



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 06:30 am BST

PDB ID : 3CR8  
Title : Hexameric APS kinase from *Thiobacillus denitrificans*  
Authors : Gay, S.C.; Segel, I.H.; Fisher, A.J.  
Deposited on : 2008-04-04  
Resolution : 2.95 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

<https://www.wwpdb.org/validation/2017/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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

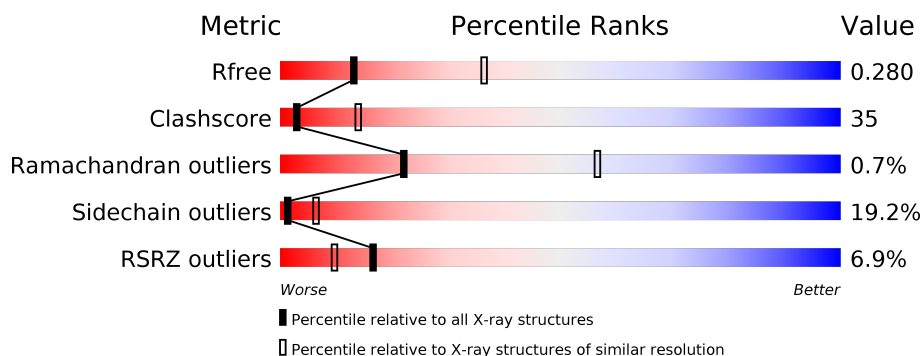
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.95 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 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	552	<div> <div>5%</div> <div> <div></div> <div>51%</div> <div>28%</div> <div>10%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	552	<div> <div>6%</div> <div> <div></div> <div>50%</div> <div>29%</div> <div>10%</div> <div>•</div> <div>11%</div> </div> </div>
1	C	552	<div> <div>7%</div> <div> <div></div> <div>48%</div> <div>32%</div> <div>9%</div> <div>11%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11339 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 Sulfate adenylyltransferase, adenylylsulfate kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	493	Total	C	N	O	S	0	0	0
			3761	2390	676	679	16			
1	B	493	Total	C	N	O	S	0	0	0
			3744	2378	671	679	16			
1	C	493	Total	C	N	O	S	0	0	0
			3748	2380	670	682	16			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	545	LEU	-	EXPRESSION TAG	UNP Q3SM86
A	546	GLU	-	EXPRESSION TAG	UNP Q3SM86
A	547	HIS	-	EXPRESSION TAG	UNP Q3SM86
A	548	HIS	-	EXPRESSION TAG	UNP Q3SM86
A	549	HIS	-	EXPRESSION TAG	UNP Q3SM86
A	550	HIS	-	EXPRESSION TAG	UNP Q3SM86
A	551	HIS	-	EXPRESSION TAG	UNP Q3SM86
A	552	HIS	-	EXPRESSION TAG	UNP Q3SM86
B	545	LEU	-	EXPRESSION TAG	UNP Q3SM86
B	546	GLU	-	EXPRESSION TAG	UNP Q3SM86
B	547	HIS	-	EXPRESSION TAG	UNP Q3SM86
B	548	HIS	-	EXPRESSION TAG	UNP Q3SM86
B	549	HIS	-	EXPRESSION TAG	UNP Q3SM86
B	550	HIS	-	EXPRESSION TAG	UNP Q3SM86
B	551	HIS	-	EXPRESSION TAG	UNP Q3SM86
B	552	HIS	-	EXPRESSION TAG	UNP Q3SM86
C	545	LEU	-	EXPRESSION TAG	UNP Q3SM86
C	546	GLU	-	EXPRESSION TAG	UNP Q3SM86
C	547	HIS	-	EXPRESSION TAG	UNP Q3SM86
C	548	HIS	-	EXPRESSION TAG	UNP Q3SM86
C	549	HIS	-	EXPRESSION TAG	UNP Q3SM86
C	550	HIS	-	EXPRESSION TAG	UNP Q3SM86
C	551	HIS	-	EXPRESSION TAG	UNP Q3SM86

*Continued on next page...*

*Continued from previous page...*

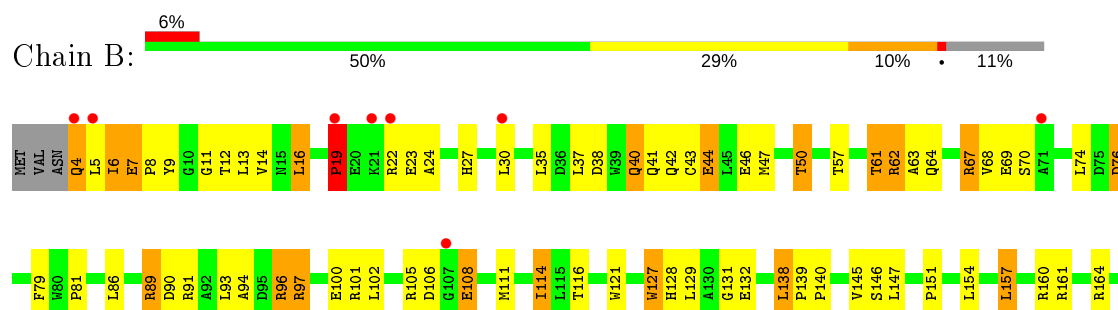
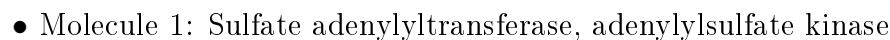
Chain	Residue	Modelled	Actual	Comment	Reference
C	552	HIS	-	EXPRESSION TAG	UNP Q3SM86

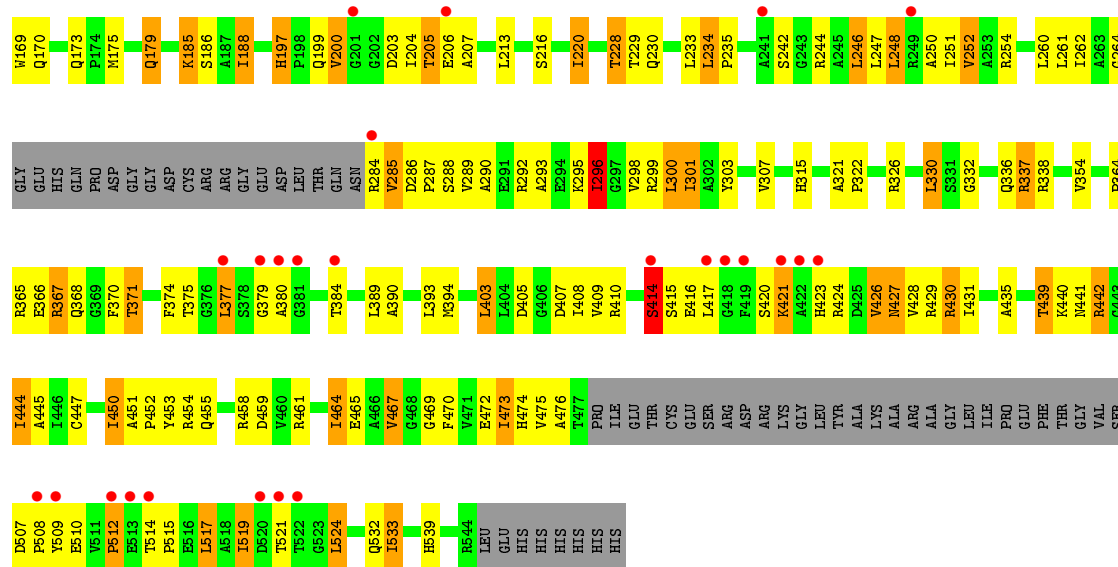
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	31	Total 31	O 31	0	0
2	B	22	Total 22	O 22	0	0
2	C	33	Total 33	O 33	0	0

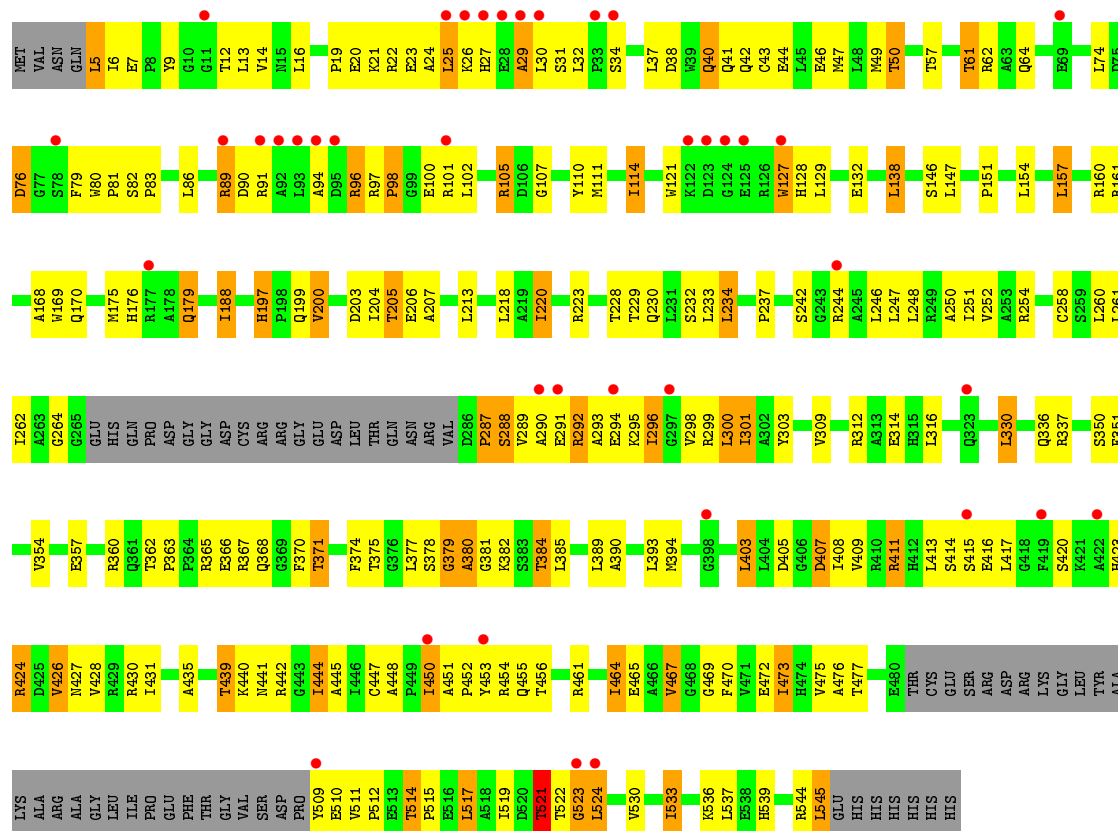


- Molecule 1: Sulfate adenylyltransferase, adenylylsulfate kinase





- Molecule 1: Sulfate adenylyltransferase, adenylylsulfate kinase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	161.07Å 227.21Å 106.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.03 – 2.95 35.03 – 2.95	Depositor EDS
% Data completeness (in resolution range)	92.1 (35.03-2.95) 92.1 (35.03-2.95)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.13 (at 2.95Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.243 , 0.282 0.238 , 0.280	Depositor DCC
$R_{free}$ test set	1889 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	62.7	Xtriage
Anisotropy	0.134	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 55.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	11339	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% 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

### 5.1 Standard geometry

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	0.67	2/3843 (0.1%)	0.87	15/5227 (0.3%)
1	B	0.68	0/3825	0.83	8/5205 (0.2%)
1	C	0.65	0/3829	0.88	12/5209 (0.2%)
All	All	0.66	2/11497 (0.0%)	0.86	35/15641 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	258	CYS	CB-SG	-5.88	1.72	1.81
1	A	447	CYS	CB-SG	-5.42	1.73	1.81

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	88	SER	CB-CA-C	11.04	131.07	110.10
1	C	522	THR	N-CA-CB	-10.76	89.85	110.30
1	C	521	THR	CB-CA-C	-10.58	83.03	111.60
1	B	416	GLU	N-CA-CB	-9.50	93.50	110.60
1	C	29	ALA	CB-CA-C	-9.35	96.08	110.10

There are no chirality outliers.

All (2) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	286	ASP	Peptide
1	C	380	ALA	Peptide

## 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	3761	0	3716	267	0
1	B	3744	0	3685	289	1
1	C	3748	0	3684	235	0
2	A	31	0	0	2	0
2	B	22	0	0	0	0
2	C	33	0	0	4	0
All	All	11339	0	11085	773	1

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 35.

The worst 5 of 773 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:414:SER:HB2	1:A:417:LEU:CD1	1.57	1.35
1:B:4:GLN:NE2	1:B:5:LEU:HB2	1.42	1.34
1:C:288:SER:O	1:C:291:GLU:HB2	1.32	1.28
1:A:244:ARG:N	1:A:244:ARG:HD2	1.40	1.22
1:C:287:PRO:O	1:C:291:GLU:HG2	1.32	1.22

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:430:ARG:NH2	1:B:430:ARG:NH2[3_455]	1.87	0.33

## 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	487/552 (88%)	441 (91%)	44 (9%)	2 (0%)	34	69
1	B	487/552 (88%)	440 (90%)	42 (9%)	5 (1%)	15	48
1	C	487/552 (88%)	438 (90%)	46 (9%)	3 (1%)	25	60
All	All	1461/1656 (88%)	1319 (90%)	132 (9%)	10 (1%)	22	56

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	366	GLU
1	C	366	GLU
1	A	98	PRO
1	B	69	GLU
1	B	366	GLU

### 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	376/449 (84%)	311 (83%)	65 (17%)	2	8
1	B	373/449 (83%)	295 (79%)	78 (21%)	1	4
1	C	373/449 (83%)	301 (81%)	72 (19%)	1	6
All	All	1122/1347 (83%)	907 (81%)	215 (19%)	1	6

5 of 215 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	197	HIS
1	B	371	THR
1	C	407	ASP
1	B	205	THR
1	B	248	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	27	HIS
1	B	173	GLN
1	C	230	GLN
1	A	539	HIS
1	B	4	GLN

### 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 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, 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	493/552 (89%)	0.26	30 (6%)	21 12	30, 55, 94, 118	11 (2%)
1	B	493/552 (89%)	0.24	33 (6%)	17 10	26, 57, 93, 129	6 (1%)
1	C	493/552 (89%)	0.24	39 (7%)	12 7	26, 57, 94, 114	7 (1%)
All	All	1479/1656 (89%)	0.24	102 (6%)	16 10	26, 56, 94, 129	24 (1%)

The worst 5 of 102 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	124	GLY	5.7
1	B	422	ALA	5.6
1	B	379	GLY	5.6
1	B	380	ALA	5.4
1	B	418	GLY	5.3

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.