



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 03:10 am GMT

PDB ID : 1ING  
Title : INFLUENZA A SUBTYPE N2 NEURAMINIDASE COMPLEXED WITH AROMATIC BANA109 INHIBITOR  
Authors : Jedrzejewski, M.J.; Luo, M.  
Deposited on : 1995-07-07  
Resolution : 2.40 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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
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)	:	recalc28949

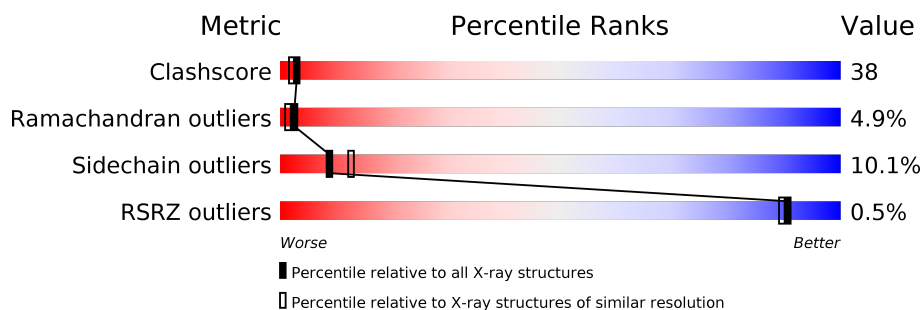
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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	388	<div> <div></div> <div>41%</div> <div>50%</div> <div>9%</div> <div>.</div> </div>
1	B	388	<div> <div>40%</div> <div>51%</div> <div>9%</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	FUL	A	477	-	-	X	-
4	NAG	A	478	-	-	X	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MAN	A	481	-	-	X	-
4	MAN	A	483	-	-	X	X
4	NAG	B	478	-	-	-	X
4	MAN	B	483	-	-	-	X
5	FUC	B	477	X	-	-	-
7	ST5	A	471	-	-	-	X
7	ST5	B	471	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 8978 atoms, of which 2292 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 INFLUENZA A SUBTYPE N2 NEURAMINIDASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	388	Total	C	H	N	O	S	0	0	0
			3745	1866	723	545	588	23			
1	B	388	Total	C	H	N	O	S	0	0	0
			3745	1866	723	545	588	23			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	339	ASP	ASN	CONFLICT	UNP P06820
B	339	ASP	ASN	CONFLICT	UNP P06820

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	2	Total	C	H	N	O	0	0
			55	16	27	2	10		
2	A	2	Total	C	H	N	O	0	0
			55	16	27	2	10		
2	B	2	Total	C	H	N	O	0	0
			55	16	27	2	10		
2	B	2	Total	C	H	N	O	0	0
			55	16	27	2	10		

- Molecule 3 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	4	Total	C	H	N	O	0	0
			96	28	47	2	19		

- Molecule 4 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	6	Total	C	H	N	O	0	0
			139	40	67	2	30		
4	B	6	Total	C	H	N	O	0	0
			139	40	67	2	30		

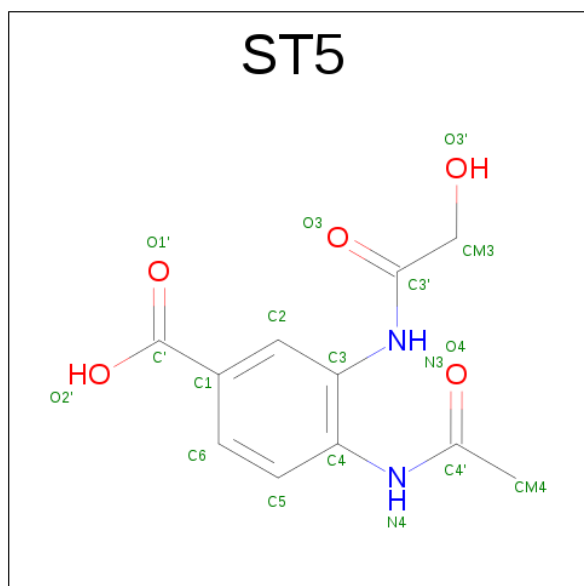
- Molecule 5 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	4	Total	C	H	N	O	0	0
			96	28	47	2	19		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 7 is 4-(ACETYLAMINO)-3-[(HYDROXYACETYL)AMINO]BENZOIC ACID (three-letter code: ST5) (formula: C<sub>11</sub>H<sub>12</sub>N<sub>2</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	H	N	O	0	0
			23	11	5	2	5		
7	B	1	Total	C	H	N	O	0	0
			23	11	5	2	5		

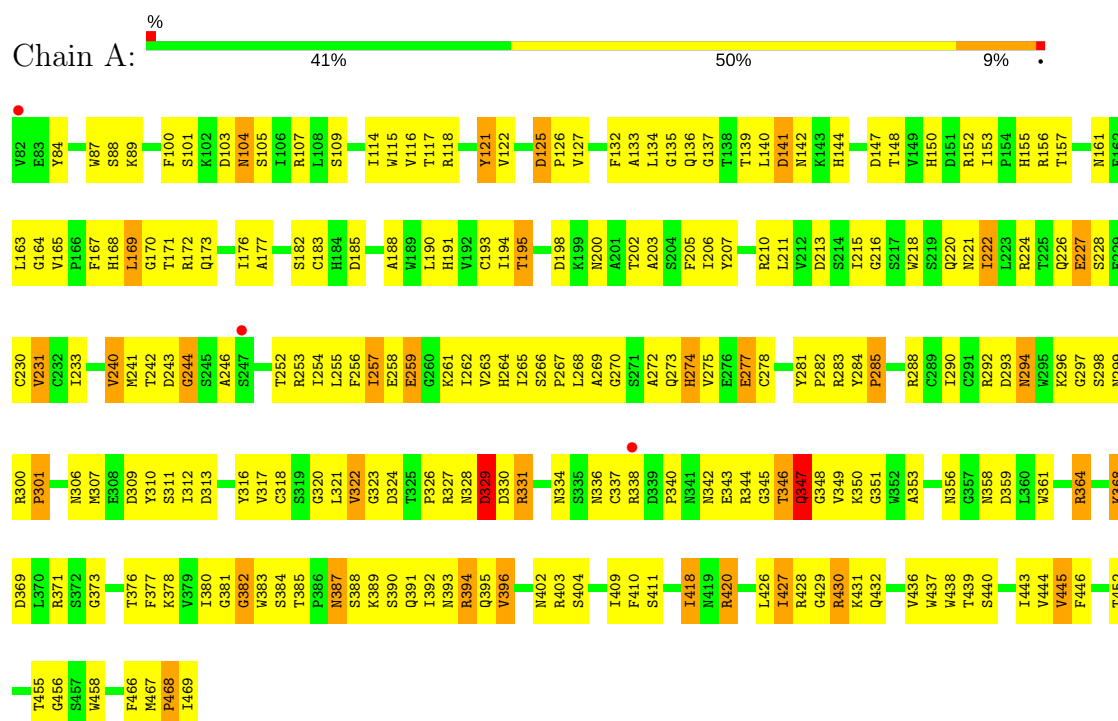
- Molecule 8 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	122	Total 366	H 244	O 122	0	0
8	B	128	Total 384	H 256	O 128	0	0

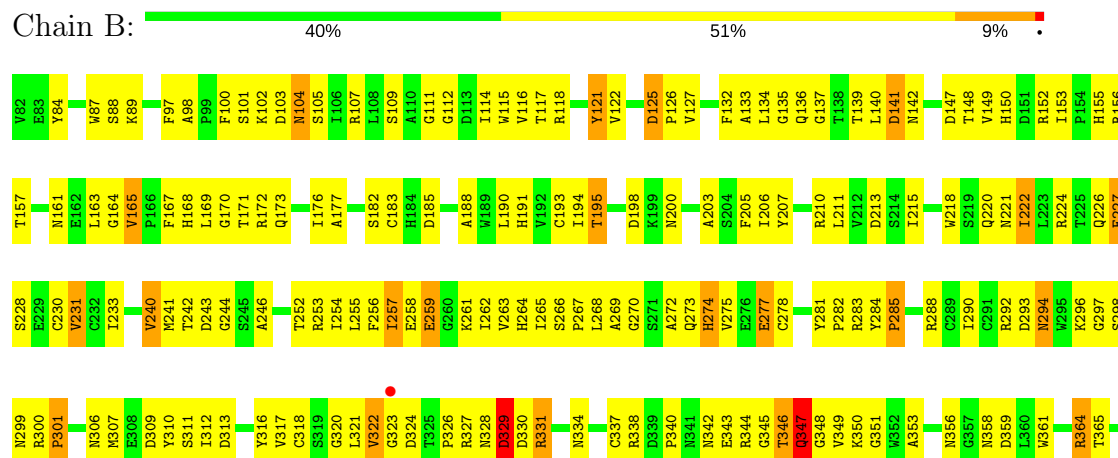
### 3 Residue-property plots

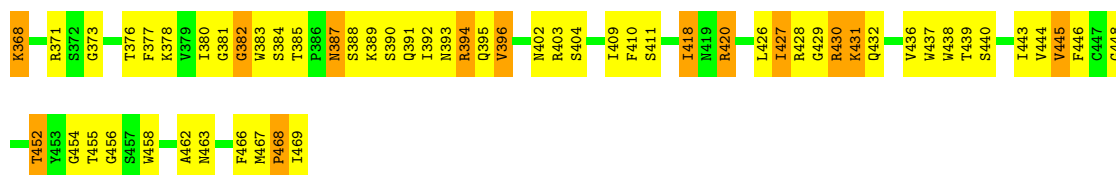
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: INFLUENZA A SUBTYPE N2 NEURAMINIDASE



#### • Molecule 1: INFLUENZA A SUBTYPE N2 NEURAMINIDASE







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.42Å 139.06Å 139.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.50 – 2.40 21.23 – 2.34	Depositor EDS
% Data completeness (in resolution range)	(Not available) (6.50-2.40) 44.6 (21.23-2.34)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.95 (at 2.36Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.163 , (Not available) 0.212 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	13.8	Xtriage
Anisotropy	2.422	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.46 , 122.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	8978	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.51% 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, CA, ST5, FUC, FUL, MAN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.57	0/3092	0.90	2/4194 (0.0%)
1	B	0.57	0/3092	0.90	2/4194 (0.0%)
All	All	0.57	0/6184	0.90	4/8388 (0.0%)

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
1	A	0	1
1	B	0	1
5	B	1	0
All	All	1	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	346	THR	N-CA-C	-5.15	97.10	111.00
1	B	346	THR	N-CA-C	-5.15	97.10	111.00
1	A	125	ASP	N-CA-C	-5.00	97.49	111.00
1	B	125	ASP	N-CA-C	-5.00	97.49	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	B	477	FUC	C1

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	121	TYR	Sidechain
1	B	121	TYR	Sidechain

## 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	3022	723	2851	229	28
1	B	3022	723	2850	258	12
2	A	56	54	50	0	0
2	B	56	54	50	0	0
3	A	49	47	43	0	10
4	A	72	67	59	21	10
4	B	72	67	61	1	6
5	B	49	47	43	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	18	5	11	2	0
7	B	18	5	11	2	0
8	A	122	244	0	18	9
8	B	128	256	0	24	21
All	All	6686	2292	6029	468	48

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 38.

The worst 5 of 468 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)
4:A:478:NAG:C6	1:B:455:THR:CB	2.15	1.23
1:B:102:LYS:HE3	8:B:488:HOH:O	1.45	1.14
4:A:478:NAG:H5	1:B:455:THR:HG21	1.32	1.09
4:A:478:NAG:C1	1:B:455:THR:HG23	1.81	1.09
1:A:136:GLN:HG3	1:A:148:THR:HG23	1.44	0.99

The worst 5 of 48 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:170:GLY:CA	8:A:602:HOH:O[4_555]	0.79	1.41
3:A:477:FUL:H3	4:A:481:MAN:H4[3_654]	0.40	1.20
1:B:368:LYS:HZ1	8:B:613:HOH:H1[3_654]	0.46	1.14
1:A:394:ARG:CA	8:B:517:HOH:O[4_555]	1.06	1.14
3:A:477:FUL:C4	4:A:481:MAN:H61[3_654]	0.59	1.01

## 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	386/388 (100%)	302 (78%)	65 (17%)	19 (5%)	2	1
1	B	386/388 (100%)	302 (78%)	65 (17%)	19 (5%)	2	1
All	All	772/776 (100%)	604 (78%)	130 (17%)	38 (5%)	2	1

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	87	TRP
1	A	274	HIS
1	A	329	ASP
1	A	356	ASN
1	A	430	ARG

### 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	338/338 (100%)	304 (90%)	34 (10%)	9	12
1	B	338/338 (100%)	304 (90%)	34 (10%)	9	12
All	All	676/676 (100%)	608 (90%)	68 (10%)	9	12

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	427	ILE
1	B	157	THR
1	B	418	ILE
1	A	445	VAL
1	B	118	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	402	ASN
1	B	104	ASN
1	B	402	ASN
1	A	465	ASN
1	B	161	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 ⓘ

28 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	$\# Z  > 2$
2	NAG	A	472	1,2	14,14,15	1.41	2 (14%)	15,19,21	2.18	3 (20%)
2	NAG	A	473	2	14,14,15	1.46	3 (21%)	15,19,21	2.35	2 (13%)
3	NAG	A	474	1,3	14,14,15	2.05	3 (21%)	15,19,21	3.04	4 (26%)
3	NAG	A	475	3	14,14,15	1.91	2 (14%)	15,19,21	2.35	5 (33%)
3	BMA	A	476	3	11,11,12	1.26	1 (9%)	13,15,17	2.32	4 (30%)
3	FUL	A	477	3,4	9,10,11	2.12	3 (33%)	13,14,16	2.10	5 (38%)
4	NAG	A	478	1,4	14,14,15	1.98	4 (28%)	15,19,21	3.03	5 (33%)
4	NAG	A	479	4	14,14,15	1.74	3 (21%)	15,19,21	2.74	4 (26%)
4	BMA	A	480	4	11,11,12	1.95	5 (45%)	13,15,17	2.81	4 (30%)
4	MAN	A	481	3,4	11,11,12	3.14	3 (27%)	13,15,17	2.79	4 (30%)
4	MAN	A	482	4	11,11,12	1.90	4 (36%)	13,15,17	2.63	5 (38%)
4	MAN	A	483	1,4	11,11,12	1.72	4 (36%)	13,15,17	1.67	3 (23%)
2	NAG	A	484	1,2	14,14,15	0.94	1 (7%)	15,19,21	2.50	4 (26%)
2	NAG	A	485	2	14,14,15	1.33	2 (14%)	15,19,21	3.68	11 (73%)
2	NAG	B	472	1,2	14,14,15	1.41	2 (14%)	15,19,21	2.18	3 (20%)
2	NAG	B	473	2	14,14,15	1.46	3 (21%)	15,19,21	2.35	2 (13%)
5	NAG	B	474	1,5	14,14,15	2.05	3 (21%)	15,19,21	3.04	4 (26%)
5	NAG	B	475	5	14,14,15	1.91	2 (14%)	15,19,21	2.35	5 (33%)
5	BMA	B	476	5	11,11,12	1.26	1 (9%)	13,15,17	2.32	4 (30%)
5	FUC	B	477	5	9,10,11	2.12	3 (33%)	13,14,16	2.10	5 (38%)
4	NAG	B	478	1,4	14,14,15	1.98	4 (28%)	15,19,21	3.03	5 (33%)
4	NAG	B	479	4	14,14,15	1.74	3 (21%)	15,19,21	2.74	4 (26%)
4	BMA	B	480	4	11,11,12	1.95	5 (45%)	13,15,17	2.81	4 (30%)
4	MAN	B	481	4	11,11,12	3.14	3 (27%)	13,15,17	2.79	4 (30%)
4	MAN	B	482	4	11,11,12	1.90	4 (36%)	13,15,17	2.63	5 (38%)
4	MAN	B	483	1,4	11,11,12	1.72	4 (36%)	13,15,17	1.67	3 (23%)
2	NAG	B	484	1,2	14,14,15	0.94	1 (7%)	15,19,21	2.50	4 (26%)
2	NAG	B	485	2	14,14,15	1.33	2 (14%)	15,19,21	3.68	11 (73%)

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
2	NAG	A	472	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	473	2	-	0/6/23/26	0/1/1/1
3	NAG	A	474	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	475	3	-	0/6/23/26	0/1/1/1
3	BMA	A	476	3	-	0/2/19/22	0/1/1/1
3	FUL	A	477	3,4	-	0/0/17/20	0/1/1/1
4	NAG	A	478	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	479	4	-	0/6/23/26	0/1/1/1
4	BMA	A	480	4	-	0/2/19/22	0/1/1/1
4	MAN	A	481	3,4	-	0/2/19/22	0/1/1/1
4	MAN	A	482	4	-	0/2/19/22	0/1/1/1
4	MAN	A	483	1,4	-	0/2/19/22	1/1/1/1
2	NAG	A	484	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	485	2	-	0/6/23/26	0/1/1/1
2	NAG	B	472	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	473	2	-	0/6/23/26	0/1/1/1
5	NAG	B	474	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	475	5	-	0/6/23/26	0/1/1/1
5	BMA	B	476	5	-	0/2/19/22	0/1/1/1
5	FUC	B	477	5	1/1/5/5	0/0/17/20	0/1/1/1
4	NAG	B	478	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	479	4	-	0/6/23/26	0/1/1/1
4	BMA	B	480	4	-	0/2/19/22	0/1/1/1
4	MAN	B	481	4	-	0/2/19/22	0/1/1/1
4	MAN	B	482	4	-	0/2/19/22	0/1/1/1
4	MAN	B	483	1,4	-	0/2/19/22	1/1/1/1
2	NAG	B	484	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	485	2	-	0/6/23/26	0/1/1/1

The worst 5 of 80 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	474	NAG	C1-C2	-6.02	1.44	1.52
5	B	474	NAG	C1-C2	-6.02	1.44	1.52
2	B	472	NAG	C1-C2	-3.35	1.47	1.52
2	A	472	NAG	C1-C2	-3.35	1.47	1.52
2	B	472	NAG	O5-C1	-2.90	1.39	1.43

The worst 5 of 126 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	479	NAG	C4-C3-C2	-6.07	102.12	111.02
4	A	479	NAG	C4-C3-C2	-6.07	102.12	111.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	472	NAG	O5-C1-C2	-5.89	103.27	111.47
2	A	472	NAG	O5-C1-C2	-5.89	103.27	111.47
2	A	485	NAG	C4-C3-C2	-5.15	103.47	111.02

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	B	477	FUC	C1

There are no torsion outliers.

All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	483	MAN	C1-C2-C3-C4-C5-O5
4	B	483	MAN	C1-C2-C3-C4-C5-O5

6 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	477	FUL	0	10
4	A	478	NAG	15	0
4	A	481	MAN	0	10
4	A	483	MAN	6	0
4	B	478	NAG	1	5
4	B	483	MAN	0	1

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\# Z  > 2$
7	ST5	A	471	-	15,18,18	0.69	0	20,24,24	1.22	2 (10%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	ST5	B	471	-	15,18,18	0.69	0	20,24,24	1.22	2 (10%)

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
7	ST5	A	471	-	-	0/9/14/14	0/1/1/1
7	ST5	B	471	-	-	0/9/14/14	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	471	ST5	C4-N4-C4'	-2.45	121.50	127.56
7	A	471	ST5	C4-N4-C4'	-2.45	121.50	127.56
7	B	471	ST5	CM4-C4'-N4	2.97	119.38	115.02
7	A	471	ST5	CM4-C4'-N4	2.97	119.38	115.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	471	ST5	2	0
7	B	471	ST5	2	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	388/388 (100%)	0.09	3 (0%) 86 84	2, 12, 29, 38	0
1	B	388/388 (100%)	0.13	1 (0%) 93 93	2, 12, 29, 38	0
All	All	776/776 (100%)	0.11	4 (0%) 90 89	2, 12, 29, 38	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	82	VAL	2.2
1	A	247	SER	2.2
1	B	323	GLY	2.1
1	A	338	ARG	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](#)

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
4	MAN	A	483	11/12	0.71	0.51	14.11	15,15,46,49	11
4	MAN	B	483	11/12	0.76	0.50	10.34	15,15,46,49	11
4	NAG	B	478	14/15	0.74	0.23	2.74	2,15,36,41	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
4	NAG	A	478	14/15	0.82	0.19	1.69	2,15,36,41	0
3	NAG	A	474	14/15	0.85	0.15	-0.24	15,23,29,37	14
2	NAG	B	484	14/15	0.78	0.18	-0.28	15,15,42,43	14
5	NAG	B	474	14/15	0.86	0.15	-0.32	15,23,29,37	14
2	NAG	A	484	14/15	0.80	0.15	-0.60	15,15,42,43	14
4	NAG	B	479	14/15	0.58	0.45	-	15,15,40,41	14
4	NAG	A	479	14/15	0.52	0.37	-	15,15,40,41	14
3	NAG	A	475	14/15	0.53	0.39	-	15,15,39,43	14
2	NAG	A	473	14/15	0.72	0.30	-	15,15,46,47	14
4	MAN	A	481	11/12	0.70	0.36	-	15,15,38,40	11
2	NAG	A	472	14/15	0.71	0.24	-	15,15,39,41	14
2	NAG	B	485	14/15	0.78	0.37	-	15,15,45,47	14
4	BMA	B	480	11/12	0.63	0.56	-	15,15,38,39	11
4	MAN	B	481	11/12	0.73	0.44	-	15,15,38,40	11
4	BMA	A	480	11/12	0.70	0.42	-	15,15,38,39	11
4	MAN	B	482	11/12	0.62	0.44	-	15,15,44,46	11
2	NAG	B	472	14/15	0.90	0.21	-	15,15,39,41	14
5	FUC	B	477	10/11	0.90	0.19	-	11,15,22,23	10
2	NAG	B	473	14/15	0.58	0.40	-	15,15,46,47	14
4	MAN	A	482	11/12	0.66	0.37	-	15,15,44,46	11
5	NAG	B	475	14/15	0.54	0.30	-	15,15,39,43	14
3	BMA	A	476	11/12	0.32	0.50	-	15,15,48,50	11
3	FUL	A	477	10/11	0.80	0.29	-	11,15,22,23	10
2	NAG	A	485	14/15	0.80	0.29	-	15,15,45,47	14
5	BMA	B	476	11/12	0.66	0.36	-	15,15,48,50	11

## 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
7	ST5	B	471	18/18	0.73	0.28	3.88	15,47,51,53	0
7	ST5	A	471	18/18	0.75	0.25	2.37	15,47,51,53	0
6	CA	A	470	1/1	0.92	0.08	-2.49	37,37,37,37	0
6	CA	B	470	1/1	0.95	0.07	-3.13	37,37,37,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.