



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 15, 2017 – 12:59 PM EDT

PDB ID : 5VTW  
Title : Crystal structure of the A/Hong Kong/1/1968 (H3N2) influenza virus hemagglutinin G225M/L226T/S228A mutant in complex with 6'-SLN  
Authors : Wu, N.C.; Wilson, I.A.  
Deposited on : unknown  
Resolution : 2.35 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20029824

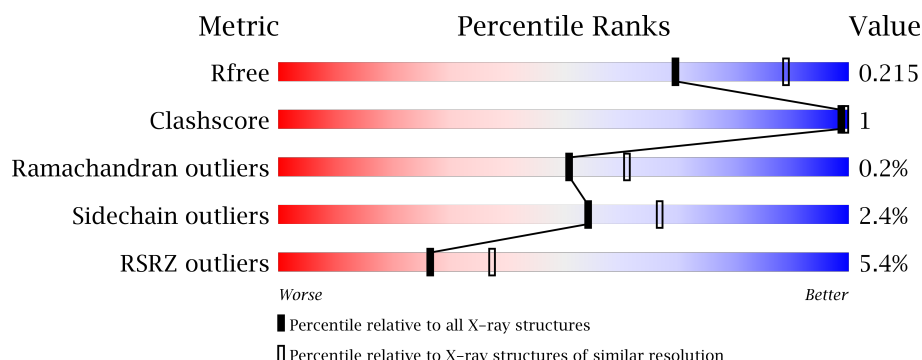
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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))
$R_{free}$	100719	1522 (2.38-2.34)
Clashscore	112137	1626 (2.38-2.34)
Ramachandran outliers	110173	1605 (2.38-2.34)
Sidechain outliers	110143	1606 (2.38-2.34)
RSRZ outliers	101464	1528 (2.38-2.34)

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	323	<div> <div>3%</div> <div>94%</div> <div>5%</div> </div>
1	C	323	<div> <div>2%</div> <div>94%</div> <div></div> </div>
1	E	323	<div> <div>2%</div> <div>94%</div> <div></div> </div>
2	B	174	<div> <div>14%</div> <div>97%</div> <div></div> </div>
2	D	174	<div> <div>10%</div> <div>95%</div> <div></div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	174	

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	A	406	-	-	-	X
3	NAG	A	408	-	-	-	X
3	NAG	E	506	-	-	-	X
5	SIA	A	410	-	-	-	X
5	SIA	C	508	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 12410 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	321	Total	C	N	O	S	0	0	0
			2481	1553	438	476	14			
1	C	317	Total	C	N	O	S	0	0	0
			2445	1532	429	470	14			
1	E	317	Total	C	N	O	S	0	0	0
			2445	1532	429	470	14			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q91MA7
A	8	ASP	-	expression tag	UNP Q91MA7
A	9	PRO	-	expression tag	UNP Q91MA7
A	10	GLY	-	expression tag	UNP Q91MA7
A	225	MET	GLY	engineered mutation	UNP Q91MA7
A	226	THR	LEU	engineered mutation	UNP Q91MA7
A	228	ALA	SER	engineered mutation	UNP Q91MA7
C	7	ALA	-	expression tag	UNP Q91MA7
C	8	ASP	-	expression tag	UNP Q91MA7
C	9	PRO	-	expression tag	UNP Q91MA7
C	10	GLY	-	expression tag	UNP Q91MA7
C	225	MET	GLY	engineered mutation	UNP Q91MA7
C	226	THR	LEU	engineered mutation	UNP Q91MA7
C	228	ALA	SER	engineered mutation	UNP Q91MA7
E	7	ALA	-	expression tag	UNP Q91MA7
E	8	ASP	-	expression tag	UNP Q91MA7
E	9	PRO	-	expression tag	UNP Q91MA7
E	10	GLY	-	expression tag	UNP Q91MA7
E	225	MET	GLY	engineered mutation	UNP Q91MA7
E	226	THR	LEU	engineered mutation	UNP Q91MA7
E	228	ALA	SER	engineered mutation	UNP Q91MA7

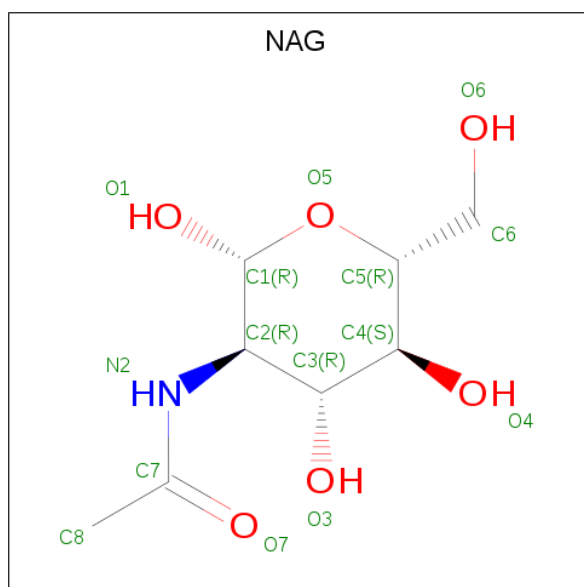
- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1391	863	243	279	6			
2	D	171	Total	C	N	O	S	0	0	0
			1382	858	241	277	6			
2	F	171	Total	C	N	O	S	0	0	0
			1382	858	241	277	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	123	GLY	ARG	conflict	UNP Q91MA7
D	123	GLY	ARG	conflict	UNP Q91MA7
F	123	GLY	ARG	conflict	UNP Q91MA7

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

Continued on next page...

*Continued from previous page...*

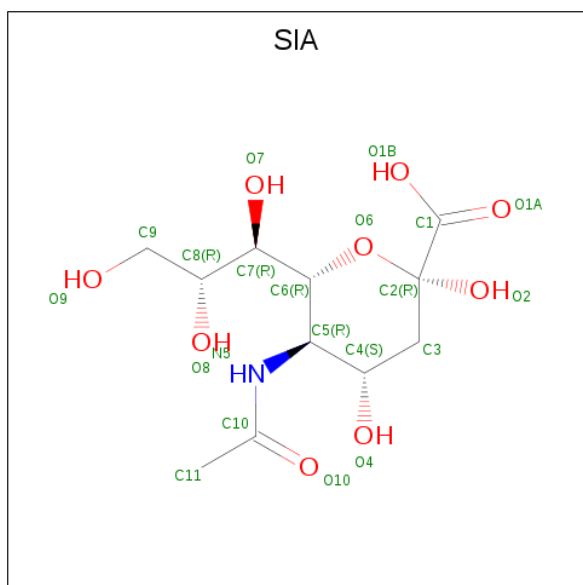
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1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

- Molecule 4 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



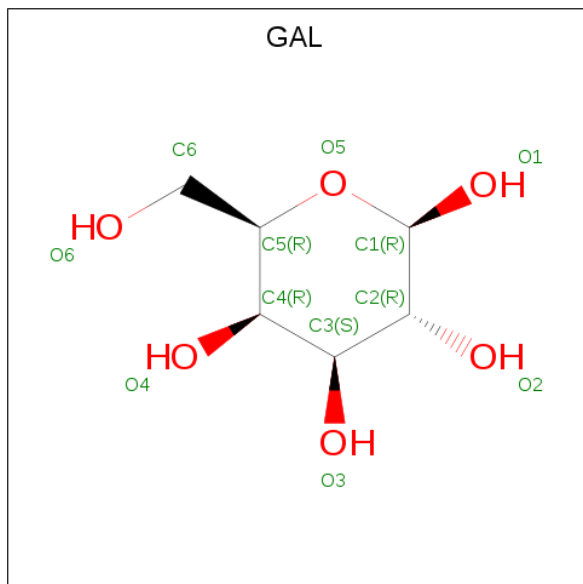
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	C	1	Total	C	O	0	0
			11	6	5		
4	E	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is O-SIALIC ACID (three-letter code: SIA) (formula:  $C_{11}H_{19}NO_9$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			21	11	1	9		
5	C	1	Total	C	N	O	0	0
			20	11	1	8		
5	E	1	Total	C	N	O	0	0
			20	11	1	8		

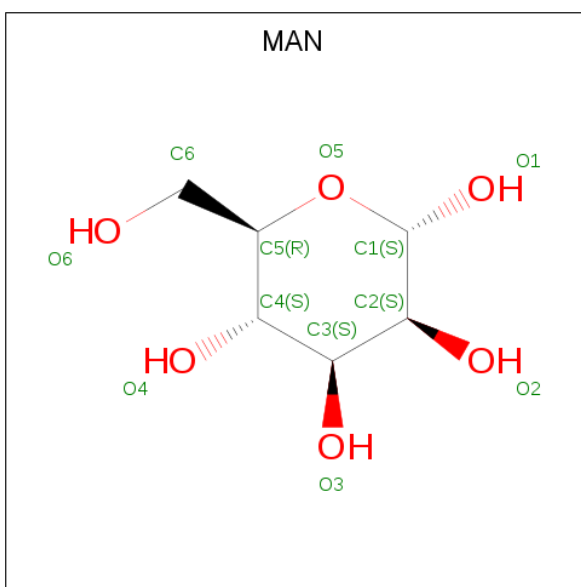
- Molecule 6 is BETA-D-GALACTOSE (three-letter code: GAL) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			12	6	6		
6	E	1	Total	C	O	0	0
			12	6	6		

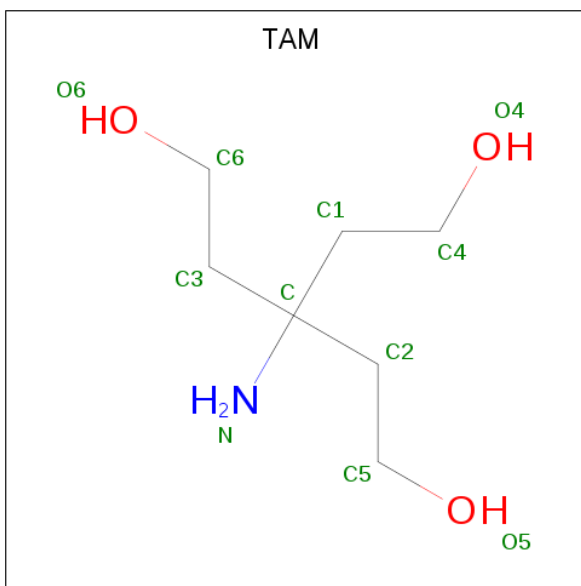
- Molecule 7 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).





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

- Molecule 8 is TRIS(HYDROXYETHYL)AMINOMETHANE (three-letter code: TAM) (formula:  $C_7H_{17}NO_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	E	1	Total	C	N	O	0	0
			11	7	1	3		

- Molecule 9 is water.

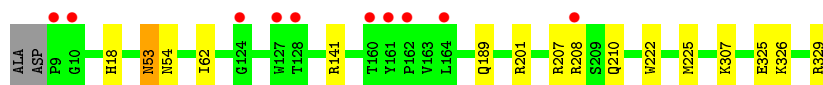
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	102	Total 102	O 102	0	0
9	B	49	Total 49	O 49	0	0
9	C	85	Total 85	O 85	0	0
9	D	46	Total 46	O 46	0	0
9	E	146	Total 146	O 146	0	0
9	F	53	Total 53	O 53	0	0

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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 HA1 chain

Chain A: 



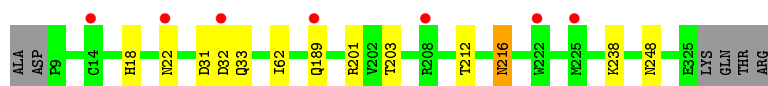
- Molecule 1: Hemagglutinin HA1 chain

Chain C: 



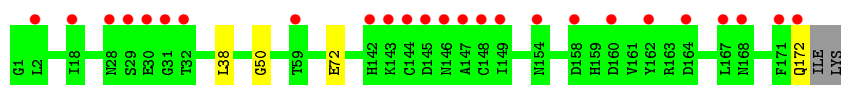
- Molecule 1: Hemagglutinin HA1 chain

Chain E: 



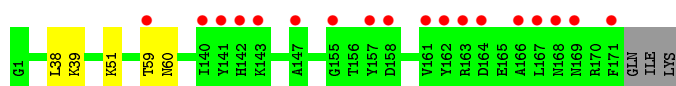
- Molecule 2: Hemagglutinin HA2 chain

Chain B: 

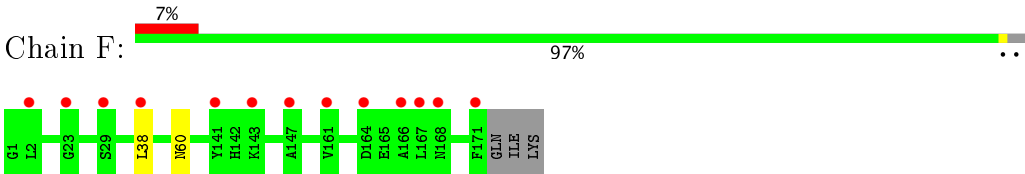


- Molecule 2: Hemagglutinin HA2 chain

Chain D: 



- Molecule 2: Hemagglutinin HA2 chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	208.65Å 131.53Å 72.73Å 90.00° 98.05° 90.00°	Depositor
Resolution (Å)	110.95 – 2.35 45.06 – 2.35	Depositor EDS
% Data completeness (in resolution range)	98.5 (110.95-2.35) 98.6 (45.06-2.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.14	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.193 , 0.213 0.196 , 0.215	Depositor DCC
$R_{free}$ test set	3928 reflections (5.21%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.3	Xtriage
Anisotropy	0.356	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 37.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12410	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, SIA, GAL, TAM, 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	0.50	0/2538	0.70	1/3457 (0.0%)
1	C	0.48	0/2502	0.71	0/3410
1	E	0.51	0/2502	0.70	1/3410 (0.0%)
2	B	0.48	0/1415	0.67	0/1902
2	D	0.48	0/1406	0.69	2/1890 (0.1%)
2	F	0.46	0/1406	0.67	0/1890
All	All	0.49	0/11769	0.70	4/15959 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	201	ARG	NE-CZ-NH2	-5.28	117.66	120.30
2	D	38	LEU	CB-CG-CD2	-5.26	102.05	111.00
2	D	38	LEU	CB-CG-CD1	5.24	119.90	111.00
1	E	31	ASP	CB-CG-OD1	5.08	122.87	118.30

There are no chirality outliers.

There are no planarity outliers.

### 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	2481	0	2433	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2445	0	2392	7	0
1	E	2445	0	2393	7	0
2	B	1391	0	1307	2	0
2	D	1382	0	1299	0	0
2	F	1382	0	1298	0	0
3	A	98	0	86	0	0
3	C	84	0	75	1	0
3	E	56	0	50	0	0
3	F	14	0	13	0	0
4	A	22	0	20	0	0
4	C	11	0	10	0	0
4	E	11	0	9	0	0
5	A	21	0	18	0	0
5	C	20	0	17	0	0
5	E	20	0	17	0	0
6	C	12	0	11	0	0
6	E	12	0	11	0	0
7	E	11	0	10	0	0
8	E	11	0	17	0	0
9	A	102	0	0	0	0
9	B	49	0	0	0	0
9	C	85	0	0	0	0
9	D	46	0	0	0	0
9	E	146	0	0	1	0
9	F	53	0	0	0	0
All	All	12410	0	11486	14	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 1.

All (14) 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:212:THR:HG21	1:E:216:ASN:HB3	1.44	0.99
2:B:72:GLU:OE2	1:E:238:LYS:HE2	1.76	0.84
1:C:212:THR:CG2	1:E:216:ASN:HB3	2.22	0.64
1:C:212:THR:HG21	1:E:216:ASN:CB	2.26	0.63
1:A:222:TRP:CZ2	1:A:225:MET:HA	2.41	0.54
1:A:53:ASN:HD22	1:A:54:ASN:N	2.05	0.54
1:E:203:THR:HG23	1:E:212:THR:HG22	1.90	0.53
1:C:53:ASN:HD21	1:C:276:THR:HG22	1.72	0.53

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:53:ASN:C	1:A:53:ASN:HD22	2.17	0.48
1:A:207:ARG:HG2	1:C:221:PRO:HB2	1.99	0.45
1:E:33:GLN:NE2	9:E:604:HOH:O	2.49	0.45
1:E:201:ARG:HG2	1:E:248:ASN:OD1	2.17	0.44
1:C:167:THR:OG1	3:C:503:NAG:H62	2.20	0.41
2:B:50:GLY:HA3	1:C:30:THR:O	2.21	0.41

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	319/323 (99%)	309 (97%)	9 (3%)	1 (0%)	44	53
1	C	315/323 (98%)	305 (97%)	9 (3%)	1 (0%)	44	53
1	E	315/323 (98%)	306 (97%)	8 (2%)	1 (0%)	44	53
2	B	170/174 (98%)	163 (96%)	7 (4%)	0	100	100
2	D	169/174 (97%)	162 (96%)	7 (4%)	0	100	100
2	F	169/174 (97%)	162 (96%)	7 (4%)	0	100	100
All	All	1457/1491 (98%)	1407 (97%)	47 (3%)	3 (0%)	51	61

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	ILE
1	C	62	ILE
1	E	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	282/283 (100%)	272 (96%)	10 (4%)	41	52
1	C	278/283 (98%)	270 (97%)	8 (3%)	48	60
1	E	278/283 (98%)	273 (98%)	5 (2%)	64	76
2	B	146/148 (99%)	144 (99%)	2 (1%)	71	82
2	D	145/148 (98%)	141 (97%)	4 (3%)	49	61
2	F	145/148 (98%)	143 (99%)	2 (1%)	71	82
All	All	1274/1293 (98%)	1243 (98%)	31 (2%)	54	67

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	53	ASN
1	A	141	ARG
1	A	189	GLN
1	A	208	ARG
1	A	210	GLN
1	A	307	LYS
1	A	325	GLU
1	A	326	LYS
1	A	329	ARG
2	B	38	LEU
2	B	172	GLN
1	C	18	HIS
1	C	22	ASN
1	C	31	ASP
1	C	141	ARG
1	C	167	THR
1	C	208	ARG
1	C	210	GLN
1	C	216	ASN
2	D	39	LYS
2	D	51	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	59	THR
2	D	60	ASN
1	E	18	HIS
1	E	22	ASN
1	E	32	ASP
1	E	189	GLN
1	E	216	ASN
2	F	38	LEU
2	F	60	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	ASN
2	D	53	ASN
1	E	296	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](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

29 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	1,3	14,14,15	0.32	0	15,19,21	1.26	2 (13%)
3	NAG	A	402	3,4	14,14,15	0.57	0	15,19,21	1.44	3 (20%)
4	BMA	A	403	3	11,11,12	0.45	0	13,15,17	0.91	1 (7%)
3	NAG	A	404	1	14,14,15	0.59	0	15,19,21	1.36	3 (20%)
3	NAG	A	405	1,3	14,14,15	0.36	0	15,19,21	1.42	2 (13%)
3	NAG	A	406	3,4	14,14,15	0.43	0	15,19,21	0.78	1 (6%)
4	BMA	A	407	3	11,11,12	0.52	0	13,15,17	1.18	2 (15%)
3	NAG	A	408	1,3	14,14,15	0.55	0	15,19,21	1.27	3 (20%)
3	NAG	A	409	3	14,14,15	0.41	0	15,19,21	1.72	2 (13%)
5	SIA	A	410	-	18,21,21	1.43	2 (11%)	19,31,31	0.68	0
3	NAG	C	501	1	14,14,15	0.61	0	15,19,21	2.51	4 (26%)
3	NAG	C	502	1	14,14,15	0.63	0	15,19,21	1.71	1 (6%)
3	NAG	C	503	1,3	14,14,15	0.43	0	15,19,21	1.34	1 (6%)
3	NAG	C	504	3,4	14,14,15	0.41	0	15,19,21	1.13	1 (6%)
4	BMA	C	505	3	11,11,12	0.93	0	13,15,17	2.49	4 (30%)
3	NAG	C	506	1,3	14,14,15	0.46	0	15,19,21	1.02	1 (6%)
3	NAG	C	507	3	14,14,15	0.41	0	15,19,21	1.54	2 (13%)
5	SIA	C	508	6	17,20,21	0.57	0	19,28,31	0.64	0
6	GAL	C	509	5	12,12,12	0.46	0	17,17,17	0.72	0
3	NAG	E	501	1	14,14,15	0.49	0	15,19,21	1.40	1 (6%)
3	NAG	E	502	1,3	14,14,15	0.26	0	15,19,21	1.17	1 (6%)
3	NAG	E	503	3,4	14,14,15	0.43	0	15,19,21	1.10	1 (6%)
4	BMA	E	504	3,7	11,11,12	0.49	0	13,15,17	1.12	2 (15%)
7	MAN	E	505	4	11,11,12	0.54	0	13,15,17	1.37	3 (23%)
3	NAG	E	506	1	14,14,15	0.52	0	15,19,21	1.08	0
8	TAM	E	507	-	7,10,10	0.68	0	9,12,12	0.90	0
5	SIA	E	508	6	17,20,21	0.48	0	19,28,31	0.94	1 (5%)
6	GAL	E	509	5	12,12,12	0.49	0	17,17,17	0.77	0
3	NAG	F	201	2	14,14,15	0.56	0	15,19,21	2.50	3 (20%)

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	1,3	-	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	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	403	3	-	0/2/19/22	0/1/1/1
3	NAG	A	404	1	-	0/6/23/26	0/1/1/1
3	NAG	A	405	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	406	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	407	3	-	0/2/19/22	0/1/1/1
3	NAG	A	408	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	409	3	-	0/6/23/26	0/1/1/1
5	SIA	A	410	-	-	0/14/38/38	0/1/1/1
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
3	NAG	C	502	1	-	0/6/23/26	0/1/1/1
3	NAG	C	503	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	504	3,4	-	0/6/23/26	0/1/1/1
4	BMA	C	505	3	-	0/2/19/22	0/1/1/1
3	NAG	C	506	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	507	3	-	0/6/23/26	0/1/1/1
5	SIA	C	508	6	-	0/14/34/38	0/1/1/1
6	GAL	C	509	5	-	0/2/22/22	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	BMA	E	504	3,7	-	0/2/19/22	0/1/1/1
7	MAN	E	505	4	-	0/2/19/22	0/1/1/1
3	NAG	E	506	1	-	0/6/23/26	0/1/1/1
8	TAM	E	507	-	-	0/12/12/12	0/0/0/0
5	SIA	E	508	6	-	0/14/34/38	0/1/1/1
6	GAL	E	509	5	-	0/2/22/22	0/1/1/1
3	NAG	F	201	2	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	410	SIA	C3-C2	2.35	1.54	1.51
5	A	410	SIA	O2-C2	4.79	1.45	1.39

All (45) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	NAG	C1-C2-N2	-3.65	104.26	110.49
3	A	405	NAG	O5-C1-C2	-2.45	108.06	111.47
3	A	409	NAG	C4-C3-C2	-2.18	107.83	111.02
3	A	408	NAG	C3-C4-C5	-2.10	106.51	110.22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	404	NAG	O5-C1-C2	-2.08	108.58	111.47
3	C	507	NAG	C4-C3-C2	-2.06	108.00	111.02
3	A	401	NAG	C1-O5-C5	2.02	114.95	112.17
3	C	501	NAG	C2-N2-C7	2.04	125.92	122.94
3	A	402	NAG	C3-C4-C5	2.05	113.82	110.22
3	A	408	NAG	O4-C4-C5	2.09	114.54	109.28
3	A	404	NAG	C1-C2-N2	2.12	114.11	110.49
3	C	506	NAG	C1-C2-N2	2.12	114.11	110.49
7	E	505	MAN	C2-C3-C4	2.14	114.61	110.88
3	A	408	NAG	C6-C5-C4	2.14	118.01	113.00
7	E	505	MAN	C3-C4-C5	2.21	114.10	110.22
3	A	402	NAG	C1-O5-C5	2.23	115.23	112.17
4	C	505	BMA	C2-C3-C4	2.29	114.86	110.88
3	A	406	NAG	C1-O5-C5	2.31	115.36	112.17
4	E	504	BMA	C3-C4-C5	2.36	114.37	110.22
3	A	404	NAG	C1-O5-C5	2.41	115.49	112.17
3	C	504	NAG	C1-C2-N2	2.43	114.65	110.49
4	A	403	BMA	C1-C2-C3	2.48	112.80	109.65
4	E	504	BMA	C1-O5-C5	2.51	115.63	112.17
4	A	407	BMA	C1-O5-C5	2.53	115.65	112.17
3	C	501	NAG	O5-C1-C2	2.56	115.04	111.47
5	E	508	SIA	O8-C8-C7	2.57	115.48	109.09
3	F	201	NAG	C4-C3-C2	2.60	114.83	111.02
4	A	407	BMA	C1-C2-C3	2.73	113.11	109.65
3	C	501	NAG	C3-C4-C5	2.78	115.12	110.22
3	E	502	NAG	C1-O5-C5	2.81	116.04	112.17
7	E	505	MAN	C1-O5-C5	2.89	116.14	112.17
3	E	503	NAG	C1-O5-C5	3.05	116.37	112.17
3	A	402	NAG	C4-C3-C2	3.40	116.00	111.02
3	F	201	NAG	C3-C4-C5	3.55	116.47	110.22
3	A	405	NAG	C1-O5-C5	3.97	117.63	112.17
4	C	505	BMA	O5-C1-C2	4.12	117.24	110.79
3	C	503	NAG	C1-O5-C5	4.27	118.06	112.17
4	C	505	BMA	C1-O5-C5	4.31	118.11	112.17
3	C	507	NAG	C1-O5-C5	4.56	118.45	112.17
3	E	501	NAG	C1-O5-C5	4.87	118.88	112.17
3	C	502	NAG	C1-O5-C5	5.14	119.25	112.17
4	C	505	BMA	C1-C2-C3	5.62	116.78	109.65
3	A	409	NAG	C1-O5-C5	5.78	120.13	112.17
3	F	201	NAG	C1-O5-C5	7.81	122.93	112.17
3	C	501	NAG	C1-O5-C5	8.08	123.30	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	503	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, 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	321/323 (99%)	0.21	10 (3%) 49 61	25, 44, 77, 98	0
1	C	317/323 (98%)	0.16	7 (2%) 62 72	28, 43, 64, 94	0
1	E	317/323 (98%)	0.14	7 (2%) 62 72	21, 35, 58, 102	0
2	B	172/174 (98%)	0.75	25 (14%) 3 5	23, 56, 85, 115	0
2	D	171/174 (98%)	0.57	18 (10%) 7 11	26, 55, 84, 100	0
2	F	171/174 (98%)	0.61	13 (7%) 15 22	24, 56, 95, 108	0
All	All	1469/1491 (98%)	0.34	80 (5%) 26 37	21, 44, 83, 115	0

All (80) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	9	PRO	6.0
2	B	29	SER	5.3
2	B	147	ALA	4.8
2	D	171	PHE	4.7
2	B	143	LYS	4.6
2	B	30	GLU	4.6
2	D	162	TYR	4.6
1	E	222	TRP	4.3
2	B	148	CYS	4.1
2	B	167	LEU	4.1
2	F	147	ALA	4.0
1	C	12	THR	3.7
2	D	168	ASN	3.7
1	C	9	PRO	3.7
1	A	164	LEU	3.5
2	B	146	ASN	3.5
2	F	23	GLY	3.5
2	B	154	ASN	3.4
1	E	22	ASN	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	22	ASN	3.2
2	B	145	ASP	3.2
2	F	168	ASN	3.2
2	B	31	GLY	3.1
2	D	167	LEU	3.1
2	F	166	ALA	3.1
2	F	141	TYR	3.1
2	F	143	LYS	3.0
2	B	164	ASP	3.0
2	D	141	TYR	2.9
2	B	168	ASN	2.9
1	C	222	TRP	2.8
2	F	2	LEU	2.8
1	C	225	MET	2.8
2	D	157	TYR	2.8
2	D	161	VAL	2.8
2	F	164	ASP	2.8
2	B	59	THR	2.8
2	D	140	ILE	2.7
2	B	160	ASP	2.7
2	D	59	THR	2.7
2	F	29	SER	2.6
2	D	164	ASP	2.6
1	A	160	THR	2.6
2	D	143	LYS	2.6
1	A	161	TYR	2.6
1	A	124	GLY	2.6
2	B	28	ASN	2.6
2	F	171	PHE	2.5
2	B	162	TYR	2.5
2	B	158	ASP	2.5
2	D	166	ALA	2.4
2	F	161	VAL	2.4
1	C	221	PRO	2.4
1	A	10	GLY	2.4
2	D	155	GLY	2.4
1	A	128	THR	2.3
1	A	162	PRO	2.3
2	B	18	ILE	2.3
1	E	208	ARG	2.3
1	A	208	ARG	2.3
2	F	167	LEU	2.2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	14	CYS	2.2
2	B	32	THR	2.2
2	D	147	ALA	2.2
2	B	142	HIS	2.2
1	E	225	MET	2.2
2	D	158	ASP	2.1
2	F	38	LEU	2.1
1	C	23	GLY	2.1
2	B	2	LEU	2.1
2	D	142	HIS	2.1
2	B	172	GLN	2.1
1	E	189	GLN	2.1
2	B	144	CYS	2.1
2	B	149	ILE	2.0
1	A	127	TRP	2.0
1	E	32	ASP	2.0
2	D	163	ARG	2.0
2	D	169	ASN	2.0
2	B	171	PHE	2.0

## 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](#)

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, 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(Å <sup>2</sup> )	Q<0.9
5	SIA	A	410	21/21	0.69	0.26	3.27	77,93,99,103	0
3	NAG	E	506	14/15	0.93	0.19	2.89	34,36,45,45	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	408	14/15	0.93	0.16	2.49	42,46,54,65	0
3	NAG	A	406	14/15	0.90	0.32	2.40	68,75,87,90	0
5	SIA	C	508	20/21	0.82	0.20	2.09	61,68,76,76	0
3	NAG	E	502	14/15	0.95	0.20	1.29	44,48,52,57	0
5	SIA	E	508	20/21	0.94	0.19	1.10	42,51,55,56	0
3	NAG	C	506	14/15	0.94	0.14	0.20	41,48,52,60	0
3	NAG	E	503	14/15	0.92	0.15	-0.17	43,47,53,57	0
3	NAG	C	504	14/15	0.90	0.14	-0.93	48,53,64,72	0
3	NAG	C	503	14/15	0.96	0.10	-1.24	38,43,46,48	0
4	BMA	A	407	11/12	0.63	0.30	-	92,97,100,101	0
3	NAG	A	402	14/15	0.87	0.41	-	82,89,92,96	0
3	NAG	A	404	14/15	0.89	0.22	-	56,62,65,66	0
4	BMA	E	504	11/12	0.83	0.19	-	67,72,75,79	0
3	NAG	C	502	14/15	0.61	0.38	-	71,80,89,91	0
3	NAG	E	501	14/15	0.82	0.21	-	56,62,66,69	0
8	TAM	E	507	11/11	0.73	0.26	-	72,79,82,83	0
7	MAN	E	505	11/12	0.77	0.33	-	81,84,86,89	0
3	NAG	C	507	14/15	0.86	0.22	-	69,75,78,78	0
6	GAL	E	509	12/12	0.87	0.52	-	70,84,92,92	0
3	NAG	A	405	14/15	0.91	0.23	-	61,65,78,79	0
4	BMA	C	505	11/12	0.48	0.33	-	78,81,85,88	0
3	NAG	A	409	14/15	0.72	0.32	-	68,81,87,91	0
4	BMA	A	403	11/12	0.81	0.30	-	93,96,99,100	0
3	NAG	F	201	14/15	0.75	0.45	-	98,104,117,121	0
3	NAG	C	501	14/15	0.85	0.23	-	59,69,75,77	0
6	GAL	C	509	12/12	0.73	0.33	-	89,99,106,108	0
3	NAG	A	401	14/15	0.93	0.21	-	55,65,69,75	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.