



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 10:57 AM GMT

PDB ID : 4H53  
Title : Influenza N2-Tyr406Asp neuraminidase in complex with beta-Neu5Ac  
Authors : Vavricka, C.J.; Liu, Y.; Kiyota, H.; Sriwilaijaroen, N.; Qi, J.; Tanaka, K.; Wu, Y.; Li, Q.; Li, Y.; Yan, J.; Suzuki, Y.; Gao, G.F.  
Deposited on : 2012-09-18  
Resolution : 1.50 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

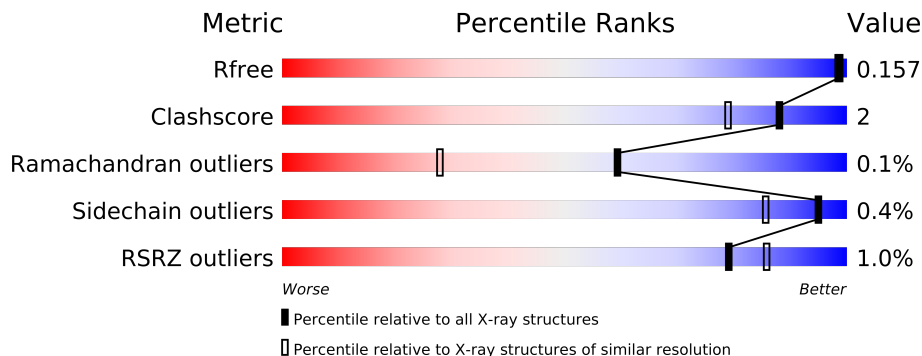
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1513 (1.50-1.50)
Clashscore	79885	1768 (1.50-1.50)
Ramachandran outliers	78287	1720 (1.50-1.50)
Sidechain outliers	78261	1718 (1.50-1.50)
RSRZ outliers	66119	1514 (1.50-1.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	388	
1	B	388	
1	C	388	
1	D	388	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
10	GOL	D	609	-	X
3	NAG	A	502	-	X
6	SLB	C	512	-	X

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 14833 atoms, of which 0 are hydrogen and 0 are deuterium.

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

- Molecule 1 is a protein called Neuraminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	0	0
			3013	1858	542	590	23			
1	B	388	Total	C	N	O	S	0	1	0
			3019	1861	543	592	23			
1	C	388	Total	C	N	O	S	0	1	0
			3019	1861	543	592	23			
1	D	388	Total	C	N	O	S	0	3	0
			3031	1867	545	596	23			

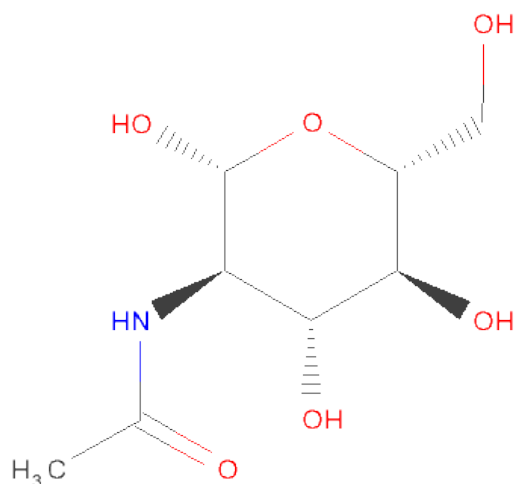
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
B	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
C	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
D	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	A	1	Total	Ca	0	0
			1	1		
2	C	1	Total	Ca	0	0
			1	1		

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 4 is a polymer of unknown type called SUGAR (7-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	7	Total	C	N	O	0	0
			83	46	2	35		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1

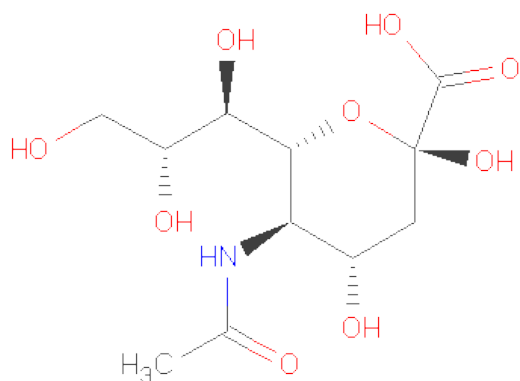
- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			32	17	1	14		
5	B	2	Total	C	N	O	0	0
			32	17	1	14		
5	C	2	Total	C	N	O	0	0
			32	17	1	14		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
B	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
C	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1

- Molecule 6 is SUGAR (5-N-ACETYL-BETA-D-NEURAMINICACID) (three-letter code: SLB) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>).



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

- Molecule 7 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	4	Total	C	N	O	0	0
			50	28	2	20		
7	B	4	Total	C	N	O	0	0
			50	28	2	20		
7	C	4	Total	C	N	O	0	0
			50	28	2	20		
7	C	4	Total	C	N	O	0	0
			50	28	2	20		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
B	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
C	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1
C	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1

- Molecule 8 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	D	2	Total	C	N	O	0	0
			28	16	2	10		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1

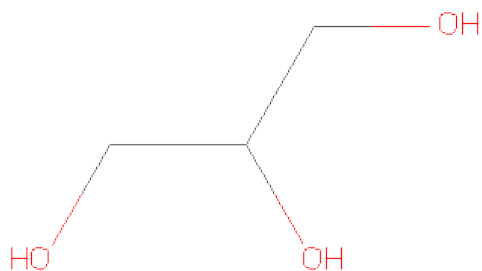
- Molecule 9 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	D	6	Total	C	N	O	0	0
			72	40	2	30		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	406	ASP	TYR	ENGINEERED MUTATION	UNP Q194T1

- Molecule 10 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 11 is water.

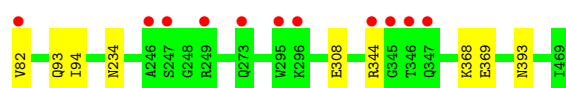
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	531	Total	O	0	0
			531	531		
11	B	564	Total	O	0	0
			564	564		
11	C	548	Total	O	0	0
			548	548		
11	D	543	Total	O	0	0
			543	543		

### 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 ( $RSRZ > 2$ ). 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: Neuraminidase

Chain A:



- Molecule 1: Neuraminidase

Chain B:



- Molecule 1: Neuraminidase

Chain C:



- Molecule 1: Neuraminidase

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.22Å 114.83Å 84.26Å 90.00° 99.71° 90.00°	Depositor
Resolution (Å)	39.05 – 1.50 39.05 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (39.05-1.50) 96.7 (39.05-1.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.87 (at 1.50Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.140 , 0.161 0.135 , 0.157	Depositor DCC
$R_{free}$ test set	12074 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	9.0	Xtriage
Anisotropy	0.309	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 29.7	EDS
Estimated twinning fraction	0.012 for l,-k,h	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 248199 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	14833	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, BMA, NAG, CA, SIA, GAL, SLB, 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	0.36	0/3085	0.60	0/4188
1	B	0.36	0/3091	0.60	0/4196
1	C	0.35	0/3091	0.61	0/4196
1	D	0.34	0/3103	0.59	0/4212
All	All	0.35	0/12370	0.60	0/16792

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3013	0	0	5	0
1	B	3019	0	0	10	0
1	C	3019	0	0	11	0
1	D	3031	0	0	5	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	14	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	83	0	70	1	0
5	A	32	0	28	0	0
5	B	32	0	28	0	0
5	C	32	0	28	0	0
6	A	21	0	18	0	0
6	B	21	0	18	0	0
6	C	21	0	17	0	0
7	B	100	0	86	0	0
7	C	100	0	86	1	0
8	D	28	0	25	0	0
9	D	72	0	61	1	0
10	D	6	0	8	0	0
11	A	531	0	0	3	0
11	B	564	0	0	8	0
11	C	548	0	0	6	0
11	D	543	0	0	5	0
All	All	14833	0	486	31	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (31) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:249:ARG:NH1	11:D:1040:HOH:O	2.15	0.78
1:C:388[B]:SER:OG	11:C:857:HOH:O	2.04	0.75
1:B:249:ARG:NH1	11:B:913:HOH:O	2.24	0.69
1:B:338:ARG:NH1	11:B:925:HOH:O	2.29	0.64
1:A:368:LYS:NZ	1:A:369:GLU:OE2	2.31	0.64
1:B:347:GLN:NE2	11:B:1140:HOH:O	2.31	0.63
1:C:249:ARG:NH1	11:C:1026:HOH:O	2.32	0.63
1:A:234:ASN:ND2	11:A:885:HOH:O	2.32	0.62
1:C:93:GLN:NE2	11:C:1133:HOH:O	2.34	0.61
1:D:234:ASN:ND2	11:D:1034:HOH:O	2.33	0.60
1:C:329:ASP:OD1	1:C:331:SER:OG	2.21	0.58
1:C:338:ARG:NE	11:C:1076:HOH:O	2.37	0.58
1:B:347:GLN:NE2	11:B:892:HOH:O	2.37	0.58
1:A:344:ARG:NE	11:A:894:HOH:O	2.38	0.57
1:C:338:ARG:NH2	11:C:1021:HOH:O	2.39	0.54
1:C:368:LYS:NZ	1:C:369:GLU:OE2	2.41	0.53
1:C:393:ASN:N	9:D:604:NAG:HO3	2.07	0.52
1:B:338:ARG:NH2	11:B:1071:HOH:O	2.44	0.51
1:B:368:LYS:NZ	1:B:369:GLU:OE2	2.46	0.49

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:93:GLN:OE1	11:B:953:HOH:O	2.20	0.48
1:D:296:LYS:NZ	11:D:1014:HOH:O	2.46	0.47
1:C:435:ARG:NH2	11:C:804:HOH:O	2.49	0.46
1:C:380:ILE:O	1:C:388[B]:SER:OG	2.34	0.45
1:A:93:GLN:OE1	11:A:883:HOH:O	2.21	0.44
1:A:393:ASN:N	7:C:507:NAG:HO3	2.16	0.44
4:A:504:NAG:HO3	1:B:393:ASN:N	2.17	0.43
1:B:199:LYS:NZ	11:B:1117:HOH:O	2.53	0.42
1:D:187:LYS:NZ	11:D:1131:HOH:O	2.53	0.41
1:D:234:ASN:ND2	11:D:1127:HOH:O	2.52	0.41
1:B:143:LYS:NZ	11:B:1080:HOH:O	2.54	0.40
1:C:288:ARG:NE	1:C:304:ASP:OD1	2.55	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	386/388 (100%)	372 (96%)	14 (4%)	0	100	100
1	B	387/388 (100%)	372 (96%)	15 (4%)	0	100	100
1	C	387/388 (100%)	371 (96%)	16 (4%)	0	100	100
1	D	389/388 (100%)	376 (97%)	12 (3%)	1 (0%)	50	20
All	All	1549/1552 (100%)	1491 (96%)	57 (4%)	1 (0%)	59	28

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	222	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	338/338 (100%)	335 (99%)	3 (1%)	87	69
1	B	339/338 (100%)	339 (100%)	0	100	100
1	C	339/338 (100%)	339 (100%)	0	100	100
1	D	341/338 (101%)	339 (99%)	2 (1%)	92	80
All	All	1357/1352 (100%)	1352 (100%)	5 (0%)	95	85

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	VAL
1	A	94	ILE
1	A	308	GLU
1	D	82	VAL
1	D	147	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

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

## 5.5 Carbohydrates ⓘ

37 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
4	NAG	A	503	1,4	12,14,15	0.54	0	15,19,21	1.09	2 (13%)
4	NAG	A	504	4	12,14,15	0.65	1 (8%)	15,19,21	0.86	0
4	BMA	A	505	4	10,11,12	0.69	0	11,15,17	0.89	1 (9%)
4	MAN	A	506	4	10,11,12	0.68	0	11,15,17	0.58	0
4	MAN	A	507	4	10,11,12	0.67	0	11,15,17	0.74	0
4	MAN	A	508	4	10,11,12	0.67	0	11,15,17	0.75	0
4	MAN	A	509	4	10,11,12	0.62	0	11,15,17	1.06	1 (9%)
5	SIA	A	510	5	20,20,21	3.53	8 (40%)	23,28,31	4.17	7 (30%)
5	GAL	A	511	5	12,12,12	0.68	0	17,17,17	1.20	2 (11%)
7	NAG	B	502	1,7	12,14,15	0.64	0	15,19,21	0.78	0
7	NAG	B	503	7	12,14,15	0.65	0	15,19,21	0.99	1 (6%)
7	BMA	B	504	7	10,11,12	0.76	0	11,15,17	0.95	1 (9%)
7	MAN	B	505	7	10,11,12	0.75	0	11,15,17	0.91	1 (9%)
7	NAG	B	506	1,7	12,14,15	0.56	0	15,19,21	0.95	1 (6%)
7	NAG	B	507	7	12,14,15	0.65	0	15,19,21	0.70	0
7	BMA	B	508	7	10,11,12	0.63	0	11,15,17	0.68	0
7	MAN	B	509	7	10,11,12	0.66	0	11,15,17	0.79	0
5	SIA	B	510	5	20,20,21	3.61	8 (40%)	23,28,31	4.60	8 (34%)
5	GAL	B	511	5	12,12,12	0.46	0	17,17,17	1.23	3 (17%)
7	NAG	C	502	1,7	12,14,15	0.67	1 (8%)	15,19,21	0.74	0
7	NAG	C	503	7	12,14,15	0.75	1 (8%)	15,19,21	1.14	1 (6%)
7	BMA	C	504	7	10,11,12	0.78	0	11,15,17	1.62	2 (18%)
7	MAN	C	505	7	10,11,12	0.82	1 (10%)	11,15,17	0.85	0
7	NAG	C	506	1,7	12,14,15	0.52	0	15,19,21	0.60	0
7	NAG	C	507	7	12,14,15	0.57	0	15,19,21	0.95	1 (6%)
7	BMA	C	508	7	10,11,12	0.74	0	11,15,17	0.61	0
7	MAN	C	509	7	10,11,12	0.73	0	11,15,17	0.73	0
5	SIA	C	510	5	20,20,21	3.47	7 (35%)	23,28,31	4.26	7 (30%)
5	GAL	C	511	5	12,12,12	0.65	0	17,17,17	1.84	5 (29%)
8	NAG	D	601	1,8	12,14,15	0.64	0	15,19,21	0.79	0
8	NAG	D	602	8	12,14,15	0.66	0	15,19,21	0.86	1 (6%)
9	NAG	D	603	9,1	12,14,15	0.63	0	15,19,21	0.82	1 (6%)
9	NAG	D	604	9	12,14,15	0.69	1 (8%)	15,19,21	0.67	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	BMA	D	605	9	10,11,12	0.81	1 (10%)	11,15,17	0.78	0
9	MAN	D	606	9	10,11,12	0.75	0	11,15,17	1.12	1 (9%)
9	MAN	D	607	9	10,11,12	0.69	0	11,15,17	0.93	1 (9%)
9	MAN	D	608	9	10,11,12	0.73	0	11,15,17	0.53	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	503	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	504	4	-	0/6/23/26	0/1/1/1
4	BMA	A	505	4	-	0/2/19/22	0/1/1/1
4	MAN	A	506	4	-	0/2/19/22	0/1/1/1
4	MAN	A	507	4	-	0/2/19/22	0/1/1/1
4	MAN	A	508	4	-	0/2/19/22	0/1/1/1
4	MAN	A	509	4	-	0/2/19/22	0/1/1/1
5	SIA	A	510	5	-	0/15/34/38	0/1/1/1
5	GAL	A	511	5	-	0/2/22/22	0/1/1/1
7	NAG	B	502	1,7	-	0/6/23/26	0/1/1/1
7	NAG	B	503	7	-	0/6/23/26	0/1/1/1
7	BMA	B	504	7	-	0/2/19/22	0/1/1/1
7	MAN	B	505	7	-	0/2/19/22	0/1/1/1
7	NAG	B	506	1,7	-	0/6/23/26	0/1/1/1
7	NAG	B	507	7	-	0/6/23/26	0/1/1/1
7	BMA	B	508	7	-	0/2/19/22	0/1/1/1
7	MAN	B	509	7	-	0/2/19/22	0/1/1/1
5	SIA	B	510	5	-	0/15/34/38	0/1/1/1
5	GAL	B	511	5	-	0/2/22/22	0/1/1/1
7	NAG	C	502	1,7	-	0/6/23/26	0/1/1/1
7	NAG	C	503	7	-	0/6/23/26	0/1/1/1
7	BMA	C	504	7	-	0/2/19/22	0/1/1/1
7	MAN	C	505	7	-	0/2/19/22	0/1/1/1
7	NAG	C	506	1,7	-	0/6/23/26	0/1/1/1
7	NAG	C	507	7	-	0/6/23/26	0/1/1/1
7	BMA	C	508	7	-	0/2/19/22	0/1/1/1
7	MAN	C	509	7	-	0/2/19/22	0/1/1/1
5	SIA	C	510	5	-	0/15/34/38	0/1/1/1
5	GAL	C	511	5	-	0/2/22/22	0/1/1/1
8	NAG	D	601	1,8	-	0/6/23/26	0/1/1/1
8	NAG	D	602	8	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	NAG	D	603	9,1	-	0/6/23/26	0/1/1/1
9	NAG	D	604	9	-	0/6/23/26	0/1/1/1
9	BMA	D	605	9	-	0/2/19/22	0/1/1/1
9	MAN	D	606	9	-	0/2/19/22	0/1/1/1
9	MAN	D	607	9	-	0/2/19/22	0/1/1/1
9	MAN	D	608	9	-	0/2/19/22	0/1/1/1

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	510	SIA	O6-C2	13.17	1.40	1.28
5	A	510	SIA	O6-C2	12.90	1.40	1.28
5	C	510	SIA	O6-C2	12.49	1.39	1.28
5	C	510	SIA	C4-C5	-6.20	1.47	1.53
5	A	510	SIA	C4-C5	-5.83	1.47	1.53
5	B	510	SIA	C4-C5	-5.72	1.47	1.53
5	A	510	SIA	C7-C6	-3.68	1.48	1.52
5	B	510	SIA	C7-C6	-3.62	1.48	1.52
5	B	510	SIA	C6-C5	-3.41	1.47	1.53
5	C	510	SIA	C7-C6	-3.39	1.48	1.52
5	C	510	SIA	C6-C5	-2.95	1.48	1.53
5	A	510	SIA	C6-C5	-2.68	1.48	1.53
5	B	510	SIA	O8-C8	-2.54	1.37	1.43
5	C	510	SIA	O8-C8	-2.25	1.38	1.43
9	D	605	BMA	O5-C5	-2.24	1.41	1.45
5	A	510	SIA	C10-N5	2.24	1.43	1.34
5	B	510	SIA	C10-N5	2.21	1.43	1.34
9	D	604	NAG	O5-C5	-2.20	1.41	1.45
7	C	503	NAG	O5-C5	-2.19	1.41	1.45
5	A	510	SIA	O8-C8	-2.17	1.38	1.43
5	A	510	SIA	C3-C2	2.13	1.54	1.46
7	C	505	MAN	O5-C5	-2.12	1.41	1.45
5	C	510	SIA	C3-C2	2.11	1.53	1.46
5	C	510	SIA	C10-N5	2.10	1.42	1.34
7	C	502	NAG	O5-C5	-2.10	1.41	1.45
5	B	510	SIA	C3-C2	2.07	1.53	1.46
5	A	510	SIA	C5-N5	2.05	1.49	1.45
4	A	504	NAG	O5-C5	-2.03	1.41	1.45
5	B	510	SIA	O7-C7	-2.03	1.38	1.43

All (48) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	510	SIA	O6-C2-C3	-20.06	98.70	124.91
5	C	510	SIA	O6-C2-C3	-18.77	100.39	124.91
5	A	510	SIA	O6-C2-C3	-18.08	101.29	124.91
5	A	510	SIA	C3-C2-C1	-4.76	94.48	121.08
5	B	510	SIA	C3-C2-C1	-4.42	96.36	121.08
5	C	510	SIA	C3-C2-C1	-4.34	96.80	121.08
7	C	504	BMA	O5-C5-C6	3.93	111.11	106.98
5	B	510	SIA	C7-C6-C5	3.79	119.83	114.24
5	C	511	GAL	O3-C3-C4	-3.66	102.14	110.35
5	B	510	SIA	O6-C6-C5	-3.55	105.84	110.29
5	A	510	SIA	C7-C6-C5	3.44	119.32	114.24
5	C	510	SIA	C7-C6-C5	3.41	119.26	114.24
7	C	504	BMA	O5-C5-C4	-3.26	106.52	110.65
5	C	511	GAL	C3-C4-C5	3.23	115.98	110.20
4	A	509	MAN	O5-C5-C6	3.11	110.24	106.98
9	D	606	MAN	O5-C5-C6	3.11	110.24	106.98
5	C	511	GAL	C1-O5-C5	-3.00	108.04	113.40
5	A	510	SIA	O4-C4-C3	2.85	115.64	109.08
5	C	511	GAL	C1-C2-C3	2.75	114.90	110.53
5	C	510	SIA	O6-C6-C5	-2.68	106.93	110.29
5	A	510	SIA	O6-C6-C5	-2.67	106.94	110.29
5	A	510	SIA	C3-C4-C5	2.65	114.05	110.41
5	B	510	SIA	O4-C4-C3	2.62	115.11	109.08
5	B	511	GAL	O3-C3-C4	2.62	116.22	110.35
7	B	503	NAG	O5-C5-C6	2.61	109.72	106.98
5	C	510	SIA	O4-C4-C3	2.58	115.03	109.08
7	B	505	MAN	O5-C5-C6	2.57	109.67	106.98
4	A	503	NAG	O5-C5-C6	2.56	109.67	106.98
5	A	511	GAL	O3-C3-C2	2.46	115.87	110.35
8	D	602	NAG	O5-C5-C6	2.42	109.52	106.98
4	A	503	NAG	C3-C2-N2	-2.40	108.11	111.76
7	B	506	NAG	O5-C5-C6	2.39	109.49	106.98
5	C	510	SIA	O9-C9-C8	2.38	116.38	111.05
5	B	510	SIA	O1B-C1-O1A	-2.27	118.19	123.35
4	A	505	BMA	C6-C5-C4	-2.23	107.61	113.00
5	A	510	SIA	O9-C9-C8	2.23	116.04	111.05
5	B	511	GAL	O3-C3-C2	2.21	115.30	110.35
5	B	511	GAL	C1-C2-C3	-2.19	107.07	110.53
5	C	511	GAL	O2-C2-C1	-2.18	105.25	109.89
5	B	510	SIA	C3-C4-C5	2.17	113.39	110.41
7	B	504	BMA	O5-C5-C6	2.16	109.25	106.98
5	B	510	SIA	O9-C9-C8	2.15	115.88	111.05
9	D	607	MAN	O3-C3-C2	-2.13	106.04	109.94

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	511	GAL	C3-C4-C5	2.11	113.97	110.20
9	D	603	NAG	C2-N2-C7	-2.09	119.58	123.09
5	C	510	SIA	C3-C4-C5	2.09	113.28	110.41
7	C	503	NAG	O5-C5-C6	2.05	109.14	106.98
7	C	507	NAG	C3-C2-N2	-2.00	108.72	111.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 3 are monoatomic - leaving 5 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	A	502	1	12,14,15	0.62	0	15,19,21	1.18	1 (6%)
6	SLB	A	512	-	21,21,21	2.73	11 (52%)	31,31,31	1.30	4 (12%)
6	SLB	B	512	-	21,21,21	2.59	9 (42%)	31,31,31	1.20	4 (12%)
6	SLB	C	512	-	21,21,21	2.71	10 (47%)	31,31,31	1.07	3 (9%)
10	GOL	D	609	-	5,5,5	0.19	0	5,5,5	0.25	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	502	1	-	0/6/23/26	0/1/1/1
6	SLB	A	512	-	-	0/20/38/38	0/1/1/1
6	SLB	B	512	-	-	0/20/38/38	0/1/1/1
6	SLB	C	512	-	-	0/20/38/38	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	GOL	D	609	-	-	0/4/4/4	0/0/0/0

All (30) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	512	SLB	C2-C1	-6.26	1.47	1.53
6	A	512	SLB	C4-C5	-6.23	1.47	1.53
6	A	512	SLB	C2-C1	-6.03	1.47	1.53
6	C	512	SLB	C4-C5	-5.97	1.47	1.53
6	B	512	SLB	C2-C1	-5.84	1.47	1.53
6	B	512	SLB	C4-C5	-5.56	1.48	1.53
6	A	512	SLB	C5-N5	3.65	1.52	1.45
6	C	512	SLB	C5-N5	3.57	1.52	1.45
6	B	512	SLB	C6-C5	-3.51	1.47	1.53
6	A	512	SLB	O6-C6	3.49	1.50	1.44
6	A	512	SLB	C6-C5	-3.49	1.47	1.53
6	C	512	SLB	C6-C5	-3.40	1.47	1.53
6	B	512	SLB	O6-C6	3.30	1.49	1.44
6	A	512	SLB	C10-N5	3.30	1.47	1.34
6	C	512	SLB	O6-C6	3.30	1.49	1.44
6	B	512	SLB	C5-N5	3.08	1.51	1.45
6	C	512	SLB	C10-N5	3.02	1.46	1.34
6	B	512	SLB	C10-N5	2.93	1.46	1.34
6	B	512	SLB	C3-C4	-2.85	1.48	1.53
6	B	512	SLB	O6-C2	-2.76	1.40	1.42
6	C	512	SLB	C7-C6	-2.72	1.49	1.52
6	C	512	SLB	C3-C4	-2.56	1.49	1.53
6	A	512	SLB	C3-C4	-2.52	1.49	1.53
6	A	512	SLB	C7-C6	-2.41	1.49	1.52
6	B	512	SLB	C7-C6	-2.38	1.49	1.52
6	C	512	SLB	O2-C2	-2.19	1.36	1.39
6	A	512	SLB	O6-C2	-2.18	1.40	1.42
6	C	512	SLB	O6-C2	-2.15	1.40	1.42
6	A	512	SLB	O7-C7	-2.13	1.37	1.43
6	A	512	SLB	O2-C2	-2.02	1.37	1.39

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	512	SLB	O1A-C1-C2	-3.96	117.97	123.48
6	B	512	SLB	O1A-C1-C2	-3.94	118.01	123.48
3	A	502	NAG	O5-C5-C6	3.71	110.87	106.98

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
6	C	512	SLB	O6-C2-C3	2.89	111.92	110.22
6	A	512	SLB	O6-C2-C3	2.70	111.81	110.22
6	A	512	SLB	O9-C9-C8	2.29	116.17	111.05
6	C	512	SLB	O9-C9-C8	2.26	116.11	111.05
6	B	512	SLB	O6-C2-C3	2.23	111.53	110.22
6	C	512	SLB	O2-C2-C1	-2.17	106.19	109.86
6	B	512	SLB	C4-C5-N5	-2.10	105.47	110.12
6	A	512	SLB	C3-C4-C5	2.07	113.12	110.72
6	B	512	SLB	C3-C4-C5	2.06	113.11	110.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	388/388 (100%)	-0.38	11 (2%)	50	56	5, 9, 21, 46	0
1	B	388/388 (100%)	-0.47	1 (0%)	91	94	5, 9, 17, 39	0
1	C	388/388 (100%)	-0.52	1 (0%)	91	94	5, 8, 18, 36	0
1	D	388/388 (100%)	-0.53	1 (0%)	91	94	5, 9, 17, 42	0
All	All	1552/1552 (100%)	-0.47	14 (0%)	79	87	5, 9, 18, 46	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	82	VAL	5.6
1	A	82	VAL	5.0
1	C	82	VAL	4.9
1	B	82	VAL	4.8
1	A	249	ARG	4.5
1	A	246	ALA	3.4
1	A	295	TRP	3.0
1	A	247	SER	3.0
1	A	345	GLY	3.0
1	A	347	GLN	2.5
1	A	296	LYS	2.3
1	A	346	THR	2.2
1	A	273	GLN	2.1
1	A	344	ARG	2.1

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

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

### 6.3 Carbohydrates

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	MAN	B	505	11/12	0.10	69.00	18,20,25,27	0
7	BMA	C	504	11/12	0.18	16.26	17,22,28,29	0
4	MAN	A	509	11/12	0.16	13.17	13,21,27,36	0
7	BMA	B	504	11/12	0.11	11.80	16,18,21,21	0
7	NAG	C	503	14/15	0.11	9.90	15,24,32,33	0
7	NAG	B	506	14/15	0.10	9.22	9,12,30,33	0
4	NAG	A	503	14/15	0.09	5.99	8,10,28,37	0
4	BMA	A	505	11/12	0.07	4.04	8,9,11,11	0
9	MAN	D	608	11/12	0.08	3.55	11,12,14,15	0
7	NAG	C	502	14/15	0.08	2.43	11,15,21,23	0
7	MAN	B	509	11/12	0.07	2.32	12,15,17,19	0
7	BMA	C	508	11/12	0.07	2.31	11,11,17,17	0
9	MAN	D	607	11/12	0.08	1.76	11,12,18,25	0
9	MAN	D	606	11/12	0.08	1.59	10,14,17,18	0
7	MAN	C	505	11/12	0.14	1.52	14,17,19,32	0
4	MAN	A	506	11/12	0.07	1.35	10,12,16,16	0
4	MAN	A	508	11/12	0.08	1.18	9,11,15,20	0
5	GAL	A	511	12/12	0.21	1.16	20,35,43,44	0
8	NAG	D	601	14/15	0.10	1.06	16,22,34,41	0
7	NAG	B	502	14/15	0.10	0.98	11,15,19,23	0
5	SIA	B	510	20/21	0.08	0.70	11,14,18,18	0
7	NAG	C	507	14/15	0.06	0.61	8,10,14,15	0
9	BMA	D	605	11/12	0.06	0.58	8,10,13,17	0
5	SIA	A	510	20/21	0.08	0.46	11,13,18,18	0
9	NAG	D	604	14/15	0.06	0.40	8,8,12,13	0
7	BMA	B	508	11/12	0.07	0.34	10,12,19,19	0
5	SIA	C	510	20/21	0.07	0.31	9,11,15,15	0
9	NAG	D	603	14/15	0.07	0.14	8,9,18,18	0
4	MAN	A	507	11/12	0.07	-0.01	11,12,14,15	0
7	NAG	C	506	14/15	0.06	-0.02	7,9,12,15	0
7	NAG	B	507	14/15	0.07	-0.04	10,12,16,16	0
7	MAN	C	509	11/12	0.06	-0.67	12,15,17,20	0
4	NAG	A	504	14/15	0.04	-0.77	7,9,14,15	0
8	NAG	D	602	14/15	0.26	-	32,46,56,56	0
7	NAG	B	503	14/15	0.17	-	17,26,34,35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GAL	C	511	12/12	0.26	-	23,38,48,49	0
5	GAL	B	511	12/12	0.36	-	20,48,59,60	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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	502	14/15	0.13	24.51	19,25,28,29	0
10	GOL	D	609	6/6	0.15	8.48	20,20,20,20	0
6	SLB	C	512	21/21	0.09	2.23	8,10,13,19	0
6	SLB	A	512	21/21	0.17	1.86	14,18,20,25	0
6	SLB	B	512	21/21	0.07	0.64	7,9,11,18	0
2	CA	A	501	1/1	0.06	-1.18	17,17,17,17	0
2	CA	B	501	1/1	0.03	-1.54	7,7,7,7	0
2	CA	C	501	1/1	0.02	-2.06	8,8,8,8	0

## 6.5 Other polymers

There are no such residues in this entry.