



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 07:23 am GMT

PDB ID : 3K39  
Title : Crystal Structure of B/Perth Neuraminidase D197E mutant in complex with Peramivir  
Authors : Oakley, A.J.; McKimm-Breschkin, J.L.  
Deposited on : 2009-10-02  
Resolution : 2.54 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



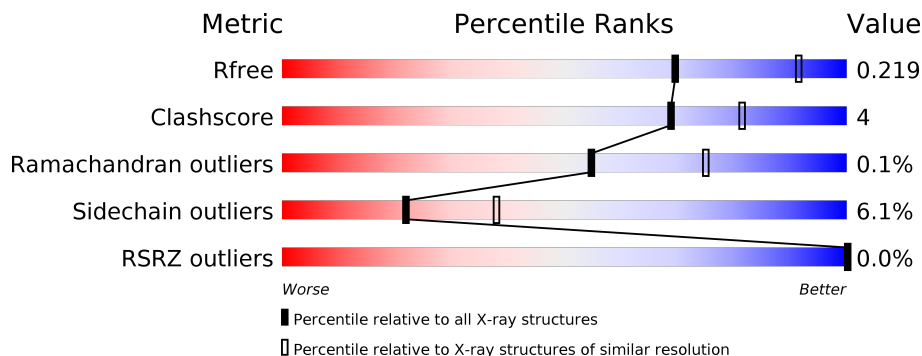
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.54 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	4993 (2.58-2.50)
Clashscore	112137	5755 (2.58-2.50)
Ramachandran outliers	110173	5652 (2.58-2.50)
Sidechain outliers	110143	5654 (2.58-2.50)
RSRZ outliers	101464	5026 (2.58-2.50)











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. 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	397	
1	B	397	
1	C	397	
1	D	397	
1	E	397	
1	F	397	

Continued on next page...



*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	397	 86% 10% ..
1	H	397	 87% 9% ..
1	I	397	 86% 10% ..
1	J	397	 87% 10% ..
1	K	397	 87% 9% ..
1	L	397	 86% 10% ..
1	M	397	 86% 11% ..
1	N	397	 88% 8% ..
1	O	397	 87% 10% ..
1	P	397	 87% 9% ..

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
2	NAG	A	900	-	-	-	X
2	NAG	B	900	-	-	-	X
2	NAG	D	900	-	-	-	X
2	NAG	E	900	-	-	-	X
2	NAG	F	900	-	-	-	X
2	NAG	G	900	-	-	-	X
2	NAG	H	900	-	-	-	X
2	NAG	I	900	-	-	-	X
2	NAG	J	900	-	-	-	X
2	NAG	M	900	-	-	-	X
2	NAG	N	900	-	-	-	X
2	NAG	O	900	-	-	-	X
2	NAG	P	900	-	-	-	X
3	CA	I	1000	-	-	-	X
4	BCZ	A	1001	-	-	-	X



## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 49796 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 Neuraminidase.

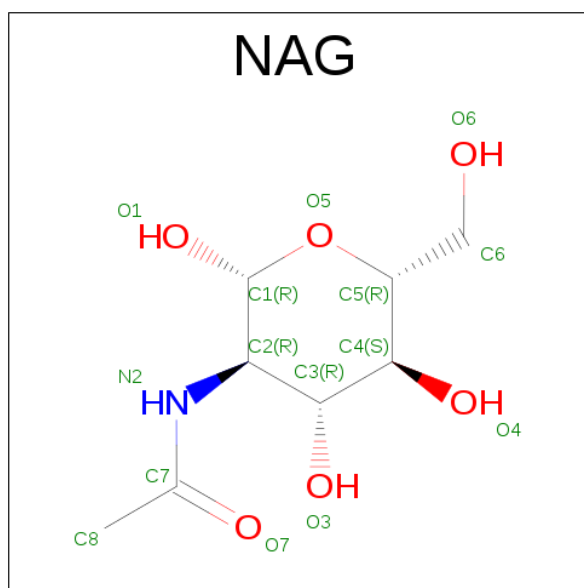
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	B	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	C	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	D	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	E	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	F	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	G	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	H	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	I	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	J	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	K	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	L	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	M	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	N	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	O	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			
1	P	389	Total	C	N	O	S	0	1	0
			3016	1892	522	573	29			



There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	197	GLU	ASP	ENGINEERED	UNP Q3S340
B	197	GLU	ASP	ENGINEERED	UNP Q3S340
C	197	GLU	ASP	ENGINEERED	UNP Q3S340
D	197	GLU	ASP	ENGINEERED	UNP Q3S340
E	197	GLU	ASP	ENGINEERED	UNP Q3S340
F	197	GLU	ASP	ENGINEERED	UNP Q3S340
G	197	GLU	ASP	ENGINEERED	UNP Q3S340
H	197	GLU	ASP	ENGINEERED	UNP Q3S340
I	197	GLU	ASP	ENGINEERED	UNP Q3S340
J	197	GLU	ASP	ENGINEERED	UNP Q3S340
K	197	GLU	ASP	ENGINEERED	UNP Q3S340
L	197	GLU	ASP	ENGINEERED	UNP Q3S340
M	197	GLU	ASP	ENGINEERED	UNP Q3S340
N	197	GLU	ASP	ENGINEERED	UNP Q3S340
O	197	GLU	ASP	ENGINEERED	UNP Q3S340
P	197	GLU	ASP	ENGINEERED	UNP Q3S340

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	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
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		
2	G	1	Total	C	N	O	0	0
			14	8	1	5		
2	H	1	Total	C	N	O	0	0
			14	8	1	5		
2	I	1	Total	C	N	O	0	0
			14	8	1	5		
2	J	1	Total	C	N	O	0	0
			14	8	1	5		
2	K	1	Total	C	N	O	0	0
			14	8	1	5		
2	L	1	Total	C	N	O	0	0
			14	8	1	5		
2	M	1	Total	C	N	O	0	0
			14	8	1	5		
2	N	1	Total	C	N	O	0	0
			14	8	1	5		
2	O	1	Total	C	N	O	0	0
			14	8	1	5		
2	P	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	1	Total	Ca	0	0
			1	1		
3	G	1	Total	Ca	0	0
			1	1		
3	J	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		
3	K	1	Total	Ca	0	0
			1	1		

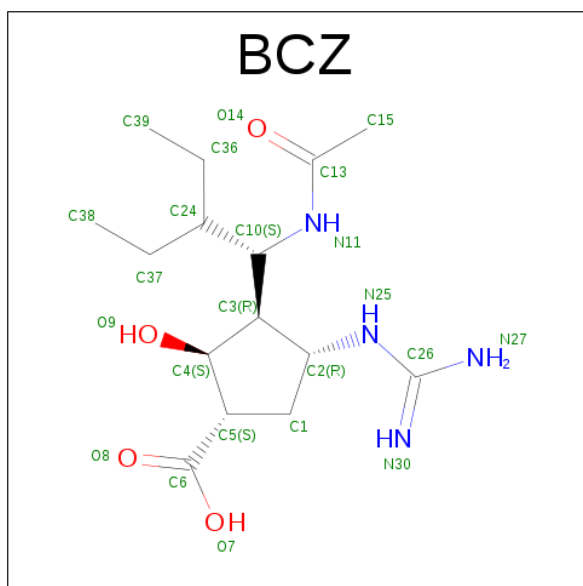
*Continued on next page...*



Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total	Ca	0	0
			1	1		
3	H	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	I	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	A	1	Total	Ca	0	0
			1	1		
3	N	1	Total	Ca	0	0
			1	1		
3	O	1	Total	Ca	0	0
			1	1		
3	L	1	Total	Ca	0	0
			1	1		
3	F	1	Total	Ca	0	0
			1	1		
3	M	1	Total	Ca	0	0
			1	1		

- Molecule 4 is 3-(1-ACETYLAMINO-2-ETHYL-BUTYL)-4-GUANIDINO-2-HYDROXY-CYCLOPENTANECARBOXYLIC ACID (three-letter code: BCZ) (formula: C<sub>15</sub>H<sub>28</sub>N<sub>4</sub>O<sub>4</sub>).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 23 15 4 4	0	0
4	B	1	Total C N O 23 15 4 4	0	0
4	C	1	Total C N O 23 15 4 4	0	0
4	D	1	Total C N O 23 15 4 4	0	0
4	E	1	Total C N O 23 15 4 4	0	0
4	F	1	Total C N O 23 15 4 4	0	0
4	G	1	Total C N O 23 15 4 4	0	0
4	H	1	Total C N O 23 15 4 4	0	0
4	I	1	Total C N O 23 15 4 4	0	0
4	J	1	Total C N O 23 15 4 4	0	0
4	K	1	Total C N O 23 15 4 4	0	0
4	L	1	Total C N O 23 15 4 4	0	0
4	M	1	Total C N O 23 15 4 4	0	0
4	N	1	Total C N O 23 15 4 4	0	0
4	O	1	Total C N O 23 15 4 4	0	0
4	P	1	Total C N O 23 15 4 4	0	0

- Molecule 5 is YTTRIUM (III) ION (three-letter code: YT3) (formula: Y).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	I	1	Total Y 1 1	0	0
5	A	1	Total Y 1 1	0	0
5	M	1	Total Y 1 1	0	0
5	E	1	Total Y 1 1	0	0



- Molecule 6 is water.

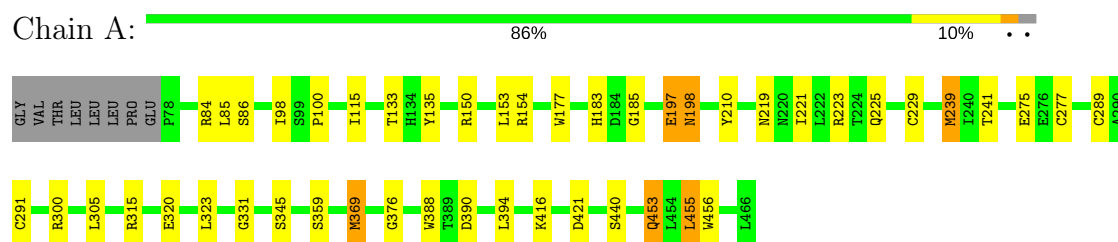
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	56	Total 56	O 56	0	0
6	B	56	Total 56	O 56	0	0
6	C	61	Total 61	O 61	0	0
6	D	56	Total 56	O 56	0	0
6	E	62	Total 62	O 62	0	0
6	F	58	Total 58	O 58	0	0
6	G	58	Total 58	O 58	0	0
6	H	57	Total 57	O 57	0	0
6	I	56	Total 56	O 56	0	0
6	J	57	Total 57	O 57	0	0
6	K	56	Total 56	O 56	0	0
6	L	59	Total 59	O 59	0	0
6	M	58	Total 58	O 58	0	0
6	N	58	Total 58	O 58	0	0
6	O	62	Total 62	O 62	0	0
6	P	58	Total 58	O 58	0	0



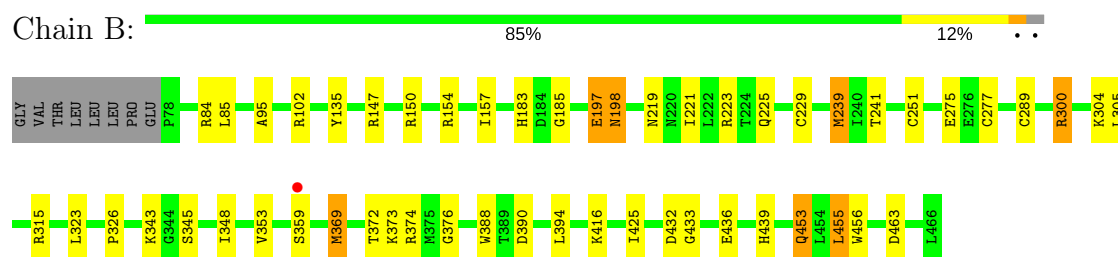
### 3 Residue-property plots

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.

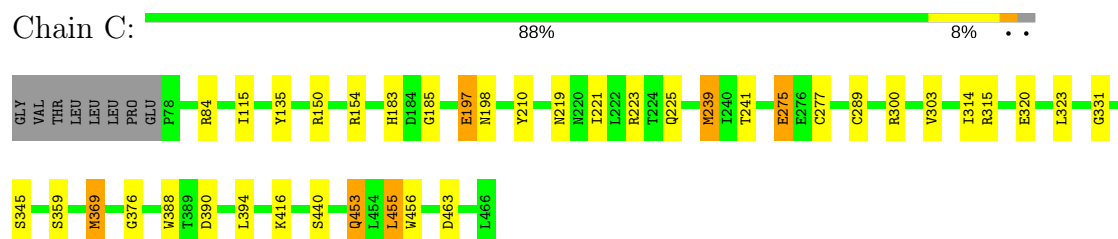
#### • Molecule 1: Neuraminidase



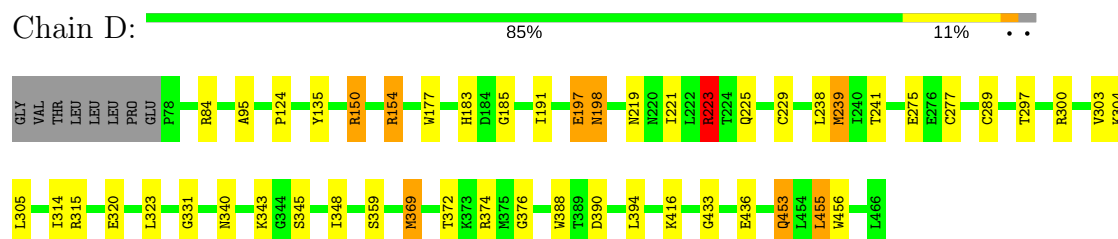
#### • Molecule 1: Neuraminidase



#### • Molecule 1: Neuraminidase

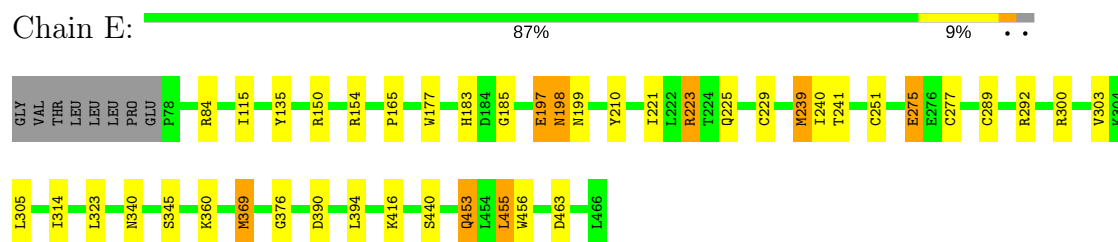


#### • Molecule 1: Neuraminidase

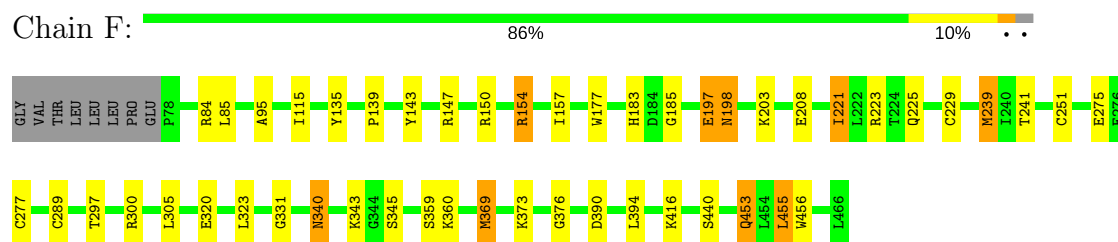




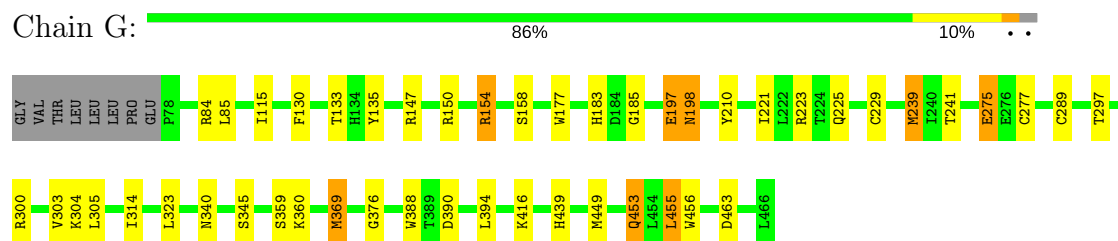
- Molecule 1: Neuraminidase



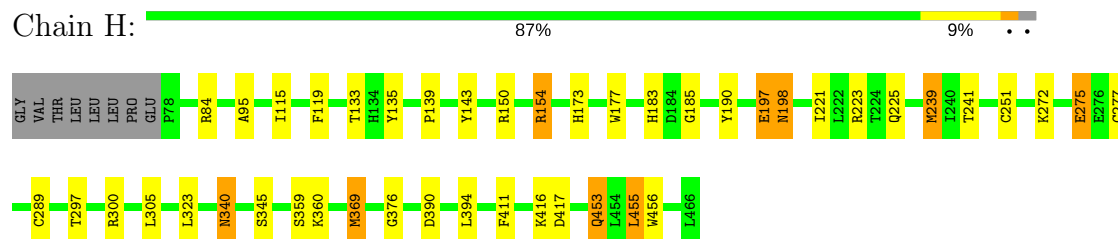
- Molecule 1: Neuraminidase



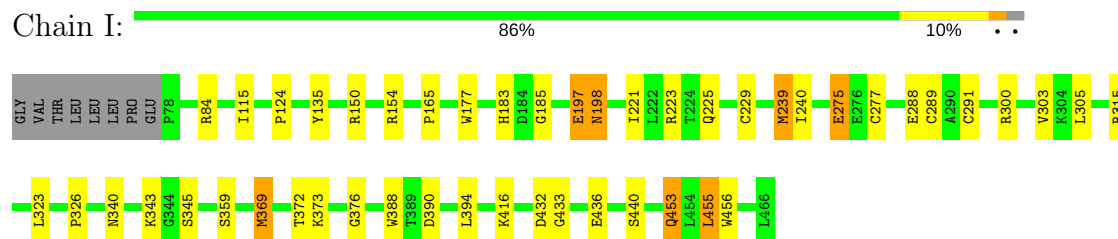
- Molecule 1: Neuraminidase



- Molecule 1: Neuraminidase

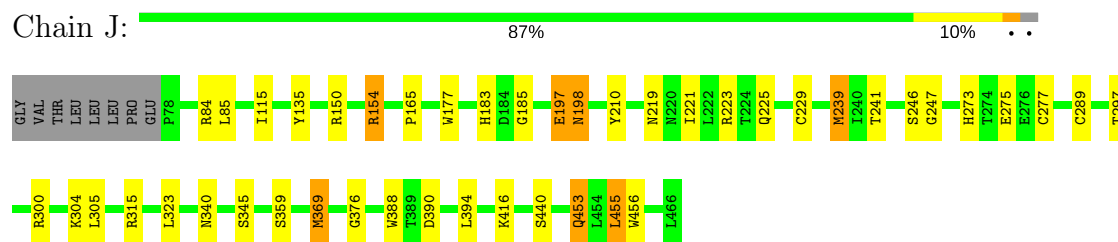


- Molecule 1: Neuraminidase

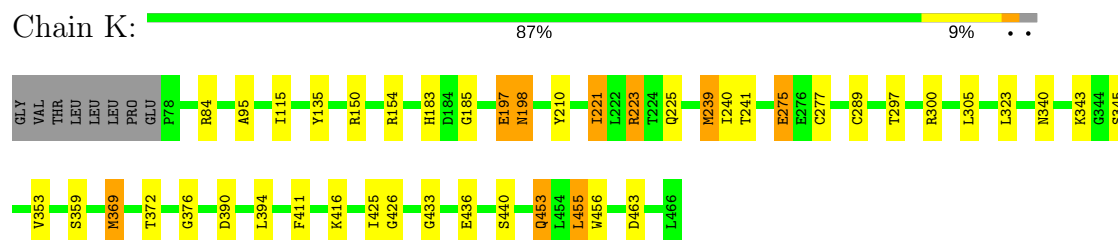


- Molecule 1: Neuraminidase

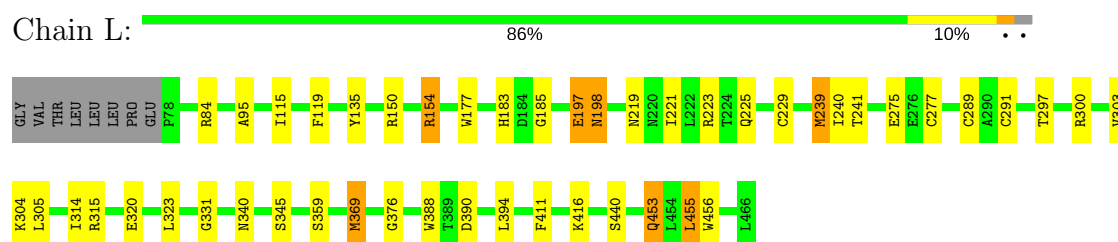




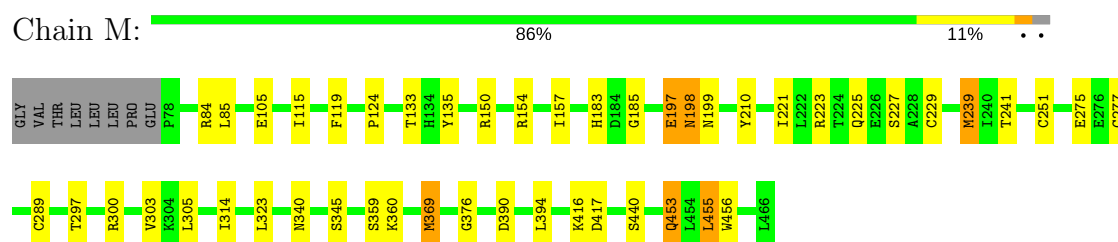
• Molecule 1: Neuraminidase



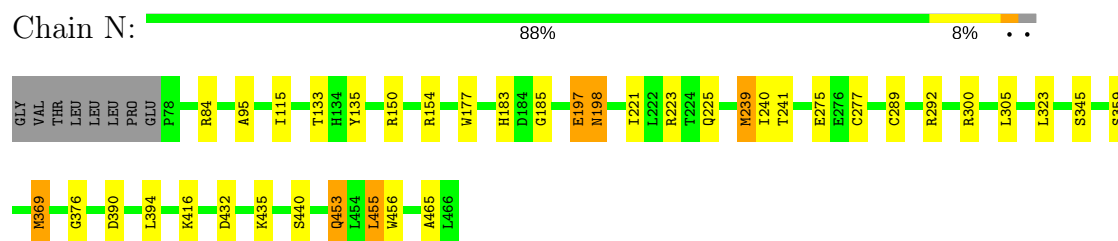
• Molecule 1: Neuraminidase



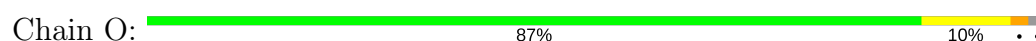
• Molecule 1: Neuraminidase



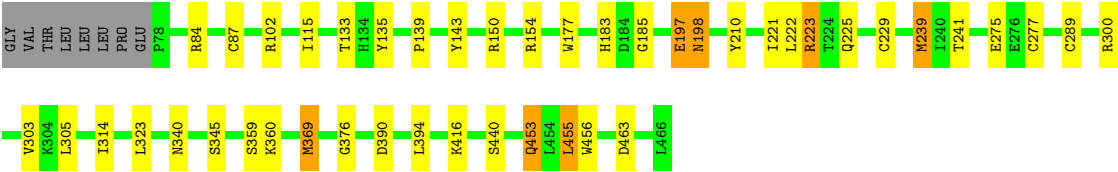
• Molecule 1: Neuraminidase



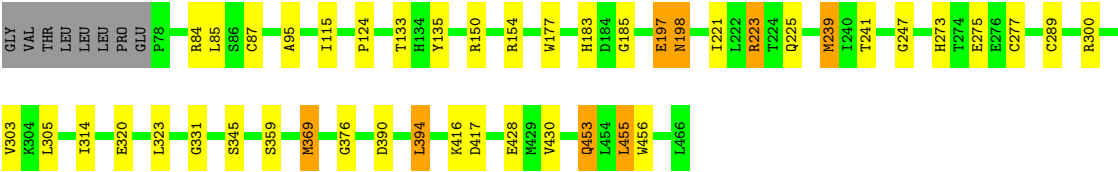
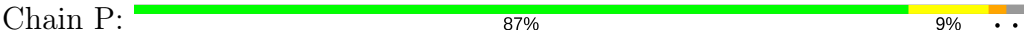
• Molecule 1: Neuraminidase







● Molecule 1: Neuraminidase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.96Å 124.87Å 125.25Å 90.03° 92.11° 91.27°	Depositor
Resolution (Å)	55.90 – 2.54 55.83 – 2.30	Depositor EDS
% Data completeness (in resolution range)	90.6 (55.90-2.54) 95.2 (55.83-2.30)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.16 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.198 , 0.214 0.215 , 0.219	Depositor DCC
$R_{free}$ test set	10923 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.4	Xtriage
Anisotropy	0.423	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 2.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.119 for h,l,-k 0.119 for h,-l,k 0.118 for h,-k,-l 0.179 for -h,k,-l 0.125 for -h,-k,l 0.099 for -h,l,k 0.189 for -h,-l,-k	Xtriage
Reported twinning fraction	0.758 for H, K, L 0.139 for -H, -L, -K 0.103 for -H, K, -L	Depositor
Outliers	0 of 286793 reflections	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	49796	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage'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.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, NAG, BCZ, YT3

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	0.76	0/3089	0.75	0/4171
1	B	0.76	0/3089	0.75	1/4171 (0.0%)
1	C	0.75	0/3089	0.75	0/4171
1	D	0.74	0/3089	0.74	1/4171 (0.0%)
1	E	0.79	0/3089	0.76	1/4171 (0.0%)
1	F	0.83	1/3089 (0.0%)	0.77	0/4171
1	G	0.79	0/3089	0.76	0/4171
1	H	0.78	0/3089	0.76	1/4171 (0.0%)
1	I	0.77	0/3089	0.76	0/4171
1	J	0.78	0/3089	0.74	0/4171
1	K	0.76	0/3089	0.74	0/4171
1	L	0.77	0/3089	0.76	0/4171
1	M	0.77	0/3089	0.74	1/4171 (0.0%)
1	N	0.80	0/3089	0.77	1/4171 (0.0%)
1	O	0.78	1/3089 (0.0%)	0.76	1/4171 (0.0%)
1	P	0.80	1/3089 (0.0%)	0.78	2/4171 (0.0%)
All	All	0.78	3/49424 (0.0%)	0.76	9/66736 (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	E	0	1
1	F	0	1
1	H	0	1
1	I	0	1
1	O	0	2
All	All	0	6



All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	P	87	CYS	CB-SG	5.64	1.91	1.82
1	F	208	GLU	CG-CD	5.15	1.59	1.51
1	O	87	CYS	CB-SG	5.06	1.90	1.82

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	N	292	ARG	NE-CZ-NH1	6.63	123.61	120.30
1	P	417	ASP	CB-CG-OD1	6.27	123.94	118.30
1	H	417	ASP	CB-CG-OD1	6.24	123.92	118.30
1	E	292	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	B	300	ARG	NE-CZ-NH2	-5.37	117.62	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	340	ASN	Peptide
1	F	340	ASN	Peptide
1	H	340	ASN	Peptide
1	I	340	ASN	Peptide
1	O	222	LEU	Peptide

## 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	3016	0	2912	25	1
1	B	3016	0	2912	47	0
1	C	3016	0	2912	22	1
1	D	3016	0	2912	23	5
1	E	3016	0	2912	23	1
1	F	3016	0	2912	47	0
1	G	3016	0	2912	28	1
1	H	3016	0	2912	23	1
1	I	3016	0	2912	47	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	3016	0	2912	22	1
1	K	3016	0	2912	24	4
1	L	3016	0	2912	28	1
1	M	3016	0	2912	25	1
1	N	3016	0	2912	43	0
1	O	3016	0	2912	22	1
1	P	3016	0	2912	25	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
2	D	14	0	13	0	0
2	E	14	0	13	0	0
2	F	14	0	13	0	0
2	G	14	0	13	0	0
2	H	14	0	13	0	0
2	I	14	0	13	0	0
2	J	14	0	13	0	0
2	K	14	0	13	0	0
2	L	14	0	13	0	0
2	M	14	0	13	0	0
2	N	14	0	13	0	0
2	O	14	0	13	0	0
2	P	14	0	13	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	J	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
3	M	1	0	0	0	0
3	N	1	0	0	0	0
3	O	1	0	0	0	0
3	P	1	0	0	0	0
4	A	23	0	27	1	0
4	B	23	0	27	1	0
4	C	23	0	27	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	23	0	27	1	0
4	E	23	0	27	1	0
4	F	23	0	27	1	0
4	G	23	0	27	2	0
4	H	23	0	27	1	0
4	I	23	0	27	2	0
4	J	23	0	27	1	0
4	K	23	0	27	1	0
4	L	23	0	27	1	0
4	M	23	0	27	2	0
4	N	23	0	27	1	0
4	O	23	0	27	1	0
4	P	23	0	27	2	0
5	A	1	0	0	0	0
5	E	1	0	0	0	0
5	I	1	0	0	0	0
5	M	1	0	0	0	0
6	A	56	0	0	1	0
6	B	56	0	0	1	0
6	C	61	0	0	1	0
6	D	56	0	0	1	0
6	E	62	0	0	2	0
6	F	58	0	0	1	0
6	G	58	0	0	2	0
6	H	57	0	0	2	0
6	I	56	0	0	2	0
6	J	57	0	0	2	0
6	K	56	0	0	1	0
6	L	59	0	0	1	0
6	M	58	0	0	2	0
6	N	58	0	0	1	0
6	O	62	0	0	1	0
6	P	58	0	0	1	0
All	All	49796	0	47232	417	9

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 4.

The worst 5 of 417 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:F:343:LYS:CE	1:N:465:ALA:HB1	1.36	1.49
1:B:372:THR:HG21	1:I:433:GLY:O	1.14	1.30
1:F:343:LYS:HZ1	1:N:465:ALA:CB	1.43	1.30
1:B:433:GLY:O	1:I:372:THR:CG2	1.80	1.30
1:F:343:LYS:NZ	1:N:465:ALA:CB	1.95	1.27

The worst 5 of 9 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:D:433:GLY:O	1:K:372:THR:CG2[1_545]	1.78	0.42
1:D:372:THR:CG2	1:K:433:GLY:O[1_545]	1.81	0.39
1:A:219:ASN:OD1	1:G:360:LYS:NZ[1_556]	1.90	0.30
1:J:219:ASN:OD1	1:M:360:LYS:NZ[1_665]	1.94	0.26
1:C:219:ASN:OD1	1:E:360:LYS:NZ[1_545]	1.96	0.24

## 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	388/397 (98%)	367 (95%)	21 (5%)	0	100	100
1	B	388/397 (98%)	370 (95%)	18 (5%)	0	100	100
1	C	388/397 (98%)	367 (95%)	21 (5%)	0	100	100
1	D	388/397 (98%)	365 (94%)	23 (6%)	0	100	100
1	E	388/397 (98%)	369 (95%)	18 (5%)	1 (0%)	44	64
1	F	388/397 (98%)	369 (95%)	18 (5%)	1 (0%)	44	64
1	G	388/397 (98%)	368 (95%)	20 (5%)	0	100	100
1	H	388/397 (98%)	368 (95%)	20 (5%)	0	100	100
1	I	388/397 (98%)	371 (96%)	17 (4%)	0	100	100
1	J	388/397 (98%)	368 (95%)	20 (5%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	K	388/397 (98%)	370 (95%)	17 (4%)	1 (0%)	44	64
1	L	388/397 (98%)	372 (96%)	16 (4%)	0	100	100
1	M	388/397 (98%)	369 (95%)	18 (5%)	1 (0%)	44	64
1	N	388/397 (98%)	369 (95%)	19 (5%)	0	100	100
1	O	388/397 (98%)	367 (95%)	21 (5%)	0	100	100
1	P	388/397 (98%)	371 (96%)	17 (4%)	0	100	100
All	All	6208/6352 (98%)	5900 (95%)	304 (5%)	4 (0%)	55	75

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	199	ASN
1	M	199	ASN
1	K	221	ILE
1	F	221	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	323/329 (98%)	303 (94%)	20 (6%)	21	37
1	B	323/329 (98%)	302 (94%)	21 (6%)	20	35
1	C	323/329 (98%)	306 (95%)	17 (5%)	26	45
1	D	323/329 (98%)	302 (94%)	21 (6%)	20	35
1	E	323/329 (98%)	304 (94%)	19 (6%)	23	39
1	F	323/329 (98%)	302 (94%)	21 (6%)	20	35
1	G	323/329 (98%)	303 (94%)	20 (6%)	21	37
1	H	323/329 (98%)	304 (94%)	19 (6%)	23	39
1	I	323/329 (98%)	305 (94%)	18 (6%)	25	42
1	J	323/329 (98%)	302 (94%)	21 (6%)	20	35

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	323/329 (98%)	304 (94%)	19 (6%)	23	39
1	L	323/329 (98%)	304 (94%)	19 (6%)	23	39
1	M	323/329 (98%)	302 (94%)	21 (6%)	20	35
1	N	323/329 (98%)	305 (94%)	18 (6%)	25	42
1	O	323/329 (98%)	303 (94%)	20 (6%)	21	37
1	P	323/329 (98%)	303 (94%)	20 (6%)	21	37
All	All	5168/5264 (98%)	4854 (94%)	314 (6%)	22	38

5 of 314 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	197	GLU
1	I	456	TRP
1	O	416	LYS
1	H	251	CYS
1	I	135	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 77 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	183	HIS
1	J	183	HIS
1	O	439	HIS
1	H	198	ASN
1	I	198	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 ⓘ

There are no carbohydrates in this entry.



## 5.6 Ligand geometry

Of 52 ligands modelled in this entry, 20 are monoatomic - leaving 32 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$
4	BCZ	A	1001	-	19,23,23	1.10	1 (5%)	22,32,32	1.39	4 (18%)
2	NAG	A	900	1	14,14,15	0.61	0	15,19,21	2.18	3 (20%)
4	BCZ	B	1001	-	19,23,23	0.83	1 (5%)	22,32,32	1.44	3 (13%)
2	NAG	B	900	1	14,14,15	0.65	0	15,19,21	2.19	4 (26%)
4	BCZ	C	1001	-	19,23,23	1.30	1 (5%)	22,32,32	1.26	3 (13%)
2	NAG	C	900	1	14,14,15	0.74	0	15,19,21	2.15	2 (13%)
4	BCZ	D	1001	-	19,23,23	1.00	1 (5%)	22,32,32	1.55	4 (18%)
2	NAG	D	900	1	14,14,15	0.92	0	15,19,21	2.04	4 (26%)
4	BCZ	E	1001	-	19,23,23	1.10	2 (10%)	22,32,32	1.52	3 (13%)
2	NAG	E	900	1	14,14,15	0.77	0	15,19,21	1.94	4 (26%)
4	BCZ	F	1001	-	19,23,23	1.20	1 (5%)	22,32,32	1.30	3 (13%)
2	NAG	F	900	1	14,14,15	0.88	1 (7%)	15,19,21	1.97	2 (13%)
4	BCZ	G	1001	-	19,23,23	0.87	1 (5%)	22,32,32	1.48	4 (18%)
2	NAG	G	900	1	14,14,15	0.54	0	15,19,21	2.22	3 (20%)
4	BCZ	H	1001	-	19,23,23	1.40	1 (5%)	22,32,32	1.32	4 (18%)
2	NAG	H	900	1	14,14,15	0.84	0	15,19,21	2.00	3 (20%)
4	BCZ	I	1001	-	19,23,23	1.10	1 (5%)	22,32,32	1.50	3 (13%)
2	NAG	I	900	1	14,14,15	0.72	0	15,19,21	2.34	4 (26%)
4	BCZ	J	1001	-	19,23,23	0.94	1 (5%)	22,32,32	1.46	4 (18%)
2	NAG	J	900	1	14,14,15	0.65	0	15,19,21	2.07	2 (13%)
4	BCZ	K	1001	-	19,23,23	1.19	1 (5%)	22,32,32	1.59	5 (22%)
2	NAG	K	900	1	14,14,15	0.68	0	15,19,21	1.91	4 (26%)
4	BCZ	L	1001	-	19,23,23	1.10	1 (5%)	22,32,32	1.42	3 (13%)
2	NAG	L	900	1	14,14,15	0.57	0	15,19,21	2.14	4 (26%)
4	BCZ	M	1001	-	19,23,23	1.40	2 (10%)	22,32,32	1.17	3 (13%)
2	NAG	M	900	1	14,14,15	0.66	0	15,19,21	2.30	4 (26%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	BCZ	N	1001	-	19,23,23	1.25	1 (5%)	22,32,32	1.29	3 (13%)
2	NAG	N	900	1	14,14,15	0.79	0	15,19,21	2.03	3 (20%)
4	BCZ	O	1001	-	19,23,23	1.02	1 (5%)	22,32,32	1.60	5 (22%)
2	NAG	O	900	1	14,14,15	0.61	0	15,19,21	2.35	3 (20%)
4	BCZ	P	1001	-	19,23,23	0.87	1 (5%)	22,32,32	1.36	3 (13%)
2	NAG	P	900	1	14,14,15	0.71	0	15,19,21	1.90	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
4	BCZ	A	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	A	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	B	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	B	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	C	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	C	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	D	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	D	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	E	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	E	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	F	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	F	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	G	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	G	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	H	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	H	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	I	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	I	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	J	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	J	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	K	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	K	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	L	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	L	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	M	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	M	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	N	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	N	900	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
4	BCZ	O	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	O	900	1	-	0/6/23/26	0/1/1/1
4	BCZ	P	1001	-	-	0/20/40/40	0/1/1/1
2	NAG	P	900	1	-	0/6/23/26	0/1/1/1

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	1001	BCZ	C5-C4	-2.76	1.50	1.53
4	B	1001	BCZ	C1-C5	-2.23	1.51	1.54
2	F	900	NAG	C1-C2	2.06	1.55	1.52
4	G	1001	BCZ	C26-N25	2.12	1.36	1.33
4	D	1001	BCZ	C26-N25	2.18	1.36	1.33

The worst 5 of 109 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	O	1001	BCZ	C15-C13-N11	-3.87	109.12	116.11
4	D	1001	BCZ	C1-C5-C6	-3.67	106.50	115.37
4	I	1001	BCZ	C15-C13-N11	-3.43	109.91	116.11
4	B	1001	BCZ	C1-C5-C6	-3.42	107.12	115.37
4	J	1001	BCZ	C15-C13-N11	-3.37	110.02	116.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

16 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1001	BCZ	1	0
4	B	1001	BCZ	1	0
4	C	1001	BCZ	1	0
4	D	1001	BCZ	1	0
4	E	1001	BCZ	1	0
4	F	1001	BCZ	1	0
4	G	1001	BCZ	2	0
4	H	1001	BCZ	1	0
4	I	1001	BCZ	2	0
4	J	1001	BCZ	1	0
4	K	1001	BCZ	1	0
4	L	1001	BCZ	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	M	1001	BCZ	2	0
4	N	1001	BCZ	1	0
4	O	1001	BCZ	1	0
4	P	1001	BCZ	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	389/397 (97%)	-0.68	0 100 100	14, 22, 30, 36	3 (0%)
1	B	389/397 (97%)	-0.64	1 (0%) 93 94	14, 22, 30, 36	3 (0%)
1	C	389/397 (97%)	-0.65	0 100 100	15, 22, 30, 36	3 (0%)
1	D	389/397 (97%)	-0.69	0 100 100	15, 22, 30, 36	3 (0%)
1	E	389/397 (97%)	-0.62	0 100 100	15, 22, 30, 36	3 (0%)
1	F	389/397 (97%)	-0.50	0 100 100	15, 22, 30, 36	3 (0%)
1	G	389/397 (97%)	-0.61	0 100 100	15, 22, 30, 36	3 (0%)
1	H	389/397 (97%)	-0.62	0 100 100	15, 22, 30, 36	3 (0%)
1	I	389/397 (97%)	-0.67	0 100 100	14, 22, 30, 36	3 (0%)
1	J	389/397 (97%)	-0.68	0 100 100	14, 22, 30, 36	3 (0%)
1	K	389/397 (97%)	-0.65	0 100 100	14, 22, 30, 36	3 (0%)
1	L	389/397 (97%)	-0.68	0 100 100	15, 22, 30, 36	3 (0%)
1	M	389/397 (97%)	-0.61	0 100 100	14, 22, 30, 36	3 (0%)
1	N	389/397 (97%)	-0.60	0 100 100	15, 22, 30, 36	3 (0%)
1	O	389/397 (97%)	-0.64	0 100 100	15, 22, 30, 36	3 (0%)
1	P	389/397 (97%)	-0.61	0 100 100	15, 22, 30, 36	3 (0%)
All	All	6224/6352 (97%)	-0.63	1 (0%) 100 100	14, 22, 30, 36	48 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	359	SER	2.5

### 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
2	NAG	N	900	14/15	0.77	0.22	7.11	33,39,40,41	0
2	NAG	F	900	14/15	0.77	0.25	6.73	33,38,41,41	0
2	NAG	H	900	14/15	0.78	0.23	5.61	33,38,41,41	0
2	NAG	P	900	14/15	0.79	0.20	4.70	33,38,40,41	0
2	NAG	O	900	14/15	0.78	0.20	4.45	34,39,41,41	0
2	NAG	E	900	14/15	0.82	0.19	4.05	33,38,40,41	0
2	NAG	A	900	14/15	0.83	0.18	3.88	33,38,41,42	0
3	CA	I	1000	1/1	0.96	0.18	3.78	53,53,53,53	0
2	NAG	M	900	14/15	0.80	0.19	3.72	33,39,40,41	0
2	NAG	I	900	14/15	0.87	0.20	3.54	34,39,41,41	0
2	NAG	B	900	14/15	0.84	0.18	3.27	33,38,40,41	0
2	NAG	J	900	14/15	0.84	0.17	3.11	33,38,41,41	0
2	NAG	D	900	14/15	0.87	0.20	2.95	33,38,41,41	0
2	NAG	G	900	14/15	0.83	0.16	2.62	33,39,40,41	0
4	BCZ	A	1001	23/23	0.93	0.14	2.18	15,25,27,29	0
2	NAG	C	900	14/15	0.86	0.17	1.85	33,38,40,41	0
4	BCZ	C	1001	23/23	0.92	0.13	1.16	15,25,27,29	0
4	BCZ	F	1001	23/23	0.92	0.16	1.07	15,25,27,29	0
4	BCZ	B	1001	23/23	0.93	0.14	1.07	15,25,27,28	0
4	BCZ	D	1001	23/23	0.94	0.12	1.03	15,25,27,29	0
4	BCZ	P	1001	23/23	0.92	0.14	0.92	15,25,27,29	0
4	BCZ	L	1001	23/23	0.94	0.13	0.92	15,25,27,28	0
4	BCZ	E	1001	23/23	0.93	0.14	0.86	15,25,27,29	0
4	BCZ	M	1001	23/23	0.94	0.13	0.85	15,25,27,29	0
2	NAG	K	900	14/15	0.86	0.17	0.84	33,38,41,42	0
4	BCZ	N	1001	23/23	0.92	0.14	0.80	15,25,27,29	0
4	BCZ	O	1001	23/23	0.94	0.13	0.67	15,25,27,29	0
4	BCZ	I	1001	23/23	0.94	0.12	0.45	15,25,27,29	0
4	BCZ	K	1001	23/23	0.94	0.12	0.43	15,25,27,28	0
3	CA	D	1000	1/1	0.95	0.14	0.34	53,53,53,53	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	BCZ	J	1001	23/23	0.95	0.12	0.32	15,25,27,29	0
4	BCZ	H	1001	23/23	0.93	0.13	0.22	15,25,27,29	0
4	BCZ	G	1001	23/23	0.93	0.13	0.10	15,25,27,29	0
2	NAG	L	900	14/15	0.89	0.14	0.10	33,38,40,41	0
3	CA	A	1000	1/1	0.94	0.09	-0.89	53,53,53,53	0
3	CA	F	1000	1/1	0.91	0.10	-0.99	53,53,53,53	0
3	CA	J	1000	1/1	0.89	0.08	-1.09	53,53,53,53	0
3	CA	O	1000	1/1	0.90	0.10	-1.18	53,53,53,53	0
3	CA	E	1000	1/1	0.95	0.08	-1.29	53,53,53,53	0
3	CA	N	1000	1/1	0.96	0.07	-1.49	53,53,53,53	0
3	CA	C	1000	1/1	0.97	0.08	-1.67	53,53,53,53	0
3	CA	P	1000	1/1	0.91	0.06	-1.74	53,53,53,53	0
3	CA	K	1000	1/1	0.94	0.08	-1.94	52,52,52,52	0
3	CA	M	1000	1/1	0.92	0.07	-2.16	53,53,53,53	0
3	CA	B	1000	1/1	0.98	0.07	-2.25	53,53,53,53	0
3	CA	H	1000	1/1	0.93	0.06	-2.28	53,53,53,53	0
3	CA	G	1000	1/1	0.98	0.06	-2.46	53,53,53,53	0
3	CA	L	1000	1/1	0.96	0.06	-2.89	53,53,53,53	0
5	YT3	A	1	1/1	0.99	0.07	-	28,28,28,28	0
5	YT3	I	3	1/1	0.99	0.07	-	29,29,29,29	0
5	YT3	M	4	1/1	0.99	0.05	-	30,30,30,30	0
5	YT3	E	2	1/1	0.99	0.04	-	28,28,28,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.