



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 09:38 am GMT

PDB ID : 4HOM  
Title : Crystal structure of porcine aminopeptidase-N complexed with substance P  
Authors : Chen, L.; Lin, Y.L.; Peng, G.; Li, F.  
Deposited on : 2012-10-22  
Resolution : 1.90 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

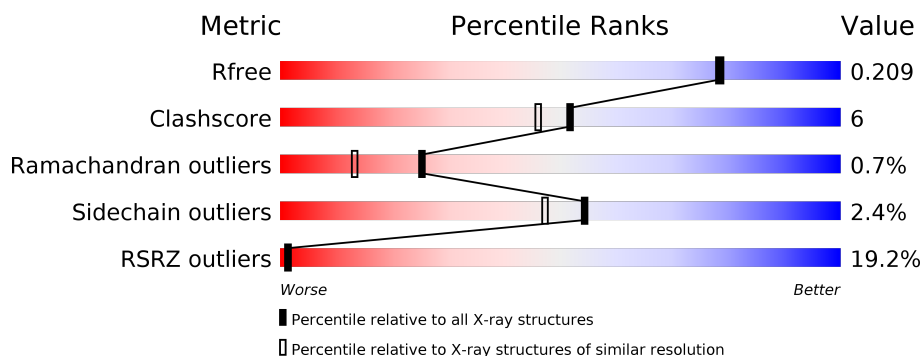
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	5047 (1.90-1.90)
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)
RSRZ outliers	101464	5100 (1.90-1.90)

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	
2	B	11	

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	1011	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9093 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	902	Total	C	N	O	S	0	0	0
			7241	4622	1210	1379	30			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	ASN	PHE	CONFLICT	UNP P15145
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 substance P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	11	Total	C	N	O	S	0	0	0
			95	63	17	14	1			

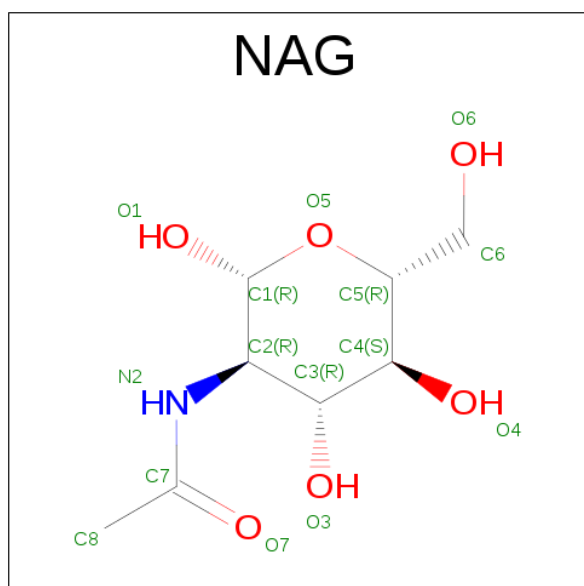
- Molecule 3 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	3	Total	C	N	O	0	0
			42	24	3	15		
3	A	3	Total	C	N	O	0	0
			42	24	3	15		
3	A	3	Total	C	N	O	0	0
			42	24	3	15		

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

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

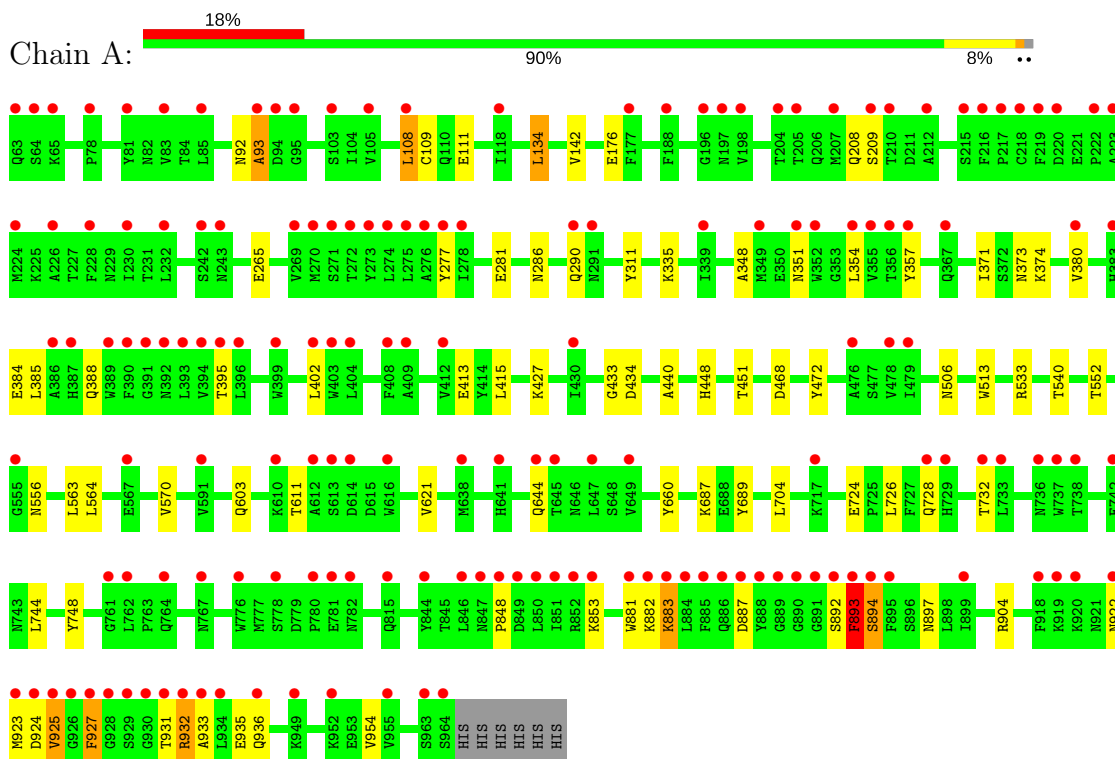
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1423	Total 1423	O 1423	0	0
7	B	11	Total 11	O 11	0	0

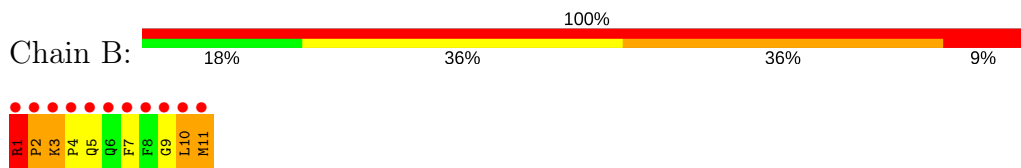
### 3 Residue-property plots

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

#### • Molecule 1: Aminopeptidase N



#### • Molecule 2: substance P



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	261.57Å 62.60Å 81.55Å 90.00° 100.27° 90.00°	Depositor
Resolution (Å)	37.18 – 1.90 37.15 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.0 (37.18-1.90) 98.0 (37.15-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.147 , 0.201 0.151 , 0.209	Depositor DCC
$R_{free}$ test set	5035 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.8	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 71.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9093	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

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

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	0.34	0/7429	0.50	0/10124
2	B	0.65	0/98	1.50	4/128 (3.1%)
All	All	0.35	0/7527	0.53	4/10252 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	1	ARG	C-N-CD	-7.80	103.43	120.60
2	B	2	PRO	N-CA-C	6.73	129.60	112.10
2	B	3	LYS	N-CA-C	6.52	128.60	111.00
2	B	2	PRO	CB-CA-C	-5.78	97.56	112.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1	ARG	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	7241	0	6994	80	0
2	B	95	0	99	33	0
3	A	126	0	111	3	0
4	A	168	0	150	3	0
5	A	28	0	26	0	0
6	A	1	0	0	0	0
7	A	1423	0	0	18	2
7	B	11	0	0	0	0
All	All	9093	0	7380	96	2

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 6.

All (96) 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:348:ALA:HB3	2:B:2:PRO:HB3	1.28	1.08
2:B:2:PRO:O	2:B:4:PRO:HD3	1.60	0.99
1:A:384:GLU:OE1	2:B:2:PRO:HA	1.67	0.92
1:A:380:VAL:HG13	7:A:1847:HOH:O	1.70	0.91
1:A:265:GLU:HG3	7:A:1351:HOH:O	1.69	0.91
1:A:451:THR:HG22	7:A:2203:HOH:O	1.75	0.86
7:A:1269:HOH:O	2:B:1:ARG:HB3	1.73	0.86
2:B:3:LYS:HB3	2:B:5:GLN:OE1	1.77	0.83
1:A:556:ASN:HD22	4:A:1018:NAG:H83	1.47	0.78
1:A:111:GLU:HG2	7:A:2218:HOH:O	1.85	0.76
1:A:348:ALA:CB	2:B:2:PRO:HB3	2.13	0.75
1:A:384:GLU:OE1	2:B:2:PRO:CA	2.35	0.75
7:A:1269:HOH:O	2:B:1:ARG:CB	2.31	0.74
2:B:2:PRO:O	2:B:4:PRO:CD	2.36	0.72
1:A:371:ILE:HG13	1:A:744:LEU:HD12	1.75	0.69
1:A:373:ASN:HB3	7:A:2199:HOH:O	1.92	0.68
2:B:1:ARG:HB2	2:B:2:PRO:HD3	1.74	0.68
1:A:371:ILE:CG1	1:A:744:LEU:HD12	2.25	0.67
1:A:472:TYR:CD2	2:B:3:LYS:HB2	2.30	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:433:GLY:HA2	2:B:7:PHE:HA	1.77	0.66
1:A:927:PHE:HD2	1:A:931:THR:HG23	1.60	0.66
1:A:92:ASN:HB3	7:A:2014:HOH:O	1.95	0.65
1:A:564:LEU:HD13	2:B:11:MET:HG3	1.79	0.64
1:A:108:LEU:HD12	1:A:109:CYS:N	2.13	0.63
1:A:704:LEU:HD21	1:A:904:ARG:HG2	1.81	0.63
1:A:176:GLU:OE1	7:A:1512:HOH:O	2.16	0.63
1:A:892:SER:O	1:A:893:PHE:HB2	2.00	0.62
1:A:611:THR:HG21	7:A:1382:HOH:O	2.01	0.61
1:A:208:GLN:HE21	1:A:209:SER:HA	1.67	0.60
1:A:904:ARG:HD2	7:A:1894:HOH:O	2.02	0.60
1:A:932:ARG:N	1:A:932:ARG:HE	2.00	0.59
1:A:552:THR:HB	1:A:611:THR:HG22	1.86	0.58
1:A:448:HIS:HE1	1:A:468:ASP:OD2	1.87	0.57
1:A:92:ASN:O	1:A:93:ALA:HB3	2.05	0.56
1:A:564:LEU:CD1	2:B:11:MET:HG3	2.37	0.55
3:A:1005:NAG:H61	3:A:1006:NAG:O5	2.08	0.54
1:A:881:TRP:C	1:A:883:LYS:H	2.11	0.54
2:B:3:LYS:CB	2:B:5:GLN:OE1	2.52	0.54
1:A:434:ASP:OD1	2:B:5:GLN:HG3	2.10	0.52
1:A:954:VAL:HG23	7:A:1573:HOH:O	2.10	0.52
1:A:563:LEU:HD11	1:A:570:VAL:CG2	2.40	0.52
1:A:660:TYR:CD1	2:B:10:LEU:HG	2.45	0.51
1:A:660:TYR:HD1	2:B:10:LEU:HG	1.75	0.51
1:A:357:TYR:HE2	1:A:380:VAL:HG12	1.76	0.51
2:B:1:ARG:HB2	2:B:2:PRO:CD	2.42	0.50
2:B:4:PRO:C	2:B:5:GLN:HG2	2.31	0.50
1:A:395:THR:O	1:A:506:ASN:HA	2.12	0.50
1:A:451:THR:HG23	1:A:540:THR:HB	1.94	0.49
1:A:892:SER:O	1:A:893:PHE:CB	2.60	0.49
1:A:848:PRO:HG3	1:A:853:LYS:HE3	1.94	0.49
1:A:281:GLU:OE1	7:A:1541:HOH:O	2.20	0.49
7:A:1441:HOH:O	2:B:10:LEU:HD13	2.13	0.48
1:A:931:THR:HG22	1:A:932:ARG:NE	2.29	0.48
1:A:440:ALA:HA	2:B:11:MET:HG2	1.96	0.48
1:A:357:TYR:CE2	1:A:380:VAL:HG12	2.49	0.47
1:A:927:PHE:CD2	1:A:931:THR:HG23	2.46	0.47
1:A:348:ALA:H	2:B:2:PRO:HG3	1.80	0.47
1:A:108:LEU:HD12	1:A:109:CYS:H	1.79	0.46
1:A:371:ILE:HG13	1:A:744:LEU:CD1	2.43	0.46
3:A:1003:NAG:H83	7:A:2191:HOH:O	2.15	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:348:ALA:N	2:B:2:PRO:HG3	2.30	0.46
1:A:402:LEU:O	1:A:402:LEU:HG	2.12	0.46
2:B:4:PRO:O	2:B:5:GLN:HG2	2.17	0.45
1:A:448:HIS:CE1	1:A:468:ASP:OD2	2.69	0.45
1:A:208:GLN:HE21	1:A:209:SER:CA	2.29	0.45
1:A:415:LEU:HD22	1:A:427:LYS:HE3	1.97	0.45
1:A:556:ASN:HD22	4:A:1018:NAG:C8	2.23	0.45
1:A:440:ALA:HB2	2:B:11:MET:H	1.81	0.45
7:A:2360:HOH:O	2:B:4:PRO:HB3	2.17	0.45
1:A:732:THR:HB	7:A:2161:HOH:O	2.17	0.44
1:A:351:ASN:HB2	1:A:354:LEU:O	2.18	0.44
1:A:563:LEU:HD11	1:A:570:VAL:HG23	2.00	0.44
1:A:924:ASP:O	1:A:925:VAL:HG22	2.18	0.44
7:A:1269:HOH:O	2:B:1:ARG:HB2	2.09	0.44
1:A:371:ILE:HG12	1:A:744:LEU:HD12	1.99	0.43
1:A:92:ASN:O	1:A:93:ALA:CB	2.66	0.43
1:A:208:GLN:HA	1:A:209:SER:HA	1.82	0.43
1:A:932:ARG:HA	1:A:935:GLU:HB2	2.00	0.43
1:A:385:LEU:HA	1:A:388:GLN:HG2	2.01	0.42
1:A:413:GLU:OE1	2:B:4:PRO:HB3	2.19	0.42
1:A:311:TYR:CE1	1:A:374:LYS:HE2	2.55	0.42
1:A:932:ARG:H	1:A:932:ARG:HE	1.67	0.42
3:A:1002:NAG:H4	3:A:1003:NAG:H2	1.81	0.41
1:A:440:ALA:CA	2:B:11:MET:HG2	2.49	0.41
1:A:621:VAL:HG11	4:A:1021:NAG:H82	2.01	0.41
1:A:689:TYR:CD1	1:A:748:TYR:HB3	2.56	0.41
1:A:472:TYR:CE2	2:B:3:LYS:HB2	2.55	0.41
1:A:894:SER:OG	1:A:897:ASN:ND2	2.53	0.41
1:A:933:ALA:HA	1:A:936:GLN:OE1	2.21	0.41
1:A:883:LYS:HG2	1:A:883:LYS:O	2.20	0.41
1:A:134:LEU:HD13	1:A:142:VAL:CG2	2.51	0.41
1:A:413:GLU:OE2	2:B:4:PRO:HA	2.21	0.41
1:A:380:VAL:HG22	2:B:4:PRO:HG2	2.03	0.41
1:A:724:GLU:O	1:A:728:GLN:HG2	2.21	0.40
1:A:687:LYS:HA	1:A:726:LEU:HD13	2.03	0.40
1:A:371:ILE:HG23	1:A:748:TYR:HE1	1.87	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:2035:HOH:O	7:A:2035:HOH:O[2_556]	1.95	0.25
7:A:2351:HOH:O	7:A:2351:HOH:O[2_556]	1.98	0.22

## 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	900/908 (99%)	869 (97%)	26 (3%)	5 (1%)	28	16
2	B	9/11 (82%)	2 (22%)	6 (67%)	1 (11%)	0	0
All	All	909/919 (99%)	871 (96%)	32 (4%)	6 (1%)	25	13

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	882	LYS
1	A	893	PHE
1	A	925	VAL
2	B	9	GLY
1	A	93	ALA
1	A	894	SER

### 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	797/803 (99%)	780 (98%)	17 (2%)	59	53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	10/10 (100%)	8 (80%)	2 (20%)	1	0
All	All	807/813 (99%)	788 (98%)	19 (2%)	54	47

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

Mol	Chain	Res	Type
1	A	108	LEU
1	A	134	LEU
1	A	277	TYR
1	A	286	ASN
1	A	290	GLN
1	A	335	LYS
1	A	513	TRP
1	A	533	ARG
1	A	603	GLN
1	A	644	GLN
1	A	883	LYS
1	A	887	ASP
1	A	893	PHE
1	A	922	ASN
1	A	923	MET
1	A	927	PHE
1	A	932	ARG
2	B	10	LEU
2	B	11	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	208	GLN
1	A	448	HIS
1	A	775	GLN
1	A	782	ASN
1	A	886	GLN
1	A	897	ASN
1	A	921	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 ⓘ

21 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	0.55	0	15,19,21	0.83	0
3	NAG	A	1002	3	14,14,15	0.50	0	15,19,21	1.43	1 (6%)
3	NAG	A	1003	3	14,14,15	0.72	0	15,19,21	1.37	2 (13%)
3	NAG	A	1004	1,3	14,14,15	0.56	0	15,19,21	0.91	0
3	NAG	A	1005	3	14,14,15	0.71	0	15,19,21	1.65	1 (6%)
3	NAG	A	1006	3	14,14,15	0.53	0	15,19,21	0.82	0
4	NAG	A	1007	1,4	14,14,15	0.50	0	15,19,21	0.60	0
4	NAG	A	1008	4	14,14,15	0.56	0	15,19,21	0.80	0
4	NAG	A	1009	1,4	14,14,15	0.48	0	15,19,21	0.91	0
4	NAG	A	1010	4	14,14,15	0.49	0	15,19,21	0.89	0
3	NAG	A	1011	1,3	14,14,15	0.46	0	15,19,21	1.15	1 (6%)
3	NAG	A	1012	3	14,14,15	0.54	0	15,19,21	1.21	1 (6%)
3	NAG	A	1013	3	14,14,15	0.51	0	15,19,21	1.86	2 (13%)
4	NAG	A	1014	1,4	14,14,15	0.54	0	15,19,21	0.90	0
4	NAG	A	1015	4	14,14,15	0.48	0	15,19,21	0.92	1 (6%)
4	NAG	A	1016	1,4	14,14,15	0.42	0	15,19,21	1.88	3 (20%)
4	NAG	A	1017	4	14,14,15	0.48	0	15,19,21	0.79	0
4	NAG	A	1018	1,4	14,14,15	0.50	0	15,19,21	1.50	4 (26%)
4	NAG	A	1019	4	14,14,15	0.44	0	15,19,21	0.97	0
4	NAG	A	1021	1,4	14,14,15	0.50	0	15,19,21	1.01	1 (6%)
4	NAG	A	1022	4	14,14,15	0.45	0	15,19,21	1.00	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	1003	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1004	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1005	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1006	3	-	0/6/23/26	0/1/1/1
4	NAG	A	1007	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1008	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1009	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1010	4	-	0/6/23/26	0/1/1/1
3	NAG	A	1011	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1012	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1013	3	-	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	1016	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1017	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1018	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1019	4	-	0/6/23/26	0/1/1/1
4	NAG	A	1021	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1022	4	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1018	NAG	O5-C1-C2	-3.09	107.17	111.47
4	A	1016	NAG	C2-N2-C7	-2.58	119.17	122.94
4	A	1016	NAG	C1-C2-N2	-2.38	106.42	110.49
4	A	1018	NAG	C2-N2-C7	2.03	125.90	122.94
4	A	1018	NAG	C8-C7-N2	2.08	119.87	116.11
4	A	1018	NAG	C1-O5-C5	2.13	115.10	112.17
3	A	1003	NAG	O5-C1-C2	2.32	114.71	111.47
3	A	1013	NAG	C3-C4-C5	2.34	114.35	110.22
4	A	1022	NAG	C1-O5-C5	2.58	115.73	112.17
4	A	1021	NAG	C1-O5-C5	2.78	116.00	112.17
4	A	1015	NAG	C1-O5-C5	2.79	116.02	112.17
3	A	1011	NAG	C1-O5-C5	2.97	116.26	112.17
3	A	1002	NAG	C4-C3-C2	3.13	115.60	111.02
3	A	1003	NAG	C4-C3-C2	3.55	116.22	111.02
3	A	1012	NAG	C4-C3-C2	3.82	116.61	111.02

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1005	NAG	C4-C3-C2	4.91	118.22	111.02
4	A	1016	NAG	C1-O5-C5	5.73	120.06	112.17
3	A	1013	NAG	C2-N2-C7	5.75	131.33	122.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1002	NAG	1	0
3	A	1003	NAG	2	0
3	A	1005	NAG	1	0
3	A	1006	NAG	1	0
4	A	1018	NAG	2	0
4	A	1021	NAG	1	0

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	1020	1	14,14,15	0.49	0	15,19,21	1.44	3 (20%)
5	NAG	A	1023	1	14,14,15	0.54	0	15,19,21	0.66	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	1020	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	A	1023	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1020	NAG	O5-C1-C2	-2.02	108.66	111.47
5	A	1020	NAG	C1-O5-C5	2.11	115.07	112.17
5	A	1020	NAG	C2-N2-C7	3.56	128.13	122.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	902/908 (99%)	0.96	164 (18%)  	16, 23, 37, 54	0
2	B	11/11 (100%)	4.13	11 (100%)  	63, 73, 84, 86	0
All	All	913/919 (99%)	0.99	175 (19%)  	16, 23, 39, 86	0

All (175) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	885	PHE	14.4
1	A	888	TYR	12.3
1	A	925	VAL	10.7
1	A	893	PHE	9.6
1	A	889	GLY	8.7
1	A	275	LEU	7.1
1	A	927	PHE	6.9
2	B	8	PHE	6.8
1	A	882	LYS	6.7
1	A	890	GLY	6.5
1	A	964	SER	6.4
1	A	932	ARG	6.2
1	A	895	PHE	6.1
1	A	923	MET	6.1
2	B	11	MET	5.8
1	A	884	LEU	5.8
1	A	198	VAL	5.6
1	A	928	GLY	5.6
1	A	216	PHE	5.6
1	A	891	GLY	5.5
1	A	274	LEU	5.4
1	A	924	ASP	5.3
1	A	929	SER	5.2
2	B	7	PHE	5.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	218	CYS	5.0
1	A	197	ASN	5.0
1	A	926	GLY	5.0
1	A	83	VAL	4.8
1	A	93	ALA	4.7
1	A	63	GLN	4.7
1	A	290	GLN	4.7
1	A	922	ASN	4.7
1	A	881	TRP	4.6
1	A	886	GLN	4.6
1	A	933	ALA	4.5
1	A	781	GLU	4.5
1	A	892	SER	4.4
1	A	212	ALA	4.3
1	A	272	THR	4.3
1	A	728	GLN	4.3
1	A	273	TYR	4.1
1	A	276	ALA	4.1
1	A	270	MET	4.1
1	A	930	GLY	4.1
1	A	732	THR	4.0
2	B	9	GLY	4.0
2	B	4	PRO	4.0
1	A	228	PHE	3.9
1	A	355	VAL	3.9
1	A	352	TRP	3.8
1	A	479	ILE	3.8
2	B	10	LEU	3.8
1	A	277	TYR	3.8
1	A	354	LEU	3.8
1	A	649	VAL	3.7
2	B	3	LYS	3.7
2	B	5	GLN	3.7
1	A	196	GLY	3.7
1	A	271	SER	3.7
1	A	780	PRO	3.6
1	A	349	MET	3.6
1	A	920	LYS	3.6
1	A	390	PHE	3.5
1	A	81	TYR	3.5
1	A	94	ASP	3.5
1	A	851	ILE	3.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	219	PHE	3.4
1	A	230	ILE	3.4
1	A	395	THR	3.3
1	A	207	MET	3.3
1	A	736	ASN	3.3
2	B	2	PRO	3.3
1	A	894	SER	3.3
1	A	848	PRO	3.2
1	A	394	VAL	3.2
2	B	6	GLN	3.2
1	A	217	PRO	3.2
1	A	118	ILE	3.1
1	A	278	ILE	3.1
1	A	351	ASN	3.1
1	A	853	LYS	3.1
1	A	931	THR	3.1
1	A	386	ALA	3.1
1	A	738	THR	3.0
1	A	737	TRP	3.0
1	A	849	ASP	3.0
1	A	215	SER	3.0
1	A	567	GLU	3.0
1	A	645	THR	2.9
1	A	613	SER	2.9
1	A	555	GLY	2.9
1	A	402	LEU	2.9
1	A	95	GLY	2.9
1	A	887	ASP	2.9
1	A	392	ASN	2.8
1	A	778	SER	2.8
1	A	782	ASN	2.8
1	A	733	LEU	2.8
1	A	103	SER	2.8
1	A	729	HIS	2.8
1	A	918	PHE	2.8
1	A	934	LEU	2.8
1	A	742	GLU	2.8
1	A	356	THR	2.8
1	A	409	ALA	2.7
1	A	641	HIS	2.7
1	A	188	PHE	2.7
1	A	204	THR	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	644	GLN	2.6
1	A	717	LYS	2.6
1	A	226	ALA	2.6
1	A	899	ILE	2.6
1	A	408	PHE	2.6
1	A	949	LYS	2.6
1	A	243	ASN	2.6
1	A	430	ILE	2.6
1	A	761	GLY	2.6
1	A	815	GLN	2.6
1	A	389	TRP	2.6
1	A	269	VAL	2.6
1	A	393	LEU	2.5
1	A	762	LEU	2.5
1	A	919	LYS	2.5
1	A	638	MET	2.5
1	A	616	TRP	2.5
1	A	177	PHE	2.5
1	A	64	SER	2.5
1	A	367	GLN	2.5
1	A	223	ALA	2.5
1	A	85	LEU	2.5
1	A	357	TYR	2.5
1	A	391	GLY	2.5
1	A	412	VAL	2.5
1	A	767	ASN	2.4
1	A	205	THR	2.4
1	A	224	MET	2.4
1	A	852	ARG	2.4
1	A	78	PRO	2.4
1	A	936	GLN	2.4
1	A	614	ASP	2.3
1	A	955	VAL	2.3
1	A	847	ASN	2.3
1	A	396	LEU	2.3
1	A	387	HIS	2.3
1	A	612	ALA	2.3
1	A	210	THR	2.2
1	A	776	TRP	2.2
1	A	105	VAL	2.2
1	A	380	VAL	2.2
1	A	591	VAL	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	108	LEU	2.2
1	A	404	LEU	2.2
1	A	846	LEU	2.2
1	A	850	LEU	2.2
1	A	339	ILE	2.2
1	A	476	ALA	2.2
1	A	232	LEU	2.1
1	A	952	LYS	2.1
1	A	220	ASP	2.1
1	A	764	GLN	2.1
1	A	610	LYS	2.1
2	B	1	ARG	2.1
1	A	65	LYS	2.1
1	A	399	TRP	2.1
1	A	963	SER	2.1
1	A	403	TRP	2.0
1	A	222	PRO	2.0
1	A	647	LEU	2.0
1	A	291	ASN	2.0
1	A	209	SER	2.0
1	A	383	HIS	2.0
1	A	844	TYR	2.0
1	A	478	VAL	2.0
1	A	883	LYS	2.0
1	A	242	SER	2.0

## 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	1011	14/15	0.84	0.13	2.64	49,58,63,63	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	1021	14/15	0.90	0.13	0.70	47,61,67,67	0
4	NAG	A	1007	14/15	0.90	0.16	-0.34	36,40,51,51	0
3	NAG	A	1004	14/15	0.85	0.11	-0.40	49,54,60,67	0
4	NAG	A	1016	14/15	0.92	0.17	-0.46	42,46,52,53	0
3	NAG	A	1001	14/15	0.88	0.13	-1.33	38,45,50,55	0
4	NAG	A	1018	14/15	0.62	0.17	-	59,70,76,81	0
4	NAG	A	1015	14/15	0.77	0.32	-	46,77,83,86	0
3	NAG	A	1002	14/15	0.90	0.23	-	59,69,80,83	0
4	NAG	A	1014	14/15	0.80	0.22	-	53,68,74,83	0
4	NAG	A	1017	14/15	0.87	0.15	-	55,60,77,78	0
3	NAG	A	1006	14/15	0.48	0.56	-	74,83,94,97	0
3	NAG	A	1013	14/15	0.63	0.37	-	65,83,87,92	0
4	NAG	A	1019	14/15	0.69	0.35	-	75,81,89,89	0
3	NAG	A	1012	14/15	0.82	0.18	-	58,68,79,85	0
4	NAG	A	1010	14/15	0.62	0.29	-	71,74,86,86	0
4	NAG	A	1022	14/15	0.86	0.25	-	62,71,91,93	0
3	NAG	A	1003	14/15	0.76	0.48	-	75,80,88,97	0
3	NAG	A	1005	14/15	0.77	0.32	-	66,73,82,88	0
4	NAG	A	1009	14/15	0.84	0.11	-	52,62,72,73	0
4	NAG	A	1008	14/15	0.81	0.17	-	53,61,73,80	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
6	ZN	A	1024	1/1	1.00	0.28	0.53	36,36,36,36	0
5	NAG	A	1023	14/15	0.83	0.25	-	64,79,88,91	0
5	NAG	A	1020	14/15	0.77	0.48	-	68,81,97,99	0

## 6.5 Other polymers

There are no such residues in this entry.