



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 7, 2017 – 05:20 AM EST

PDB ID : 4GTX  
Title : Crystal structure of mouse Enpp1 in complex with TMP  
Authors : Kato, K.; Nishimasu, H.; Ishitani, R.; Nureki, O.  
Deposited on : unknown  
Resolution : 3.20 Å(reported)

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

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

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.7.2 (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

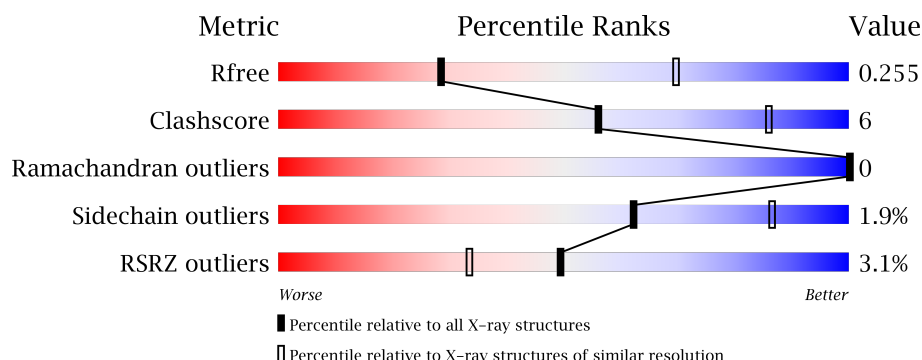
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*



The reported resolution of this entry is 3.20 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	1015 (3.22-3.18)
Clashscore	112137	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)
RSRZ outliers	101464	1020 (3.22-3.18)

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	823	 .% 73% 13% 14%
1	B	823	 4% 70% 14% 15%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 11211 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, Alkaline phosphodiesterase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	706	Total	C	N	O	S	0	0	0
			5534	3562	917	1026	29			
1	B	697	Total	C	N	O	S	0	0	0
			5473	3522	898	1024	29			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	ARG	LYS	ENGINEERED MUTATION	UNP Q9R1E6
B	59	ARG	LYS	ENGINEERED MUTATION	UNP Q9R1E6

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

*Continued on next page...*

Continued from previous page...

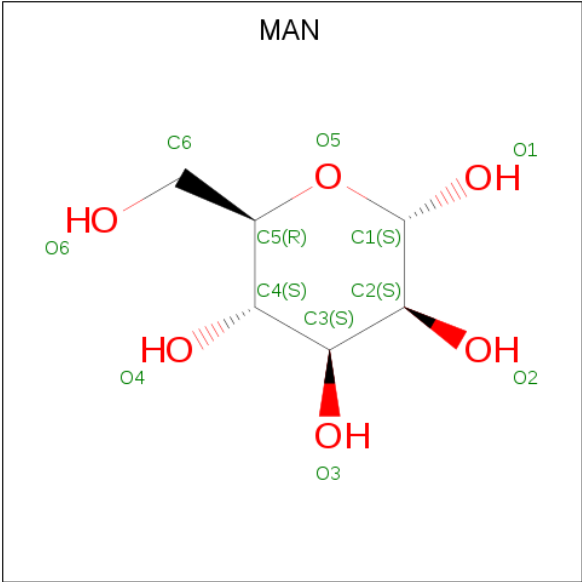
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

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



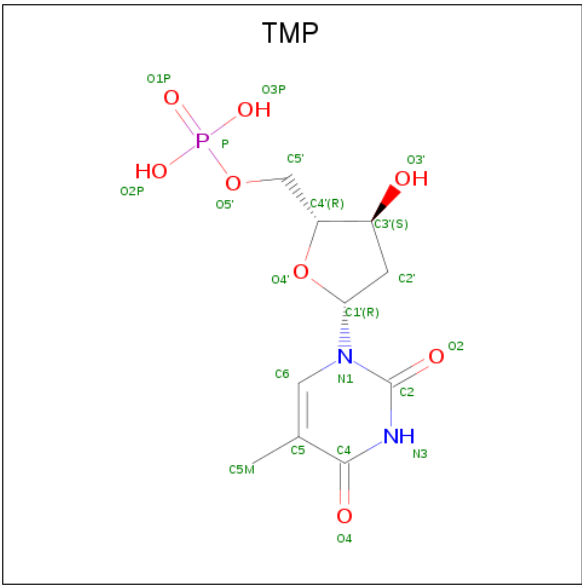
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula:  $C_6H_{12}O_6$ ).



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 THYMIDINE-5'-PHOSPHATE (three-letter code: TMP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>2</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			21	10	2	8	1		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			21	10	2	8	1		

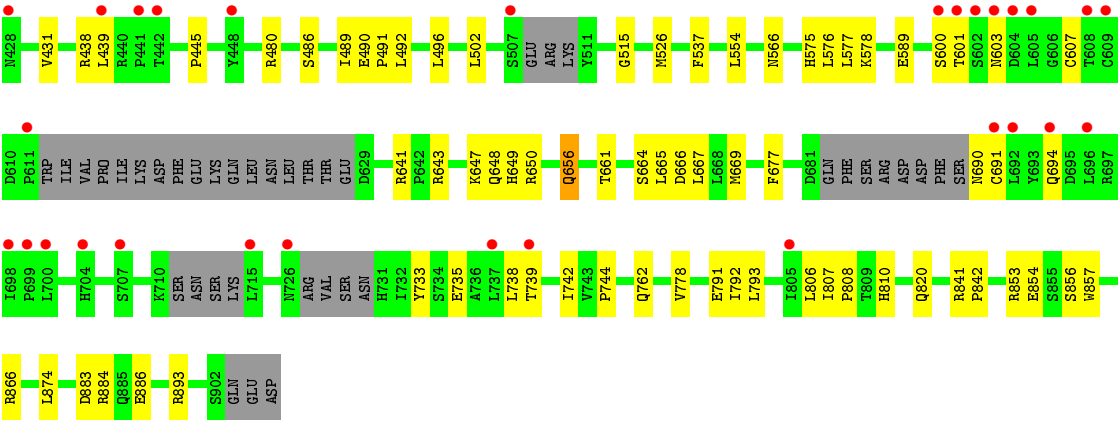
- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total	Zn	0	0
			2	2		
6	A	2	Total	Zn	0	0
			2	2		

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	1	Total	Ca	0	0
			1	1		
7	A	1	Total	Ca	0	0
			1	1		







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.72Å 105.72Å 174.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.09 – 3.20 49.09 – 3.20	Depositor EDS
% Data completeness (in resolution range)	97.1 (49.09-3.20) 97.1 (49.09-3.20)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.52 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.7.2_869	Depositor
R, $R_{free}$	0.218 , 0.265 0.214 , 0.255	Depositor DCC
$R_{free}$ test set	1763 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	56.3	Xtriage
Anisotropy	0.322	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 45.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.086 for -h,-k,l 0.099 for h,-h-k,-l 0.086 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	11211	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% 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: TMP, ZN, BMA, NAG, 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	0.23	0/5698	0.43	0/7781
1	B	0.23	0/5633	0.43	0/7691
All	All	0.23	0/11331	0.43	0/15472

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	5534	0	5219	61	0
1	B	5473	0	5147	64	0
2	A	56	0	50	1	0
2	B	56	0	51	1	0
3	A	11	0	8	0	0
4	A	33	0	29	1	0
5	A	21	0	13	2	0
5	B	21	0	13	2	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
7	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
7	B	1	0	0	0	0
All	All	11211	0	10530	126	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 6.

The worst 5 of 126 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:225:THR:HG21	1:A:576:LEU:HG	1.69	0.74
1:B:603:ASN:H	1:B:694:GLN:HE22	1.37	0.72
1:B:225:THR:HG21	1:B:576:LEU:HG	1.72	0.70
1:A:308:ASP:OD1	1:A:308:ASP:N	2.24	0.69
1:A:715:LEU:O	1:A:801:ARG:NH2	2.23	0.69

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	698/823 (85%)	675 (97%)	23 (3%)	0	100	100
1	B	685/823 (83%)	661 (96%)	24 (4%)	0	100	100
All	All	1383/1646 (84%)	1336 (97%)	47 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	600/755 (80%)	588 (98%)	12 (2%)	60	86
1	B	595/755 (79%)	584 (98%)	11 (2%)	64	87
All	All	1195/1510 (79%)	1172 (98%)	23 (2%)	62	86

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	874	LEU
1	B	197	PHE
1	B	792	ILE
1	A	886	GLU
1	B	236	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	741	ASN
1	A	769	ASN
1	B	694	GLN
1	A	658	GLN
1	B	656	GLN

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 20 ligands modelled in this entry, 6 are monoatomic - leaving 14 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
2	NAG	A	1001	1,2	14,14,15	0.57	0	15,19,21	1.02	1 (6%)
2	NAG	A	1002	3,2	14,14,15	0.50	0	15,19,21	0.69	0
3	BMA	A	1003	2,4	11,11,12	0.55	0	13,15,17	0.93	1 (7%)
4	MAN	A	1004	3	11,11,12	0.68	0	13,15,17	0.64	0
4	MAN	A	1005	3,4	11,11,12	0.50	0	13,15,17	2.51	4 (30%)
4	MAN	A	1006	4	11,11,12	0.53	0	13,15,17	0.83	1 (7%)
2	NAG	A	1007	1	14,14,15	0.50	0	15,19,21	0.91	1 (6%)
2	NAG	A	1008	1	14,14,15	0.49	0	15,19,21	1.02	1 (6%)
5	TMP	A	1009	6	17,22,22	2.14	4 (23%)	24,33,33	2.73	5 (20%)
2	NAG	B	1001	1,2	14,14,15	0.57	0	15,19,21	1.28	2 (13%)
2	NAG	B	1002	2	14,14,15	0.45	0	15,19,21	1.18	2 (13%)
2	NAG	B	1003	1	14,14,15	0.57	0	15,19,21	0.99	1 (6%)
2	NAG	B	1004	1	14,14,15	0.49	0	15,19,21	0.94	1 (6%)
5	TMP	B	1005	6	17,22,22	2.16	4 (23%)	24,33,33	2.82	5 (20%)

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
2	NAG	A	1001	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	1002	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	1003	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	1004	3	-	0/2/19/22	0/1/1/1
4	MAN	A	1005	3,4	-	0/2/19/22	0/1/1/1
4	MAN	A	1006	4	-	0/2/19/22	1/1/1/1
2	NAG	A	1007	1	-	0/6/23/26	0/1/1/1
2	NAG	A	1008	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	TMP	A	1009	6	-	0/6/22/22	0/2/2/2
2	NAG	B	1001	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	1002	2	-	0/6/23/26	0/1/1/1
2	NAG	B	1003	1	-	0/6/23/26	0/1/1/1
2	NAG	B	1004	1	-	0/6/23/26	0/1/1/1
5	TMP	B	1005	6	-	0/6/22/22	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1009	TMP	C3'-C4'	-2.45	1.46	1.53
5	B	1005	TMP	C3'-C4'	-2.37	1.46	1.53
5	A	1009	TMP	C6-C5	3.04	1.47	1.40
5	B	1005	TMP	C6-C5	3.05	1.48	1.40
5	A	1009	TMP	C2-N3	3.15	1.44	1.38

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1005	TMP	C5-C4-N3	-6.32	118.28	125.24
5	A	1009	TMP	C5-C4-N3	-6.20	118.41	125.24
5	B	1005	TMP	C5-C6-N1	-3.19	118.70	122.15
5	A	1009	TMP	C5-C6-N1	-3.13	118.76	122.15
2	B	1002	NAG	C4-C3-C2	-2.16	107.86	111.02

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1006	MAN	C1-C2-C3-C4-C5-O5

7 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	NAG	1	0
2	A	1002	NAG	1	0
4	A	1005	MAN	1	0
4	A	1006	MAN	1	0
5	A	1009	TMP	2	0
2	B	1003	NAG	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1005	TMP	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	706/823 (85%)	-0.07	12 (1%)	70 57	42, 77, 118, 161	0
1	B	697/823 (84%)	0.02	31 (4%)	35 22	35, 81, 123, 178	0
All	All	1403/1646 (85%)	-0.03	43 (3%)	49 33	35, 79, 121, 178	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	600	SER	4.8
1	B	441	PRO	4.2
1	B	698	ILE	4.1
1	A	712	ASN	3.4
1	A	411	GLY	3.3

### 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](#)

There are no carbohydrates in this entry.

### 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( $\text{\AA}^2$ )	Q<0.9
2	NAG	B	1001	14/15	0.95	0.21	-0.20	58,67,77,94	0
6	ZN	B	1006	1/1	0.98	0.17	-0.30	73,73,73,73	0
6	ZN	A	1011	1/1	0.99	0.14	-0.50	70,70,70,70	0
2	NAG	A	1001	14/15	0.95	0.19	-0.64	49,59,66,74	0
5	TMP	B	1005	21/21	0.96	0.17	-0.65	45,76,85,94	0
6	ZN	A	1010	1/1	0.99	0.15	-0.94	72,72,72,72	0
5	TMP	A	1009	21/21	0.97	0.15	-0.95	56,73,82,99	0
6	ZN	B	1007	1/1	0.99	0.14	-1.59	68,68,68,68	0
7	CA	B	1008	1/1	0.98	0.15	-1.71	92,92,92,92	0
7	CA	A	1012	1/1	0.97	0.12	-1.84	68,68,68,68	0
2	NAG	A	1007	14/15	0.80	0.34	-	101,111,120,121	0
4	MAN	A	1005	11/12	0.82	0.19	-	89,99,108,110	0
2	NAG	A	1008	14/15	0.92	0.31	-	93,104,117,118	0
2	NAG	B	1003	14/15	0.78	0.57	-	99,123,131,131	0
2	NAG	A	1002	14/15	0.94	0.20	-	62,82,90,92	0
2	NAG	B	1002	14/15	0.84	0.25	-	97,104,110,110	0
4	MAN	A	1004	11/12	0.88	0.14	-	70,84,99,99	0
4	MAN	A	1006	11/12	0.75	0.18	-	102,112,121,123	0
3	BMA	A	1003	11/12	0.86	0.14	-	88,101,116,117	0
2	NAG	B	1004	14/15	0.86	0.26	-	95,108,114,117	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.