



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:01 AM GMT

PDB ID : 3D12  
Title : Crystal Structures of Nipah Virus G Attachment Glycoprotein in Complex with its Receptor Ephrin-B3  
Authors : Xu, K.; Rajashankar, K.R.; Chan, Y.P.; Himanen, P.; Broder, C.C.; Nikolov, D.B.  
Deposited on : 2008-05-02  
Resolution : 3.00 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

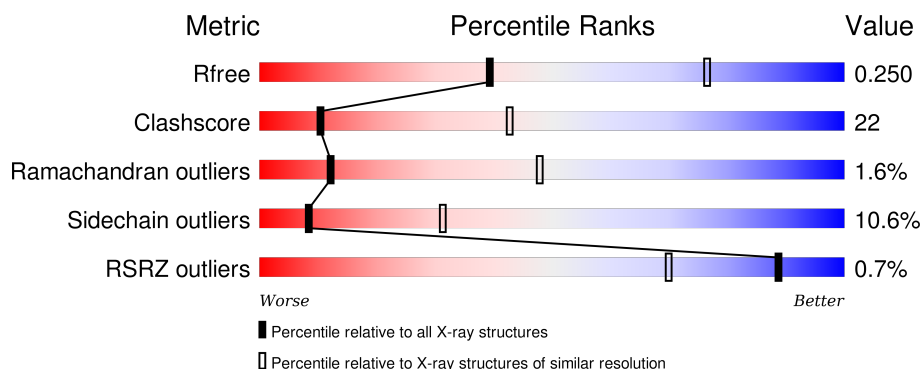
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	428	<div> <div>%</div> <div> <div></div> <div>57%</div> <div>36%</div> <div>6%</div> </div> </div>
1	D	428	<div> <div>58%</div> <div>35%</div> <div>7%</div> </div>
2	B	141	<div> <div>%</div> <div> <div></div> <div>57%</div> <div>37%</div> <div>6%</div> </div> </div>
2	E	141	<div> <div>53%</div> <div>40%</div> <div>6%</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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NGA	A	1362	X	-	-	-
4	GL0	A	1363	X	-	-	-
4	GLC	A	1365	X	-	-	-
4	BGC	A	1366	X	-	-	-
4	GLC	A	1367	X	-	-	-
5	NAG	D	1310	X	-	-	-
6	NAG	A	1313	X	-	-	-
6	NAG	A	1314	X	-	-	-
6	NAG	D	1413	X	-	-	-
6	NAG	D	1414	X	-	-	-
7	SO4	A	1368	-	-	X	-
7	SO4	A	1375	-	-	-	X
7	SO4	B	202	-	-	-	X
7	SO4	D	1474	-	-	-	X
9	NGZ	D	1462	X	-	-	-
9	GL0	D	1463	X	-	-	-
9	GLC	D	1465	X	-	-	-
9	BGC	D	1466	X	-	-	-
9	GXL	D	1467	X	-	-	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 9598 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin-neuraminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	428	Total	C	N	O	S	0	0	0
			3370	2146	568	635	21			
1	D	428	Total	C	N	O	S	0	0	0
			3370	2146	568	635	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	603	ALA	-	EXPRESSION TAG	UNP Q9IH62
D	603	ALA	-	EXPRESSION TAG	UNP Q9IH62

- Molecule 2 is a protein called Ephrin-B3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	141	Total	C	N	O	S	0	0	0
			1132	720	201	206	5			
2	E	141	Total	C	N	O	S	0	0	0
			1132	720	201	206	5			

- Molecule 3 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	5	Total	C	N	O	0	0
			61	34	2	25		

- 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		

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

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

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

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	B	1	Total	O	S	0	0
			5	4	1		
7	B	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	E	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	D	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 9 is a polymer of unknown type called SUGAR (7-MER).

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

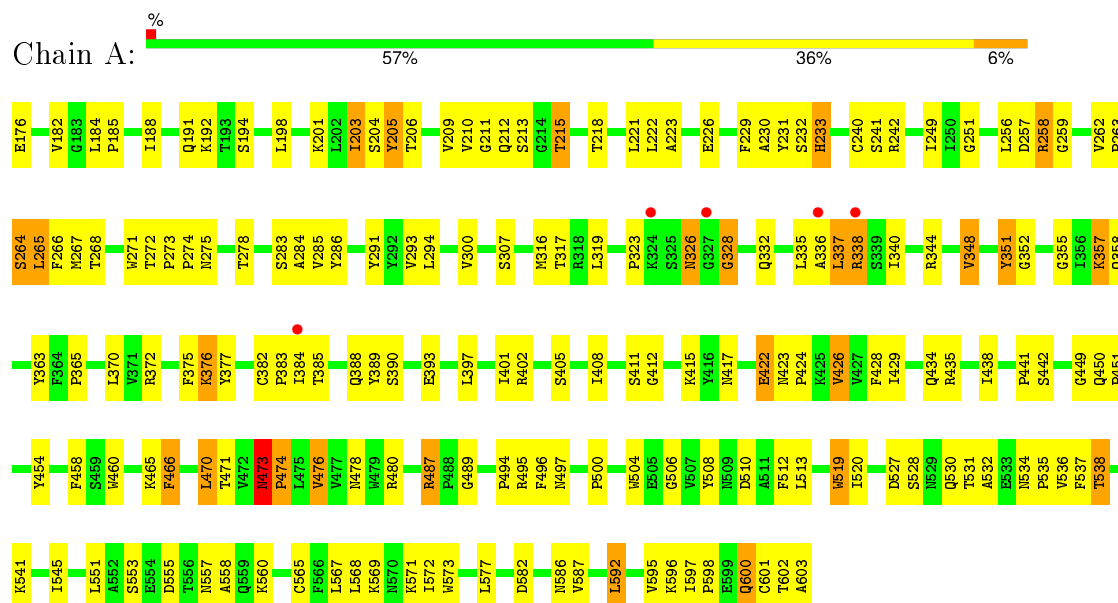
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	46	Total	O	0	0
			46	46		
10	B	9	Total	O	0	0
			9	9		
10	D	44	Total	O	0	0
			44	44		
10	E	13	Total	O	0	0
			13	13		

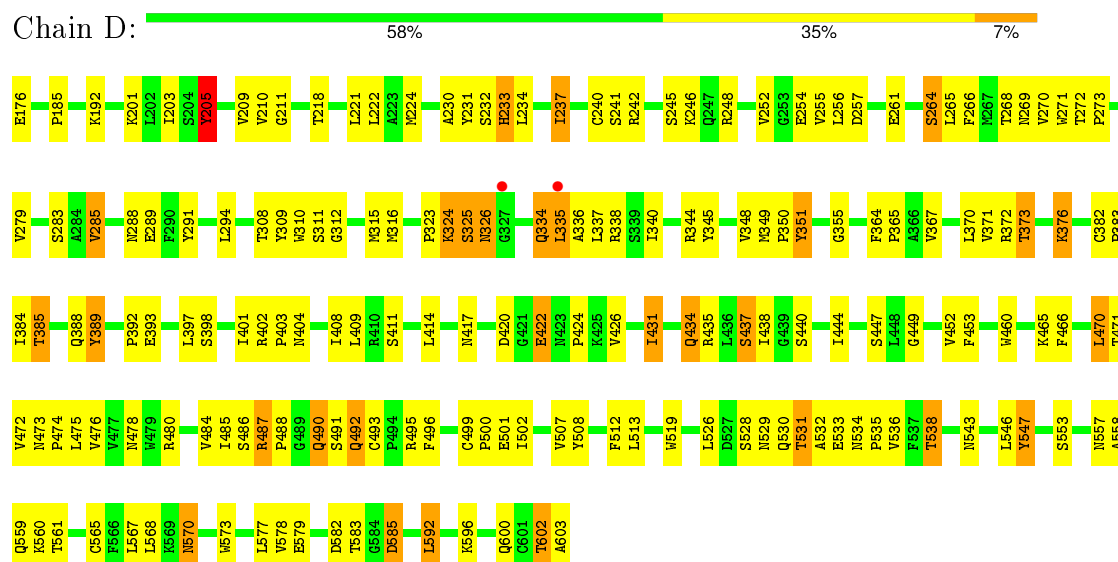
### 3 Residue-property plots

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

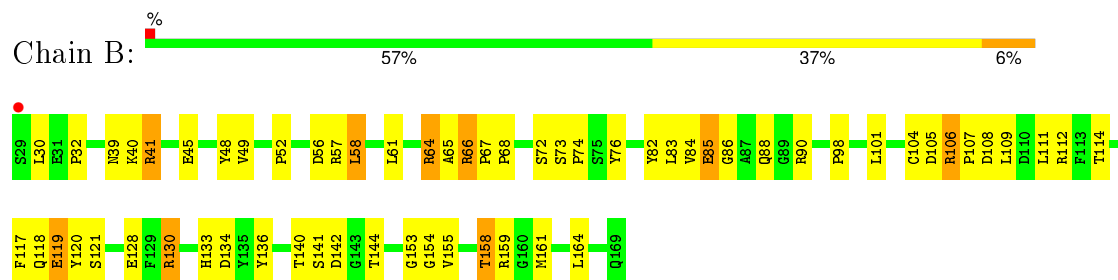


#### • Molecule 1: Hemagglutinin-neuraminidase

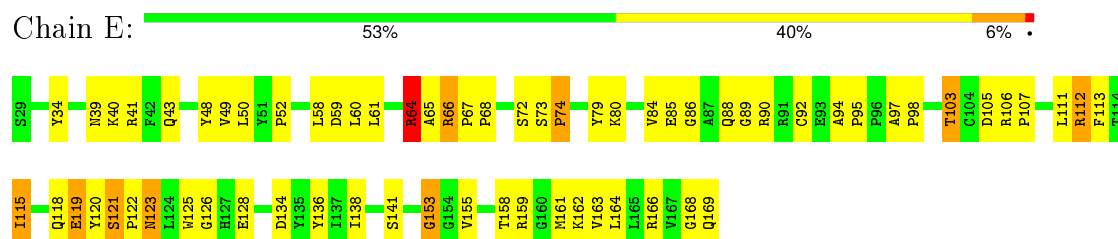




- Molecule 2: Ephrin-B3



- Molecule 2: Ephrin-B3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	189.49Å 189.49Å 277.05Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.34 – 3.00 39.34 – 3.00	Depositor EDS
% Data completeness (in resolution range)	91.3 (39.34-3.00) 91.4 (39.34-3.00)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.57 (at 3.01Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.221 , 0.265 0.206 , 0.250	Depositor DCC
$R_{free}$ test set	2911 reflections (5.69%)	DCC
Wilson B-factor (Å <sup>2</sup> )	62.2	Xtriage
Anisotropy	0.466	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 58204 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	9598	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: BGC, NAG, GXL, GLC, LXZ, NGA, BMA, SO4, GL0, MAN, NGZ, LXB

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	0.46	1/3451 (0.0%)	0.54	0/4694
1	D	0.45	1/3451 (0.0%)	0.54	0/4694
2	B	0.37	0/1165	0.54	0/1582
2	E	0.39	0/1165	0.54	0/1582
All	All	0.44	2/9232 (0.0%)	0.54	0/12552

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	D	0	1
4	A	17	0
6	A	2	0
6	D	3	0
9	D	9	0
All	All	31	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	176	GLU	CD-OE2	7.35	1.33	1.25
1	D	176	GLU	CD-OE2	7.19	1.33	1.25

There are no bond angle outliers.

5 of 31 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	1366	BGC	C5
4	A	1362	NGA	C2,C5,C3,C1,C4
4	A	1363	GL0	C1
4	A	1365	GLC	C2,C5,C3,C1,C4
4	A	1367	GLC	C2,C5,C3,C1,C4

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	473	ASN	Peptide
1	D	473	ASN	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	3370	0	3313	145	0
1	D	3370	0	3313	143	0
2	B	1132	0	1094	53	0
2	E	1132	0	1094	66	0
3	A	61	0	52	1	0
4	A	83	0	70	1	0
5	A	14	0	13	0	0
5	D	14	0	13	0	0
6	A	28	0	25	2	0
6	D	28	0	25	1	0
7	A	45	0	0	4	0
7	B	10	0	0	1	0
7	D	50	0	0	0	0
7	E	5	0	0	1	0
8	D	61	0	52	3	0
9	D	83	0	70	5	0
10	A	46	0	0	2	0
10	B	9	0	0	0	0
10	D	44	0	0	2	0
10	E	13	0	0	1	0
All	All	9598	0	9134	401	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 22.

The worst 5 of 401 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:A:376:LYS:H	1:A:376:LYS:HD3	1.26	0.97
1:D:403:PRO:HG3	1:D:502:ILE:HD11	1.44	0.97
1:A:376:LYS:N	1:A:376:LYS:HD3	1.81	0.94
1:D:531:THR:HG21	1:D:533:GLU:OE2	1.71	0.90
2:B:86:GLY:HA2	2:B:136:TYR:HD1	1.37	0.90

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	426/428 (100%)	382 (90%)	40 (9%)	4 (1%)	21	64
1	D	426/428 (100%)	372 (87%)	50 (12%)	4 (1%)	21	64
2	B	139/141 (99%)	120 (86%)	13 (9%)	6 (4%)	3	19
2	E	139/141 (99%)	118 (85%)	17 (12%)	4 (3%)	6	29
All	All	1130/1138 (99%)	992 (88%)	120 (11%)	18 (2%)	12	48

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	64	ARG
2	E	64	ARG
1	A	338	ARG
2	B	45	GLU
2	B	153	GLY

### 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	382/382 (100%)	344 (90%)	38 (10%)	10	35
1	D	382/382 (100%)	335 (88%)	47 (12%)	6	25
2	B	121/121 (100%)	111 (92%)	10 (8%)	14	46
2	E	121/121 (100%)	109 (90%)	12 (10%)	10	35
All	All	1006/1006 (100%)	899 (89%)	107 (11%)	8	31

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	201	LYS
1	D	308	THR
2	E	64	ARG
1	D	205	TYR
1	D	264	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	326	ASN
1	D	334	GLN
2	E	123	ASN
1	D	332	GLN
1	D	423	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 ⓘ

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
3	NAG	A	1138	1,3	14,14,15	1.82	4 (28%)	15,19,21	2.16	4 (26%)
3	NAG	A	1139	3	14,14,15	2.47	6 (42%)	15,19,21	1.36	2 (13%)
3	BMA	A	1140	3	11,11,12	1.96	4 (36%)	14,15,17	1.66	3 (21%)
3	BMA	A	1141	3	11,11,12	1.62	3 (27%)	14,15,17	1.78	3 (21%)
3	MAN	A	1142	3	11,11,12	1.62	3 (27%)	14,15,17	1.75	3 (21%)
6	NAG	A	1313	1,6	14,14,15	1.85	4 (28%)	15,19,21	2.13	3 (20%)
6	NAG	A	1314	6	14,14,15	2.16	4 (28%)	15,19,21	1.60	4 (26%)
4	LXZ	A	1361	1,4	14,14,15	2.28	4 (28%)	15,19,21	1.63	3 (20%)
4	NGA	A	1362	4	14,14,15	2.24	6 (42%)	15,19,21	1.85	3 (20%)
4	GL0	A	1363	4	11,11,12	2.36	4 (36%)	14,15,17	2.24	5 (35%)
4	BGC	A	1364	4	11,11,12	1.60	3 (27%)	14,15,17	1.44	2 (14%)
4	GLC	A	1365	4	11,11,12	1.62	3 (27%)	14,15,17	1.32	3 (21%)
4	BGC	A	1366	4	11,11,12	1.30	2 (18%)	14,15,17	2.36	4 (28%)
4	GLC	A	1367	4	11,11,12	1.49	3 (27%)	14,15,17	1.58	2 (14%)
8	NAG	D	1238	1,8	14,14,15	1.92	6 (42%)	15,19,21	1.47	2 (13%)
8	NAG	D	1239	8	14,14,15	2.40	7 (50%)	15,19,21	1.55	3 (20%)
8	BMA	D	1240	8	11,11,12	2.11	5 (45%)	14,15,17	1.58	4 (28%)
8	BGC	D	1241	8	11,11,12	1.84	3 (27%)	14,15,17	2.59	6 (42%)
8	MAN	D	1242	8	11,11,12	1.75	3 (27%)	14,15,17	1.22	2 (14%)
6	NAG	D	1413	1,6	14,14,15	1.88	4 (28%)	15,19,21	1.73	3 (20%)
6	NAG	D	1414	6	14,14,15	2.25	4 (28%)	15,19,21	1.47	2 (13%)
9	LXB	D	1461	9,1	14,14,15	2.26	3 (21%)	15,19,21	1.57	4 (26%)
9	NGZ	D	1462	9	14,14,15	2.37	7 (50%)	15,19,21	1.61	2 (13%)
9	GL0	D	1463	9	11,11,12	2.36	4 (36%)	14,15,17	2.68	6 (42%)
9	BGC	D	1464	9	11,11,12	1.53	2 (18%)	14,15,17	1.29	2 (14%)
9	GLC	D	1465	9	11,11,12	1.56	1 (9%)	14,15,17	1.63	2 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	BGC	D	1466	9	11,11,12	1.29	1 (9%)	14,15,17	2.57	4 (28%)
9	GXL	D	1467	9	11,11,12	1.50	3 (27%)	14,15,17	1.11	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	1138	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1139	3	-	0/6/23/26	0/1/1/1
3	BMA	A	1140	3	-	0/2/19/22	0/1/1/1
3	BMA	A	1141	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1142	3	-	0/2/19/22	0/1/1/1
6	NAG	A	1313	1,6	1/1/5/7	0/6/23/26	0/1/1/1
6	NAG	A	1314	6	1/1/5/7	0/6/23/26	0/1/1/1
4	LXZ	A	1361	1,4	-	0/6/23/26	0/1/1/1
4	NGA	A	1362	4	5/5/5/7	0/6/23/26	0/1/1/1
4	GL0	A	1363	4	1/1/4/5	0/2/19/22	0/1/1/1
4	BGC	A	1364	4	-	0/2/19/22	0/1/1/1
4	GLC	A	1365	4	5/5/4/5	0/2/19/22	0/1/1/1
4	BGC	A	1366	4	1/1/4/5	0/2/19/22	0/1/1/1
4	GLC	A	1367	4	5/5/4/5	0/2/19/22	0/1/1/1
8	NAG	D	1238	1,8	-	0/6/23/26	0/1/1/1
8	NAG	D	1239	8	-	0/6/23/26	0/1/1/1
8	BMA	D	1240	8	-	0/2/19/22	0/1/1/1
8	BGC	D	1241	8	-	0/2/19/22	0/1/1/1
8	MAN	D	1242	8	-	0/2/19/22	1/1/1/1
6	NAG	D	1413	1,6	2/2/5/7	0/6/23/26	0/1/1/1
6	NAG	D	1414	6	1/1/5/7	0/6/23/26	0/1/1/1
9	LXB	D	1461	9,1	-	0/6/23/26	0/1/1/1
9	NGZ	D	1462	9	1/1/5/7	0/6/23/26	0/1/1/1
9	GL0	D	1463	9	1/1/4/5	0/2/19/22	0/1/1/1
9	BGC	D	1464	9	-	0/2/19/22	0/1/1/1
9	GLC	D	1465	9	5/5/4/5	0/2/19/22	0/1/1/1
9	BGC	D	1466	9	1/1/4/5	0/2/19/22	0/1/1/1
9	GXL	D	1467	9	1/1/4/5	0/2/19/22	0/1/1/1

The worst 5 of 106 bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	1414	NAG	C8-C7	2.02	1.54	1.50
4	A	1365	GLC	C1-C2	2.03	1.57	1.52
4	A	1366	BGC	C1-C2	2.03	1.57	1.52
6	A	1314	NAG	C8-C7	2.04	1.54	1.50
8	D	1239	NAG	C3-C2	2.05	1.57	1.52

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1138	NAG	C2-N2-C7	-4.43	117.35	123.04
8	D	1238	NAG	C2-N2-C7	-3.38	118.69	123.04
3	A	1138	NAG	C6-C5-C4	-3.36	104.72	113.02
6	A	1313	NAG	C2-N2-C7	-3.35	118.73	123.04
8	D	1239	NAG	C2-N2-C7	-3.03	119.14	123.04

5 of 31 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
9	D	1465	GLC	C2
9	D	1465	GLC	C5
9	D	1465	GLC	C3
9	D	1465	GLC	C1
9	D	1465	GLC	C4

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	D	1242	MAN	C1-C2-C3-C4-C5-O5

13 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1140	BMA	1	0
3	A	1141	BMA	1	0
6	A	1313	NAG	1	0
6	A	1314	NAG	1	0
4	A	1362	NGA	1	0
4	A	1367	GLC	1	0
8	D	1238	NAG	1	0
8	D	1240	BMA	2	0
8	D	1241	BGC	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	1414	NAG	1	0
9	D	1461	LXB	4	0
9	D	1463	GL0	1	0
9	D	1464	BGC	1	0

## 5.6 Ligand geometry

24 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	NAG	A	1210	1	14,14,15	1.63	3 (21%)	15,19,21	1.56	3 (20%)
7	SO4	A	1368	-	4,4,4	0.26	0	6,6,6	0.32	0
7	SO4	A	1369	-	4,4,4	0.25	0	6,6,6	0.27	0
7	SO4	A	1370	-	4,4,4	0.10	0	6,6,6	0.18	0
7	SO4	A	1371	-	4,4,4	0.04	0	6,6,6	0.10	0
7	SO4	A	1372	-	4,4,4	0.17	0	6,6,6	0.09	0
7	SO4	A	1373	-	4,4,4	0.12	0	6,6,6	0.19	0
7	SO4	A	1374	-	4,4,4	0.09	0	6,6,6	0.20	0
7	SO4	A	1375	-	4,4,4	0.25	0	6,6,6	0.40	0
7	SO4	A	1376	-	4,4,4	0.08	0	6,6,6	0.20	0
7	SO4	B	202	-	4,4,4	0.06	0	6,6,6	0.11	0
7	SO4	B	203	-	4,4,4	0.06	0	6,6,6	0.13	0
5	NAG	D	1310	1	14,14,15	1.66	2 (14%)	15,19,21	1.07	1 (6%)
7	SO4	D	1468	-	4,4,4	0.12	0	6,6,6	0.19	0
7	SO4	D	1469	-	4,4,4	0.12	0	6,6,6	0.08	0
7	SO4	D	1470	-	4,4,4	0.08	0	6,6,6	0.09	0
7	SO4	D	1471	-	4,4,4	0.11	0	6,6,6	0.12	0
7	SO4	D	1472	-	4,4,4	0.09	0	6,6,6	0.14	0
7	SO4	D	1473	-	4,4,4	0.07	0	6,6,6	0.12	0
7	SO4	D	1474	-	4,4,4	0.21	0	6,6,6	0.40	0
7	SO4	D	1475	-	4,4,4	0.09	0	6,6,6	0.10	0
7	SO4	D	1476	-	4,4,4	0.07	0	6,6,6	0.22	0
7	SO4	D	1477	-	4,4,4	0.03	0	6,6,6	0.17	0
7	SO4	E	190	-	4,4,4	0.07	0	6,6,6	0.23	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
5	NAG	A	1210	1	-	0/6/23/26	0/1/1/1
7	SO4	A	1368	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1369	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1370	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1371	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1372	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1373	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1374	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1375	-	-	0/0/0/0	0/0/0/0
7	SO4	A	1376	-	-	0/0/0/0	0/0/0/0
7	SO4	B	202	-	-	0/0/0/0	0/0/0/0
7	SO4	B	203	-	-	0/0/0/0	0/0/0/0
5	NAG	D	1310	1	1/1/5/7	0/6/23/26	0/1/1/1
7	SO4	D	1468	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1469	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1470	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1471	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1472	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1473	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1474	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1475	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1476	-	-	0/0/0/0	0/0/0/0
7	SO4	D	1477	-	-	0/0/0/0	0/0/0/0
7	SO4	E	190	-	-	0/0/0/0	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1210	NAG	C8-C7	2.14	1.54	1.50
5	A	1210	NAG	C1-C2	3.15	1.56	1.52
5	D	1310	NAG	C1-C2	3.29	1.57	1.52
5	D	1310	NAG	C7-N2	3.34	1.47	1.34
5	A	1210	NAG	C7-N2	3.48	1.47	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1210	NAG	C6-C5-C4	-2.13	107.76	113.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	1310	NAG	C1-O5-C5	2.24	115.09	112.25
5	A	1210	NAG	C8-C7-N2	3.01	121.87	116.11
5	A	1210	NAG	C1-O5-C5	3.59	116.80	112.25

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	D	1310	NAG	C1

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1368	SO4	3	0
7	A	1371	SO4	1	0
7	B	203	SO4	1	0
7	E	190	SO4	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, 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	428/428 (100%)	-0.24	5 (1%) 81 55	43, 59, 85, 110	0
1	D	428/428 (100%)	-0.28	2 (0%) 91 76	45, 61, 86, 104	0
2	B	141/141 (100%)	-0.19	1 (0%) 89 70	45, 72, 94, 106	0
2	E	141/141 (100%)	-0.28	0 100 100	48, 70, 97, 99	0
All	All	1138/1138 (100%)	-0.25	8 (0%) 89 70	43, 63, 90, 110	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	336	ALA	3.3
1	A	324	LYS	2.6
1	A	338	ARG	2.4
1	A	327	GLY	2.2
1	A	384	ILE	2.1

### 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( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	1138	14/15	0.94	0.15	-0.93	55,59,65,67	0
3	NAG	A	1139	14/15	0.96	0.14	-	60,67,72,73	0
9	GLC	D	1465	11/12	0.88	0.41	-	98,107,112,113	0
3	MAN	A	1142	11/12	0.84	0.33	-	87,91,96,96	0
9	BGC	D	1466	11/12	0.93	0.19	-	79,82,87,89	0
9	LXB	D	1461	14/15	0.94	0.15	-	73,76,80,85	0
9	GXL	D	1467	11/12	0.92	0.20	-	69,79,90,91	0
8	NAG	D	1239	14/15	0.80	0.31	-	90,103,113,116	0
4	NGA	A	1362	14/15	0.93	0.18	-	71,80,84,91	0
6	NAG	D	1414	14/15	0.85	0.37	-	89,102,109,110	0
3	BMA	A	1141	11/12	0.91	0.29	-	86,90,93,93	0
4	GL0	A	1363	11/12	0.93	0.27	-	93,97,99,101	0
9	NGZ	D	1462	14/15	0.94	0.25	-	70,92,99,106	0
4	GLC	A	1365	11/12	0.87	0.43	-	87,97,101,101	0
6	NAG	A	1313	14/15	0.92	0.21	-	62,79,89,97	0
8	MAN	D	1242	11/12	0.56	0.76	-	108,118,121,122	0
4	LXZ	A	1361	14/15	0.91	0.16	-	65,70,80,81	0
8	BGC	D	1241	11/12	0.69	0.60	-	75,98,105,106	0
4	BGC	A	1366	11/12	0.91	0.21	-	75,81,87,90	0
8	NAG	D	1238	14/15	0.89	0.26	-	85,90,100,104	0
6	NAG	A	1314	14/15	0.88	0.36	-	85,95,100,100	0
3	BMA	A	1140	11/12	0.95	0.16	-	75,81,86,89	0
4	GLC	A	1367	11/12	0.94	0.21	-	69,75,82,90	0
4	BGC	A	1364	11/12	0.87	0.39	-	93,103,106,106	0
9	GL0	D	1463	11/12	0.83	0.34	-	110,116,119,121	0
8	BMA	D	1240	11/12	0.78	0.48	-	101,110,116,117	0
9	BGC	D	1464	11/12	0.79	0.40	-	113,120,124,125	0
6	NAG	D	1413	14/15	0.91	0.18	-	72,83,92,104	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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	SO4	D	1474	5/5	0.85	0.40	5.19	90,95,109,114	0
7	SO4	B	202	5/5	0.94	0.36	2.80	102,106,112,126	0
7	SO4	A	1375	5/5	0.84	0.38	2.77	82,83,103,104	0
7	SO4	A	1368	5/5	0.95	0.21	1.24	58,60,64,67	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
7	SO4	A	1369	5/5	0.96	0.16	0.72	57,66,73,79	0
7	SO4	D	1468	5/5	0.94	0.12	-1.23	81,86,90,97	0
7	SO4	D	1476	5/5	0.96	0.14	-1.57	86,89,99,105	0
7	SO4	D	1471	5/5	0.82	0.59	-	84,97,109,120	0
7	SO4	A	1370	5/5	0.89	0.28	-	116,121,126,139	0
7	SO4	D	1472	5/5	0.85	0.37	-	98,103,119,121	0
5	NAG	D	1310	14/15	0.82	0.35	-	95,105,108,108	0
5	NAG	A	1210	14/15	0.87	0.30	-	88,92,97,97	0
7	SO4	A	1372	5/5	0.83	0.38	-	81,92,102,116	0
7	SO4	D	1469	5/5	0.81	0.34	-	124,128,134,144	0
7	SO4	A	1371	5/5	0.85	0.32	-	108,110,117,129	0
7	SO4	E	190	5/5	0.89	0.37	-	105,109,123,130	0
7	SO4	B	203	5/5	0.88	0.33	-	108,108,125,130	0
7	SO4	D	1477	5/5	0.74	0.32	-	105,107,117,130	0
7	SO4	D	1470	5/5	0.76	0.30	-	109,117,128,134	0
7	SO4	A	1374	5/5	0.87	0.31	-	92,99,110,121	0
7	SO4	A	1373	5/5	0.83	0.31	-	93,99,118,124	0
7	SO4	D	1475	5/5	0.94	0.23	-	88,88,100,108	0
7	SO4	D	1473	5/5	0.83	0.45	-	105,111,123,131	0
7	SO4	A	1376	5/5	0.96	0.22	-	86,88,100,108	0

## 6.5 Other polymers

There are no such residues in this entry.