



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 23, 2017 – 02:57 PM EST

PDB ID : 5HMG  
Title : REFINEMENT OF THE INFLUENZA VIRUS HEMAGGLUTININ BY SIMULATED ANNEALING  
Authors : Weis, W.I.; Bruenger, A.T.; Skehel, J.J.; Wiley, D.C.  
Deposited on : 1989-09-11  
Resolution : 3.20 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](mailto: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-20028442  
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-20028442

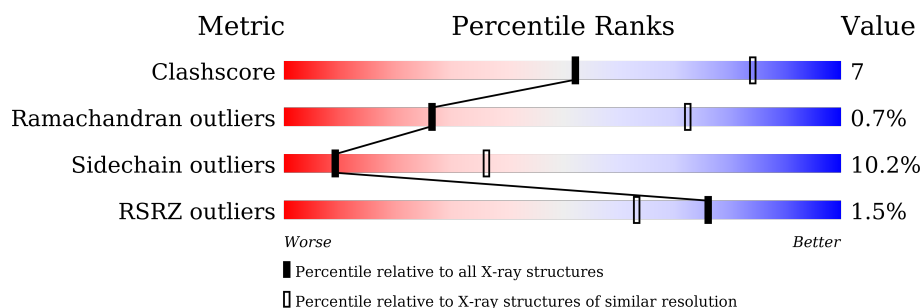
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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))
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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  $\geq 3$ , 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	328	<div> <div>2%</div> <div>72%</div> <div>24%</div> <div>.</div> </div>
1	C	328	<div> <div>2%</div> <div>72%</div> <div>24%</div> <div>..</div> </div>
1	E	328	<div> <div>2%</div> <div>73%</div> <div>24%</div> <div>.</div> </div>
2	B	175	<div> <div>%</div> <div>70%</div> <div>25%</div> <div>..</div> </div>
2	D	175	<div> <div>%</div> <div>70%</div> <div>26%</div> <div>..</div> </div>
2	F	175	<div> <div>%</div> <div>73%</div> <div>23%</div> <div>..</div> </div>

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	SIA	A	329	-	-	-	X
3	SIA	C	329	-	-	-	X
3	SIA	E	329	-	-	-	X
4	NAG	A	330	-	-	-	X
4	NAG	A	348	-	-	-	X
4	NAG	C	330	-	-	-	X
4	NAG	C	348	-	-	-	X
4	NAG	E	330	-	-	-	X
4	NAG	E	348	-	-	-	X
5	NAG	A	340	-	-	-	X
5	MAN	A	341	X	-	-	-
5	NAG	C	340	-	-	-	X
5	MAN	C	341	X	-	-	-
5	MAN	E	341	X	-	-	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12234 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 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	0	0
			2532	1581	445	493	13			
1	C	328	Total	C	N	O	S	0	0	0
			2532	1581	445	493	13			
1	E	328	Total	C	N	O	S	0	0	0
			2532	1581	445	493	13			

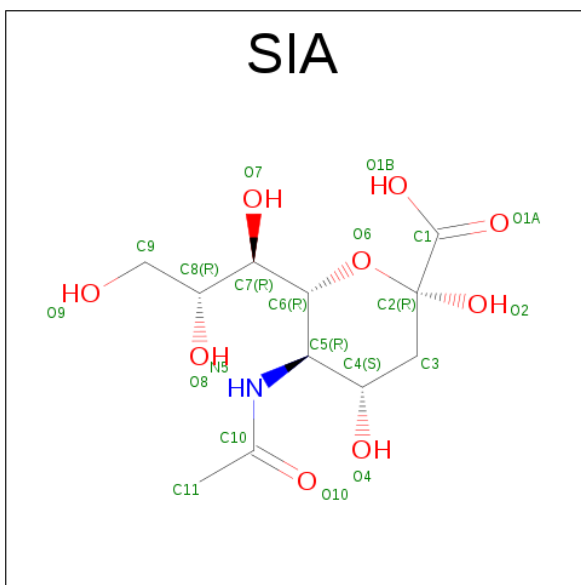
- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	175	Total	C	N	O	S	0	0	0
			1417	880	250	281	6			
2	D	175	Total	C	N	O	S	0	0	0
			1417	880	250	281	6			
2	F	175	Total	C	N	O	S	0	0	0
			1417	880	250	281	6			

There are 3 discrepancies between the modelled and reference sequences:

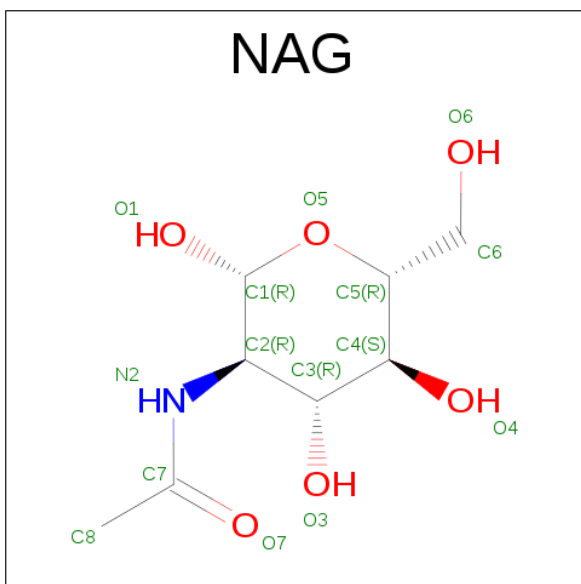
Chain	Residue	Modelled	Actual	Comment	Reference
B	112	GLY	ASP	CONFLICT	UNP P03437
D	112	GLY	ASP	CONFLICT	UNP P03437
F	112	GLY	ASP	CONFLICT	UNP P03437

- Molecule 3 is SUGAR (O-SIALIC ACID) (three-letter code: SIA) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			21	11	1	9		
3	C	1	Total	C	N	O	0	0
			21	11	1	9		
3	E	1	Total	C	N	O	0	0
			21	11	1	9		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	8	Total	O	0	0
			8	8		
6	B	5	Total	O	0	0
			5	5		
6	C	8	Total	O	0	0
			8	8		

*Continued on next page...*

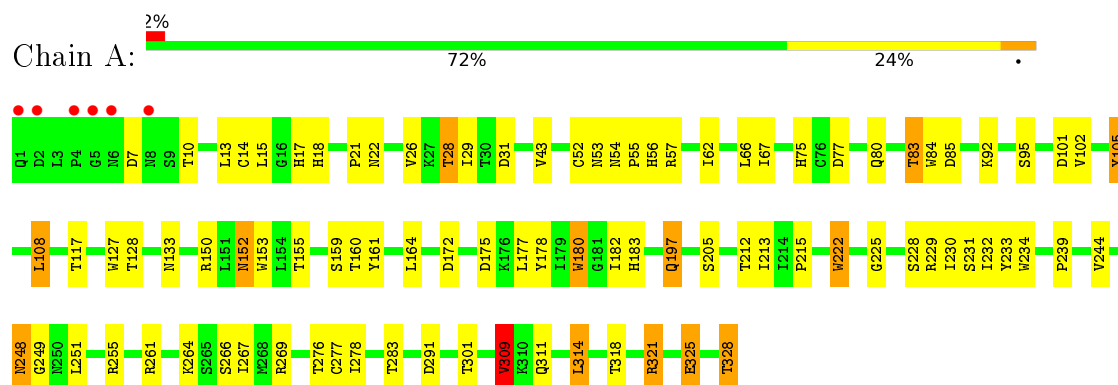
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	5	Total	O	0	0
			5	5		
6	E	9	Total	O	0	0
			9	9		
6	F	4	Total	O	0	0
			4	4		

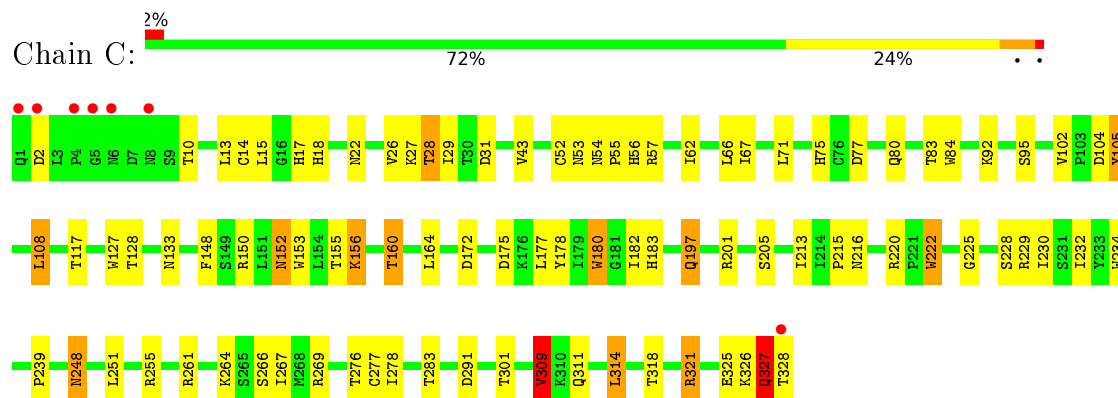
### 3 Residue-property plots [i](#)

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.

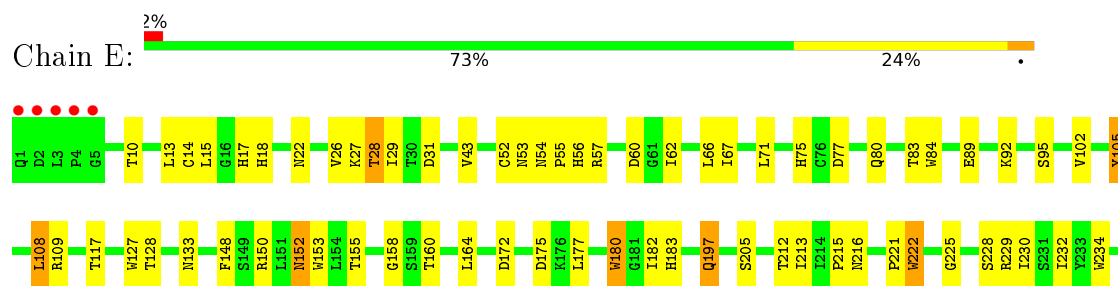
#### • Molecule 1: Hemagglutinin HA1 chain



#### • Molecule 1: Hemagglutinin HA1 chain



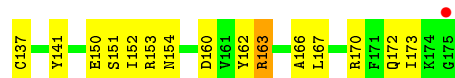
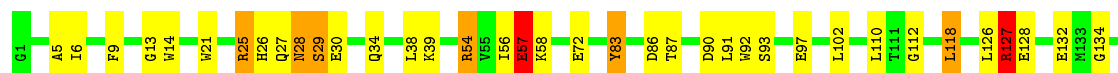
#### • Molecule 1: Hemagglutinin HA1 chain



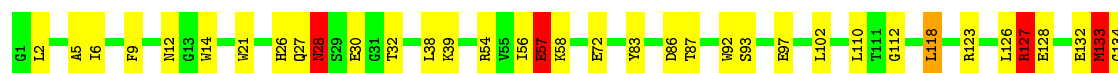




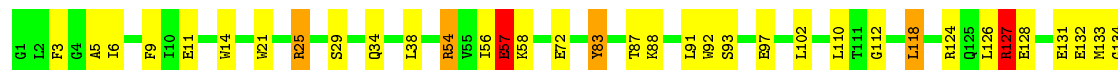
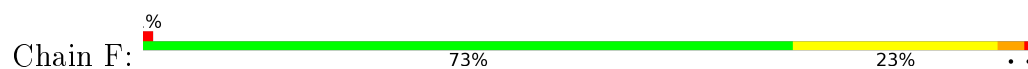
• Molecule 2: Hemagglutinin HA2 chain



• Molecule 2: Hemagglutinin HA2 chain



• Molecule 2: Hemagglutinin HA2 chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	162.60Å 162.60Å 177.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.20 7.00 – 3.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-3.20) 78.9 (7.00-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.220 , (Not available) (Not available) , (Not available)	Depositor DCC
$R_{free}$ test set	NotAvailable	DCC
Wilson B-factor (Å <sup>2</sup> )	41.4	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 46.7	EDS
L-test for twinning <sup>1</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.056 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	12234	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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, 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	0.86	0/2589	1.41	29/3527 (0.8%)
1	C	0.87	0/2589	1.43	32/3527 (0.9%)
1	E	0.88	0/2589	1.40	28/3527 (0.8%)
2	B	0.91	0/1441	1.48	22/1933 (1.1%)
2	D	0.93	0/1441	1.47	24/1933 (1.2%)
2	F	0.92	0/1441	1.47	22/1933 (1.1%)
All	All	0.89	0/12090	1.44	157/16380 (1.0%)

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
5	A	1	0
5	C	1	0
5	E	1	0
All	All	3	0

There are no bond length outliers.

The worst 5 of 157 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	255	ARG	NE-CZ-NH1	-9.95	115.32	120.30
2	B	170	ARG	NE-CZ-NH1	-9.36	115.62	120.30
2	B	25	ARG	NE-CZ-NH1	-9.35	115.63	120.30
1	A	255	ARG	NE-CZ-NH1	-9.12	115.74	120.30
1	C	321	ARG	NE-CZ-NH1	-9.09	115.75	120.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	341	MAN	C1
5	C	341	MAN	C1
5	E	341	MAN	C1

There are no planarity outliers.

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2532	0	2473	40	0
1	C	2532	0	2473	39	0
1	E	2532	0	2473	38	0
2	B	1417	0	1344	28	0
2	D	1417	0	1344	28	0
2	F	1417	0	1344	25	0
3	A	21	0	18	1	0
3	C	21	0	18	1	0
3	E	21	0	18	1	0
4	A	42	0	39	0	0
4	B	14	0	13	0	0
4	C	42	0	39	0	0
4	D	14	0	13	0	0
4	E	42	0	39	0	0
4	F	14	0	13	0	0
5	A	39	0	34	0	0
5	C	39	0	34	0	0
5	E	39	0	34	0	0
6	A	8	0	0	0	0
6	B	5	0	0	0	0
6	C	8	0	0	0	0
6	D	5	0	0	0	0
6	E	9	0	0	0	0
6	F	4	0	0	0	0
All	All	12234	0	11763	161	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 161 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)
2:B:26:HIS:HD2	2:B:28:ASN:HB3	1.47	0.78
1:A:10:THR:HG22	2:B:141:TYR:HA	1.70	0.73
1:E:10:THR:HG22	2:F:141:TYR:HA	1.70	0.73
2:B:132:GLU:HG2	2:B:134:GLY:H	1.53	0.72
2:D:132:GLU:HG2	2:D:134:GLY:H	1.55	0.72

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	326/328 (99%)	307 (94%)	18 (6%)	1 (0%)	46	85
1	C	326/328 (99%)	311 (95%)	13 (4%)	2 (1%)	30	75
1	E	326/328 (99%)	312 (96%)	12 (4%)	2 (1%)	30	75
2	B	173/175 (99%)	161 (93%)	9 (5%)	3 (2%)	11	52
2	D	173/175 (99%)	162 (94%)	10 (6%)	1 (1%)	30	75
2	F	173/175 (99%)	163 (94%)	9 (5%)	1 (1%)	30	75
All	All	1497/1509 (99%)	1416 (95%)	71 (5%)	10 (1%)	26	72

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	29	SER
2	B	30	GLU
1	C	327	GLN
1	A	62	ILE
1	C	62	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	289/289 (100%)	257 (89%)	32 (11%)	8	32
1	C	289/289 (100%)	258 (89%)	31 (11%)	8	34
1	E	289/289 (100%)	258 (89%)	31 (11%)	8	34
2	B	148/148 (100%)	134 (90%)	14 (10%)	11	40
2	D	148/148 (100%)	134 (90%)	14 (10%)	11	40
2	F	148/148 (100%)	136 (92%)	12 (8%)	15	51
All	All	1311/1311 (100%)	1177 (90%)	134 (10%)	9	36

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	150	ARG
1	C	321	ARG
2	F	29	SER
1	C	156	LYS
1	C	248	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	248	ASN
1	C	327	GLN
1	E	152	ASN
1	C	152	ASN
2	D	49	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 ⓘ

9 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	339	1,5	14,14,15	0.77	1 (7%)	15,19,21	0.72	0
5	NAG	A	340	5	14,14,15	0.42	0	15,19,21	0.95	1 (6%)
5	MAN	A	341	5	11,11,12	0.84	0	15,15,17	1.28	2 (13%)
5	NAG	C	339	1,5	14,14,15	0.85	1 (7%)	15,19,21	0.77	0
5	NAG	C	340	5	14,14,15	0.46	0	15,19,21	0.95	1 (6%)
5	MAN	C	341	5	11,11,12	0.68	0	15,15,17	1.35	2 (13%)
5	NAG	E	339	1,5	14,14,15	0.61	0	15,19,21	0.78	1 (6%)
5	NAG	E	340	5	14,14,15	0.47	0	15,19,21	0.99	1 (6%)
5	MAN	E	341	5	11,11,12	0.78	0	15,15,17	1.41	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	339	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	340	5	-	0/6/23/26	0/1/1/1
5	MAN	A	341	5	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	C	339	1,5	-	0/6/23/26	0/1/1/1
5	NAG	C	340	5	-	0/6/23/26	0/1/1/1
5	MAN	C	341	5	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	E	339	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	340	5	-	0/6/23/26	0/1/1/1
5	MAN	E	341	5	1/1/4/5	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	339	NAG	O5-C1	-2.50	1.39	1.43
5	A	339	NAG	O5-C1	-2.09	1.40	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	341	MAN	C1-C2-C3	-2.66	106.33	109.55
5	C	340	NAG	C4-C3-C2	-2.60	107.31	111.34
5	E	340	NAG	C4-C3-C2	-2.53	107.42	111.34
5	A	340	NAG	C4-C3-C2	-2.38	107.65	111.34
5	E	339	NAG	C1-O5-C5	2.04	115.14	112.14

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	341	MAN	C1
5	C	341	MAN	C1
5	E	341	MAN	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

15 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	SIA	A	329	-	18,21,21	1.11	1 (5%)	18,31,31	1.44	3 (16%)
4	NAG	A	330	1	14,14,15	0.78	0	15,19,21	1.01	0
4	NAG	A	334	1	14,14,15	0.87	0	15,19,21	1.33	1 (6%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	348	1	14,14,15	0.79	1 (7%)	15,19,21	1.08	0
4	NAG	B	401	2	14,14,15	0.62	0	15,19,21	1.09	1 (6%)
3	SIA	C	329	-	18,21,21	1.12	1 (5%)	18,31,31	1.49	3 (16%)
4	NAG	C	330	1	14,14,15	0.72	0	15,19,21	1.09	0
4	NAG	C	334	1	14,14,15	0.85	0	15,19,21	1.25	1 (6%)
4	NAG	C	348	1	14,14,15	0.68	0	15,19,21	1.04	0
4	NAG	D	401	2	14,14,15	0.71	0	15,19,21	1.15	1 (6%)
3	SIA	E	329	-	18,21,21	1.11	1 (5%)	18,31,31	1.35	0
4	NAG	E	330	1	14,14,15	0.87	1 (7%)	15,19,21	1.01	0
4	NAG	E	334	1	14,14,15	0.86	0	15,19,21	1.31	1 (6%)
4	NAG	E	348	1	14,14,15	0.64	0	15,19,21	1.11	0
4	NAG	F	401	2	14,14,15	0.67	0	15,19,21	1.11	1 (6%)

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	329	-	-	0/14/38/38	0/1/1/1
4	NAG	A	330	1	-	0/6/23/26	0/1/1/1
4	NAG	A	334	1	-	0/6/23/26	0/1/1/1
4	NAG	A	348	1	-	0/6/23/26	0/1/1/1
4	NAG	B	401	2	-	0/6/23/26	0/1/1/1
3	SIA	C	329	-	-	0/14/38/38	0/1/1/1
4	NAG	C	330	1	-	0/6/23/26	0/1/1/1
4	NAG	C	334	1	-	0/6/23/26	0/1/1/1
4	NAG	C	348	1	-	0/6/23/26	0/1/1/1
4	NAG	D	401	2	-	0/6/23/26	0/1/1/1
3	SIA	E	329	-	-	0/14/38/38	0/1/1/1
4	NAG	E	330	1	-	0/6/23/26	0/1/1/1
4	NAG	E	334	1	-	0/6/23/26	0/1/1/1
4	NAG	E	348	1	-	0/6/23/26	0/1/1/1
4	NAG	F	401	2	-	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	A	348	NAG	C1-C2	2.20	1.55	1.52
4	E	330	NAG	C4-C5	2.41	1.58	1.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	329	SIA	C3-C2	4.06	1.56	1.51
3	C	329	SIA	C3-C2	4.13	1.57	1.51
3	E	329	SIA	C3-C2	4.15	1.57	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	329	SIA	C8-C7-C6	-2.18	108.48	113.08
3	A	329	SIA	C8-C7-C6	-2.05	108.75	113.08
3	A	329	SIA	C2-C3-C4	2.02	114.39	110.40
3	C	329	SIA	C11-C10-N5	2.06	120.04	116.10
3	A	329	SIA	C7-C6-C5	2.22	117.12	114.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	329	SIA	1	0
3	C	329	SIA	1	0
3	E	329	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	328/328 (100%)	-0.84	6 (1%) 71 58	3, 22, 51, 133	0
1	C	328/328 (100%)	-0.85	7 (2%) 67 52	2, 21, 47, 129	0
1	E	328/328 (100%)	-0.91	6 (1%) 71 58	3, 21, 46, 131	0
2	B	175/175 (100%)	-0.91	1 (0%) 90 84	2, 16, 47, 98	0
2	D	175/175 (100%)	-0.99	1 (0%) 90 84	2, 15, 46, 99	0
2	F	175/175 (100%)	-0.96	1 (0%) 90 84	2, 15, 46, 98	0
All	All	1509/1509 (100%)	-0.90	22 (1%) 76 63	2, 20, 50, 133	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	5	GLY	8.2
1	E	2	ASP	7.2
1	A	6	ASN	5.5
1	A	8	ASN	5.0
2	D	175	GLY	5.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
5	NAG	C	340	14/15	0.90	0.22	3.01	48,50,54,56	0
5	NAG	A	340	14/15	0.92	0.22	2.12	48,51,54,54	0
5	NAG	C	339	14/15	0.88	0.21	1.77	38,42,44,45	0
5	NAG	E	340	14/15	0.92	0.19	1.50	48,51,53,55	0
5	NAG	A	339	14/15	0.94	0.18	1.42	38,41,43,45	0
5	NAG	E	339	14/15	0.97	0.12	-0.30	37,40,41,44	0
5	MAN	A	341	11/12	0.87	0.26	-	57,59,60,60	0
5	MAN	C	341	11/12	0.86	0.32	-	60,62,64,65	0
5	MAN	E	341	11/12	0.83	0.30	-	59,61,63,64	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	330	14/15	0.90	0.26	7.95	38,42,44,44	0
3	SIA	C	329	21/21	0.80	0.35	6.00	49,50,50,50	21
4	NAG	E	330	14/15	0.90	0.25	4.74	39,43,44,44	0
4	NAG	E	348	14/15	0.86	0.27	4.61	36,40,42,42	0
4	NAG	A	348	14/15	0.89	0.20	4.16	34,39,40,40	0
3	SIA	E	329	21/21	0.87	0.28	3.97	49,49,50,50	21
3	SIA	A	329	21/21	0.83	0.30	3.56	49,49,50,50	21
4	NAG	C	330	14/15	0.93	0.18	3.52	38,41,43,43	0
4	NAG	C	348	14/15	0.92	0.20	2.89	35,39,41,42	0
4	NAG	F	401	14/15	0.85	0.30	-	50,54,56,56	0
4	NAG	E	334	14/15	0.90	0.20	-	38,40,43,44	0
4	NAG	D	401	14/15	0.87	0.28	-	51,55,56,57	0
4	NAG	A	334	14/15	0.88	0.22	-	37,41,42,43	0
4	NAG	C	334	14/15	0.92	0.18	-	38,41,42,43	0
4	NAG	B	401	14/15	0.90	0.23	-	51,54,56,56	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.