



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:54 PM GMT

PDB ID : 1R5J
Title : Crystal Structure of a Phosphotransacetylase from *Streptococcus pyogenes*
Authors : Xu, Q.S.; Shin, D.H.; Pufan, R.; Yokota, H.; Kim, R.; Kim, S.H.; Berkeley Structural Genomics Center (BSGC)
Deposited on : 2003-10-10
Resolution : 2.70 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

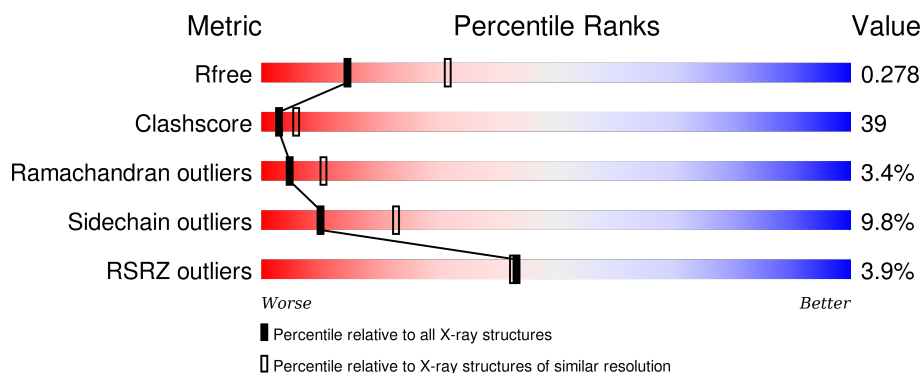
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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. 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	337	<div> <div>5%</div> <div>47%</div> <div>42%</div> <div>9%</div> <div>•</div> </div>
1	B	337	<div> <div>2%</div> <div>50%</div> <div>39%</div> <div>8%</div> <div>••</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5034 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 putative phosphotransacetylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	Se	0	0	0
			2502	1582	424	485	1	10			
1	B	329	Total	C	N	O	S	Se	0	0	0
			2502	1582	424	485	1	10			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
A	-4	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
A	-3	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
A	-2	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
A	-1	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
A	0	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	19	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	83	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	103	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	113	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	117	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	123	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	157	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	204	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
A	288	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	-5	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
B	-4	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
B	-3	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
B	-2	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
B	-1	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
B	0	GLY	-	CLONING ARTIFACT	UNP Q99ZQ5
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	19	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	83	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	103	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	113	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	117	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	123	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	157	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	204	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5
B	288	MSE	MET	MODIFIED RESIDUE	UNP Q99ZQ5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	12	Total O 12 12	0	0
2	B	18	Total O 18 18	0	0

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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.

- Chain A:
-
- 5% 47% 42% 9%
- GLY GLY GLY GLY GLY M1 S2 I3 R4 F7 R11 L15 G16 K17 N18 M19 K20 L21 P24 E25 G26 N27 D28 E29 R30 V31 V32 R33 L38 R39 F40 L43 L44 E45 P46 T47 L48 L49 A50 G51 S52 E54 R55 R56 N57 L58 L59 T60 R61 L62 L63 L64 A65 D66 Q67 D68 T69 I70 I71 P74 Y77 A78 D79 F80 R81 I84 R83 K84 E85 A86 R87 R88 V88 E89 R90 R91 K92 R93 R94 R95 R96 I97 R98 D99 A100 D101 R102 M103 L104 R105 L106 V107 M108 I109 F110 G111 G112 M113 L114 V115 M116 M117 G118 G119 L120 Y124 L125 G126 A127 L128 T131 A132 V135 R136 P137 A138 L139 Q140 I141 I142 P146 F147 I148 S149 R150 T151 V154 M158 R159 E160 M161 T162 S163 A164 R165 R166 V167 F168 A169 D170 D173 M174 T175 D176 P177 T178 L179 Q180 E181 E184 T182 R201 L202 G203 R204 L205 T209 Q217 V230 T225 E226 I227 M232 P233 D234 L237 D238 G239 E240 L241 V248 P249 E250 T251 A252 A253 K255 L254 A261 V261 A262 G263 Q264 A265 M266 T267 F268 V269 F270 T271 D272 L273 I274 Q274 Y280 R285 L286 F289 T292 G293 L296 Q297 V303 R304 D305 L306

- Chain B:
-
- 2% 50% 39% 8%
- GLY GLY GLY GLY GLY M1 S2 T3 R4 S5 L6 R11 L15 G16 R17 M18 M19 K20 L21 P24 N27 D28 E29 R30 R33 L43 L44 E45 P46 L47 L48 L49 G50 Q51 S52 E53 E54 V55 R56 N57 L58 L59 T60 K61 L62 L63 F64 A65 D66 Q67 D68

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 3 2	Depositor
Cell constants a, b, c, α , β , γ	173.72Å 173.72Å 173.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.93 – 2.70 19.93 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.93-2.70) 100.0 (19.93-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.57 (at 2.71Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.229 , 0.281 0.227 , 0.278	Depositor DCC
R_{free} test set	2537 reflections (10.10%)	DCC
Wilson B-factor (Å ²)	67.7	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 56.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 25109 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5034	wwPDB-VP
Average B, all atoms (Å ²)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.62	5/2529 (0.2%)	0.78	4/3400 (0.1%)
1	B	0.63	3/2529 (0.1%)	0.79	3/3400 (0.1%)
All	All	0.63	8/5058 (0.2%)	0.78	7/6800 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	92	LYS	CE-NZ	7.80	1.68	1.49
1	A	96	THR	CB-CG2	6.92	1.75	1.52
1	A	94	LYS	CB-CG	6.36	1.69	1.52
1	B	92	LYS	CE-NZ	6.35	1.65	1.49
1	B	94	LYS	CD-CE	5.98	1.66	1.51

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	169	ALA	N-CA-C	5.92	126.98	111.00
1	A	239	GLY	N-CA-C	5.75	127.49	113.10
1	B	169	ALA	N-CA-C	5.65	126.25	111.00
1	A	201	LYS	N-CA-C	-5.36	96.54	111.00
1	B	201	LYS	N-CA-C	-5.33	96.62	111.00

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	2502	0	2542	216	0
1	B	2502	0	2542	180	0
2	A	12	0	0	1	0
2	B	18	0	0	1	0
All	All	5034	0	5084	393	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 39.

The worst 5 of 393 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:96:THR:CG2	1:A:96:THR:CB	1.75	1.63
1:A:92:LYS:NZ	1:A:92:LYS:CE	1.68	1.56
1:A:169:ALA:CB	1:A:269:VAL:HA	1.71	1.19
1:B:169:ALA:CB	1:B:269:VAL:HA	1.74	1.18
1:A:92:LYS:N	1:A:92:LYS:HD2	1.58	1.15

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	327/337 (97%)	284 (87%)	32 (10%)	11 (3%)	5	10
1	B	327/337 (97%)	284 (87%)	32 (10%)	11 (3%)	5	10
All	All	654/674 (97%)	568 (87%)	64 (10%)	22 (3%)	5	10

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	92	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	293	GLY
1	B	293	GLY
1	B	78	ALA
1	B	92	LYS

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	265/256 (104%)	239 (90%)	26 (10%)	10	23
1	B	265/256 (104%)	239 (90%)	26 (10%)	10	23
All	All	530/512 (104%)	478 (90%)	52 (10%)	10	23

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	285	ARG
1	B	58	LEU
1	B	285	ARG
1	A	308	ARG
1	B	1	MSE

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	266	ASN
1	B	57	ASN
1	B	264	GLN
1	A	264	GLN
1	B	232	ASN

5.3.3 RNA ⓘ

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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, 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	319/337 (94%)	0.22	18 (5%) 28 26	42, 74, 121, 138	0
1	B	319/337 (94%)	-0.10	7 (2%) 65 66	40, 61, 106, 131	0
All	All	638/674 (94%)	0.06	25 (3%) 43 43	40, 67, 116, 138	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	95	ALA	5.8
1	A	105	ARG	5.3
1	A	93	GLY	4.0
1	A	51	GLN	3.6
1	A	146	PRO	3.5

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.