



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 02:54 PM GMT

PDB ID : 3TY5
Title : Crystal Structure of C. thermocellum PNKP Ligase domain in complex with ATP
Authors : Smith, P.; Wang, L.; Shuman, S.
Deposited on : 2011-09-23
Resolution : 2.40 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

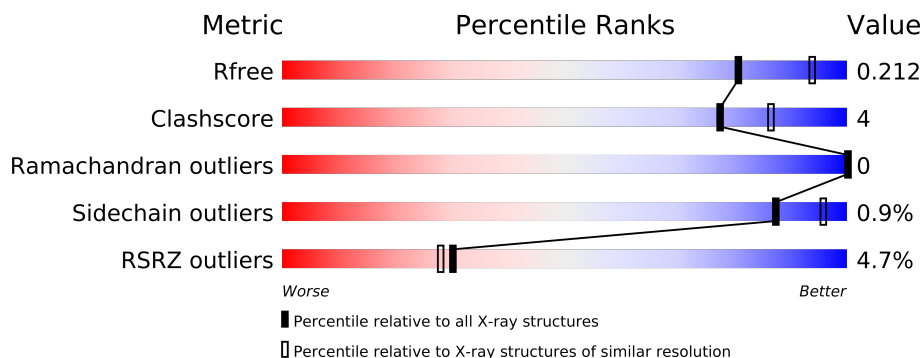
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.40 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}	66092	2207 (2.40-2.40)
Clashscore	79885	2789 (2.40-2.40)
Ramachandran outliers	78287	2736 (2.40-2.40)
Sidechain outliers	78261	2737 (2.40-2.40)
RSRZ outliers	66119	2210 (2.40-2.40)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	413	
1	B	413	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	ADP	B	1001	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6381 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Polynucleotide 2',3'-cyclic phosphate phosphodiesterase / polynucleotide 5'-hydroxyl-kinase / polynucleotide 3'-phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	S	0	0	0
			3057	1935	535	564	23			
1	B	381	Total	C	N	O	S	0	0	0
			3043	1928	529	565	21			

There are 44 discrepancies between the modelled and reference sequences:

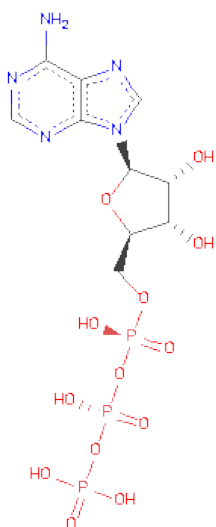
Chain	Residue	Modelled	Actual	Comment	Reference
A	458	MET	-	EXPRESSION TAG	UNP A3DJ38
A	459	GLY	-	EXPRESSION TAG	UNP A3DJ38
A	460	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	461	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	462	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	463	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	464	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	465	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	466	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	467	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	468	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	469	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	470	SER	-	EXPRESSION TAG	UNP A3DJ38
A	471	SER	-	EXPRESSION TAG	UNP A3DJ38
A	472	GLY	-	EXPRESSION TAG	UNP A3DJ38
A	473	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	474	ILE	-	EXPRESSION TAG	UNP A3DJ38
A	475	GLU	-	EXPRESSION TAG	UNP A3DJ38
A	476	GLY	-	EXPRESSION TAG	UNP A3DJ38
A	477	ARG	-	EXPRESSION TAG	UNP A3DJ38
A	478	HIS	-	EXPRESSION TAG	UNP A3DJ38
A	529	GLY	GLU	ENGINEERED MUTATION	UNP A3DJ38
B	458	MET	-	EXPRESSION TAG	UNP A3DJ38
B	459	GLY	-	EXPRESSION TAG	UNP A3DJ38

Continued on next page...

Continued from previous page...

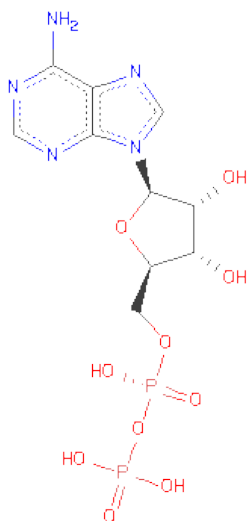
Chain	Residue	Modelled	Actual	Comment	Reference
B	460	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	461	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	462	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	463	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	464	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	465	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	466	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	467	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	468	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	469	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	470	SER	-	EXPRESSION TAG	UNP A3DJ38
B	471	SER	-	EXPRESSION TAG	UNP A3DJ38
B	472	GLY	-	EXPRESSION TAG	UNP A3DJ38
B	473	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	474	ILE	-	EXPRESSION TAG	UNP A3DJ38
B	475	GLU	-	EXPRESSION TAG	UNP A3DJ38
B	476	GLY	-	EXPRESSION TAG	UNP A3DJ38
B	477	ARG	-	EXPRESSION TAG	UNP A3DJ38
B	478	HIS	-	EXPRESSION TAG	UNP A3DJ38
B	529	GLY	GLU	ENGINEERED MUTATION	UNP A3DJ38

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



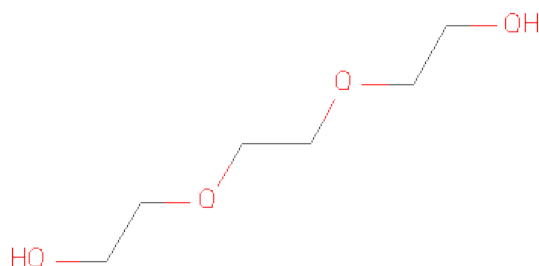
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	B	1	27	10	5	10	2	0	0

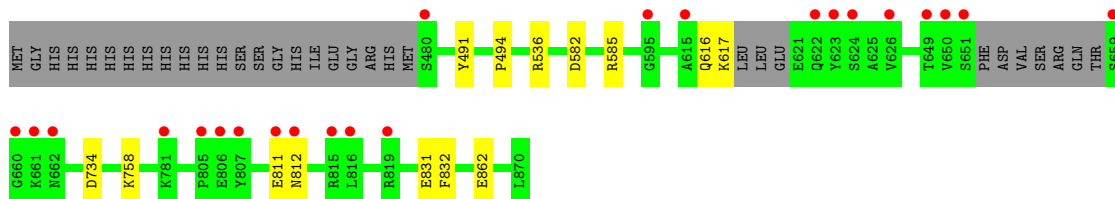
- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	B	1	10	6	4	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	117	Total 117	O 117	0	0
5	B	96	Total 96	O 96	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	63.99Å 94.38Å 75.31Å 90.00° 93.51° 90.00°	Depositor
Resolution (Å)	44.33 – 2.40 39.97 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.7 (44.33-2.40) 99.2 (39.97-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.32 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, R_{free}	0.178 , 0.225 0.170 , 0.212	Depositor DCC
R_{free} test set	1571 reflections (5.11%)	DCC
Wilson B-factor (Å ²)	48.9	Xtriage
Anisotropy	0.167	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 40.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 30766 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6381	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PGE, ATP, ADP

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	0.29	0/3122	0.46	0/4213
1	B	0.28	0/3107	0.43	0/4192
All	All	0.28	0/6229	0.45	0/8405

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3057	0	0	11	0
1	B	3043	0	0	11	0
2	A	31	0	0	1	0
3	B	27	0	0	0	0
4	B	10	0	14	5	0
5	A	117	0	0	2	0
5	B	96	0	0	4	0
All	All	6381	0	14	23	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including

hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 23 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:491:TYR:CE1	4:B:1:PGE:H5	2.15	0.82
4:B:1:PGE:H1	5:B:75:HOH:O	1.93	0.68
1:B:832:PHE:O	4:B:1:PGE:H6	2.01	0.60
1:A:817:ARG:O	1:A:820:ALA:N	2.35	0.59
1:A:700:LYS:NZ	1:A:735:SER:O	2.38	0.56

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	377/413 (91%)	362 (96%)	15 (4%)	0	100	100
1	B	375/413 (91%)	364 (97%)	11 (3%)	0	100	100
All	All	752/826 (91%)	726 (96%)	26 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	329/360 (91%)	323 (98%)	6 (2%)	71	88
1	B	328/360 (91%)	328 (100%)	0	100	100
All	All	657/720 (91%)	651 (99%)	6 (1%)	87	96

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	607	GLU
1	A	814	GLU
1	A	733	ASP
1	A	570	PHE
1	A	809	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

3 ligands 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	$\# Z > 2$
2	ATP	A	1001	-	33,33,33	1.76	8 (24%)	52,52,52	3.48	27 (51%)
4	PGE	B	1	-	9,9,9	0.56	0	8,8,8	1.33	1 (12%)
3	ADP	B	1001	-	29,29,29	2.53	11 (37%)	45,45,45	2.52	16 (35%)

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	ATP	A	1001	-	-	0/22/38/38	0/1/3/3
4	PGE	B	1	-	-	0/7/7/7	0/0/0/0
3	ADP	B	1001	-	-	0/16/32/32	0/1/3/3

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1001	ADP	C5-C4	6.29	1.54	1.40
3	B	1001	ADP	PB-O3A	5.92	1.70	1.60
3	B	1001	ADP	PA-O3A	5.53	1.69	1.59
2	A	1001	ATP	O4'-C1'	4.59	1.48	1.41
2	A	1001	ATP	PG-O3B	4.10	1.67	1.60

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	ATP	O4'-C1'-N9	15.94	123.26	108.44
2	A	1001	ATP	N3-C2-N1	-7.41	122.51	128.71
2	A	1001	ATP	N3-C4-N9	6.47	137.11	125.43
2	A	1001	ATP	N6-C6-N1	6.28	131.70	119.36
3	B	1001	ADP	N3-C2-N1	-5.84	123.83	128.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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 ²)	Q<0.9
1	A	381/413 (92%)	0.00	12 (3%) 47 44	27, 51, 98, 148	0
1	B	381/413 (92%)	0.08	23 (6%) 21 19	33, 56, 111, 149	0
All	All	762/826 (92%)	0.04	35 (4%) 30 29	27, 53, 106, 149	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	805	PRO	4.4
1	B	622	GLN	4.3
1	A	647	ASN	4.0
1	B	623	TYR	3.9
1	A	478	HIS	3.8

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, 95th 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	ADP	B	1001	27/27	0.37	5.43	91,137,167,252	0
4	PGE	B	1	10/10	0.22	1.50	55,63,69,70	0
2	ATP	A	1001	31/31	0.28	0.96	74,79,89,91	0

6.5 Other polymers ⓘ

There are no such residues in this entry.