



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 17, 2023 – 01:12 PM EST

PDB ID : 4W8N
Title : The crystal structure of hemagglutinin from a swine influenza virus (A/swine/Missouri/2124514/2006)
Authors : Yang, H.; Carney, P.J.; Tumpey, T.M.; Stevens, J.
Deposited on : 2014-08-25
Resolution : 2.90 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

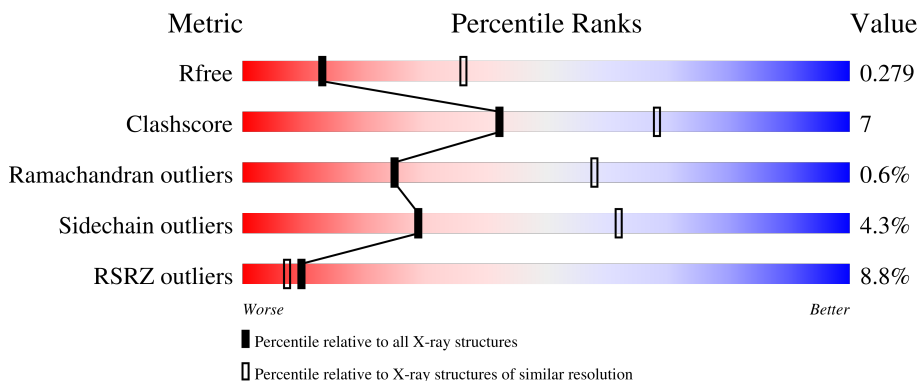
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	330	 2% 82% 15% ..
1	C	330	 6% 82% 16% ..
1	E	330	 18% 83% 14% .
2	B	179	 2% 75% 16% . . 6%
2	D	179	 4% 75% 16% 5% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	179	

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	E	402	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11829 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
			Total	C	N	O	S			
1	A	321	2520	1589	441	477	13	1	0	0
1	C	326	2563	1614	450	486	13	1	0	0
1	E	319	2506	1582	439	472	13	1	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	expression tag	UNP A9YN66
A	-3	ASP	-	expression tag	UNP A9YN66
A	-2	LEU	-	expression tag	UNP A9YN66
A	-1	GLY	-	expression tag	UNP A9YN66
A	0	SER	-	expression tag	UNP A9YN66
C	-4	ALA	-	expression tag	UNP A9YN66
C	-3	ASP	-	expression tag	UNP A9YN66
C	-2	LEU	-	expression tag	UNP A9YN66
C	-1	GLY	-	expression tag	UNP A9YN66
C	0	SER	-	expression tag	UNP A9YN66
E	-4	ALA	-	expression tag	UNP A9YN66
E	-3	ASP	-	expression tag	UNP A9YN66
E	-2	LEU	-	expression tag	UNP A9YN66
E	-1	GLY	-	expression tag	UNP A9YN66
E	0	SER	-	expression tag	UNP A9YN66

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	169	1363	843	236	275	9	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	172	1386	860	239	278	9	0	0	0
2	F	164	1337	827	231	270	9	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	173	SER	-	expression tag	UNP A9YN66
B	174	GLY	-	expression tag	UNP A9YN66
B	175	ARG	-	expression tag	UNP A9YN66
B	176	LEU	-	expression tag	UNP A9YN66
B	177	VAL	-	expression tag	UNP A9YN66
B	178	PRO	-	expression tag	UNP A9YN66
B	179	ARG	-	expression tag	UNP A9YN66
D	173	SER	-	expression tag	UNP A9YN66
D	174	GLY	-	expression tag	UNP A9YN66
D	175	ARG	-	expression tag	UNP A9YN66
D	176	LEU	-	expression tag	UNP A9YN66
D	177	VAL	-	expression tag	UNP A9YN66
D	178	PRO	-	expression tag	UNP A9YN66
D	179	ARG	-	expression tag	UNP A9YN66
F	173	SER	-	expression tag	UNP A9YN66
F	174	GLY	-	expression tag	UNP A9YN66
F	175	ARG	-	expression tag	UNP A9YN66
F	176	LEU	-	expression tag	UNP A9YN66
F	177	VAL	-	expression tag	UNP A9YN66
F	178	PRO	-	expression tag	UNP A9YN66
F	179	ARG	-	expression tag	UNP A9YN66

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).

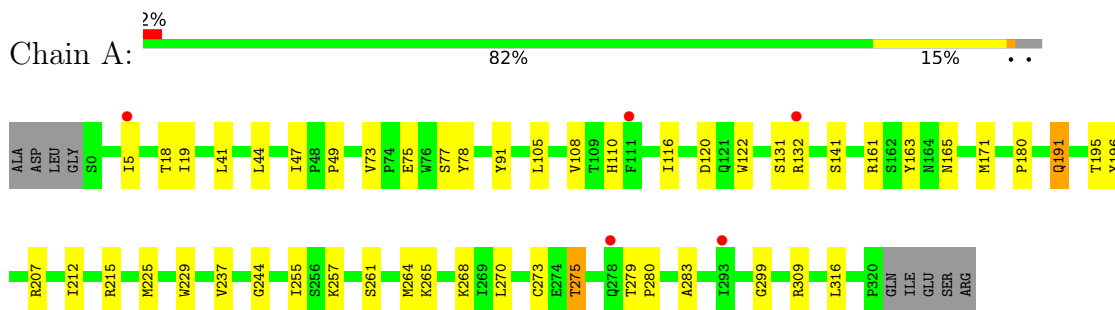


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
	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

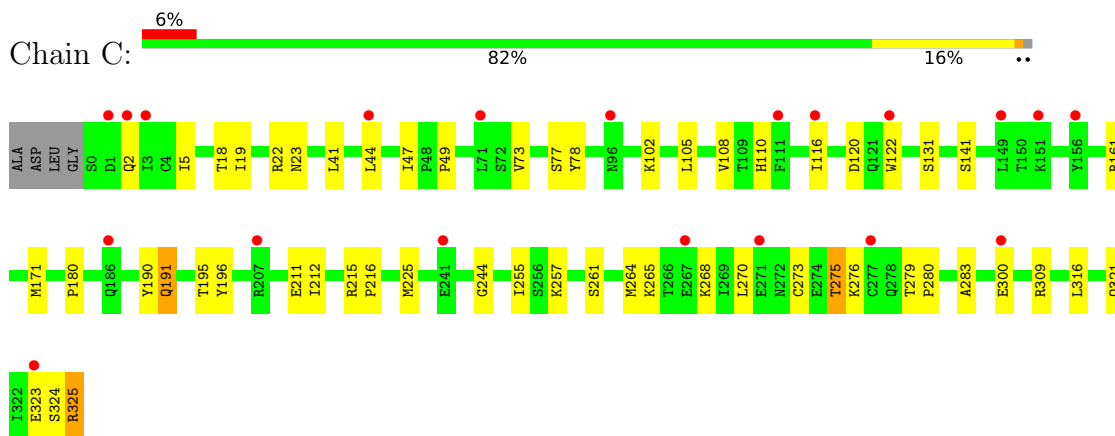
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

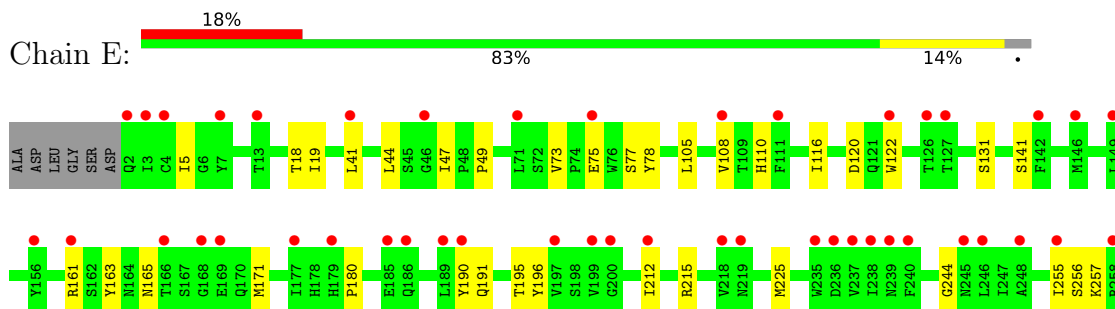
- Molecule 1: Hemagglutinin

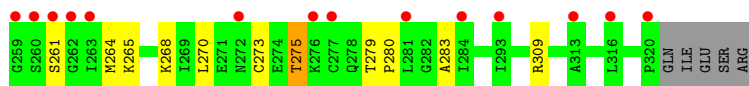


- Molecule 1: Hemagglutinin

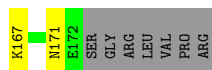
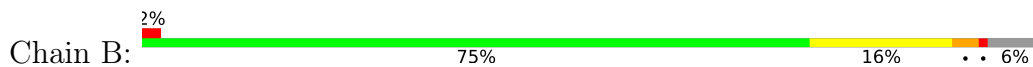


- Molecule 1: Hemagglutinin

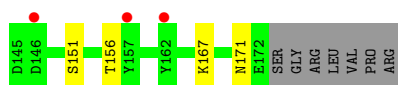
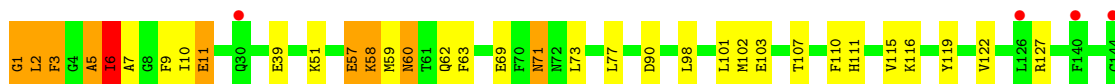
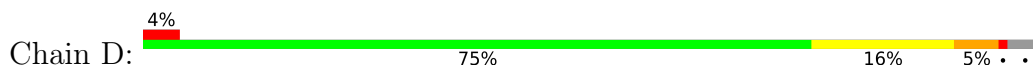




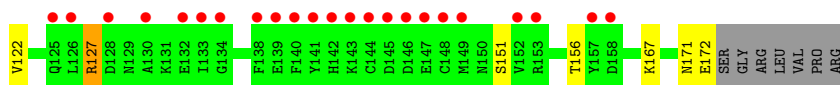
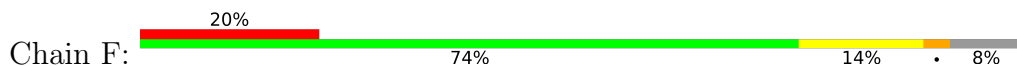
• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	70.29Å 160.03Å 170.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.90 44.90 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-2.90) 99.7 (44.90-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.82 (at 2.91Å)	Xtrriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.233 , 0.262 0.250 , 0.279	Depositor DCC
R_{free} test set	2170 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	86.8	Xtrriage
Anisotropy	0.124	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 56.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	11829	wwPDB-VP
Average B, all atoms (Å ²)	109.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.65	1/2579 (0.0%)	0.73	1/3505 (0.0%)
1	C	0.57	1/2622 (0.0%)	0.73	2/3562 (0.1%)
1	E	0.57	1/2565 (0.0%)	0.73	4/3486 (0.1%)
2	B	0.54	0/1387	0.66	1/1863 (0.1%)
2	D	0.54	0/1411	0.71	1/1895 (0.1%)
2	F	0.50	0/1361	0.65	1/1828 (0.1%)
All	All	0.57	3/11925 (0.0%)	0.71	10/16139 (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	A	0	1
1	C	0	1
2	B	0	3
2	D	0	4
2	F	0	2
All	All	0	11

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	190	TYR	C-N	-10.93	1.08	1.34
1	C	190	TYR	C-N	-6.15	1.20	1.34
1	A	229	TRP	CB-CG	-5.02	1.41	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	273	CYS	CA-CB-SG	-12.42	91.65	114.00
1	A	273	CYS	CA-CB-SG	-11.04	94.13	114.00
1	E	190	TYR	C-N-CA	8.50	142.96	121.70
1	E	190	TYR	O-C-N	-8.09	109.75	122.70
1	E	273	CYS	CA-CB-SG	-6.83	101.72	114.00

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	191	GLN	Mainchain
2	B	5	ALA	Peptide
2	B	57	GLU	Peptide
2	B	58	LYS	Peptide
1	C	191	GLN	Mainchain

5.2 Too-close contacts [i](#)

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	2520	0	2491	31	6
1	C	2563	0	2532	38	6
1	E	2506	0	2478	27	0
2	B	1363	0	1262	25	0
2	D	1386	0	1289	40	0
2	F	1337	0	1236	21	0
3	A	42	0	39	0	0
3	B	14	0	13	0	0
3	C	56	0	52	0	0
3	E	42	0	39	0	0
All	All	11829	0	11431	160	6

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 160 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:325:ARG:O	2:D:1:GLY:N	2.00	0.94
1:E:19:ILE:HD11	2:F:102:MET:HG2	1.53	0.91
1:E:255:ILE:HD12	1:E:255:ILE:O	1.73	0.89
1:A:255:ILE:O	1:A:255:ILE:HD12	1.73	0.88
2:D:1:GLY:HA3	2:D:2:LEU:HB2	1.54	0.87

The worst 5 of 6 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:132:ARG:NH2	1:C:323:GLU:OE2[2_454]	1.09	1.11
1:A:132:ARG:NE	1:C:323:GLU:OE2[2_454]	1.70	0.50
1:A:132:ARG:NH2	1:C:323:GLU:CD[2_454]	1.81	0.39
1:A:132:ARG:CZ	1:C:323:GLU:CD[2_454]	1.93	0.27
1:A:132:ARG:CZ	1:C:323:GLU:OE1[2_454]	2.18	0.02

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	319/330 (97%)	304 (95%)	15 (5%)	0	100	100
1	C	324/330 (98%)	308 (95%)	16 (5%)	0	100	100
1	E	317/330 (96%)	303 (96%)	14 (4%)	0	100	100
2	B	167/179 (93%)	145 (87%)	19 (11%)	3 (2%)	8	29
2	D	170/179 (95%)	147 (86%)	19 (11%)	4 (2%)	6	22
2	F	162/179 (90%)	144 (89%)	16 (10%)	2 (1%)	13	40
All	All	1459/1527 (96%)	1351 (93%)	99 (7%)	9 (1%)	25	58

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	6	ILE
2	D	6	ILE
2	B	60	ASN
2	D	60	ASN
2	F	60	ASN

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	284/291 (98%)	273 (96%)	11 (4%)	32 66
1	C	289/291 (99%)	277 (96%)	12 (4%)	30 63
1	E	282/291 (97%)	271 (96%)	11 (4%)	32 66
2	B	144/154 (94%)	137 (95%)	7 (5%)	25 57
2	D	146/154 (95%)	139 (95%)	7 (5%)	25 58
2	F	143/154 (93%)	136 (95%)	7 (5%)	25 57
All	All	1288/1335 (96%)	1233 (96%)	55 (4%)	29 62

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	309	ARG
2	D	116	LYS
2	F	172	GLU
2	F	57	GLU
1	C	325	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	95	ASN
1	E	2	GLN
2	F	95	ASN
1	E	110	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	95	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

10 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	401	1	14,14,15	0.61	0	17,19,21	2.00	5 (29%)
3	NAG	E	403	1	14,14,15	0.84	1 (7%)	17,19,21	2.76	4 (23%)
3	NAG	A	403	1	14,14,15	0.40	0	17,19,21	1.75	4 (23%)
3	NAG	A	402	1	14,14,15	0.69	0	17,19,21	1.60	3 (17%)
3	NAG	C	403	1	14,14,15	0.71	0	17,19,21	1.70	4 (23%)
3	NAG	B	201	2	14,14,15	0.57	0	17,19,21	1.38	2 (11%)
3	NAG	E	401	1	14,14,15	0.49	0	17,19,21	1.74	3 (17%)
3	NAG	C	401	1	14,14,15	0.51	0	17,19,21	1.33	2 (11%)
3	NAG	C	402	1	14,14,15	0.77	0	17,19,21	0.96	0
3	NAG	C	404	1	14,14,15	0.65	0	17,19,21	1.82	5 (29%)

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	401	1	-	0/6/23/26	0/1/1/1
3	NAG	E	403	1	-	2/6/23/26	0/1/1/1
3	NAG	A	403	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	A	402	1	-	2/6/23/26	0/1/1/1
3	NAG	C	403	1	-	2/6/23/26	0/1/1/1
3	NAG	B	201	2	-	2/6/23/26	0/1/1/1
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
3	NAG	C	401	1	-	2/6/23/26	0/1/1/1
3	NAG	C	402	1	-	1/6/23/26	0/1/1/1
3	NAG	C	404	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	403	NAG	C1-C2	2.34	1.55	1.52

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	403	NAG	C1-O5-C5	9.65	125.26	112.19
3	A	403	NAG	C1-O5-C5	5.38	119.48	112.19
3	A	401	NAG	C1-O5-C5	5.35	119.43	112.19
3	C	403	NAG	C1-O5-C5	4.88	118.81	112.19
3	E	401	NAG	C1-O5-C5	4.76	118.64	112.19

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	401	NAG	O5-C5-C6-O6
3	B	201	NAG	O5-C5-C6-O6
3	B	201	NAG	C4-C5-C6-O6
3	C	403	NAG	O5-C5-C6-O6
3	E	403	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates i

There are no monosaccharides in this entry.

5.6 Ligand geometry

11 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	401	1	14,14,15	0.61	0	17,19,21	2.00	5 (29%)
3	NAG	E	403	1	14,14,15	0.84	1 (7%)	17,19,21	2.76	4 (23%)
3	NAG	A	403	1	14,14,15	0.40	0	17,19,21	1.75	4 (23%)
3	NAG	A	402	1	14,14,15	0.69	0	17,19,21	1.60	3 (17%)
3	NAG	C	403	1	14,14,15	0.71	0	17,19,21	1.70	4 (23%)
3	NAG	B	201	2	14,14,15	0.57	0	17,19,21	1.38	2 (11%)
3	NAG	E	401	1	14,14,15	0.49	0	17,19,21	1.74	3 (17%)
3	NAG	E	402	1	14,14,15	1.13	1 (7%)	17,19,21	2.55	6 (35%)
3	NAG	C	401	1	14,14,15	0.51	0	17,19,21	1.33	2 (11%)
3	NAG	C	402	1	14,14,15	0.77	0	17,19,21	0.96	0
3	NAG	C	404	1	14,14,15	0.65	0	17,19,21	1.82	5 (29%)

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	401	1	-	0/6/23/26	0/1/1/1
3	NAG	E	403	1	-	2/6/23/26	0/1/1/1
3	NAG	A	403	1	-	0/6/23/26	0/1/1/1
3	NAG	A	402	1	-	2/6/23/26	0/1/1/1
3	NAG	C	403	1	-	2/6/23/26	0/1/1/1
3	NAG	B	201	2	-	2/6/23/26	0/1/1/1
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
3	NAG	E	402	1	1/1/5/7	4/6/23/26	0/1/1/1
3	NAG	C	401	1	-	2/6/23/26	0/1/1/1
3	NAG	C	402	1	-	1/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	404	1	-	1/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	402	NAG	C1-C2	3.20	1.57	1.52
3	E	403	NAG	C1-C2	2.34	1.55	1.52

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	403	NAG	C1-O5-C5	9.65	125.26	112.19
3	E	402	NAG	O5-C1-C2	-7.38	99.63	111.29
3	A	403	NAG	C1-O5-C5	5.38	119.48	112.19
3	A	401	NAG	C1-O5-C5	5.35	119.43	112.19
3	C	403	NAG	C1-O5-C5	4.88	118.81	112.19

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	E	402	NAG	C1

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	402	NAG	O5-C5-C6-O6
3	E	402	NAG	C4-C5-C6-O6
3	B	201	NAG	O5-C5-C6-O6
3	C	401	NAG	O5-C5-C6-O6
3	B	201	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	190:TYR	C	191:GLN	N	1.19
1	E	190:TYR	C	191:GLN	N	1.08

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, 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	321/330 (97%)	0.19	5 (1%) 72 71	36, 81, 115, 132	1 (0%)
1	C	326/330 (98%)	0.51	20 (6%) 21 17	26, 106, 152, 180	1 (0%)
1	E	319/330 (96%)	1.01	59 (18%) 1 0	39, 124, 175, 212	1 (0%)
2	B	169/179 (94%)	0.28	3 (1%) 68 67	59, 105, 154, 170	0
2	D	172/179 (96%)	0.39	7 (4%) 37 32	60, 108, 163, 183	0
2	F	164/179 (91%)	1.06	36 (21%) 0 0	61, 133, 187, 216	0
All	All	1471/1527 (96%)	0.57	130 (8%) 10 7	26, 105, 168, 216	3 (0%)

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	140	PHE	7.9
1	E	320	PRO	7.8
1	E	71	LEU	6.5
1	E	260	SER	5.9
1	E	190	TYR	5.9

6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	B-factors(Å ²)	Q<0.9
3	NAG	A	401	14/15	0.62	0.34	153,165,178,181	0
3	NAG	E	401	14/15	0.73	0.31	107,148,162,172	0
3	NAG	E	403	14/15	0.74	0.18	106,135,141,147	0
3	NAG	C	401	14/15	0.75	0.23	131,140,147,149	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	NAG	B	201	14/15	0.77	0.34	136,180,196,197	0
3	NAG	C	402	14/15	0.79	0.17	106,137,162,164	0
3	NAG	C	404	14/15	0.82	0.17	99,119,133,133	0
3	NAG	C	403	14/15	0.87	0.18	95,112,127,131	0
3	NAG	A	402	14/15	0.88	0.11	96,115,120,121	0
3	NAG	A	403	14/15	0.90	0.12	70,89,97,103	0

6.3 Carbohydrates [i](#)

There are no monosaccharides 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. 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	B-factors(Å ²)	Q<0.9
3	NAG	A	401	14/15	0.62	0.34	153,165,178,181	0
3	NAG	E	402	14/15	0.64	0.21	150,169,176,178	0
3	NAG	E	401	14/15	0.73	0.31	107,148,162,172	0
3	NAG	E	403	14/15	0.74	0.18	106,135,141,147	0
3	NAG	C	401	14/15	0.75	0.23	131,140,147,149	0
3	NAG	B	201	14/15	0.77	0.34	136,180,196,197	0
3	NAG	C	402	14/15	0.79	0.17	106,137,162,164	0
3	NAG	C	404	14/15	0.82	0.17	99,119,133,133	0
3	NAG	C	403	14/15	0.87	0.18	95,112,127,131	0
3	NAG	A	402	14/15	0.88	0.11	96,115,120,121	0
3	NAG	A	403	14/15	0.90	0.12	70,89,97,103	0

6.5 Other polymers [i](#)

There are no such residues in this entry.