W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Jul 6, 2017 – 08:21 PM EDT

PDB ID	:	1NCC
Title	:	CRYSTAL STRUCTURES OF TWO MUTANT NEURAMINIDASE-AN TIBODY COMPLEXES WITH AMINO ACID SUBSTITUTIONS IN THE INTERFACE
Authors	:	Tulip, W.R.; Varghese, J.N.; Colman, P.M.
Deposited on	:	unknown
Resolution	:	2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	Percentile Ranks	Value
Clashscore		14
Ramachandran outliers		2.2%
Sidechain outliers		14.6%
RSRZ outliers		1.1%

Worse .... *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
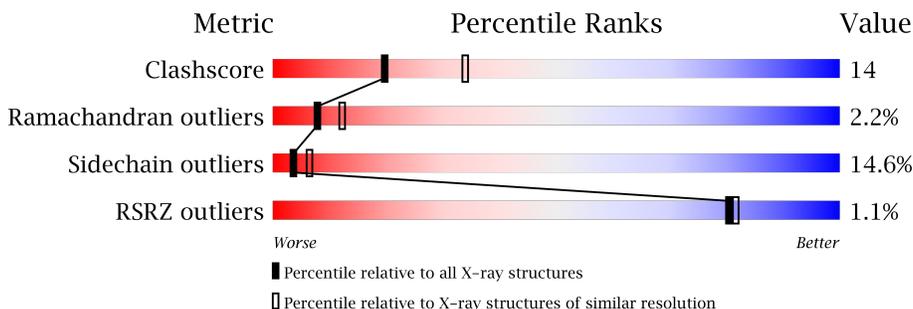
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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 >=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 <=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	N	389	53% 35% 11% •
2	L	214	.% / 52% 36% 11% •
3	H	221	3% / 55% 39% 5% •

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
7	CA	N	1	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6508 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 INFLUENZA A SUBTYPE N9 NEURAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	N	389	3078	1920	542	593	23	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	368	ARG	ILE	CONFLICT	UNP P03472

- Molecule 2 is a protein called IGG2A-KAPPA NC41 FAB (LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	214	1667	1043	280	336	8	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	20	THR	SER	CONFLICT	EMBL Y11589
L	21	ILE	VAL	CONFLICT	EMBL Y11589
L	28	ASP	ILE	CONFLICT	EMBL Y11589
L	30	SER	GLY	CONFLICT	EMBL Y11589
L	32	ALA	ASN	CONFLICT	EMBL Y11589
L	34	VAL	ALA	CONFLICT	EMBL Y11589
L	46	LEU	ALA	CONFLICT	EMBL Y11589
L	50	TRP	SER	CONFLICT	EMBL Y11589
L	53	THR	TYR	CONFLICT	EMBL Y11589
L	55	HIS	TYR	CONFLICT	EMBL Y11589
L	56	ILE	SER	CONFLICT	EMBL Y11589
L	63	ALA	THR	CONFLICT	EMBL Y11589
L	71	TYR	PHE	CONFLICT	EMBL Y11589
L	77	SER	ASN	CONFLICT	EMBL Y11589
L	80	ALA	SER	CONFLICT	EMBL Y11589

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	85	LEU	GLU	CONFLICT	EMBL Y11589
L	87	TYR	PHE	CONFLICT	EMBL Y11589
L	91	HIS	TYR	CONFLICT	EMBL Y11589
L	92	TYR	ASN	CONFLICT	EMBL Y11589
L	93	SER	ARG	CONFLICT	EMBL Y11589
L	94	PRO	TYR	CONFLICT	EMBL Y11589

- Molecule 3 is a protein called IGG2A-KAPPA NC41 FAB (HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	H	221	1662	1048	273	334	7	0	0	0

- Molecule 4 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
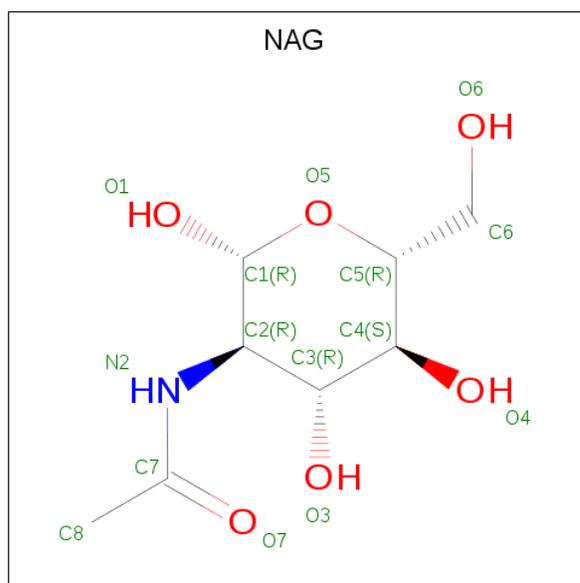
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	N	1	14	8	1	5	0	0
4	N	1	14	8	1	5	0	0
4	N	1	14	8	1	5	0	0
4	N	1	14	8	1	5	0	0

- Molecule 5 is BETA-D-MANNOSE (three-letter code: BMA) (formula: $C_6H_{12}O_6$).
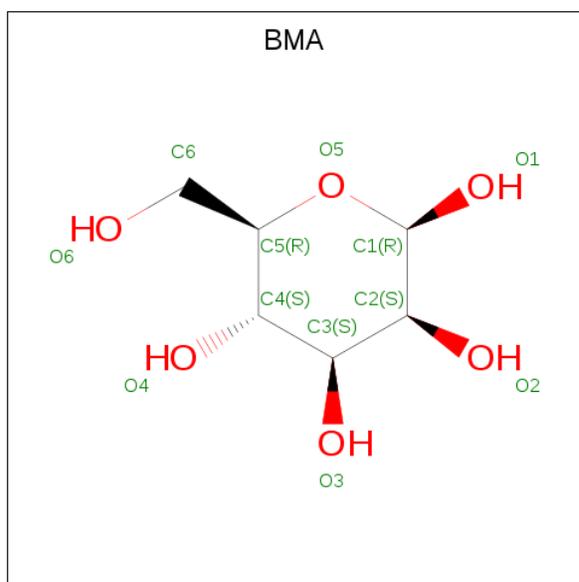
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	N	1	Total 11	C 6	O 5	0	0

- Molecule 6 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: $C_6H_{12}O_6$).
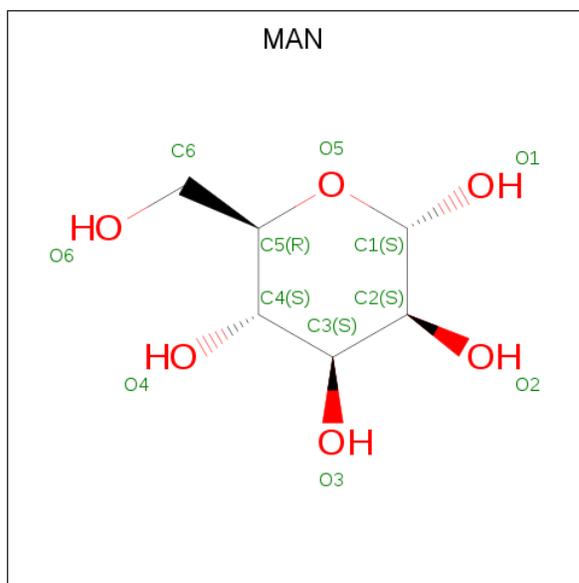
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	N	1	Total 11	C 6	O 5	0	0
6	N	1	Total 11	C 6	O 5	0	0
6	N	1	Total 11	C 6	O 5	0	0

• Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	N	1	Total	Ca	0	0
			1	1		

3 Residue-property plots (i)

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

- Molecule 1: INFLUENZA A SUBTYPE N9 NEURAMINIDASE
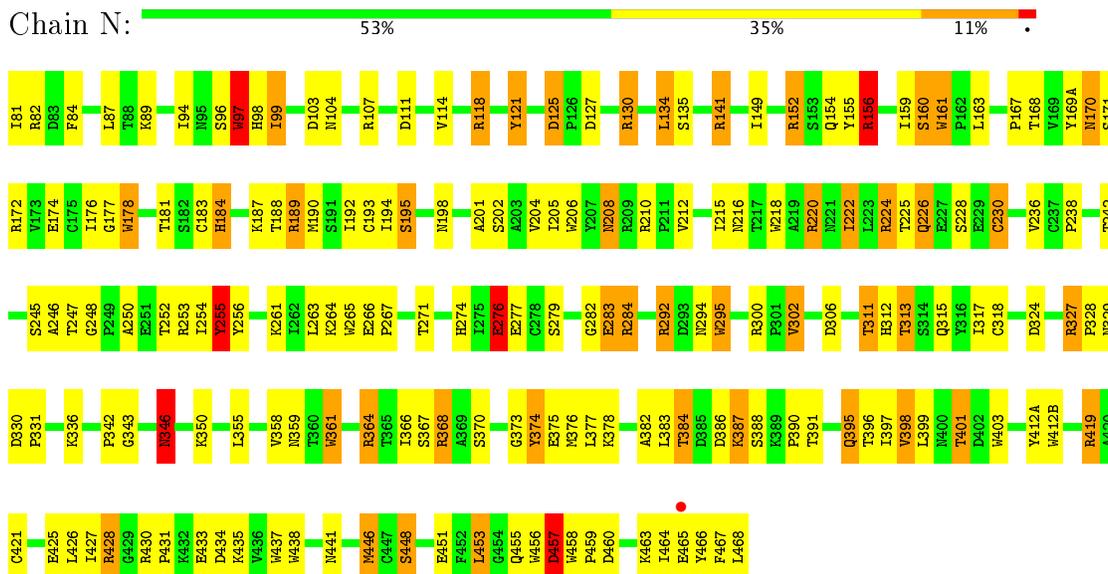
Chain N:

- Molecule 2: IGG2A-KAPPA NC41 FAB (LIGHT CHAIN)
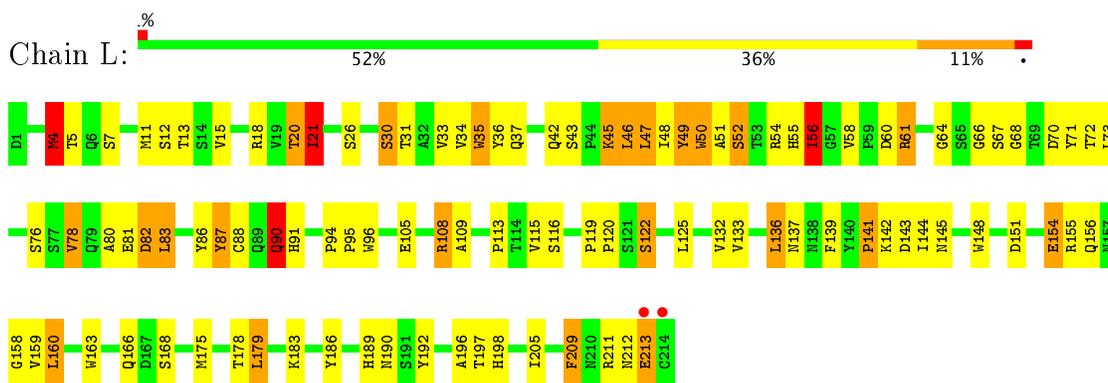
Chain L:

- Molecule 3: IGG2A-KAPPA NC41 FAB (HEAVY CHAIN)

Chain H:
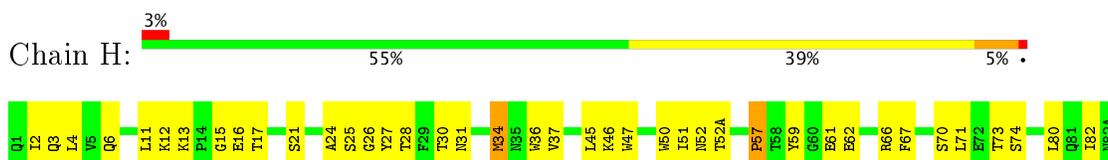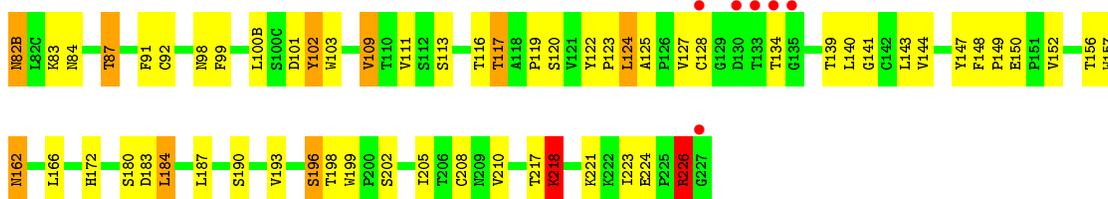
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α, β, γ	167.00Å 167.00Å 124.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 − 2.50 8.00 − 2.50	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-2.50) 49.7 (8.00-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$	-	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.212 , (Not available) 0.224 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	16.1	Xtriage
Anisotropy	0.723	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.19 , 48.5	EDS
L-test for twinning[1]	$<|L|> = 0.31, <L^2> = 0.15$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	6508	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Theoretical values of $<|L|>$, $<L^2>$ 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: CA, BMA, NAG, 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	N	1.16	6/3161 (0.2%)	1.99	113/4304 (2.6%)
2	L	1.00	0/1708	1.88	42/2323 (1.8%)
3	H	1.02	0/1704	1.91	45/2323 (1.9%)
All	All	1.09	6/6573 (0.1%)	1.94	200/8950 (2.2%)

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	N	0	6
2	L	0	3
All	All	0	9

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	N	97	TRP	CG-CD2	-6.30	1.32	1.43
1	N	327	ARG	CZ-NH2	5.77	1.40	1.33
1	N	284	ARG	CZ-NH2	5.49	1.40	1.33
1	N	456	TRP	NE1-CE2	-5.23	1.30	1.37
1	N	283	GLU	CD-OE2	-5.08	1.20	1.25

The worst 5 of 200 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	N	284	ARG	CD-NE-CZ	-11.46	107.56	123.60
3	H	36	TRP	CD1-CG-CD2	10.48	114.68	106.30
1	N	130	ARG	NE-CZ-NH2	-10.33	115.14	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	N	218	TRP	CD1-CG-CD2	10.24	114.49	106.30
3	H	34	MET	CG-SD-CE	-10.18	83.92	100.20

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	N	155	TYR	Sidechain
1	N	248	GLY	Peptide
1	N	255	TYR	Sidechain
1	N	284	ARG	Sidechain
1	N	327	ARG	Sidechain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	N	3078	0	2905	101	0
2	L	1667	0	1598	43	0
3	H	1662	0	1611	46	0
4	N	56	0	50	1	0
5	N	11	0	9	0	0
6	N	33	0	28	1	0
7	N	1	0	0	0	0
All	All	6508	0	6201	182	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 14.

The worst 5 of 182 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:N:426:LEU:HD23	1:N:460:ASP:HB2	1.40	1.03
2:L:145:ASN:HB3	2:L:197:THR:HB	1.53	0.91
1:N:426:LEU:HD23	1:N:460:ASP:CB	2.00	0.91

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:N:188:THR:HG21	1:N:208:ASN:HB2	1.60	0.82
1:N:426:LEU:HD23	1:N:460:ASP:CA	2.10	0.81

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	N	387/389 (100%)	336 (87%)	48 (12%)	3 (1%)	22	39
2	L	212/214 (99%)	187 (88%)	19 (9%)	6 (3%)	6	8
3	H	219/221 (99%)	180 (82%)	30 (14%)	9 (4%)	3	4
All	All	818/824 (99%)	703 (86%)	97 (12%)	18 (2%)	8	12

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	154	GLU
2	L	158	GLY
3	H	16	GLU
3	H	87	THR
3	H	162	ASN

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	N	342/342 (100%)	300 (88%)	42 (12%)	5	10
2	L	190/190 (100%)	153 (80%)	37 (20%)	1	3
3	H	187/187 (100%)	161 (86%)	26 (14%)	4	7
All	All	719/719 (100%)	614 (85%)	105 (15%)	3	6

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	20	THR
2	L	58	VAL
3	H	190	SER
2	L	21	ILE
2	L	43	SER

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

Mol	Chain	Res	Type
1	N	392	GLN
1	N	395	GLN
3	H	6	GLN
1	N	359	ASN
2	L	161	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	NAG	N	469(A)	1,4	14,14,15	0.59	0	15,19,21	1.73	5 (33%)
4	NAG	N	470(B)	5,4	14,14,15	0.80	0	15,19,21	2.08	3 (20%)
5	BMA	N	471(C)	4,6	11,11,12	1.79	2 (18%)	13,15,17	1.56	3 (23%)
6	MAN	N	472(D)	5,6	11,11,12	1.16	1 (9%)	13,15,17	3.51	8 (61%)
6	MAN	N	473(E)	6	11,11,12	1.19	1 (9%)	13,15,17	3.21	6 (46%)
6	MAN	N	474(F)	6	11,11,12	1.63	4 (36%)	13,15,17	2.11	1 (7%)
4	NAG	N	475(A)	1	14,14,15	1.41	2 (14%)	15,19,21	1.78	5 (33%)
4	NAG	N	476(A)	1	14,14,15	0.88	1 (7%)	15,19,21	1.69	4 (26%)

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	N	469(A)	1,4	-	0/6/23/26	0/1/1/1
4	NAG	N	470(B)	5,4	-	0/6/23/26	0/1/1/1
5	BMA	N	471(C)	4,6	-	0/2/19/22	0/1/1/1
6	MAN	N	472(D)	5,6	-	0/2/19/22	0/1/1/1
6	MAN	N	473(E)	6	-	0/2/19/22	1/1/1/1
6	MAN	N	474(F)	6	-	0/2/19/22	0/1/1/1
4	NAG	N	475(A)	1	-	0/6/23/26	0/1/1/1
4	NAG	N	476(A)	1	-	0/6/23/26	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	N	475(A)	NAG	C2-N2	-3.73	1.39	1.46
6	N	474(F)	MAN	O5-C1	2.01	1.47	1.43
4	N	476(A)	NAG	C4-C3	2.02	1.57	1.52
6	N	472(D)	MAN	C6-C5	2.09	1.59	1.51
4	N	475(A)	NAG	O5-C1	2.17	1.47	1.43

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	N	472(D)	MAN	O2-C2-C3	-4.29	101.76	110.17
4	N	469(A)	NAG	C1-C2-N2	-3.78	104.03	110.49
4	N	470(B)	NAG	O7-C7-C8	-3.72	115.28	122.06
4	N	476(A)	NAG	C4-C3-C2	-3.54	105.83	111.02
6	N	473(E)	MAN	C1-C2-C3	-3.01	105.83	109.65

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	N	473(E)	MAN	C1-C2-C3-C4-C5-O5

3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	N	472(D)	MAN	1	0
6	N	473(E)	MAN	1	0
4	N	475(A)	NAG	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB($Å^2$)	Q<0.9
1	N	389/389 (100%)	-1.17	1 (0%)	93	94	2, 3, 18, 39	0
2	L	214/214 (100%)	-0.77	2 (0%)	84	85	2, 19, 38, 68	0
3	H	221/221 (100%)	-0.58	6 (2%)	55	58	2, 24, 44, 67	0
All	All	824/824 (100%)	-0.91	9 (1%)	80	81	2, 12, 37, 68	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	H	135	GLY	5.8
3	H	130	ASP	5.3
3	H	134	THR	4.5
2	L	214	CYS	4.2
2	L	213	GLU	3.5

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, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å²)	Q<0.9
7	CA	N	1	1/1	0.95	0.17	9.75	44,44,44,44	0
6	MAN	N	472(D)	11/12	0.96	0.10	1.19	6,7,9,11	0
4	NAG	N	469(A)	14/15	0.97	0.10	0.98	8,11,16,17	0
5	BMA	N	471(C)	11/12	0.97	0.09	0.61	5,7,8,9	0
6	MAN	N	474(F)	11/12	0.97	0.10	0.35	13,14,15,16	0
4	NAG	N	475(A)	14/15	0.97	0.10	0.30	19,22,24,25	0
4	NAG	N	470(B)	14/15	0.97	0.09	-	8,11,12,13	0
6	MAN	N	473(E)	11/12	0.96	0.10	-	10,13,15,17	0
4	NAG	N	476(A)	14/15	0.93	0.13	-	20,22,26,27	0

6.5 Other polymers (i)

There are no such residues in this entry.