



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 10, 2023 – 02:14 PM EDT

PDB ID : 4KXX  
Title : Human transketolase in covalent complex with donor ketose D-sedoheptulose-7-phosphate  
Authors : Neumann, P.; Luedtke, S.; Ficner, R.; Tittmann, K.  
Deposited on : 2013-05-28  
Resolution : 1.03 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.35.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

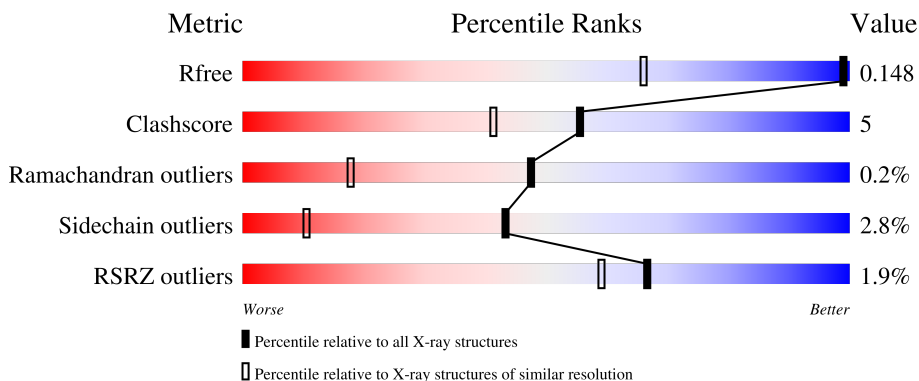
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.03 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))
$R_{free}$	130704	1596 (1.10-0.98)
Clashscore	141614	1677 (1.10-0.98)
Ramachandran outliers	138981	1591 (1.10-0.98)
Sidechain outliers	138945	1589 (1.10-0.98)
RSRZ outliers	127900	1557 (1.10-0.98)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	637	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6152 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.

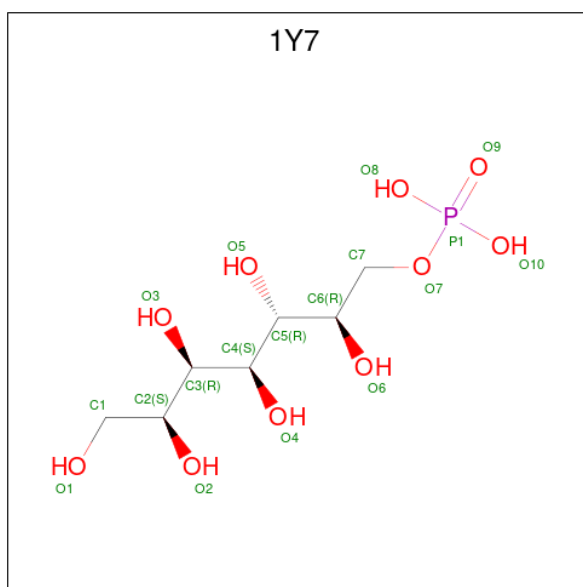
- Molecule 1 is a protein called Transketolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	621	5255	3305	922	999	29	0	63	0

There are 14 discrepancies between the modelled and reference sequences:

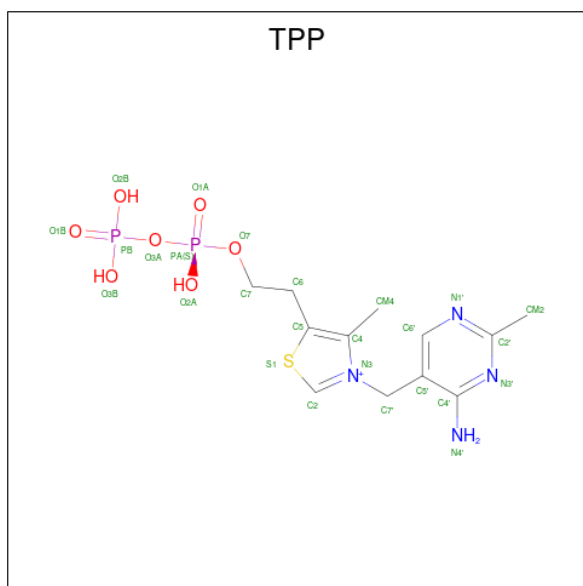
Chain	Residue	Modelled	Actual	Comment	Reference
A	624	LEU	-	expression tag	UNP P29401
A	625	VAL	-	expression tag	UNP P29401
A	626	PRO	-	expression tag	UNP P29401
A	627	ARG	-	expression tag	UNP P29401
A	628	GLY	-	expression tag	UNP P29401
A	629	SER	-	expression tag	UNP P29401
A	630	LEU	-	expression tag	UNP P29401
A	631	GLU	-	expression tag	UNP P29401
A	632	HIS	-	expression tag	UNP P29401
A	633	HIS	-	expression tag	UNP P29401
A	634	HIS	-	expression tag	UNP P29401
A	635	HIS	-	expression tag	UNP P29401
A	636	HIS	-	expression tag	UNP P29401
A	637	HIS	-	expression tag	UNP P29401

- Molecule 2 is (2R,3R,4S,5R,6S)-2,3,4,5,6,7-hexahydroxyheptyl dihydrogen phosphate (three-letter code: 1Y7) (formula: C<sub>7</sub>H<sub>17</sub>O<sub>10</sub>P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
2	A	1	18	7	10	1	0	0

- Molecule 3 is THIAMINE DIPHOSPHATE (three-letter code: TPP) (formula:  $C_{12}H_{19}N_4O_7P_2S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
3	A	1	26	12	4	7	2	1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 8 4 4	0	1
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 8 4 4	0	1
4	A	1	Total C O 8 4 4	0	1
4	A	1	Total C O 8 4 4	0	1
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

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

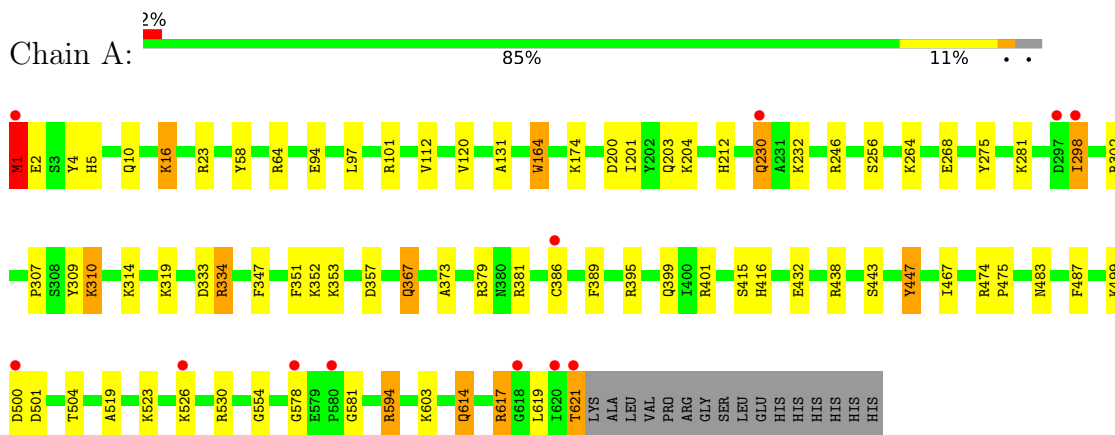
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	785	Total O 787 787	0	2

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

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Transketolase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	113.67Å 86.10Å 73.02Å 90.00° 125.43° 90.00°	Depositor
Resolution (Å)	30.00 – 1.03 18.53 – 1.03	Depositor EDS
% Data completeness (in resolution range)	99.1 (30.00-1.03) 99.1 (18.53-1.03)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 1.03Å)	Xtrriage
Refinement program	SHELX, SHELXL	Depositor
R, $R_{free}$	0.117 , 0.148 0.123 , 0.148	Depositor DCC
$R_{free}$ test set	13950 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	7.0	Xtrriage
Anisotropy	0.201	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 70.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	6152	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% 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: NA, 1Y7, EDO, TPP, MG

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	0.70	2/5358 (0.0%)	1.24	49/7244 (0.7%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	395[A]	ARG	CZ-NH1	-6.65	1.24	1.33
1	A	395[B]	ARG	CZ-NH1	-6.65	1.24	1.33

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	334	ARG	NE-CZ-NH1	16.05	128.33	120.30
1	A	395[A]	ARG	NE-CZ-NH2	12.09	126.35	120.30
1	A	395[B]	ARG	NE-CZ-NH2	12.09	126.35	120.30
1	A	334	ARG	NE-CZ-NH2	-11.19	114.71	120.30
1	A	395[A]	ARG	NE-CZ-NH1	-10.06	115.27	120.30
1	A	395[B]	ARG	NE-CZ-NH1	-10.06	115.27	120.30
1	A	23	ARG	NE-CZ-NH2	-9.75	115.42	120.30
1	A	367[A]	GLN	OE1-CD-NE2	9.54	143.85	121.90
1	A	367[B]	GLN	OE1-CD-NE2	9.54	143.85	121.90
1	A	447	TYR	CB-CG-CD1	8.48	126.09	121.00
1	A	501	ASP	CB-CG-OD2	-8.10	111.01	118.30
1	A	1	MET	C-N-CA	8.01	141.73	121.70
1	A	530	ARG	NE-CZ-NH2	-7.90	116.35	120.30
1	A	447	TYR	CB-CG-CD2	-7.67	116.39	121.00
1	A	23	ARG	NE-CZ-NH1	6.93	123.77	120.30
1	A	200[A]	ASP	CB-CG-OD1	6.79	124.42	118.30
1	A	200[B]	ASP	CB-CG-OD1	6.79	124.42	118.30
1	A	200[A]	ASP	CB-CG-OD2	-6.77	112.21	118.30
1	A	200[B]	ASP	CB-CG-OD2	-6.77	112.21	118.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	487	PHE	CB-CG-CD1	-6.55	116.22	120.80
1	A	379	ARG	CD-NE-CZ	6.17	132.24	123.60
1	A	381	ARG	NE-CZ-NH2	-6.10	117.25	120.30
1	A	101	ARG	NE-CZ-NH1	6.02	123.31	120.30
1	A	367[A]	GLN	CG-CD-NE2	-5.95	102.41	116.70
1	A	367[B]	GLN	CG-CD-NE2	-5.95	102.41	116.70
1	A	164	TRP	CA-CB-CG	-5.95	102.39	113.70
1	A	246	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	A	246	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	A	389	PHE	CB-CG-CD1	5.83	124.88	120.80
1	A	16	LYS	CD-CE-NZ	5.82	125.08	111.70
1	A	351	PHE	CB-CG-CD2	5.74	124.82	120.80
1	A	357	ASP	CB-CG-OD2	-5.73	113.15	118.30
1	A	401	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	A	438	ARG	NE-CZ-NH1	5.67	123.14	120.30
1	A	389	PHE	CB-CG-CD2	-5.64	116.85	120.80
1	A	334	ARG	CG-CD-NE	5.60	123.55	111.80
1	A	438	ARG	NE-CZ-NH2	-5.57	117.52	120.30
1	A	268	GLU	OE1-CD-OE2	5.54	129.95	123.30
1	A	401	ARG	NE-CZ-NH1	5.29	122.95	120.30
1	A	64	ARG	NE-CZ-NH1	-5.26	117.67	120.30
1	A	58	TYR	CB-CG-CD1	5.22	124.13	121.00
1	A	594	ARG	NE-CZ-NH1	5.21	122.91	120.30
1	A	4	TYR	CB-CG-CD2	5.18	124.11	121.00
1	A	302	ARG	CD-NE-CZ	5.18	130.86	123.60
1	A	504	THR	OG1-CB-CG2	-5.12	98.23	110.00
1	A	333	ASP	CB-CG-OD2	-5.11	113.70	118.30
1	A	617	ARG	CD-NE-CZ	5.05	130.66	123.60
1	A	474[A]	ARG	CD-NE-CZ	5.02	130.63	123.60
1	A	474[B]	ARG	CD-NE-CZ	5.02	130.63	123.60

There are no chirality outliers.

There are no planarity outliers.

## 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	5255	0	5224	51	0
2	A	18	0	15	1	0
3	A	26	0	15	0	0
4	A	64	0	96	9	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0
7	A	787	0	0	27	0
All	All	6152	0	5350	55	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 5.

All (55) 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:204[A]:LYS:HE2	7:A:8716:HOH:O	1.83	0.79
4:A:1016:EDO:H22	7:A:8150:HOH:O	1.84	0.78
4:A:1016:EDO:H12	7:A:8638:HOH:O	1.82	0.78
1:A:617:ARG:O	1:A:621:THR:HG23	1.87	0.75
1:A:174[B]:LYS:HG3	7:A:8501:HOH:O	1.86	0.74
1:A:519:ALA:O	1:A:523:LYS:HG3	1.86	0.74
1:A:334:ARG:HG2	7:A:8726:HOH:O	1.86	0.73
1:A:16:LYS:HE2	7:A:8409:HOH:O	1.92	0.69
1:A:256[B]:SER:HA	4:A:1008[B]:EDO:H22	1.78	0.65
1:A:204[B]:LYS:HD2	7:A:8749:HOH:O	1.98	0.64
1:A:230[B]:GLN:HB3	7:A:8520:HOH:O	1.99	0.63
1:A:353:LYS:HD2	7:A:8426:HOH:O	2.00	0.60
1:A:204[B]:LYS:HD3	7:A:8566:HOH:O	2.01	0.59
1:A:201[A]:ILE:HD13	7:A:8749:HOH:O	2.03	0.57
1:A:164:TRP:CD1	4:A:1005:EDO:H12	2.41	0.56
1:A:603:LYS:HE3	7:A:8515:HOH:O	2.06	0.56
1:A:367[A]:GLN:NE2	7:A:8626:HOH:O	2.41	0.53
1:A:619:LEU:HD23	7:A:8743:HOH:O	2.10	0.51
1:A:367[A]:GLN:HG2	1:A:399:GLN:OE1	2.11	0.50
1:A:352:LYS:HE2	7:A:8734:HOH:O	2.11	0.50
1:A:275[A]:TYR:OH	1:A:281:LYS:HE2	2.10	0.50
1:A:204[A]:LYS:HD3	7:A:8384:HOH:O	2.12	0.49
1:A:264[A]:LYS:HG3	7:A:8628:HOH:O	2.13	0.49
1:A:203[B]:GLN:HG3	1:A:203[B]:GLN:HE21	1.71	0.49
1:A:475[A]:PRO:HB3	7:A:8233:HOH:O	2.13	0.49
4:A:1008[B]:EDO:H22	7:A:8659:HOH:O	2.12	0.49
1:A:1:MET:HB2	1:A:5:HIS:NE2	2.28	0.48

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:256[B]:SER:HA	4:A:1008[B]:EDO:C2	2.42	0.48
1:A:10[B]:GLN:HG3	7:A:8415:HOH:O	2.12	0.48
1:A:97:LEU:HD11	4:A:1007[A]:EDO:H12	1.96	0.47
1:A:298:ILE:HG23	1:A:298:ILE:O	2.15	0.45
1:A:443:SER:HA	1:A:467:ILE:O	2.16	0.45
1:A:307:PRO:O	1:A:483[B]:ASN:ND2	2.50	0.45
1:A:614:GLN:NE2	7:A:8740:HOH:O	2.49	0.45
1:A:415[B]:SER:O	1:A:416:HIS:HB2	2.17	0.45
1:A:614:GLN:NE2	7:A:8383:HOH:O	2.48	0.44
1:A:578:GLY:HA3	7:A:8335:HOH:O	2.16	0.44
1:A:298:ILE:O	1:A:298:ILE:CG2	2.65	0.44
1:A:120[B]:VAL:HG11	4:A:1003[B]:EDO:H12	1.98	0.43
1:A:212:HIS:NE2	1:A:232:LYS:HE3	2.33	0.43
1:A:319[A]:LYS:HE3	1:A:347:PHE:CE1	2.54	0.43
1:A:230[A]:GLN:NE2	1:A:232:LYS:HE2	2.34	0.42
1:A:309[A]:TYR:C	1:A:310:LYS:HE3	2.40	0.42
1:A:319[A]:LYS:NZ	7:A:8370:HOH:O	2.49	0.42
2:A:1001:1Y7:H16	2:A:1001:1Y7:O4	2.19	0.42
1:A:94[B]:GLU:HG3	7:A:8509:HOH:O	2.18	0.42
1:A:373:ALA:HB2	1:A:386:CYS:SG	2.59	0.42
1:A:554:GLY:HA3	1:A:581:GLY:O	2.20	0.42
1:A:319[A]:LYS:HE2	7:A:8674:HOH:O	2.19	0.41
1:A:578:GLY:N	7:A:8215:HOH:O	2.38	0.41
1:A:212:HIS:CD2	1:A:232:LYS:HE3	2.56	0.41
1:A:131:ALA:HB2	4:A:1003[B]:EDO:C1	2.51	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	683/637 (107%)	662 (97%)	20 (3%)	1 (0%)	51   18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	GLU

### 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	564/518 (109%)	549 (97%)	15 (3%)	44 11

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	112	VAL
1	A	230[A]	GLN
1	A	230[B]	GLN
1	A	298	ILE
1	A	310	LYS
1	A	314	LYS
1	A	432	GLU
1	A	447	TYR
1	A	499	LYS
1	A	500	ASP
1	A	526	LYS
1	A	594	ARG
1	A	614	GLN
1	A	621	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	189	GLN
1	A	233	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 [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 2 are monoatomic - leaving 18 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	A	1011	-	3,3,3	0.46	0	2,2,2	0.48	0
2	1Y7	A	1001	3	17,17,17	2.71	6 (35%)	24,24,24	1.85	8 (33%)
4	EDO	A	1014	-	3,3,3	0.74	0	2,2,2	1.46	0
4	EDO	A	1016	-	3,3,3	0.22	0	2,2,2	0.21	0
4	EDO	A	1007[B]	-	3,3,3	0.52	0	2,2,2	1.21	0
4	EDO	A	1008[B]	-	3,3,3	0.48	0	2,2,2	1.00	0
4	EDO	A	1003[A]	-	3,3,3	0.32	0	2,2,2	0.64	0
4	EDO	A	1009[A]	-	3,3,3	0.37	0	2,2,2	0.28	0
4	EDO	A	1009[B]	-	3,3,3	0.46	0	2,2,2	0.51	0
4	EDO	A	1010	-	3,3,3	0.40	0	2,2,2	0.36	0
3	TPP	A	1002	2,5	22,27,27	2.10	3 (13%)	29,40,40	1.41	3 (10%)
4	EDO	A	1004	-	3,3,3	0.61	0	2,2,2	0.92	0
4	EDO	A	1006	-	3,3,3	0.20	0	2,2,2	0.33	0
4	EDO	A	1007[A]	-	3,3,3	0.53	0	2,2,2	1.07	0
4	EDO	A	1008[A]	-	3,3,3	0.35	0	2,2,2	0.67	0
4	EDO	A	1012	-	3,3,3	0.83	0	2,2,2	0.33	0
4	EDO	A	1005	-	3,3,3	0.67	0	2,2,2	0.94	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	A	1003[B]	-	3,3,3	0.44	0	2,2,2	0.65	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	A	1011	-	-	0/1/1/1	-
2	1Y7	A	1001	3	-	5/24/24/24	-
4	EDO	A	1014	-	-	0/1/1/1	-
4	EDO	A	1016	-	-	1/1/1/1	-
4	EDO	A	1007[B]	-	-	1/1/1/1	-
4	EDO	A	1008[B]	-	-	0/1/1/1	-
4	EDO	A	1003[A]	-	-	1/1/1/1	-
4	EDO	A	1009[A]	-	-	0/1/1/1	-
4	EDO	A	1009[B]	-	-	0/1/1/1	-
4	EDO	A	1010	-	-	0/1/1/1	-
3	TPP	A	1002	2,5	-	2/16/17/17	0/2/2/2
4	EDO	A	1004	-	-	0/1/1/1	-
4	EDO	A	1006	-	-	0/1/1/1	-
4	EDO	A	1007[A]	-	-	1/1/1/1	-
4	EDO	A	1008[A]	-	-	0/1/1/1	-
4	EDO	A	1012	-	-	0/1/1/1	-
4	EDO	A	1005	-	-	1/1/1/1	-
4	EDO	A	1003[B]	-	-	0/1/1/1	-

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1002	TPP	C6-C5	-7.87	1.47	1.50
2	A	1001	1Y7	C2-C3	7.56	1.67	1.53
2	A	1001	1Y7	C4-C5	-4.88	1.43	1.53
3	A	1002	TPP	CM4-C4	3.86	1.57	1.49
2	A	1001	1Y7	O2-C2	-3.71	1.35	1.43
2	A	1001	1Y7	P1-O7	2.90	1.69	1.60
2	A	1001	1Y7	O5-C5	-2.59	1.36	1.43
3	A	1002	TPP	C4-N3	2.49	1.41	1.39
2	A	1001	1Y7	O3-C3	2.46	1.48	1.43

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1002	TPP	CM4-C4-N3	-4.64	116.62	122.53
2	A	1001	1Y7	O2-C2-C3	-4.61	97.89	109.10
3	A	1002	TPP	C5-C4-N3	4.03	115.65	107.57
2	A	1001	1Y7	C2-C3-C4	3.07	117.27	112.47
2	A	1001	1Y7	C6-C5-C4	3.05	117.24	112.47
2	A	1001	1Y7	C1-C2-C3	-2.75	106.45	112.41
2	A	1001	1Y7	O2-C2-C1	2.54	115.10	109.14
2	A	1001	1Y7	O7-P1-O9	-2.37	99.81	106.47
2	A	1001	1Y7	C3-C4-C5	2.20	116.30	113.26
2	A	1001	1Y7	O1-C1-C2	-2.15	106.38	111.07
3	A	1002	TPP	O3B-PB-O2B	2.03	115.39	107.64

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1007[A]	EDO	O1-C1-C2-O2
2	A	1001	1Y7	O5-C5-C6-C7
2	A	1001	1Y7	C3-C4-C5-O5
2	A	1001	1Y7	C4-C5-C6-C7
3	A	1002	TPP	PA-O3A-PB-O3B
2	A	1001	1Y7	C3-C4-C5-C6
4	A	1005	EDO	O1-C1-C2-O2
3	A	1002	TPP	C4-C5-C6-C7
4	A	1016	EDO	O1-C1-C2-O2
2	A	1001	1Y7	O5-C5-C6-O6
4	A	1003[A]	EDO	O1-C1-C2-O2
4	A	1007[B]	EDO	O1-C1-C2-O2

There are no ring outliers.

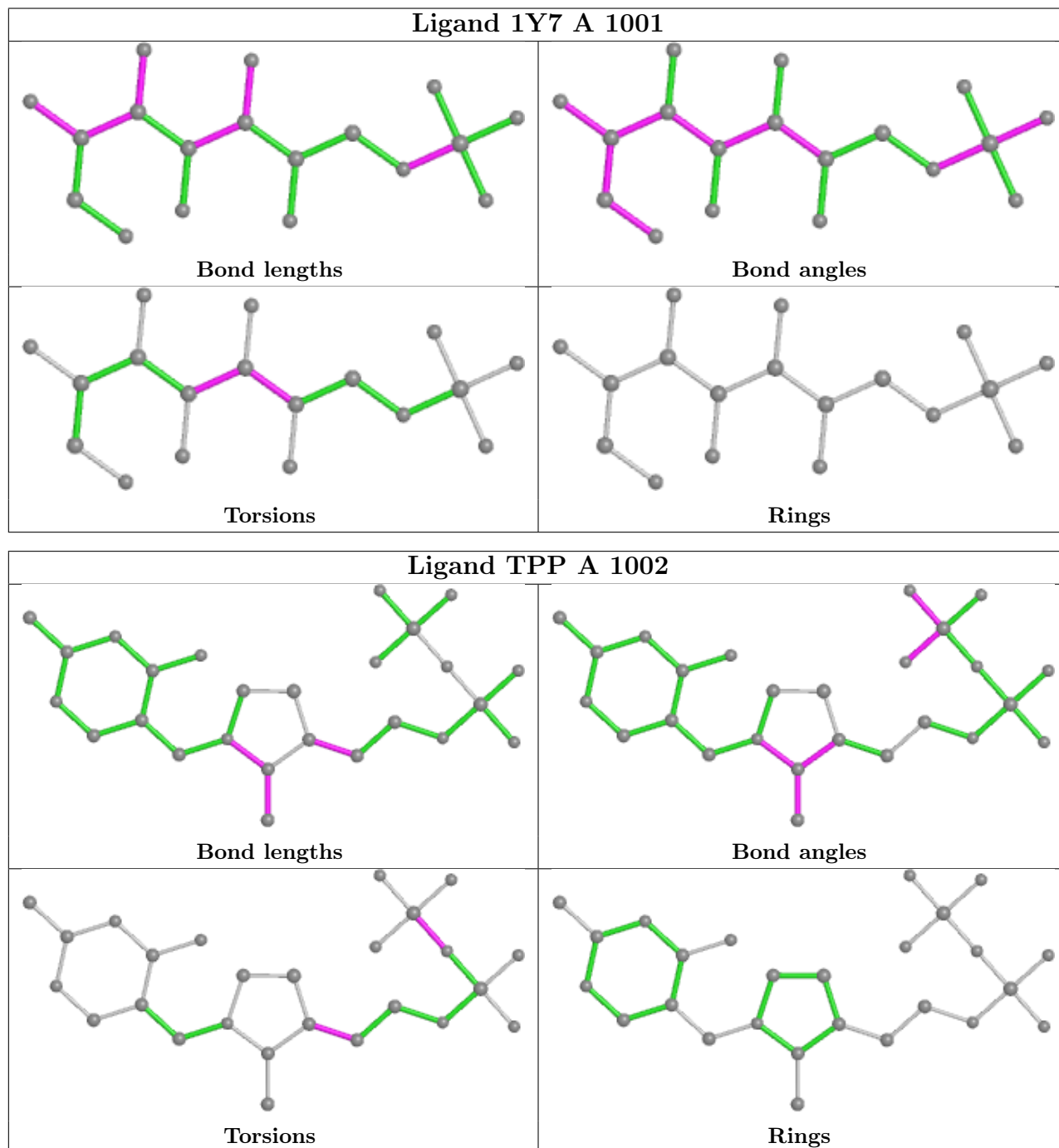
6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	1Y7	1	0
4	A	1016	EDO	2	0
4	A	1008[B]	EDO	3	0
4	A	1007[A]	EDO	1	0
4	A	1005	EDO	1	0
4	A	1003[B]	EDO	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In



addition, ligands with molecular weight  $> 250$  and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	621/637 (97%)	-0.42	12 (1%) 66 59	3, 9, 37, 83	1 (0%)

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	386	CYS	9.8
1	A	298	ILE	6.5
1	A	1	MET	4.8
1	A	578	GLY	4.4
1	A	620	ILE	3.6
1	A	526	LYS	2.6
1	A	621	THR	2.5
1	A	580	PRO	2.1
1	A	500	ASP	2.1
1	A	297	ASP	2.0
1	A	618	GLY	2.0
1	A	230[A]	GLN	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](#)

There are no monosaccharides 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. The B-factors column lists the minimum,

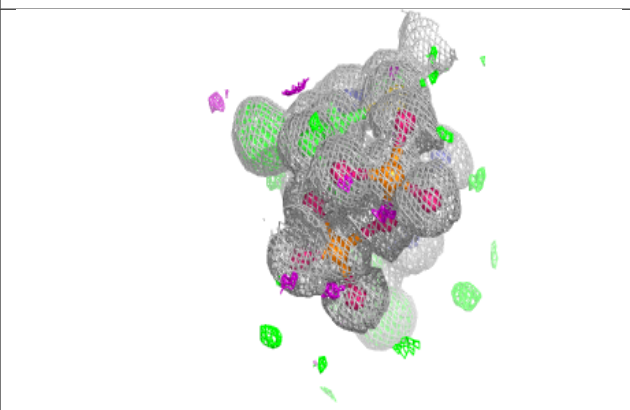
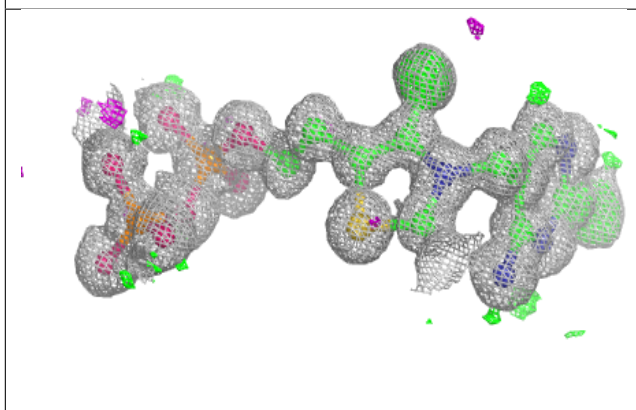
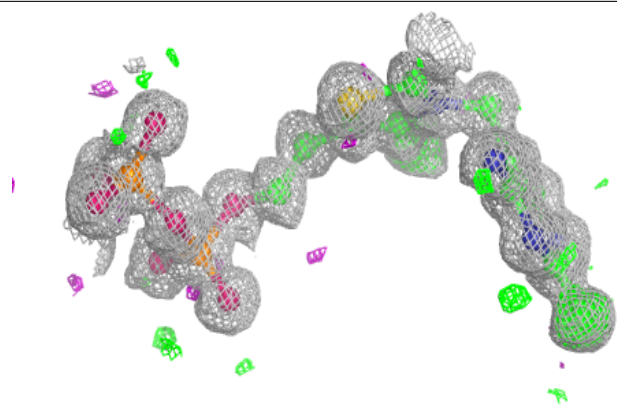
median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	EDO	A	1014	4/4	0.88	0.14	19,20,23,24	0
4	EDO	A	1006	4/4	0.90	0.08	22,23,28,32	0
4	EDO	A	1004	4/4	0.95	0.07	16,19,21,24	0
4	EDO	A	1005	4/4	0.96	0.09	10,17,18,18	0
4	EDO	A	1003[B]	4/4	0.96	0.13	6,13,17,26	4
4	EDO	A	1007[A]	4/4	0.96	0.10	14,27,37,43	4
4	EDO	A	1007[B]	4/4	0.96	0.10	9,10,14,17	4
4	EDO	A	1003[A]	4/4	0.96	0.13	14,25,29,32	4
4	EDO	A	1016	4/4	0.96	0.09	15,17,20,29	0
4	EDO	A	1008[A]	4/4	0.97	0.12	6,14,14,47	4
4	EDO	A	1008[B]	4/4	0.97	0.12	0,7,9,40	4
4	EDO	A	1009[B]	4/4	0.98	0.08	10,11,12,34	4
4	EDO	A	1012	4/4	0.98	0.05	9,10,12,13	0
3	TPP	A	1002	26/26	0.98	0.05	4,5,7,23	0
4	EDO	A	1009[A]	4/4	0.98	0.08	7,11,23,25	4
2	1Y7	A	1001	18/18	0.99	0.06	5,9,12,13	18
4	EDO	A	1010	4/4	0.99	0.04	9,13,13,14	0
4	EDO	A	1011	4/4	0.99	0.04	7,9,9,11	0
5	MG	A	1013	1/1	1.00	0.02	3,3,3,3	0
6	NA	A	1015	1/1	1.00	0.02	8,8,8,8	0

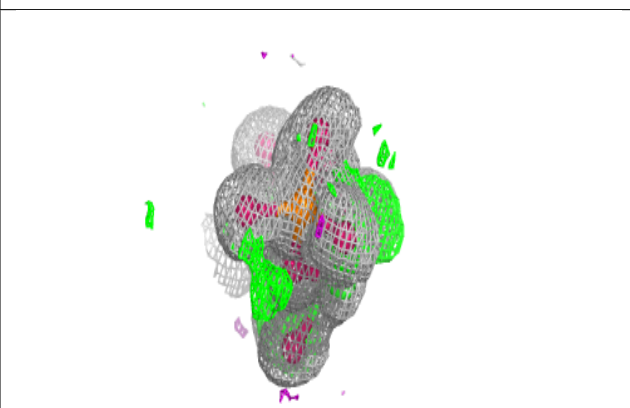
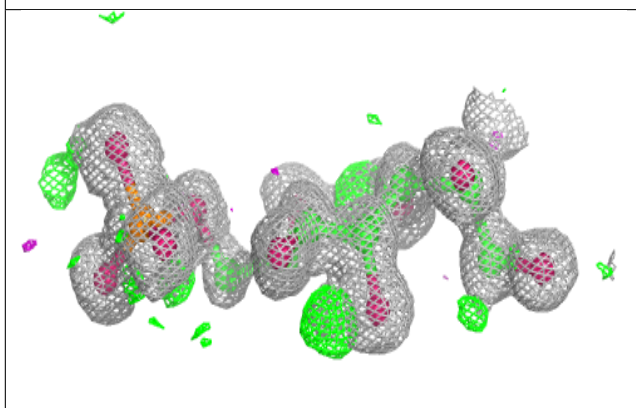
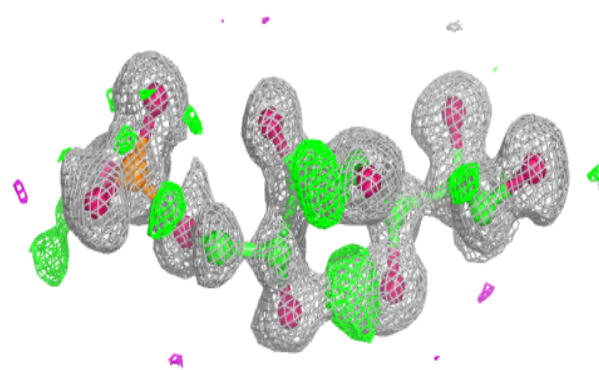
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around TPP A 1002:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around 1Y7 A 1001:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.