



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:24 PM GMT

PDB ID : 4OU3  
Title : Crystal structure of porcine aminopeptidase N complexed with CNGRCG tumor-homing peptide  
Authors : Liu, C.; Yang, Y.; Chen, L.; Lin, Y.-L.; Li, F.  
Deposited on : 2014-02-14  
Resolution : 1.95 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 (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

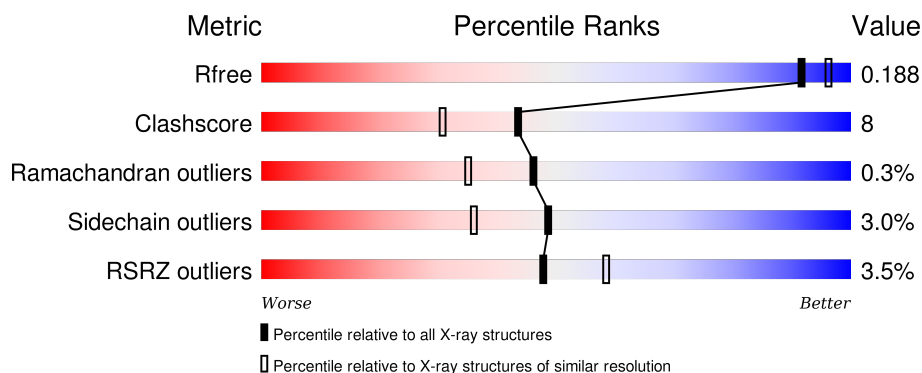
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	908	<div> <div>3%</div> <div>87%</div> <div>11%</div> <div>..</div> </div>
2	B	6	<div> <div>17%</div> <div>83%</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
4	NAG	A	1004	-	-	-	X
6	SO4	A	1020	-	-	-	X
6	SO4	A	1021	-	-	-	X
6	SO4	A	1022	-	-	X	X
6	SO4	A	1023	-	-	-	X
6	SO4	A	1024	-	-	-	X
6	SO4	A	1026	-	-	-	X
6	SO4	A	1027	-	-	X	-
6	SO4	A	1028	-	-	-	X
6	SO4	A	1029	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8498 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
			Total	C	N	O	S			
1	A	902	7249	4626	1212	1381	30	0	1	0

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 tumor-homing peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	6	40	20	10	8	2	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

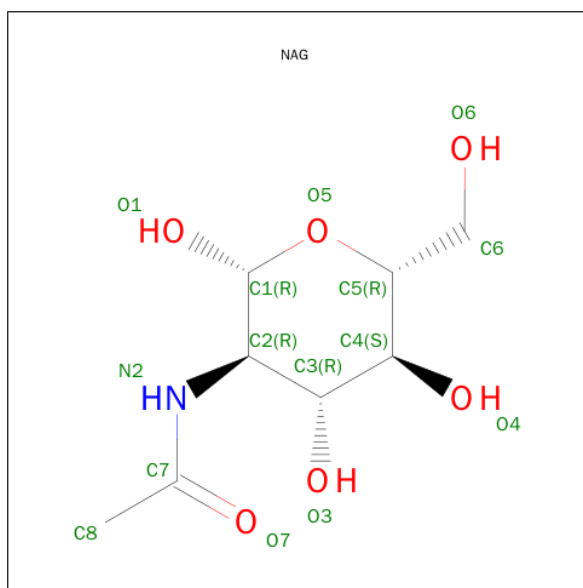
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		

*Continued on next page...*

Continued from previous page...

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

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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		

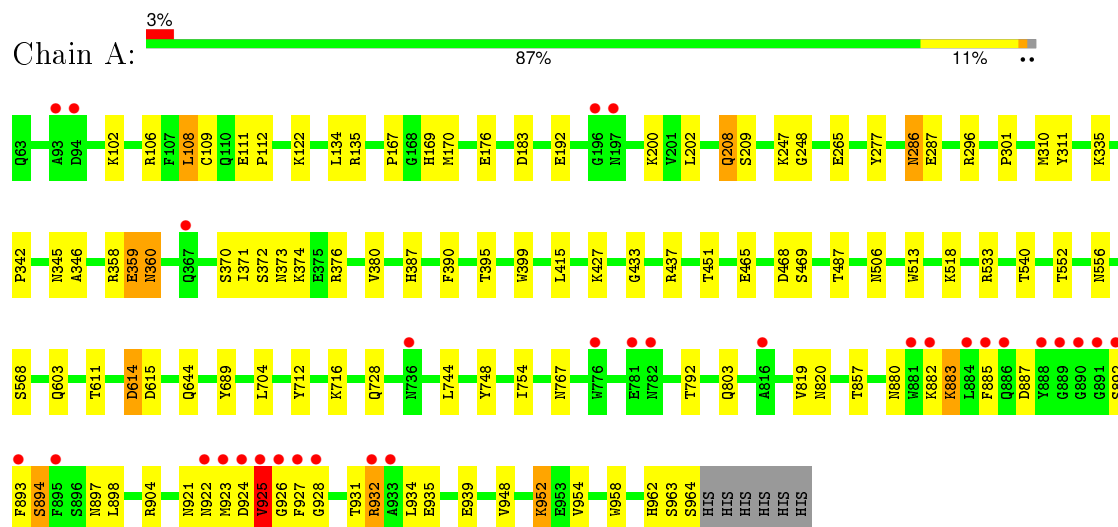
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	904	Total	O	0	0
			904	904		
7	B	2	Total	O	0	0
			2	2		

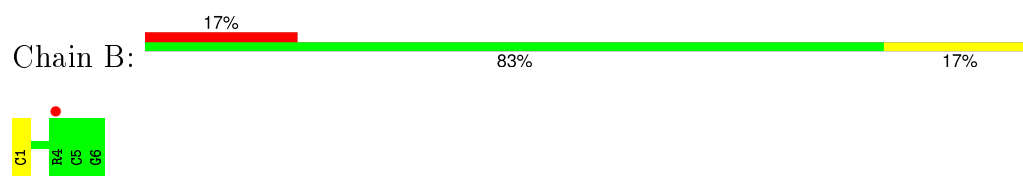
### 3 Residue-property plots [i](#)

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



#### • Molecule 2: tumor-homing peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	260.32Å 62.88Å 82.02Å 90.00° 100.59° 90.00°	Depositor
Resolution (Å)	50.00 – 1.95 47.68 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.8 (50.00-1.95) 97.8 (47.68-1.95)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.56 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.140 , 0.190 0.145 , 0.188	Depositor DCC
$R_{free}$ test set	4702 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.8	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 48.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 93588 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8498	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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 [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	0.64	0/7437	0.73	3/10135 (0.0%)
2	B	1.11	0/39	1.34	0/48
All	All	0.64	0/7476	0.74	3/10183 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	359	GLU	CB-CA-C	-5.95	98.50	110.40
1	A	135	ARG	NE-CZ-NH2	-5.69	117.46	120.30
1	A	176	GLU	CG-CD-OE2	-5.20	107.90	118.30

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	7249	0	7000	117	0
2	B	40	0	35	1	0
3	A	1	0	0	0	0
4	A	224	0	200	3	0
5	A	28	0	26	1	0
6	A	50	0	0	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	904	0	0	43	0
7	B	2	0	0	0	0
All	All	8498	0	7261	120	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 8.

All (120) 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:857:THR:HG22	1:A:893:PHE:CZ	1.65	1.31
1:A:106:ARG:HD3	7:A:1905:HOH:O	1.43	1.16
1:A:310:MET:HE3	7:A:1730:HOH:O	1.48	1.11
1:A:857:THR:HG22	1:A:893:PHE:CE1	1.89	1.06
1:A:857:THR:CG2	1:A:893:PHE:CE1	2.38	1.06
1:A:857:THR:CG2	1:A:893:PHE:CZ	2.38	1.05
1:A:451:THR:HG22	7:A:1848:HOH:O	1.56	1.03
1:A:924:ASP:O	1:A:925:VAL:HG13	1.59	1.02
1:A:265:GLU:HG3	7:A:1247:HOH:O	1.60	1.02
1:A:465:GLU:HG2	7:A:1824:HOH:O	1.59	1.00
1:A:346:ALA:HB3	7:A:1906:HOH:O	1.66	0.96
1:A:358:ARG:HG2	6:A:1022:SO4:O3	1.70	0.92
1:A:108:LEU:HD12	1:A:109:CYS:N	1.85	0.91
1:A:894:SER:HB3	7:A:1797:HOH:O	1.71	0.91
4:A:1015:NAG:H83	7:A:1564:HOH:O	1.68	0.91
1:A:183:ASP:HB2	7:A:1862:HOH:O	1.77	0.84
1:A:468:ASP:HB2	6:A:1027:SO4:O4	1.78	0.83
1:A:373:ASN:HD21	1:A:820:ASN:HD22	1.28	0.80
1:A:893:PHE:HZ	1:A:898:LEU:HD22	1.47	0.80
1:A:247:LYS:HB2	7:A:1991:HOH:O	1.83	0.78
1:A:894:SER:CB	7:A:1797:HOH:O	2.32	0.73
1:A:373:ASN:HD21	1:A:820:ASN:ND2	1.86	0.73
1:A:893:PHE:CZ	1:A:898:LEU:HD22	2.24	0.73
1:A:170:MET:SD	7:A:1905:HOH:O	2.47	0.71
1:A:924:ASP:O	1:A:925:VAL:CG1	2.35	0.71
1:A:924:ASP:C	1:A:925:VAL:HG13	2.11	0.71
1:A:927:PHE:HB2	1:A:931:THR:HA	1.74	0.69
1:A:415:LEU:HD22	1:A:427:LYS:HE2	1.73	0.69
1:A:857:THR:HG23	1:A:893:PHE:CE1	2.28	0.67
1:A:346:ALA:CB	7:A:1906:HOH:O	2.32	0.67
1:A:102:LYS:HE3	7:A:1879:HOH:O	1.96	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:208:GLN:HE21	1:A:209:SER:HA	1.62	0.64
1:A:102:LYS:HG3	7:A:1879:HOH:O	1.97	0.64
1:A:469:SER:N	6:A:1027:SO4:O4	2.24	0.63
1:A:552:THR:HB	1:A:611:THR:HG22	1.80	0.63
1:A:611:THR:HG21	7:A:1340:HOH:O	1.99	0.62
1:A:932:ARG:C	1:A:934:LEU:N	2.52	0.62
1:A:924:ASP:O	1:A:925:VAL:HG22	2.00	0.61
1:A:108:LEU:HD12	1:A:109:CYS:H	1.64	0.61
1:A:892:SER:O	1:A:893:PHE:HB3	2.00	0.60
1:A:170:MET:HB3	7:A:1905:HOH:O	2.01	0.60
1:A:372:SER:HB3	7:A:1392:HOH:O	2.02	0.59
1:A:518:LYS:HE3	7:A:1638:HOH:O	2.01	0.59
1:A:556:ASN:HD22	5:A:1016:NAG:H83	1.67	0.59
1:A:954:VAL:HG23	7:A:1841:HOH:O	2.01	0.59
1:A:248:GLY:N	7:A:1991:HOH:O	2.36	0.59
1:A:359:GLU:CD	7:A:1927:HOH:O	2.40	0.59
1:A:345:ASN:HB2	7:A:1923:HOH:O	2.04	0.56
1:A:376:ARG:NH2	6:A:1029:SO4:O4	2.39	0.55
1:A:342:PRO:HB3	7:A:1927:HOH:O	2.07	0.55
1:A:857:THR:CG2	1:A:893:PHE:CD1	2.90	0.55
1:A:932:ARG:C	1:A:934:LEU:H	2.09	0.54
1:A:926:GLY:C	1:A:927:PHE:CD1	2.81	0.54
1:A:819:VAL:HG23	7:A:1670:HOH:O	2.07	0.54
1:A:712:TYR:CE2	1:A:716:LYS:HD2	2.42	0.54
1:A:310:MET:HB3	7:A:1730:HOH:O	2.07	0.53
1:A:803:GLN:NE2	7:A:1922:HOH:O	2.41	0.53
1:A:704:LEU:HD21	1:A:904:ARG:HG2	1.90	0.53
1:A:921:ASN:OD1	1:A:921:ASN:O	2.27	0.53
1:A:286:ASN:HB3	1:A:296:ARG:HG2	1.90	0.53
1:A:380:VAL:HG13	7:A:1461:HOH:O	2.09	0.53
1:A:924:ASP:O	1:A:925:VAL:CB	2.56	0.52
4:A:1012:NAG:H81	7:A:1949:HOH:O	2.10	0.52
1:A:932:ARG:O	1:A:935:GLU:N	2.42	0.52
4:A:1015:NAG:C8	7:A:1564:HOH:O	2.39	0.51
1:A:857:THR:CG2	1:A:893:PHE:CE2	2.93	0.51
1:A:857:THR:HG21	1:A:893:PHE:CZ	2.42	0.51
1:A:370:SER:O	1:A:373:ASN:HB3	2.11	0.51
1:A:552:THR:HB	1:A:611:THR:CG2	2.42	0.50
1:A:122:LYS:CE	7:A:1984:HOH:O	2.58	0.50
1:A:857:THR:HG21	1:A:893:PHE:CE2	2.47	0.50
1:A:122:LYS:HE3	7:A:1984:HOH:O	2.11	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:433:GLY:O	1:A:437:ARG:HD2	2.11	0.50
1:A:373:ASN:ND2	1:A:820:ASN:ND2	2.59	0.50
1:A:358:ARG:HD2	1:A:360:ASN:OD1	2.11	0.50
1:A:451:THR:CG2	7:A:1848:HOH:O	2.34	0.49
1:A:451:THR:HG23	1:A:540:THR:HB	1.95	0.49
1:A:614:ASP:N	1:A:614:ASP:OD1	2.32	0.49
1:A:108:LEU:HD12	1:A:108:LEU:C	2.34	0.48
1:A:932:ARG:HH21	1:A:932:ARG:CG	2.26	0.48
1:A:108:LEU:HD13	1:A:169:HIS:O	2.13	0.48
1:A:880:ASN:O	1:A:883:LYS:HD2	2.14	0.48
1:A:399:TRP:CZ3	6:A:1028:SO4:O1	2.67	0.48
1:A:893:PHE:O	1:A:894:SER:CB	2.61	0.48
1:A:415:LEU:CD2	1:A:427:LYS:HE2	2.42	0.48
1:A:894:SER:OG	1:A:897:ASN:ND2	2.47	0.48
1:A:371:ILE:CG1	1:A:744:LEU:HD12	2.45	0.47
1:A:111:GLU:HG2	7:A:1541:HOH:O	2.15	0.46
1:A:963:SER:O	1:A:964:SER:HB3	2.16	0.46
1:A:924:ASP:C	1:A:925:VAL:CG1	2.80	0.46
1:A:712:TYR:CD2	1:A:716:LYS:HD2	2.51	0.46
1:A:958:TRP:O	1:A:962:HIS:HD2	1.99	0.45
1:A:754:ILE:HG22	1:A:792:THR:HG21	1.98	0.45
1:A:948:VAL:O	1:A:952:LYS:HB2	2.16	0.45
1:A:883:LYS:O	1:A:887:ASP:HB2	2.18	0.44
1:A:924:ASP:O	1:A:925:VAL:CG2	2.64	0.44
1:A:371:ILE:HG13	1:A:744:LEU:HD12	1.99	0.44
1:A:202:LEU:C	1:A:202:LEU:HD12	2.38	0.44
1:A:927:PHE:CD1	1:A:927:PHE:N	2.86	0.43
1:A:311:TYR:CE1	1:A:374:LYS:HE3	2.53	0.43
1:A:962:HIS:HE1	7:A:1914:HOH:O	2.01	0.43
1:A:122:LYS:HD3	7:A:1885:HOH:O	2.17	0.43
1:A:932:ARG:HD3	7:A:1597:HOH:O	2.18	0.43
1:A:106:ARG:CD	7:A:1905:HOH:O	2.26	0.42
1:A:387:HIS:HA	1:A:390:PHE:O	2.18	0.42
1:A:208:GLN:HG2	2:B:1:CYS:HB2	2.00	0.42
1:A:112:PRO:HD3	1:A:167:PRO:HG3	2.01	0.42
1:A:611:THR:OG1	1:A:615:ASP:CB	2.67	0.42
1:A:468:ASP:CB	6:A:1027:SO4:O4	2.59	0.42
1:A:395:THR:O	1:A:506:ASN:HA	2.20	0.42
1:A:893:PHE:O	1:A:894:SER:HB3	2.20	0.42
1:A:928:GLY:O	1:A:931:THR:OG1	2.31	0.41
1:A:904:ARG:NH2	7:A:1764:HOH:O	2.52	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:301:PRO:HD2	7:A:1145:HOH:O	2.20	0.41
1:A:122:LYS:HE2	7:A:1984:HOH:O	2.21	0.41
1:A:728:GLN:HG3	7:A:1644:HOH:O	2.20	0.41
1:A:192:GLU:HA	1:A:200:LYS:O	2.21	0.41
1:A:310:MET:CE	7:A:1730:HOH:O	2.32	0.41
1:A:360:ASN:HB3	6:A:1022:SO4:O2	2.21	0.40
1:A:689:TYR:CD1	1:A:748:TYR:HB3	2.57	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	901/908 (99%)	872 (97%)	26 (3%)	3 (0%)	46	35
2	B	4/6 (67%)	4 (100%)	0	0	100	100
All	All	905/914 (99%)	876 (97%)	26 (3%)	3 (0%)	46	35

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	894	SER
1	A	925	VAL
1	A	487	THR

### 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	798/803 (99%)	774 (97%)	24 (3%)	48	36
2	B	4/4 (100%)	4 (100%)	0	100	100
All	All	802/807 (99%)	778 (97%)	24 (3%)	48	36

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	108	LEU
1	A	134	LEU
1	A	208	GLN
1	A	277	TYR
1	A	286	ASN
1	A	287	GLU
1	A	335	LYS
1	A	360	ASN
1	A	513	TRP
1	A	533	ARG
1	A	568	SER
1	A	603	GLN
1	A	614	ASP
1	A	644	GLN
1	A	767	ASN
1	A	882	LYS
1	A	883	LYS
1	A	885	PHE
1	A	922	ASN
1	A	923	MET
1	A	925	VAL
1	A	932	ARG
1	A	939	GLU
1	A	952	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	286	ASN
1	A	360	ASN
1	A	767	ASN
1	A	820	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	897	ASN
1	A	962	HIS

### 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 ⓘ

16 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	$\# Z  > 2$
4	NAG	A	1002	1,4	14,14,15	0.70	0	15,19,21	1.59	3 (20%)
4	NAG	A	1003	4	14,14,15	0.61	0	15,19,21	1.93	4 (26%)
4	NAG	A	1004	1,4	14,14,15	0.56	0	15,19,21	1.20	2 (13%)
4	NAG	A	1005	4	14,14,15	0.63	0	15,19,21	1.49	2 (13%)
4	NAG	A	1006	1,4	14,14,15	0.55	0	15,19,21	1.66	4 (26%)
4	NAG	A	1007	4	14,14,15	0.72	0	15,19,21	1.09	1 (6%)
4	NAG	A	1008	1,4	14,14,15	0.81	1 (7%)	15,19,21	1.13	2 (13%)
4	NAG	A	1009	4	14,14,15	0.61	0	15,19,21	1.07	1 (6%)
4	NAG	A	1010	1,4	14,14,15	0.81	1 (7%)	15,19,21	1.27	1 (6%)
4	NAG	A	1011	4	14,14,15	0.58	0	15,19,21	1.22	1 (6%)
4	NAG	A	1012	1,4	14,14,15	0.68	0	15,19,21	1.66	4 (26%)
4	NAG	A	1013	4	14,14,15	0.70	0	15,19,21	0.99	1 (6%)
4	NAG	A	1014	1,4	14,14,15	0.49	0	15,19,21	2.11	5 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	1015	4	14,14,15	0.84	1 (7%)	15,19,21	1.47	2 (13%)
4	NAG	A	1017	1,4	14,14,15	0.94	0	15,19,21	1.59	4 (26%)
4	NAG	A	1018	4	14,14,15	0.50	0	15,19,21	2.18	5 (33%)

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	1002	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1003	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1004	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1005	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1006	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1007	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1008	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1009	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1010	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1011	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1012	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1013	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1014	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1015	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1017	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1018	4	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1010	NAG	O4-C4	-2.21	1.37	1.43
4	A	1008	NAG	C1-C2	2.30	1.55	1.52
4	A	1015	NAG	C1-C2	2.44	1.55	1.52

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1018	NAG	C2-N2-C7	-5.17	116.39	123.04
4	A	1006	NAG	C2-N2-C7	-3.34	118.75	123.04
4	A	1014	NAG	C2-N2-C7	-3.21	118.91	123.04
4	A	1014	NAG	C3-C4-C5	-2.85	105.23	110.20

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1018	NAG	C4-C3-C2	-2.78	106.90	111.23
4	A	1012	NAG	O3-C3-C4	-2.77	104.09	110.34
4	A	1013	NAG	C4-C3-C2	-2.70	107.03	111.23
4	A	1002	NAG	O3-C3-C2	-2.54	104.09	109.11
4	A	1006	NAG	C1-O5-C5	-2.52	109.05	112.25
4	A	1004	NAG	O7-C7-C8	-2.52	117.44	122.06
4	A	1006	NAG	O7-C7-C8	-2.48	117.50	122.06
4	A	1017	NAG	C3-C4-C5	-2.37	106.06	110.20
4	A	1007	NAG	C3-C4-C5	-2.37	106.07	110.20
4	A	1005	NAG	O4-C4-C3	-2.30	105.15	110.34
4	A	1017	NAG	C2-N2-C7	-2.29	120.09	123.04
4	A	1015	NAG	C3-C4-C5	-2.27	106.24	110.20
4	A	1003	NAG	C1-O5-C5	-2.21	109.44	112.25
4	A	1008	NAG	O4-C4-C3	-2.13	105.54	110.34
4	A	1017	NAG	C6-C5-C4	-2.13	107.76	113.02
4	A	1002	NAG	O7-C7-C8	-2.11	118.19	122.06
4	A	1010	NAG	C2-N2-C7	-2.10	120.34	123.04
4	A	1003	NAG	O7-C7-C8	-2.05	118.30	122.06
4	A	1018	NAG	O3-C3-C4	-2.05	105.73	110.34
4	A	1008	NAG	O7-C7-C8	-2.01	118.37	122.06
4	A	1014	NAG	O4-C4-C5	2.10	114.81	109.24
4	A	1004	NAG	C2-N2-C7	2.16	125.82	123.04
4	A	1012	NAG	C2-N2-C7	2.33	126.03	123.04
4	A	1012	NAG	C4-C3-C2	2.41	114.98	111.23
4	A	1018	NAG	C3-C2-N2	2.41	116.34	110.56
4	A	1014	NAG	C4-C3-C2	2.59	115.26	111.23
4	A	1009	NAG	C1-O5-C5	2.71	115.68	112.25
4	A	1006	NAG	O5-C5-C6	2.77	113.34	107.35
4	A	1003	NAG	C8-C7-N2	2.79	121.44	116.11
4	A	1017	NAG	O5-C5-C6	2.83	113.48	107.35
4	A	1012	NAG	O5-C5-C6	2.98	113.80	107.35
4	A	1011	NAG	C4-C3-C2	3.15	116.12	111.23
4	A	1018	NAG	C1-O5-C5	3.58	116.79	112.25
4	A	1002	NAG	C1-O5-C5	3.82	117.10	112.25
4	A	1015	NAG	O5-C5-C6	4.00	116.01	107.35
4	A	1005	NAG	C4-C3-C2	4.20	117.76	111.23
4	A	1003	NAG	C4-C3-C2	4.55	118.30	111.23
4	A	1014	NAG	C1-O5-C5	5.86	119.69	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1012	NAG	1	0
4	A	1015	NAG	2	0

## 5.6 Ligand geometry

Of 13 ligands modelled in this entry, 1 is monoatomic - leaving 12 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	$\# Z  > 2$
5	NAG	A	1016	1	14,14,15	0.58	0	15,19,21	2.13	4 (26%)
5	NAG	A	1019	1	14,14,15	0.60	0	15,19,21	1.33	1 (6%)
6	SO4	A	1020	-	4,4,4	1.24	0	6,6,6	0.76	0
6	SO4	A	1021	-	4,4,4	0.99	0	6,6,6	0.46	0
6	SO4	A	1022	-	4,4,4	1.16	0	6,6,6	0.37	0
6	SO4	A	1023	-	4,4,4	1.17	0	6,6,6	0.45	0
6	SO4	A	1024	-	4,4,4	1.23	0	6,6,6	0.33	0
6	SO4	A	1025	-	4,4,4	1.01	0	6,6,6	0.54	0
6	SO4	A	1026	-	4,4,4	1.46	1 (25%)	6,6,6	0.69	0
6	SO4	A	1027	-	4,4,4	1.39	1 (25%)	6,6,6	0.56	0
6	SO4	A	1028	-	4,4,4	0.86	0	6,6,6	2.08	1 (16%)
6	SO4	A	1029	-	4,4,4	0.79	0	6,6,6	0.35	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	1016	1	-	0/6/23/26	0/1/1/1
5	NAG	A	1019	1	-	0/6/23/26	0/1/1/1
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

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
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
6	SO4	A	1027	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1028	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1029	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1027	SO4	O3-S	2.04	1.54	1.47
6	A	1026	SO4	O2-S	2.36	1.55	1.47

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1028	SO4	O2-S-O1	-4.48	95.31	109.50
5	A	1016	NAG	O7-C7-C8	-2.85	116.84	122.06
5	A	1016	NAG	O5-C5-C6	3.48	114.87	107.35
5	A	1019	NAG	C1-O5-C5	3.50	116.69	112.25
5	A	1016	NAG	C8-C7-N2	4.02	123.80	116.11
5	A	1016	NAG	C2-N2-C7	4.82	129.23	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1016	NAG	1	0
6	A	1022	SO4	2	0
6	A	1027	SO4	3	0
6	A	1028	SO4	1	0
6	A	1029	SO4	1	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	902/908 (99%)	-0.09	31 (3%) 49 60	21, 42, 82, 198	0
2	B	6/6 (100%)	0.92	1 (16%) 2 3	34, 39, 46, 56	0
All	All	908/914 (99%)	-0.08	32 (3%) 48 58	21, 42, 82, 198	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	893	PHE	19.4
1	A	885	PHE	14.2
1	A	891	GLY	11.5
1	A	888	TYR	11.0
1	A	889	GLY	9.9
1	A	925	VAL	8.3
1	A	886	GLN	6.7
1	A	923	MET	5.7
1	A	927	PHE	5.4
1	A	932	ARG	4.5
1	A	882	LYS	4.3
1	A	928	GLY	3.9
1	A	924	ASP	3.6
1	A	197	ASN	3.6
1	A	926	GLY	3.5
1	A	933	ALA	3.3
1	A	892	SER	3.2
1	A	890	GLY	3.0
1	A	922	ASN	2.9
1	A	93	ALA	2.9
1	A	196	GLY	2.9
1	A	367	GLN	2.8
1	A	884	LEU	2.7
1	A	895	PHE	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	781	GLU	2.5
1	A	881	TRP	2.4
2	B	4	ARG	2.3
1	A	782	ASN	2.2
1	A	776	TRP	2.2
1	A	736	ASN	2.2
1	A	816	ALA	2.1
1	A	94	ASP	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
4	NAG	A	1004	14/15	0.95	0.13	7.44	40,56,82,88	0
4	NAG	A	1010	14/15	0.96	0.16	1.65	56,72,112,119	0
4	NAG	A	1017	14/15	0.92	0.11	1.30	56,99,169,169	0
4	NAG	A	1006	14/15	0.98	0.07	-0.60	29,41,55,58	0
4	NAG	A	1002	14/15	0.98	0.08	-0.75	32,40,48,50	0
4	NAG	A	1014	14/15	0.97	0.07	-1.07	34,42,50,51	0
4	NAG	A	1013	14/15	0.89	0.26	-	56,89,111,138	0
4	NAG	A	1007	14/15	0.97	0.17	-	49,64,79,91	0
4	NAG	A	1012	14/15	0.83	0.14	-	56,112,171,179	0
4	NAG	A	1015	14/15	0.91	0.19	-	54,77,130,151	0
4	NAG	A	1003	14/15	0.98	0.17	-	58,85,187,216	0
4	NAG	A	1008	14/15	0.94	0.09	-	49,76,108,114	0
4	NAG	A	1009	14/15	0.86	0.22	-	65,103,163,188	0
4	NAG	A	1011	14/15	0.92	0.18	-	60,102,131,145	0
4	NAG	A	1018	14/15	0.80	0.23	-	85,107,125,137	0
4	NAG	A	1005	14/15	0.81	0.41	-	85,113,151,153	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(Å <sup>2</sup> )	Q<0.9
6	SO4	A	1028	5/5	0.92	0.21	14.08	40,41,60,65	0
6	SO4	A	1029	5/5	0.95	0.24	10.04	64,69,76,79	0
6	SO4	A	1021	5/5	0.95	0.26	8.94	64,70,73,85	0
6	SO4	A	1023	5/5	0.93	0.22	7.27	56,57,73,76	0
6	SO4	A	1026	5/5	0.91	0.24	6.95	64,64,81,81	0
6	SO4	A	1022	5/5	0.94	0.23	4.56	62,65,81,86	0
6	SO4	A	1024	5/5	0.96	0.29	4.11	60,63,75,91	0
6	SO4	A	1020	5/5	0.97	0.18	2.45	51,56,68,77	0
6	SO4	A	1025	5/5	0.94	0.21	1.01	64,71,74,87	0
3	ZN	A	1001	1/1	1.00	0.11	-0.33	30,30,30,30	0
6	SO4	A	1027	5/5	0.93	0.23	-	60,66,75,76	0
5	NAG	A	1016	14/15	0.84	0.17	-	87,115,160,185	0
5	NAG	A	1019	14/15	0.85	0.36	-	87,105,127,141	0

## 6.5 Other polymers

There are no such residues in this entry.