3-C4-C5	-3.13	104.64	110.23
3	D	202	NAG	C4-C3-C2	-3.05	106.61	111.34
3	B	201	NAG	O7-C7-C8	-2.89	116.76	122.07

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	E	504	MAN	C1
4	C	504	MAN	C1
4	A	504	MAN	C1

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	504	MAN	C1-C2-C3-C4-C5-O5

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	202	NAG	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

21 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
3	NAG	A	501	1	14,14,15	1.01	1 (7%)	15,19,21	1.80	4 (26%)
3	NAG	A	502	1,3	14,14,15	0.64	0	15,19,21	1.31	1 (6%)
3	NAG	A	503	3,4	14,14,15	0.74	0	15,19,21	1.29	1 (6%)
4	MAN	A	504	3	11,11,12	0.74	0	15,15,17	1.58	3 (20%)
3	NAG	A	505	1	14,14,15	0.61	0	15,19,21	1.49	3 (20%)
3	NAG	B	201	3,2	14,14,15	1.00	1 (7%)	15,19,21	1.96	4 (26%)
3	NAG	B	202	3	14,14,15	0.71	0	15,19,21	1.63	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NAG	C	501	1	14,14,15	0.57	0	15,19,21	2.42	3 (20%)
3	NAG	C	502	1,3	14,14,15	0.94	1 (7%)	15,19,21	2.12	3 (20%)
3	NAG	C	503	3,4	14,14,15	0.77	0	15,19,21	1.15	0
4	MAN	C	504	3	11,11,12	0.67	0	15,15,17	2.81	5 (33%)
3	NAG	C	505	1	14,14,15	0.70	0	15,19,21	1.60	2 (13%)
3	NAG	D	201	3,2	14,14,15	0.65	0	15,19,21	1.20	3 (20%)
3	NAG	D	202	3	14,14,15	0.57	0	15,19,21	2.30	3 (20%)
3	NAG	E	501	1	14,14,15	0.88	1 (7%)	15,19,21	1.49	1 (6%)
3	NAG	E	502	1,3	14,14,15	0.60	0	15,19,21	2.52	6 (40%)
3	NAG	E	503	3,4	14,14,15	0.70	0	15,19,21	1.41	2 (13%)
4	MAN	E	504	3	11,11,12	0.90	1 (9%)	15,15,17	1.67	4 (26%)
3	NAG	E	505	1	14,14,15	0.78	0	15,19,21	1.26	1 (6%)
3	NAG	F	201	3,2	14,14,15	0.52	0	15,19,21	1.27	2 (13%)
3	NAG	F	202	3	14,14,15	0.68	0	15,19,21	1.59	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
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1
3	NAG	A	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	503	3,4	-	0/6/23/26	0/1/1/1
4	MAN	A	504	3	1/1/4/5	0/2/19/22	1/1/1/1
3	NAG	A	505	1	-	0/6/23/26	0/1/1/1
3	NAG	B	201	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	202	3	-	0/6/23/26	0/1/1/1
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
3	NAG	C	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	503	3,4	-	0/6/23/26	0/1/1/1
4	MAN	C	504	3	1/1/4/5	0/2/19/22	0/1/1/1
3	NAG	C	505	1	-	0/6/23/26	0/1/1/1
3	NAG	D	201	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	202	3	-	0/6/23/26	0/1/1/1
3	NAG	E	501	1	-	0/6/23/26	0/1/1/1
3	NAG	E	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	503	3,4	-	0/6/23/26	0/1/1/1
4	MAN	E	504	3	1/1/4/5	0/2/19/22	0/1/1/1
3	NAG	E	505	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
3	NAG	F	201	3,2	-	0/6/23/26	0/1/1/1
3	NAG	F	202	3	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	504	MAN	C2-C3	2.09	1.55	1.52
3	C	502	NAG	C1-C2	2.21	1.55	1.52
3	E	501	NAG	C1-C2	2.40	1.55	1.52
3	B	201	NAG	C1-C2	2.62	1.56	1.52
3	A	501	NAG	C1-C2	2.95	1.56	1.52

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	502	NAG	O4-C4-C3	-3.53	102.39	110.36
4	E	504	MAN	O5-C1-C2	-3.15	105.86	110.89
4	C	504	MAN	C3-C4-C5	-3.13	104.64	110.23
3	D	202	NAG	C4-C3-C2	-3.05	106.61	111.34
3	B	201	NAG	O7-C7-C8	-2.89	116.76	122.07

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	E	504	MAN	C1
4	A	504	MAN	C1
4	C	504	MAN	C1

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	504	MAN	C1-C2-C3-C4-C5-O5

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	202	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	324/334 (97%)	-0.31	5 (1%) 76 68	38, 57, 89, 190	0
1	C	324/334 (97%)	-0.28	6 (1%) 70 59	32, 51, 92, 162	0
1	E	324/334 (97%)	-0.34	3 (0%) 85 79	32, 49, 88, 134	0
2	B	176/181 (97%)	-0.22	0 100 100	34, 57, 83, 99	0
2	D	176/181 (97%)	0.03	5 (2%) 56 44	33, 79, 144, 197	0
2	F	176/181 (97%)	0.09	5 (2%) 56 44	34, 82, 125, 149	0
All	All	1500/1545 (97%)	-0.21	24 (1%) 74 65	32, 56, 114, 197	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	71	ILE	6.8
1	C	68	ASP	6.3
1	A	69	GLU	6.1
2	D	173	ILE	5.9
1	C	67	CYS	5.5

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

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
3	NAG	D	201	14/15	0.65	0.38	4.39	131,139,152,159	0
3	NAG	F	201	14/15	0.72	0.28	2.95	123,146,153,160	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	C	502	14/15	0.94	0.15	-0.39	60,66,75,89	0
3	NAG	A	502	14/15	0.92	0.11	-0.53	69,72,82,95	0
3	NAG	E	502	14/15	0.89	0.15	-0.96	66,74,88,100	0
3	NAG	A	503	14/15	0.81	0.22	-	102,113,125,137	0
3	NAG	B	201	14/15	0.79	0.32	-	102,115,119,123	0
3	NAG	E	503	14/15	0.77	0.36	-	98,109,123,137	0
3	NAG	C	503	14/15	0.90	0.26	-	83,103,121,130	0
3	NAG	C	501	14/15	0.87	0.14	-	72,87,93,99	0
3	NAG	D	202	14/15	0.69	0.44	-	123,143,156,159	0
3	NAG	A	501	14/15	0.81	0.22	-	82,99,111,111	0
3	NAG	A	505	14/15	0.78	0.23	-	98,123,132,138	0
3	NAG	E	505	14/15	0.59	0.42	-	107,133,147,147	0
4	MAN	E	504	11/12	0.69	0.41	-	106,129,137,141	0
3	NAG	F	202	14/15	0.67	0.39	-	123,138,146,149	0
4	MAN	C	504	11/12	0.55	0.33	-	116,135,142,150	0
3	NAG	B	202	14/15	0.71	0.28	-	108,129,134,134	0
4	MAN	A	504	11/12	0.66	0.27	-	111,131,136,145	0
3	NAG	C	505	14/15	0.80	0.23	-	90,103,114,117	0
3	NAG	E	501	14/15	0.73	0.21	-	100,110,120,121	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

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(\AA^2)	Q<0.9
3	NAG	D	201	14/15	0.65	0.38	4.39	131,139,152,159	0
3	NAG	F	201	14/15	0.72	0.28	2.95	123,146,153,160	0
3	NAG	C	502	14/15	0.94	0.15	-0.39	60,66,75,89	0
3	NAG	A	502	14/15	0.92	0.11	-0.53	69,72,82,95	0
3	NAG	E	502	14/15	0.89	0.15	-0.96	66,74,88,100	0
3	NAG	A	503	14/15	0.81	0.22	-	102,113,125,137	0
3	NAG	F	202	14/15	0.67	0.39	-	123,138,146,149	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	MAN	E	504	11/12	0.69	0.41	-	106,129,137,141	0
3	NAG	E	503	14/15	0.77	0.36	-	98,109,123,137	0
3	NAG	B	202	14/15	0.71	0.28	-	108,129,134,134	0
3	NAG	E	505	14/15	0.59	0.42	-	107,133,147,147	0
4	MAN	A	504	11/12	0.66	0.27	-	111,131,136,145	0
3	NAG	E	501	14/15	0.73	0.21	-	100,110,120,121	0
3	NAG	A	501	14/15	0.81	0.22	-	82,99,111,111	0
3	NAG	C	503	14/15	0.90	0.26	-	83,103,121,130	0
3	NAG	C	501	14/15	0.87	0.14	-	72,87,93,99	0
4	MAN	C	504	11/12	0.55	0.33	-	116,135,142,150	0
3	NAG	A	505	14/15	0.78	0.23	-	98,123,132,138	0
3	NAG	D	202	14/15	0.69	0.44	-	123,143,156,159	0
3	NAG	C	505	14/15	0.80	0.23	-	90,103,114,117	0
3	NAG	B	201	14/15	0.79	0.32	-	102,115,119,123	0

6.5 Other polymers [i](#)

There are no such residues in this entry.