



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 09:45 am GMT

PDB ID : 4NZ8  
Title : Crystal structure of porcine aminopeptidase-N complexed with cleaved poly-alanine  
Authors : Chen, L.; Lin, Y.L.; Peng, G.; Li, F.  
Deposited on : 2013-12-11  
Resolution : 2.00 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/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

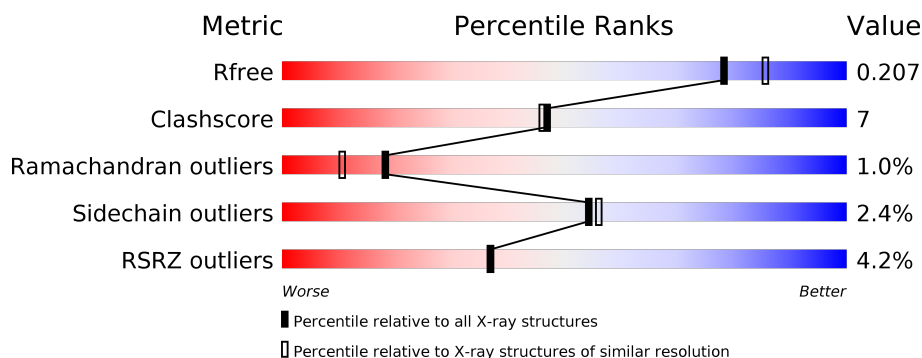
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	907	<div> <div>4%</div> <div>88%</div> <div>11%</div> <div>..</div> </div>
2	B	6	<div> <div>17%</div> <div>17%</div> <div>83%</div> </div>
2	C	6	<div> <div>67%</div> <div>67%</div> <div>17%</div> <div>17%</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	NAG	A	1010	-	-	-	X
4	NAG	A	1003	-	-	-	X
4	NAG	A	1008	-	-	-	X
6	SO4	A	1020	-	-	-	X
6	SO4	A	1023	-	-	-	X
6	SO4	A	1025	-	-	-	X
6	SO4	A	1026	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8363 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 Aminopeptidase N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	901	Total	C	N	O	S	0	1	0
			7239	4622	1209	1378	30			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	107	PHE	LEU	CONFLICT	UNP P15145
A	964	SER	-	EXPRESSION TAG	UNP P15145
A	965	HIS	-	EXPRESSION TAG	UNP P15145
A	966	HIS	-	EXPRESSION TAG	UNP P15145
A	967	HIS	-	EXPRESSION TAG	UNP P15145
A	968	HIS	-	EXPRESSION TAG	UNP P15145
A	969	HIS	-	EXPRESSION TAG	UNP P15145
A	970	HIS	-	EXPRESSION TAG	UNP P15145

- Molecule 2 is a protein called CLEAVED poly-Ala.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	1	Total	C	N	O	0	0	0
			5	3	1	1			
2	C	6	Total	C	N	O	0	0	0
			31	18	6	7			

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

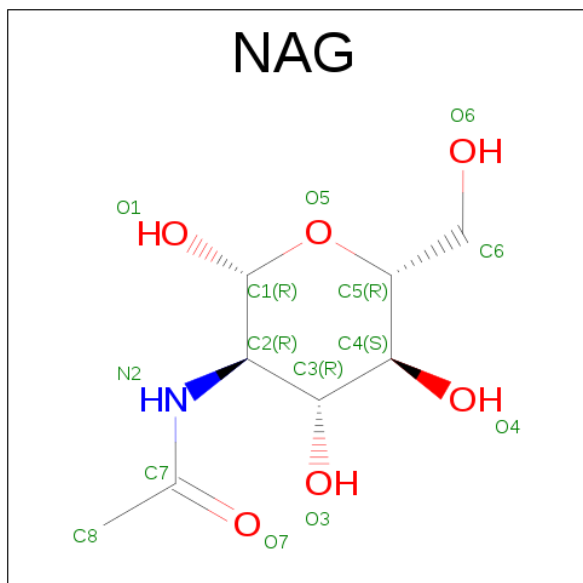
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	A	2	Total	C	N	O	0	0
			28	16	2	10		

*Continued on next page...*

Continued from previous page...

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

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

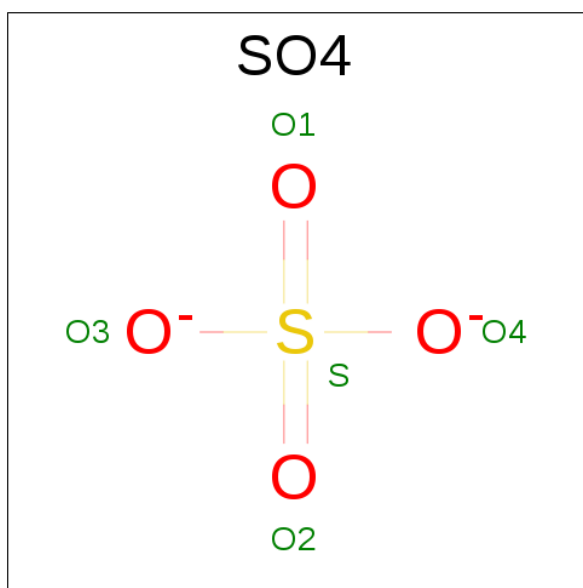


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Zn	0	0
			1	1		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	824	Total	O	0	0
			824	824		

*Continued on next page...*

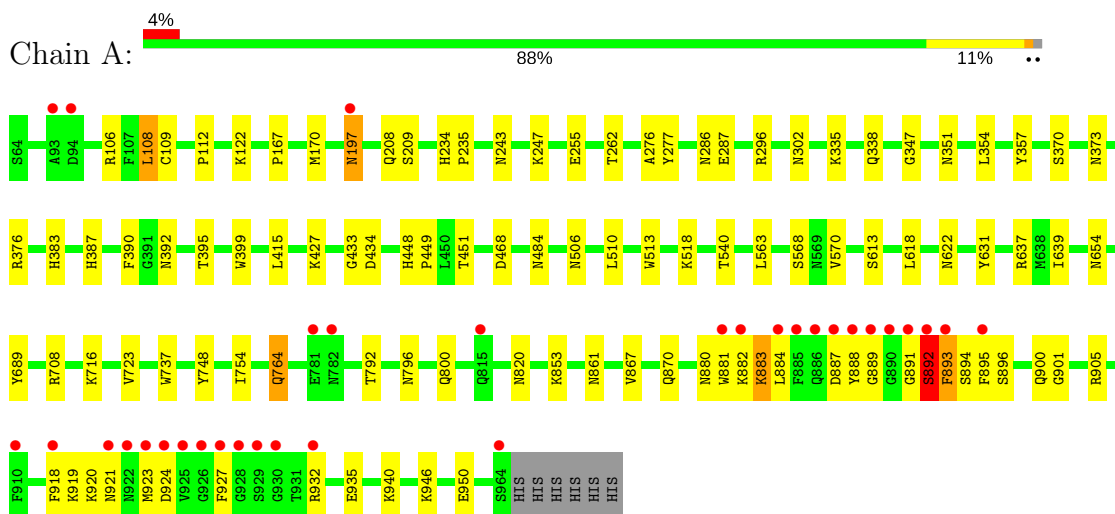
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	3	Total	O	0	0
			3	3		

### 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.

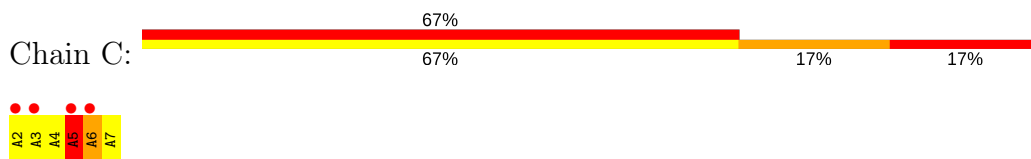
- Molecule 1: Aminopeptidase N



- Molecule 2: CLEAVED poly-Ala



- Molecule 2: CLEAVED poly-Ala





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	261.98Å 62.88Å 81.86Å 90.00° 100.39° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00 49.70 – 1.98	Depositor EDS
% Data completeness (in resolution range)	96.6 (50.00-2.00) 96.6 (49.70-1.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.150 , 0.208 0.150 , 0.207	Depositor DCC
$R_{free}$ test set	4395 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtriage
Anisotropy	0.232	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 59.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8363	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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: ZN, NAG, SO4

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	0.62	0/7427	0.70	1/10122 (0.0%)
2	B	1.27	0/4	1.23	0/4
2	C	1.37	0/30	1.52	1/39 (2.6%)
All	All	0.63	0/7461	0.71	2/10165 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	637	ARG	NE-CZ-NH2	5.38	122.99	120.30
2	C	5	ALA	N-CA-C	5.19	125.01	111.00

There are no chirality outliers.

There are no planarity outliers.

### 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	A	7239	0	6995	89	0
2	B	5	0	7	3	0
2	C	31	0	29	20	0
3	A	140	0	125	1	0
4	A	70	0	65	0	0
5	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	50	0	0	2	0
7	A	824	0	0	26	0
7	C	3	0	0	1	0
All	All	8363	0	7221	104	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 7.

All (104) 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:484:ASN:HB3	7:A:1895:HOH:O	1.44	1.17
1:A:618:LEU:HG	7:A:1885:HOH:O	1.45	1.12
1:A:286:ASN:HB3	7:A:1823:HOH:O	1.62	1.00
7:A:1654:HOH:O	2:C:4:ALA:HB3	1.67	0.93
1:A:891:GLY:HA3	7:A:1585:HOH:O	1.74	0.86
1:A:881:TRP:HZ2	1:A:927:PHE:HD1	1.25	0.83
7:A:1882:HOH:O	2:C:5:ALA:HB3	1.79	0.81
1:A:881:TRP:CE3	1:A:918:PHE:HZ	1.96	0.81
1:A:376:ARG:NH1	2:C:2:ALA:HB3	1.96	0.80
1:A:881:TRP:CZ2	1:A:927:PHE:HD1	1.99	0.80
7:A:1654:HOH:O	2:C:4:ALA:CB	2.25	0.79
1:A:622:ASN:H	1:A:654:ASN:HD21	1.28	0.78
2:C:6:ALA:HB2	7:C:103:HOH:O	1.84	0.77
1:A:888:TYR:OH	1:A:895:PHE:CD1	2.40	0.74
1:A:861:ASN:OD1	1:A:901:GLY:HA3	1.88	0.73
1:A:376:ARG:HH12	2:C:2:ALA:HB3	1.53	0.70
7:A:1804:HOH:O	2:C:7:ALA:HB2	1.90	0.70
7:A:1812:HOH:O	2:C:7:ALA:CB	2.38	0.70
1:A:302:ASN:ND2	6:A:1026:SO4:O4	2.25	0.70
2:C:4:ALA:O	2:C:5:ALA:HB2	1.90	0.69
1:A:255:GLU:HG3	7:A:1865:HOH:O	1.94	0.67
1:A:108:LEU:HD12	1:A:109:CYS:N	2.08	0.67
1:A:448:HIS:HE1	1:A:468:ASP:OD2	1.79	0.66
1:A:881:TRP:CZ2	1:A:927:PHE:CD1	2.84	0.65
1:A:881:TRP:HZ2	1:A:927:PHE:CD1	2.11	0.65
1:A:197:ASN:N	1:A:197:ASN:OD1	2.30	0.65
1:A:881:TRP:CE3	1:A:918:PHE:CZ	2.84	0.64
1:A:892:SER:OG	1:A:893:PHE:N	2.31	0.64
2:C:4:ALA:O	2:C:5:ALA:CB	2.46	0.63
1:A:208:GLN:HE21	1:A:209:SER:HA	1.64	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:243:ASN:HD21	1:A:276:ALA:H	1.47	0.62
7:A:1882:HOH:O	2:C:5:ALA:CA	2.49	0.60
1:A:919:LYS:O	1:A:921:ASN:N	2.34	0.60
1:A:434:ASP:OD1	2:C:5:ALA:HB1	2.01	0.60
1:A:415:LEU:HD22	1:A:427:LYS:HE2	1.84	0.59
1:A:881:TRP:CZ2	1:A:895:PHE:CE2	2.90	0.59
1:A:919:LYS:C	1:A:921:ASN:H	2.05	0.58
1:A:373:ASN:HD21	1:A:820:ASN:HD22	1.51	0.58
7:A:1812:HOH:O	2:C:7:ALA:HB1	2.02	0.58
1:A:618:LEU:HD13	1:A:639:ILE:HD13	1.86	0.57
1:A:882:LYS:C	1:A:884:LEU:H	2.06	0.57
1:A:234:HIS:HE1	1:A:262:THR:OG1	1.88	0.56
1:A:622:ASN:H	1:A:654:ASN:ND2	2.02	0.56
1:A:754:ILE:HG22	1:A:792:THR:HG21	1.85	0.56
7:A:1882:HOH:O	2:C:5:ALA:CB	2.46	0.56
1:A:351:ASN:HB2	1:A:354:LEU:O	2.06	0.56
1:A:870:GLN:HE21	1:A:905:ARG:HD2	1.72	0.55
1:A:618:LEU:HD13	1:A:639:ILE:CD1	2.37	0.54
1:A:881:TRP:CZ2	1:A:895:PHE:CZ	2.94	0.54
1:A:112:PRO:HD3	1:A:167:PRO:HG3	1.89	0.54
1:A:708:ARG:HD3	7:A:1165:HOH:O	2.07	0.53
1:A:563:LEU:HD11	1:A:570:VAL:CG2	2.39	0.52
1:A:106:ARG:NH1	3:A:1004:NAG:O3	2.39	0.52
1:A:881:TRP:CZ3	1:A:918:PHE:CZ	2.98	0.52
1:A:880:ASN:O	1:A:883:LYS:HD2	2.10	0.52
1:A:881:TRP:CZ3	1:A:918:PHE:HZ	2.26	0.52
1:A:448:HIS:HD2	1:A:449:PRO:O	1.93	0.52
1:A:870:GLN:HE21	1:A:905:ARG:HH21	1.57	0.52
1:A:946:LYS:HE3	1:A:950:GLU:OE2	2.10	0.51
7:A:1794:HOH:O	2:C:3:ALA:HA	2.09	0.51
7:A:1893:HOH:O	2:B:1:ALA:HB1	2.10	0.51
1:A:383:HIS:CE1	2:C:3:ALA:HB1	2.47	0.50
1:A:399:TRP:CZ3	6:A:1025:SO4:O1	2.64	0.50
1:A:618:LEU:CG	7:A:1885:HOH:O	2.25	0.50
1:A:347:GLY:HA3	2:C:2:ALA:N	2.26	0.50
1:A:433:GLY:O	2:C:5:ALA:O	2.28	0.50
1:A:881:TRP:CZ2	1:A:895:PHE:HE2	2.30	0.49
1:A:881:TRP:CH2	1:A:895:PHE:CZ	2.99	0.49
1:A:764:GLN:HE21	1:A:764:GLN:H	1.61	0.48
1:A:518:LYS:HD3	7:A:1676:HOH:O	2.13	0.48
7:A:1893:HOH:O	2:B:1:ALA:CB	2.63	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:376:ARG:HH12	2:C:2:ALA:CB	2.25	0.46
1:A:108:LEU:HD12	1:A:109:CYS:H	1.78	0.46
1:A:737:TRP:CZ3	1:A:754:ILE:HD11	2.51	0.46
1:A:286:ASN:ND2	7:A:1823:HOH:O	2.45	0.45
1:A:919:LYS:C	1:A:921:ASN:N	2.70	0.45
1:A:243:ASN:ND2	1:A:276:ALA:H	2.15	0.45
1:A:882:LYS:C	1:A:884:LEU:N	2.70	0.45
1:A:370:SER:O	1:A:373:ASN:HB3	2.16	0.44
1:A:618:LEU:HD12	1:A:631:TYR:OH	2.18	0.44
1:A:940:LYS:HE2	7:A:1386:HOH:O	2.18	0.44
1:A:208:GLN:HE21	1:A:209:SER:CA	2.31	0.44
1:A:286:ASN:HB3	1:A:296:ARG:HG2	1.99	0.44
1:A:708:ARG:HD2	1:A:905:ARG:HD3	1.99	0.44
1:A:881:TRP:CH2	1:A:895:PHE:CE2	3.06	0.43
1:A:887:ASP:C	1:A:888:TYR:CD1	2.92	0.43
1:A:618:LEU:CB	7:A:1885:HOH:O	2.64	0.43
1:A:338:GLN:NE2	1:A:357:TYR:OH	2.51	0.43
1:A:122:LYS:NZ	7:A:1400:HOH:O	2.52	0.43
7:A:1579:HOH:O	2:B:1:ALA:HB3	2.18	0.42
1:A:884:LEU:HA	1:A:888:TYR:CE1	2.55	0.42
1:A:106:ARG:HG2	1:A:170:MET:HE1	2.02	0.42
1:A:387:HIS:HA	1:A:390:PHE:O	2.20	0.42
1:A:247:LYS:HE2	7:A:1736:HOH:O	2.20	0.41
1:A:689:TYR:CD1	1:A:748:TYR:HB3	2.55	0.41
1:A:796:ASN:ND2	1:A:800:GLN:OE1	2.53	0.41
1:A:392:ASN:HD22	1:A:392:ASN:N	2.18	0.41
1:A:376:ARG:NH1	2:C:2:ALA:CB	2.77	0.41
1:A:395:THR:O	1:A:506:ASN:HA	2.20	0.41
1:A:286:ASN:CB	7:A:1823:HOH:O	2.40	0.41
1:A:510:LEU:HA	1:A:510:LEU:HD23	1.94	0.40
1:A:764:GLN:NE2	1:A:764:GLN:H	2.19	0.40
1:A:451:THR:HB	1:A:540:THR:O	2.21	0.40
1:A:234:HIS:HD2	1:A:235:PRO:O	2.05	0.40

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	900/907 (99%)	870 (97%)	23 (3%)	7 (1%)	22	15
2	C	4/6 (67%)	0	2 (50%)	2 (50%)	0	0
All	All	904/913 (99%)	870 (96%)	25 (3%)	9 (1%)	18	10

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	892	SER
1	A	920	LYS
2	C	5	ALA
2	C	6	ALA
1	A	894	SER
1	A	923	MET
1	A	889	GLY
1	A	924	ASP
1	A	893	PHE

### 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	797/802 (99%)	778 (98%)	19 (2%)	54	56

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	108	LEU
1	A	197	ASN
1	A	277	TYR
1	A	287	GLU
1	A	335	LYS
1	A	513	TRP
1	A	568	SER
1	A	613	SER
1	A	716	LYS
1	A	723	VAL
1	A	764	GLN
1	A	853	LYS
1	A	867	VAL
1	A	883	LYS
1	A	892	SER
1	A	896	SER
1	A	900	GLN
1	A	932	ARG
1	A	935	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	206	GLN
1	A	208	GLN
1	A	234	HIS
1	A	243	ASN
1	A	338	GLN
1	A	373	ASN
1	A	392	ASN
1	A	448	HIS
1	A	654	ASN
1	A	722	GLN
1	A	764	GLN
1	A	775	GLN
1	A	796	ASN
1	A	800	GLN
1	A	839	ASN
1	A	870	GLN
1	A	897	ASN
1	A	951	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 ⓘ

10 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	1001	1,3	14,14,15	0.50	0	15,19,21	1.15	1 (6%)
3	NAG	A	1002	3	14,14,15	0.95	1 (7%)	15,19,21	1.97	4 (26%)
3	NAG	A	1004	1,3	14,14,15	0.66	0	15,19,21	1.65	3 (20%)
3	NAG	A	1005	3	14,14,15	0.66	0	15,19,21	1.47	2 (13%)
3	NAG	A	1006	1,3	14,14,15	1.02	1 (7%)	15,19,21	1.13	1 (6%)
3	NAG	A	1007	3	14,14,15	0.95	2 (14%)	15,19,21	1.43	2 (13%)
3	NAG	A	1010	1,3	14,14,15	0.77	0	15,19,21	1.95	5 (33%)
3	NAG	A	1011	3	14,14,15	0.82	0	15,19,21	1.23	1 (6%)
3	NAG	A	1013	1,3	14,14,15	0.67	0	15,19,21	1.13	1 (6%)
3	NAG	A	1014	3	14,14,15	0.67	0	15,19,21	1.22	1 (6%)

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	1001	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1002	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1004	1,3	-	0/6/23/26	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	1005	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1006	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1007	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1010	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1011	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1013	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1014	3	-	0/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1002	NAG	O5-C1	-2.42	1.39	1.43
3	A	1006	NAG	O5-C1	-2.05	1.40	1.43
3	A	1007	NAG	O6-C6	2.11	1.51	1.42
3	A	1007	NAG	C6-C5	2.45	1.60	1.51

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1002	NAG	O5-C1-C2	-3.56	106.52	111.47
3	A	1004	NAG	C3-C4-C5	-3.30	104.40	110.22
3	A	1004	NAG	C2-N2-C7	-3.06	118.48	122.94
3	A	1010	NAG	O5-C1-C2	-2.98	107.33	111.47
3	A	1014	NAG	O5-C1-C2	-2.85	107.51	111.47
3	A	1004	NAG	O5-C1-C2	-2.85	107.51	111.47
3	A	1002	NAG	O3-C3-C4	-2.57	104.76	110.36
3	A	1006	NAG	C1-O5-C5	-2.45	108.78	112.17
3	A	1010	NAG	C2-N2-C7	-2.37	119.49	122.94
3	A	1011	NAG	O5-C1-C2	-2.32	108.25	111.47
3	A	1007	NAG	O6-C6-C5	-2.31	103.58	111.34
3	A	1005	NAG	C2-N2-C7	-2.21	119.72	122.94
3	A	1001	NAG	O4-C4-C3	-2.05	105.91	110.36
3	A	1010	NAG	O3-C3-C4	-2.04	105.91	110.36
3	A	1010	NAG	C4-C3-C2	2.04	114.01	111.02
3	A	1013	NAG	C1-O5-C5	2.64	115.80	112.17
3	A	1002	NAG	C3-C4-C5	2.86	115.25	110.22
3	A	1007	NAG	C1-O5-C5	3.26	116.67	112.17
3	A	1005	NAG	C1-O5-C5	3.72	117.30	112.17
3	A	1002	NAG	C4-C3-C2	4.13	117.08	111.02
3	A	1010	NAG	C1-O5-C5	5.43	119.65	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1004	NAG	1	0

## 5.6 Ligand geometry

Of 16 ligands modelled in this entry, 1 is monoatomic - leaving 15 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	A	1003	1	14,14,15	0.62	0	15,19,21	1.55	2 (13%)
4	NAG	A	1008	1	14,14,15	0.59	0	15,19,21	1.26	1 (6%)
4	NAG	A	1009	1	14,14,15	0.47	0	15,19,21	1.26	2 (13%)
4	NAG	A	1012	1	14,14,15	0.57	0	15,19,21	1.49	2 (13%)
4	NAG	A	1015	1	14,14,15	0.56	0	15,19,21	1.42	2 (13%)
6	SO4	A	1017	-	4,4,4	0.65	0	6,6,6	0.75	0
6	SO4	A	1018	-	4,4,4	0.56	0	6,6,6	0.43	0
6	SO4	A	1019	-	4,4,4	0.66	0	6,6,6	0.20	0
6	SO4	A	1020	-	4,4,4	0.77	0	6,6,6	0.48	0
6	SO4	A	1021	-	4,4,4	0.75	0	6,6,6	0.51	0
6	SO4	A	1022	-	4,4,4	0.83	0	6,6,6	0.52	0
6	SO4	A	1023	-	4,4,4	0.74	0	6,6,6	0.35	0
6	SO4	A	1024	-	4,4,4	0.58	0	6,6,6	0.38	0
6	SO4	A	1025	-	4,4,4	0.69	0	6,6,6	1.42	2 (33%)
6	SO4	A	1026	-	4,4,4	0.62	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
4	NAG	A	1003	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1008	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1009	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1012	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1015	1	-	0/6/23/26	0/1/1/1
6	SO4	A	1017	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1018	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1019	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1020	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1021	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1022	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1023	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1024	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1025	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1026	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1003	NAG	O5-C1-C2	-4.07	105.80	111.47
4	A	1009	NAG	O5-C1-C2	-2.97	107.34	111.47
4	A	1012	NAG	O5-C1-C2	-2.75	107.65	111.47
4	A	1003	NAG	O7-C7-C8	-2.41	117.67	122.06
6	A	1025	SO4	O4-S-O1	-2.08	97.81	109.26
6	A	1025	SO4	O4-S-O2	2.30	121.95	109.26
4	A	1015	NAG	C4-C3-C2	2.52	114.72	111.02
4	A	1009	NAG	C1-O5-C5	2.80	116.03	112.17
4	A	1015	NAG	C1-O5-C5	2.99	116.29	112.17
4	A	1008	NAG	C1-O5-C5	3.46	116.94	112.17
4	A	1012	NAG	C4-C3-C2	4.20	117.17	111.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1025	SO4	1	0
6	A	1026	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

### 6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	901/907 (99%)	-0.12	33 (3%) 42 43	28, 47, 82, 172	0
2	B	1/6 (16%)	3.66	1 (100%) 0 0	50, 50, 50, 50	0
2	C	6/6 (100%)	3.13	4 (66%) 0 1	35, 63, 67, 71	0
All	All	908/919 (98%)	-0.09	38 (4%) 37 37	28, 47, 82, 172	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	885	PHE	12.6
1	A	888	TYR	10.9
1	A	890	GLY	10.1
1	A	889	GLY	8.6
1	A	891	GLY	8.1
1	A	923	MET	7.9
1	A	925	VAL	7.8
2	C	2	ALA	6.4
1	A	924	ASP	5.6
1	A	886	GLN	5.5
1	A	932	ARG	5.1
1	A	927	PHE	5.0
1	A	926	GLY	4.8
1	A	893	PHE	4.7
1	A	882	LYS	4.5
1	A	922	ASN	4.5
1	A	881	TRP	4.3
1	A	928	GLY	4.0
1	A	884	LEU	4.0
1	A	94	ASP	3.9
1	A	892	SER	3.9
1	A	93	ALA	3.8
2	B	1	ALA	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	C	3	ALA	3.7
1	A	197	ASN	3.5
1	A	895	PHE	3.2
2	C	5	ALA	3.2
1	A	887	ASP	3.0
1	A	964	SER	3.0
1	A	781	GLU	2.9
1	A	929	SER	2.8
1	A	815	GLN	2.5
1	A	918	PHE	2.5
1	A	782	ASN	2.3
1	A	930	GLY	2.3
2	C	6	ALA	2.2
1	A	921	ASN	2.2
1	A	910	PHE	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(Å <sup>2</sup> )	Q<0.9
3	NAG	A	1010	14/15	0.96	0.12	3.03	56,69,141,164	0
3	NAG	A	1013	14/15	0.95	0.12	1.09	76,115,149,166	0
3	NAG	A	1004	14/15	0.98	0.10	1.08	50,63,81,85	0
3	NAG	A	1001	14/15	0.97	0.10	0.77	51,68,88,95	0
3	NAG	A	1005	14/15	0.94	0.19	-	67,96,132,137	0
3	NAG	A	1011	14/15	0.86	0.16	-	77,115,156,159	0
3	NAG	A	1002	14/15	0.93	0.20	-	92,119,141,148	0
3	NAG	A	1006	14/15	0.93	0.11	-	62,74,130,133	0
3	NAG	A	1007	14/15	0.77	0.24	-	109,155,204,261	0
3	NAG	A	1014	14/15	0.88	0.20	-	103,145,158,161	0

## 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
4	NAG	A	1003	14/15	0.93	0.12	9.13	63,79,137,170	0
6	SO4	A	1026	5/5	0.96	0.28	4.67	78,84,95,96	0
6	SO4	A	1020	5/5	0.89	0.20	4.25	68,70,87,87	0
4	NAG	A	1008	14/15	0.90	0.19	3.32	72,83,102,104	0
6	SO4	A	1025	5/5	0.94	0.16	2.72	49,55,76,76	0
6	SO4	A	1023	5/5	0.93	0.16	2.16	75,79,87,88	0
5	ZN	A	1016	1/1	0.99	0.19	0.13	43,43,43,43	0
6	SO4	A	1019	5/5	0.98	0.12	-0.78	69,75,81,84	0
6	SO4	A	1018	5/5	0.97	0.08	-1.00	72,75,79,79	0
6	SO4	A	1017	5/5	0.98	0.08	-1.07	63,66,70,77	0
6	SO4	A	1021	5/5	0.94	0.29	-	76,86,90,95	0
4	NAG	A	1015	14/15	0.83	0.39	-	133,164,196,220	0
4	NAG	A	1012	14/15	0.83	0.16	-	83,133,197,205	0
6	SO4	A	1024	5/5	0.95	0.14	-	73,73,90,94	0
4	NAG	A	1009	14/15	0.86	0.16	-	64,100,180,189	0
6	SO4	A	1022	5/5	0.91	0.19	-	78,82,88,88	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.