



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 30, 2017 – 11:41 PM EDT

PDB ID : 3NKO  
Title : Crystal structure of mouse autotaxin in complex with 16:0-LPA  
Authors : Nishimasu, H.; Ishitani, R.; Mihara, E.; Takagi, J.; Aoki, J.; Nureki, O.  
Deposited on : unknown  
Resolution : 1.75 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 : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

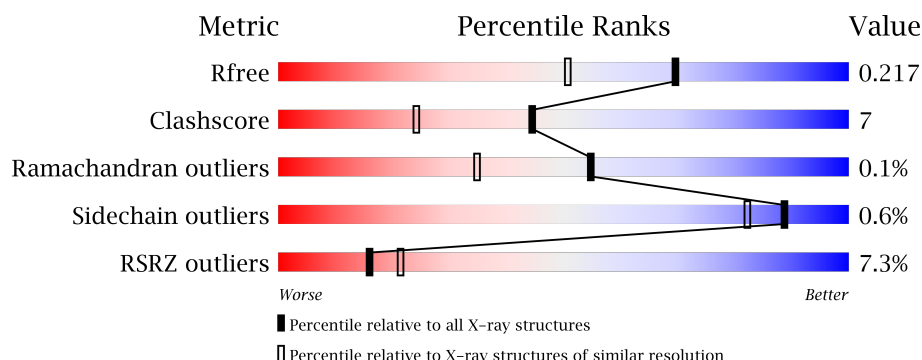
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	1762 (1.76-1.76)
Clashscore	112137	1889 (1.76-1.76)
Ramachandran outliers	110173	1868 (1.76-1.76)
Sidechain outliers	110143	1868 (1.76-1.76)
RSRZ outliers	101464	1770 (1.76-1.76)

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	831	<div> <div>7%</div> <div>82%</div> <div>12%</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	SCN	A	1008	-	-	-	X
11	EDO	A	1010	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
11	EDO	A	1011	-	-	-	X
11	EDO	A	1013	-	-	-	X
11	EDO	A	1016	-	-	-	X
11	EDO	A	1020	-	-	-	X
11	EDO	A	1024	-	-	X	X
11	EDO	A	1025	-	-	X	-
11	EDO	A	1026	-	-	X	X
11	EDO	A	1028	-	-	X	X
4	MAN	A	904	-	-	-	X
9	NKO	A	1007	-	-	-	X

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 7232 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 Ectonucleotide pyrophosphatase/phosphodiesterase family member 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	781	Total	C	N	O	S	0	5	0
			6248	3970	1066	1163	49			

There are 12 discrepancies between the modelled and reference sequences:

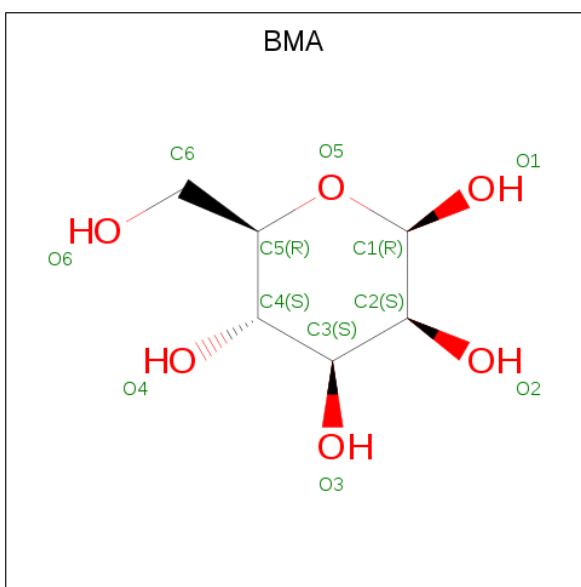
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	SEE REMARK 999	UNP Q9R1E6
A	?	-	VAL	SEE REMARK 999	UNP Q9R1E6
A	?	-	GLU	SEE REMARK 999	UNP Q9R1E6
A	?	-	PRO	SEE REMARK 999	UNP Q9R1E6
A	859	SER	-	EXPRESSION TAG	UNP Q9R1E6
A	860	ARG	-	EXPRESSION TAG	UNP Q9R1E6
A	861	GLU	-	EXPRESSION TAG	UNP Q9R1E6
A	862	ASN	-	EXPRESSION TAG	UNP Q9R1E6
A	863	LEU	-	EXPRESSION TAG	UNP Q9R1E6
A	864	TYR	-	EXPRESSION TAG	UNP Q9R1E6
A	865	PHE	-	EXPRESSION TAG	UNP Q9R1E6
A	866	GLN	-	EXPRESSION TAG	UNP Q9R1E6

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



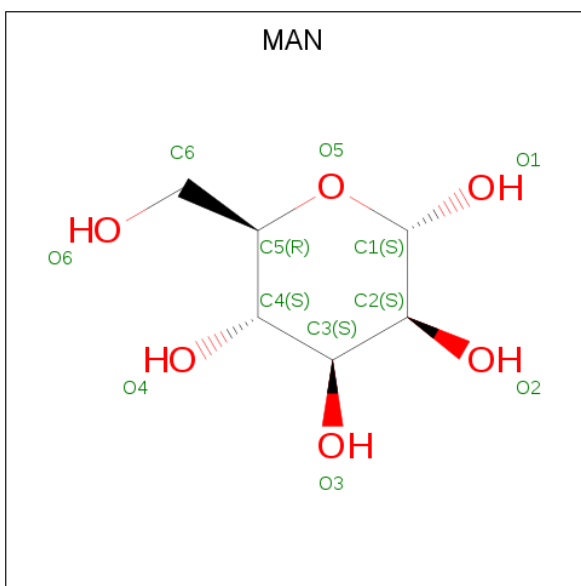
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Zn	0	0
			2	2		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

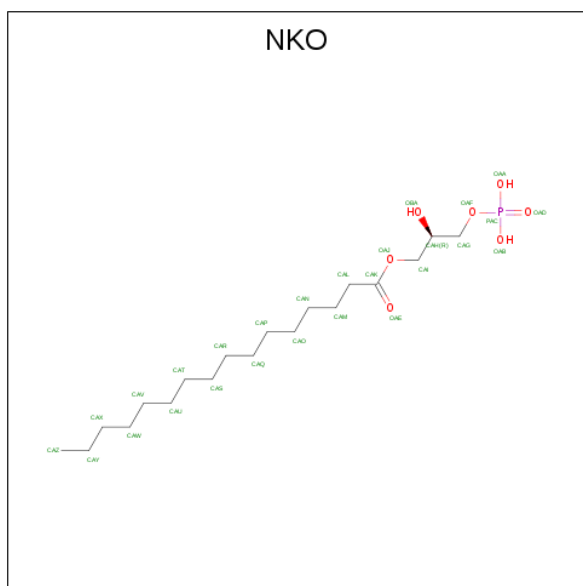
- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Na	0	0
			1	1		

- Molecule 8 is POTASSIUM ION (three-letter code: K) (formula: K).

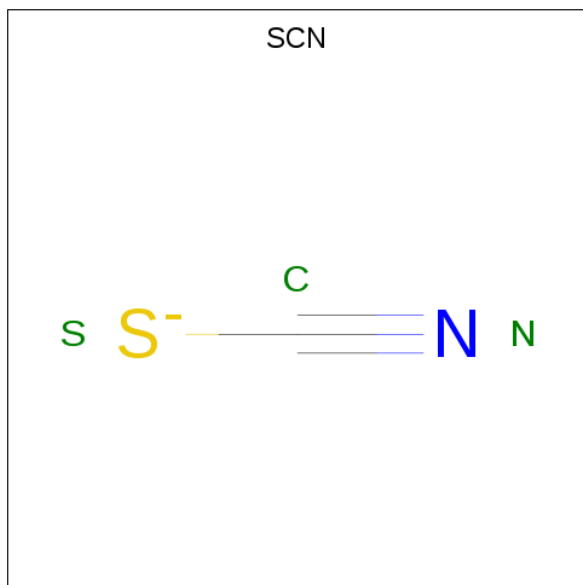
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	K	0	0
			1	1		

- Molecule 9 is (2R)-2-hydroxy-3-(phosphonoxy)propyl hexadecanoate (three-letter code: NKO) (formula: C<sub>19</sub>H<sub>39</sub>O<sub>7</sub>P).



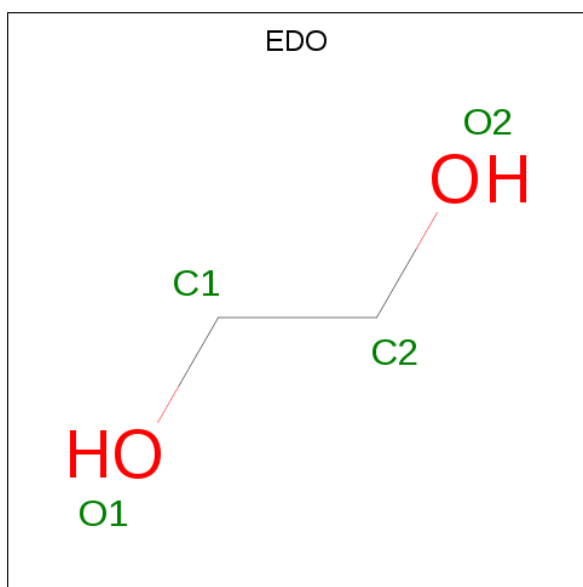
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C O P 27 19 7 1	0	0
9	A	1	Total C 8 8	0	0

- Molecule 10 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	1	Total C N S 3 1 1 1	0	0
10	A	1	Total C N S 3 1 1 1	0	0

- Molecule 11 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		
11	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	A	1	Total 4	C 2	O 2	0	0
11	A	1	Total 4	C 2	O 2	0	0
11	A	1	Total 4	C 2	O 2	0	0
11	A	1	Total 4	C 2	O 2	0	0
11	A	1	Total 4	C 2	O 2	0	0
11	A	1	Total 4	C 2	O 2	0	0

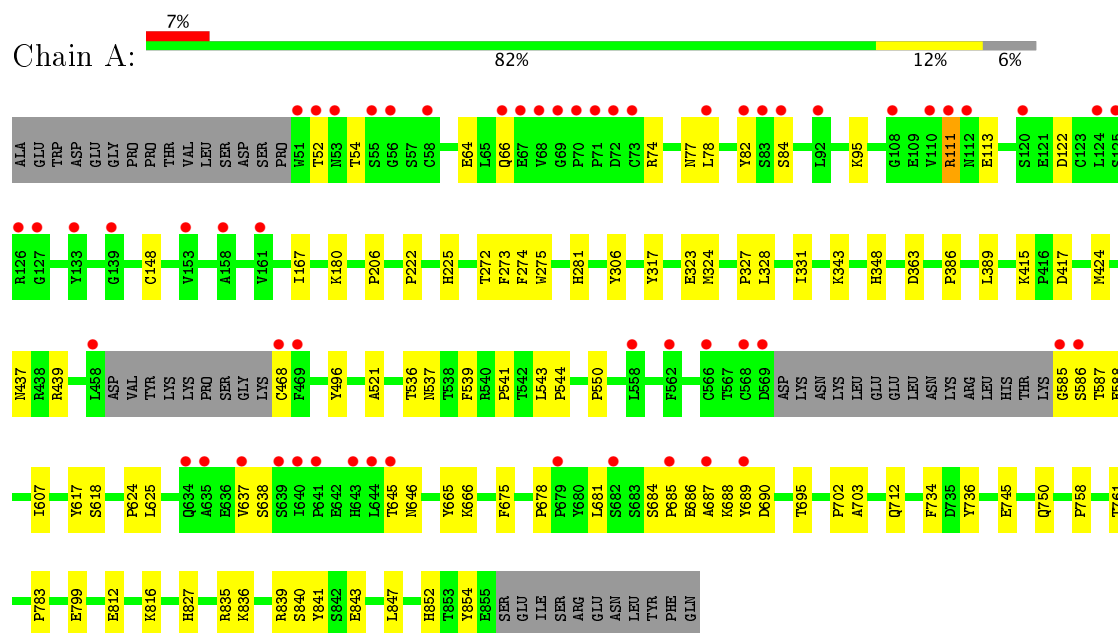
- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	730	Total 730	O 730	0	0

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

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

- Molecule 1: Ectonucleotide pyrophosphatase/phosphodiesterase family member 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.55Å 94.40Å 75.24Å 90.00° 94.98° 90.00°	Depositor
Resolution (Å)	26.45 – 1.75 49.62 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.4 (26.45-1.75) 99.4 (49.62-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.34 (at 1.75Å)	Xtriage
Refinement program	PHENIX 1.6 _289	Depositor
R, $R_{free}$	0.184 , 0.224 0.178 , 0.217	Depositor DCC
$R_{free}$ test set	4285 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.5	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 54.1	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	0.96	EDS
Total number of atoms	7232	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 4.48% 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, BMA, NAG, NA, K, EDO, NKO, SCN, CA, MAN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.47	0/6444	0.59	0/8765

There are no bond length outliers.

There are no bond angle outliers.

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	6248	0	5926	91	0
2	A	84	0	74	1	0
3	A	11	0	9	0	0
4	A	33	0	28	0	0
5	A	2	0	0	0	0
6	A	1	0	0	0	0
7	A	1	0	0	0	0
8	A	1	0	0	0	0
9	A	35	0	52	7	0
10	A	6	0	0	1	0
11	A	80	0	120	29	0
12	A	730	0	0	14	0
All	All	7232	0	6209	95	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 (95) 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:424:MET:HG3	12:A:1299:HOH:O	1.59	1.02
1:A:363:ASP:HB3	12:A:1467:HOH:O	1.63	0.97
1:A:437:ASN:HD21	11:A:1024:EDO:H22	1.41	0.85
1:A:585:GLY:N	1:A:587:THR:H	1.73	0.85
1:A:702:PRO:HB2	11:A:1027:EDO:H22	1.61	0.81
1:A:272:THR:HG22	1:A:274:PHE:H	1.44	0.79
1:A:852:HIS:HD2	1:A:854:TYR:CZ	2.01	0.79
1:A:389:LEU:HD12	12:A:1093:HOH:O	1.83	0.79
1:A:306:TYR:HB2	9:A:1006:NKO:HASA	1.66	0.77
1:A:675:PHE:H	11:A:1012:EDO:H11	1.50	0.74
1:A:585:GLY:HA2	12:A:1495:HOH:O	1.88	0.74
1:A:389:LEU:CD1	12:A:1093:HOH:O	2.35	0.73
1:A:836:LYS:HE2	11:A:1026:EDO:H21	1.71	0.71
1:A:78:LEU:O	1:A:82:TYR:HD1	1.73	0.70
1:A:675:PHE:H	11:A:1012:EDO:C1	2.07	0.66
1:A:839:ARG:HE	11:A:1028:EDO:H21	1.62	0.65
1:A:111:ARG:NH1	1:A:122:ASP:OD1	2.30	0.65
1:A:437:ASN:ND2	11:A:1024:EDO:H22	2.14	0.62
1:A:95:LYS:HE3	1:A:113:GLU:O	2.01	0.61
1:A:835:ARG:HE	11:A:1026:EDO:C2	2.14	0.60
1:A:835:ARG:HH21	11:A:1026:EDO:H11	1.66	0.60
1:A:839:ARG:NE	11:A:1028:EDO:H21	2.17	0.60
1:A:281:HIS:ND1	11:A:1025:EDO:H21	2.17	0.59
1:A:685:PRO:O	1:A:689:TYR:HD2	1.87	0.58
1:A:835:ARG:HE	11:A:1026:EDO:H22	1.69	0.57
1:A:839:ARG:HE	11:A:1028:EDO:C2	2.17	0.57
1:A:758:PRO:HD3	11:A:1019:EDO:H22	1.87	0.56
1:A:734:PHE:CD2	11:A:1013:EDO:H11	2.41	0.56
1:A:343:LYS:HA	1:A:348:HIS:CD2	2.41	0.56
1:A:167:ILE:HD13	9:A:1006:NKO:HAZB	1.87	0.56
1:A:206:PRO:HB3	1:A:389:LEU:HD13	1.89	0.55
1:A:74:ARG:HH22	1:A:78:LEU:CB	2.19	0.55
1:A:585:GLY:HA3	12:A:1590:HOH:O	2.06	0.55
1:A:348:HIS:CD2	1:A:348:HIS:H	2.24	0.55
11:A:1011:EDO:H21	12:A:982:HOH:O	2.07	0.55
1:A:835:ARG:HB3	11:A:1026:EDO:H22	1.90	0.54
1:A:678:PRO:HB3	1:A:712:GLN:HB3	1.90	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:703:ALA:N	11:A:1027:EDO:H21	2.23	0.54
1:A:468:CYS:O	12:A:1310:HOH:O	2.18	0.53
1:A:281:HIS:ND1	11:A:1025:EDO:C2	2.71	0.52
1:A:64:GLU:HB2	1:A:66:GLN:HE21	1.74	0.52
1:A:827:HIS:HE1	12:A:897:HOH:O	1.92	0.51
1:A:550:PRO:HB2	1:A:607:ILE:HG12	1.92	0.51
1:A:222:PRO:HA	1:A:225:HIS:CE1	2.46	0.50
1:A:306:TYR:CG	9:A:1006:NKO:HAQA	2.46	0.50
1:A:331:ILE:CG1	11:A:1025:EDO:H11	2.42	0.50
1:A:586:SER:HA	12:A:1495:HOH:O	2.12	0.49
1:A:702:PRO:HB2	11:A:1027:EDO:C2	2.38	0.49
1:A:389:LEU:HD11	10:A:1008:SCN:S	2.53	0.49
1:A:273:PHE:HE2	9:A:1006:NKO:HAZA	1.77	0.49
1:A:617:TYR:CE1	1:A:695:THR:HG22	2.48	0.48
1:A:687:ALA:O	1:A:690:ASP:HB2	2.14	0.48
1:A:799:GLU:HG3	12:A:1508:HOH:O	2.13	0.47
1:A:645:THR:HG22	1:A:646:ASN:ND2	2.29	0.47
1:A:827:HIS:HD2	12:A:1298:HOH:O	1.96	0.47
1:A:77:ASN:OD1	1:A:275:TRP:HA	2.14	0.47
1:A:585:GLY:CA	1:A:587:THR:H	2.28	0.47
1:A:417:ASP:OD1	1:A:417:ASP:N	2.44	0.47
1:A:273:PHE:CE2	9:A:1006:NKO:HAZA	2.50	0.47
1:A:539:PHE:O	1:A:541:PRO:HD3	2.15	0.47
1:A:439:ARG:CZ	11:A:1024:EDO:H21	2.45	0.46
1:A:840:SER:OG	1:A:843:GLU:HG3	2.15	0.46
1:A:761:THR:HA	11:A:1015:EDO:H12	1.97	0.46
1:A:437:ASN:HD21	11:A:1024:EDO:C2	2.18	0.46
1:A:74:ARG:HD3	1:A:84:SER:OG	2.16	0.46
9:A:1007:NKO:CAS	12:A:1600:HOH:O	2.63	0.46
1:A:541:PRO:HB2	1:A:841:TYR:CE2	2.50	0.45
1:A:617:TYR:HA	1:A:624:PRO:HA	1.98	0.45
1:A:521:ALA:O	11:A:1018:EDO:H22	2.16	0.45
1:A:736:TYR:OH	1:A:750:GLN:HB3	2.17	0.45
1:A:585:GLY:HA2	1:A:586:SER:HA	1.65	0.45
1:A:812:GLU:O	1:A:816:LYS:HG3	2.17	0.44
9:A:1006:NKO:HAR	9:A:1006:NKO:HAOA	1.60	0.44
1:A:148[B]:CYS:SG	1:A:496:TYR:HE2	2.41	0.43
1:A:618:SER:HB2	1:A:625:LEU:HD21	2.00	0.43
1:A:783:PRO:HB3	11:A:1028:EDO:H12	2.00	0.43
1:A:331:ILE:HG12	11:A:1025:EDO:H11	2.01	0.43
1:A:543:LEU:HD12	1:A:544:PRO:HD2	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:684:SER:O	1:A:687:ALA:N	2.51	0.42
1:A:665:TYR:O	1:A:666:LYS:C	2.58	0.42
1:A:637:VAL:HG22	1:A:681:LEU:CD2	2.49	0.42
1:A:745:GLU:HG3	12:A:1522:HOH:O	2.19	0.42
1:A:852:HIS:HD2	1:A:854:TYR:CE1	2.37	0.42
1:A:52:THR:HG21	1:A:84:SER:HA	2.02	0.41
1:A:415:LYS:HB3	1:A:415:LYS:HE2	1.64	0.41
1:A:324:MET:C	1:A:327:PRO:HD2	2.41	0.41
1:A:536:THR:O	1:A:537:ASN:HB2	2.21	0.41
1:A:550:PRO:CB	1:A:607:ILE:HG12	2.50	0.41
1:A:638:SER:OG	1:A:688:LYS:NZ	2.53	0.41
1:A:272:THR:HG21	1:A:274:PHE:O	2.20	0.41
1:A:317:TYR:CG	1:A:323:GLU:HB2	2.56	0.40
1:A:847:LEU:HD12	11:A:1020:EDO:H11	2.04	0.40
1:A:331:ILE:HG13	11:A:1025:EDO:H11	2.02	0.40
1:A:180:LYS:HD2	1:A:328:LEU:HD13	2.03	0.40
2:A:898:NAG:H2	2:A:898:NAG:H83	1.92	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	780/831 (94%)	752 (96%)	27 (4%)	1 (0%)	55	35

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	386	PRO

### 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	695/756 (92%)	691 (99%)	4 (1%)	89	82

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	THR
1	A	111	ARG
1	A	588	GLU
1	A	686	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	66	GLN
1	A	134	GLN
1	A	337	GLN
1	A	348	HIS
1	A	374	ASN
1	A	398	ASN
1	A	646	ASN
1	A	667	ASN
1	A	827	HIS

### 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

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 39 ligands modelled in this entry, 5 are monoatomic - leaving 34 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	$\# Z  > 2$
9	NKO	A	1006	5	26,26,26	0.66	1 (3%)	29,30,30	1.42	3 (10%)
9	NKO	A	1007	-	7,7,26	0.55	0	6,6,30	1.56	3 (50%)
10	SCN	A	1008	-	1,2,2	0.82	0	0,1,1	0.00	-
10	SCN	A	1009	-	1,2,2	0.98	0	0,1,1	0.00	-
11	EDO	A	1010	-	3,3,3	0.48	0	2,2,2	0.49	0
11	EDO	A	1011	-	3,3,3	0.67	0	2,2,2	0.24	0
11	EDO	A	1012	-	3,3,3	0.31	0	2,2,2	0.62	0
11	EDO	A	1013	-	3,3,3	0.58	0	2,2,2	0.29	0
11	EDO	A	1014	-	3,3,3	0.57	0	2,2,2	0.15	0
11	EDO	A	1015	-	3,3,3	0.40	0	2,2,2	0.37	0
11	EDO	A	1016	-	3,3,3	0.58	0	2,2,2	0.20	0
11	EDO	A	1017	-	3,3,3	0.61	0	2,2,2	0.14	0
11	EDO	A	1018	-	3,3,3	0.65	0	2,2,2	0.25	0
11	EDO	A	1019	-	3,3,3	0.40	0	2,2,2	0.44	0
11	EDO	A	1020	-	3,3,3	0.46	0	2,2,2	0.50	0
11	EDO	A	1021	-	3,3,3	0.45	0	2,2,2	0.47	0
11	EDO	A	1022	-	3,3,3	0.49	0	2,2,2	0.47	0
11	EDO	A	1023	-	3,3,3	0.52	0	2,2,2	0.40	0
11	EDO	A	1024	-	3,3,3	0.48	0	2,2,2	0.20	0
11	EDO	A	1025	-	3,3,3	0.43	0	2,2,2	0.99	0
11	EDO	A	1026	-	3,3,3	0.52	0	2,2,2	0.21	0
11	EDO	A	1027	-	3,3,3	0.45	0	2,2,2	0.40	0
11	EDO	A	1028	-	3,3,3	0.36	0	2,2,2	0.37	0
11	EDO	A	1029	-	3,3,3	0.47	0	2,2,2	0.38	0
2	NAG	A	898	1,2	14,14,15	0.55	0	15,19,21	0.88	1 (6%)
2	NAG	A	899	2	14,14,15	0.55	0	15,19,21	0.66	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	A	900	1,2	14,14,15	0.60	0	15,19,21	1.55	3 (20%)
2	NAG	A	901	3,2	14,14,15	0.72	1 (7%)	15,19,21	1.05	0
3	BMA	A	902	2,4	11,11,12	0.64	0	13,15,17	1.31	1 (7%)
4	MAN	A	903	3,4	11,11,12	0.60	0	13,15,17	1.11	1 (7%)
4	MAN	A	904	4	11,11,12	0.61	0	13,15,17	1.04	1 (7%)
4	MAN	A	905	4	11,11,12	0.60	0	13,15,17	1.14	2 (15%)
2	NAG	A	907	1,2	14,14,15	0.60	0	15,19,21	1.26	2 (13%)
2	NAG	A	908	2	14,14,15	0.68	0	15,19,21	0.75	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
9	NKO	A	1006	5	-	0/26/26/26	0/0/0/0
9	NKO	A	1007	-	-	0/5/5/26	0/0/0/0
10	SCN	A	1008	-	-	0/0/0/0	0/0/0/0
10	SCN	A	1009	-	-	0/0/0/0	0/0/0/0
11	EDO	A	1010	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1011	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1012	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1013	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1014	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1015	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1016	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1017	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1018	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1019	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1020	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1021	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1022	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1023	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1024	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1025	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1026	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1027	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1028	-	-	0/1/1/1	0/0/0/0
11	EDO	A	1029	-	-	0/1/1/1	0/0/0/0
2	NAG	A	898	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	899	2	-	0/6/23/26	0/1/1/1

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	900	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	901	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	902	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	903	3,4	-	0/2/19/22	0/1/1/1
4	MAN	A	904	4	-	0/2/19/22	0/1/1/1
4	MAN	A	905	4	-	0/2/19/22	0/1/1/1
2	NAG	A	907	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	908	2	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	901	NAG	O5-C1	-2.24	1.40	1.43
9	A	1006	NKO	PAC-OAB	2.05	1.63	1.54

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	902	BMA	O5-C1-C2	-3.78	104.87	110.79
2	A	900	NAG	O5-C1-C2	-3.41	106.72	111.47
9	A	1006	NKO	CAQ-CAP-CAO	-3.31	97.41	114.45
2	A	907	NAG	C6-C5-C4	-2.64	106.82	113.00
4	A	904	MAN	C1-O5-C5	-2.58	108.61	112.17
2	A	900	NAG	O7-C7-C8	-2.48	117.55	122.06
4	A	903	MAN	C2-C3-C4	-2.42	106.65	110.88
4	A	905	MAN	O2-C2-C3	-2.16	105.93	110.17
9	A	1006	NKO	CAX-CAW-CAV	-2.15	103.38	114.45
2	A	907	NAG	O5-C1-C2	-2.12	108.53	111.47
4	A	905	MAN	C2-C3-C4	-2.02	107.35	110.88
9	A	1007	NKO	CAX-CAW-CAV	-2.02	104.06	114.45
9	A	1007	NKO	CAS-CAT-CAU	2.12	129.98	113.42
2	A	898	NAG	C1-O5-C5	2.22	115.22	112.17
2	A	900	NAG	C1-C2-N2	2.26	114.34	110.49
9	A	1007	NKO	CAW-CAV-CAU	2.30	126.30	114.45
9	A	1006	NKO	OAA-PAC-OAF	3.58	116.25	106.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

16 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	1006	NKO	6	0
9	A	1007	NKO	1	0
10	A	1008	SCN	1	0
11	A	1011	EDO	1	0
11	A	1012	EDO	2	0
11	A	1013	EDO	1	0
11	A	1015	EDO	1	0
11	A	1018	EDO	1	0
11	A	1019	EDO	1	0
11	A	1020	EDO	1	0
11	A	1024	EDO	4	0
11	A	1025	EDO	5	0
11	A	1026	EDO	5	0
11	A	1027	EDO	3	0
11	A	1028	EDO	4	0
2	A	898	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 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	781/831 (93%)	0.39	57 (7%) 16 21	14, 31, 56, 71	0

All (57) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	458	LEU	6.2
1	A	110	VAL	5.8
1	A	51	TRP	5.7
1	A	644	LEU	5.1
1	A	585	GLY	5.0
1	A	78	LEU	4.9
1	A	637	VAL	4.8
1	A	640	ILE	4.7
1	A	469	PHE	4.6
1	A	689	TYR	4.3
1	A	83	SER	4.2
1	A	124	LEU	4.1
1	A	52	THR	4.1
1	A	69	GLY	4.0
1	A	70	PRO	4.0
1	A	641	PRO	3.8
1	A	55	SER	3.7
1	A	71	PRO	3.6
1	A	125	SER	3.5
1	A	568	CYS	3.5
1	A	586	SER	3.3
1	A	158	ALA	3.3
1	A	82	TYR	3.2
1	A	68	VAL	3.2
1	A	58	CYS	3.2
1	A	127	GLY	3.1
1	A	161	VAL	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	566	CYS	2.9
1	A	53	ASN	2.9
1	A	56	GLY	2.8
1	A	111	ARG	2.8
1	A	569	ASP	2.8
1	A	679	PRO	2.7
1	A	634	GLN	2.6
1	A	108	GLY	2.6
1	A	139	GLY	2.6
1	A	558	LEU	2.6
1	A	66	GLN	2.5
1	A	92	LEU	2.5
1	A	120	SER	2.5
1	A	126	ARG	2.3
1	A	73	CYS	2.3
1	A	639	SER	2.3
1	A	643	HIS	2.3
1	A	645	THR	2.2
1	A	635	ALA	2.2
1	A	72	ASP	2.2
1	A	468	CYS	2.1
1	A	153	VAL	2.1
1	A	67	GLU	2.1
1	A	562	PHE	2.1
1	A	685	PRO	2.1
1	A	133	TYR	2.1
1	A	682	SER	2.1
1	A	687	ALA	2.1
1	A	84	SER	2.0
1	A	112	ASN	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 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
4	MAN	A	904	11/12	0.91	0.19	7.50	30,34,41,47	0
9	NKO	A	1007	8/27	0.82	0.16	6.78	32,35,38,39	0
11	EDO	A	1013	4/4	0.94	0.17	6.72	34,35,35,37	0
10	SCN	A	1008	3/3	0.87	0.16	5.58	18,18,33,41	0
11	EDO	A	1010	4/4	0.96	0.17	5.02	36,36,36,38	0
11	EDO	A	1016	4/4	0.94	0.14	4.80	28,29,32,33	0
11	EDO	A	1024	4/4	0.94	0.12	3.25	21,30,33,35	0
11	EDO	A	1028	4/4	0.95	0.14	2.86	26,34,35,40	0
11	EDO	A	1020	4/4	0.94	0.13	2.43	28,33,39,43	0
11	EDO	A	1011	4/4	0.86	0.16	2.11	26,32,33,35	0
11	EDO	A	1026	4/4	0.79	0.17	2.03	44,44,47,52	0
11	EDO	A	1022	4/4	0.95	0.10	1.43	34,37,38,43	0
9	NKO	A	1006	27/27	0.96	0.14	1.14	19,38,47,50	0
11	EDO	A	1017	4/4	0.90	0.10	0.89	26,27,30,35	0
11	EDO	A	1019	4/4	0.96	0.12	0.76	37,39,42,45	0
11	EDO	A	1012	4/4	0.91	0.15	0.70	31,33,36,40	0
2	NAG	A	907	14/15	0.92	0.09	0.52	27,34,38,47	0
11	EDO	A	1014	4/4	0.93	0.11	0.19	30,32,33,35	0
2	NAG	A	898	14/15	0.84	0.22	-0.09	58,61,62,63	0
11	EDO	A	1015	4/4	0.95	0.09	-0.10	28,29,35,36	0
10	SCN	A	1009	3/3	0.99	0.09	-0.19	25,25,33,43	0
11	EDO	A	1025	4/4	0.94	0.11	-0.70	29,30,31,33	0
5	ZN	A	1002	1/1	1.00	0.09	-0.82	18,18,18,18	0
6	CA	A	1003	1/1	0.99	0.08	-1.03	25,25,25,25	0
2	NAG	A	900	14/15	0.98	0.09	-1.26	15,20,24,24	0
8	K	A	1005	1/1	0.96	0.07	-2.14	39,39,39,39	0
11	EDO	A	1023	4/4	0.99	0.06	-2.28	23,24,24,26	0
5	ZN	A	1001	1/1	1.00	0.08	-5.09	20,20,20,20	0
2	NAG	A	908	14/15	0.85	0.23	-	47,52,56,61	0
7	NA	A	1004	1/1	0.98	0.09	-	27,27,27,27	0
11	EDO	A	1027	4/4	0.95	0.23	-	35,38,39,43	0
3	BMA	A	902	11/12	0.86	0.12	-	39,43,53,54	0
2	NAG	A	899	14/15	0.84	0.20	-	60,65,67,70	0
4	MAN	A	905	11/12	0.95	0.09	-	23,28,36,37	0
11	EDO	A	1021	4/4	0.91	0.17	-	38,40,40,49	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
11	EDO	A	1018	4/4	0.84	0.16	-	23,29,31,37	0
11	EDO	A	1029	4/4	0.95	0.13	-	33,33,41,43	0
2	NAG	A	901	14/15	0.96	0.08	-	24,32,38,41	0
4	MAN	A	903	11/12	0.84	0.14	-	36,40,43,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.