



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 07:42 AM UTC

PDB ID : 5EP8 / pdb_00005ep8
Title : X-Ray Structure of the Complex Pyrimidine-nucleoside phosphorylase from
Bacillus subtilis with Sulfate Ion
Authors : Lashkov, A.A.; Balaev, V.V.; Gabdoulkhakov, A.G.; Betzel, C.; Mikhailov,
A.M.
Deposited on : 2015-11-11
Resolution : 2.66 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

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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

i

X-RAY DIFFRACTION

A.

the following graphic. The table shows the number of entries on which the scores are based.



Percentile relative to all X-ray structures

Percentile relative to X-ray structures of similar resolution

MetricWhole archive
(#Entries)

Similar resolution
(#Entries, resolution range(Å))

 R_{free}

Clashscore

Ramachandran outliers

Sidechain outliers

RSRZ outliers

180053

190562

187476

187428

180081

1110 (2.66-2.66)

1141 (2.66-2.66)

1126 (2.66-2.66)

1126 (2.66-2.66)

1110 (2.66-2.66)

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 ≥ 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

433

93%

7%

1

B

433

94%

5%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6510 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.

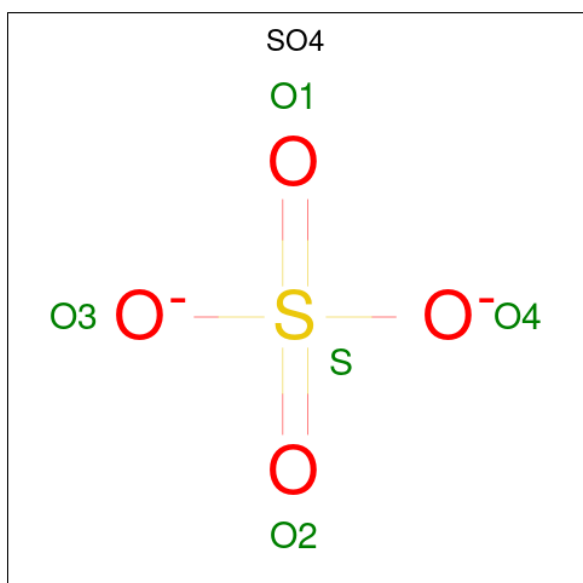
- Molecule 1 is a protein called Pyrimidine-nucleoside phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	433	Total	C	N	O	S	0	1	0
			3237	2033	548	640	16			
1	B	433	Total	C	N	O	S	0	0	0
			3235	2031	548	640	16			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	ARG	GLN	conflict	UNP P39142
A	128	ASP	MET	conflict	UNP P39142
A	137	ARG	ASP	conflict	UNP P39142
B	48	ARG	GLN	conflict	UNP P39142
B	128	ASP	MET	conflict	UNP P39142
B	137	ARG	ASP	conflict	UNP P39142

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Na	0	0
			1	1		

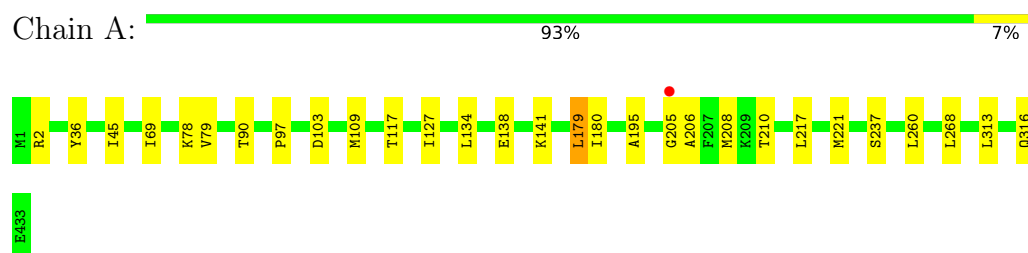
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	15	Total	O	0	0
			15	15		
4	B	12	Total	O	0	0
			12	12		

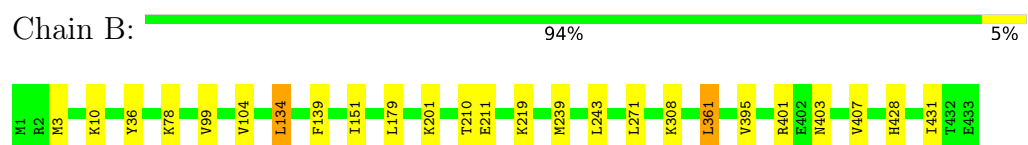
3 Residue-property plots

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: Pyrimidine-nucleoside phosphorylase



- Molecule 1: Pyrimidine-nucleoside phosphorylase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	80.46Å 91.35Å 109.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.22 – 2.66 70.22 – 2.66	Depositor EDS
% Data completeness (in resolution range)	99.4 (70.22-2.66) 99.4 (70.22-2.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 2.65Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.215 , 0.293 0.219 , 0.291	Depositor DCC
R_{free} test set	1171 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å ²)	65.0	Xtriage
Anisotropy	0.550	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 47.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6510	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 36.13 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.2394e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, NA

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	0.54	0/3272	0.86	0/4412
1	B	0.53	0/3267	0.85	0/4405
All	All	0.54	0/6539	0.85	0/8817

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

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	3237	0	3338	9	0
1	B	3235	0	3336	10	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	B	1	0	0	0	0
4	A	15	0	0	0	0
4	B	12	0	0	0	0
All	All	6510	0	6674	18	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 1.

All (18) 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:205:GLY:O	1:A:206:ALA:HB3	1.96	0.64
1:B:401:ARG:NH1	1:B:403:ASN:O	2.42	0.53
1:A:205:GLY:O	1:A:206:ALA:CB	2.57	0.52
1:A:79:VAL:HG12	1:A:195:ALA:HB3	1.92	0.51
1:A:217:LEU:HG	1:A:221:MET:HE2	1.96	0.48
1:A:127:ILE:HD13	1:A:313:LEU:HD13	1.98	0.45
1:B:201:LYS:HG2	1:B:239:MET:HG3	1.98	0.45
1:B:211:GLU:OE1	1:B:428:HIS:ND1	2.50	0.45
1:B:99:VAL:HG12	1:B:104:VAL:HB	1.98	0.44
1:B:361:LEU:HD13	1:B:407:VAL:HG21	2.00	0.44
1:A:97:PRO:HG2	1:A:260:LEU:HD21	2.00	0.44
1:A:45:ILE:HG21	1:A:179:LEU:HD13	2.00	0.43
1:B:243:LEU:HD21	1:B:271:LEU:HD23	1.99	0.43
1:A:36:TYR:HA	1:B:36:TYR:HA	2.01	0.43
1:B:134:LEU:CD1	1:B:151:ILE:HG22	2.49	0.42
1:B:139:PHE:CD1	1:B:151:ILE:HG23	2.55	0.42
1:A:69:ILE:HG13	1:A:109:MET:HE1	2.03	0.41
1:B:219:LYS:HG3	1:B:431:ILE:HD13	2.04	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	432/433 (100%)	418 (97%)	13 (3%)	1 (0%)	43	60
1	B	431/433 (100%)	410 (95%)	21 (5%)	0	100	100
All	All	863/866 (100%)	828 (96%)	34 (4%)	1 (0%)	48	66

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ARG

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	344/345 (100%)	327 (95%)	17 (5%)	22	38
1	B	344/345 (100%)	335 (97%)	9 (3%)	40	63
All	All	688/690 (100%)	662 (96%)	26 (4%)	31	49

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	78	LYS
1	A	90	THR
1	A	103	ASP
1	A	117	THR
1	A	134	LEU
1	A	138	GLU
1	A	141	LYS
1	A	179	LEU
1	A	180	ILE
1	A	208	MET
1	A	210	THR
1	A	237	SER
1	A	268	LEU
1	A	316	GLN
1	A	387[A]	LYS
1	A	387[B]	LYS
1	A	395	VAL
1	B	3	MET
1	B	10	LYS
1	B	78	LYS
1	B	134	LEU
1	B	179	LEU
1	B	210	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	308	LYS
1	B	361	LEU
1	B	395	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	ASN
1	A	64	ASN
1	A	153	GLN
1	A	249	ASN
1	A	414	ASN
1	B	26	ASN
1	B	64	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	SO4	A	501	-	4,4,4	0.42	0	6,6,6	0.12	0
2	SO4	B	501	-	4,4,4	0.45	0	6,6,6	0.18	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	433/433 (100%)	-0.26	1 (0%) 91 90	38, 74, 103, 127	1 (0%)
1	B	433/433 (100%)	-0.26	0 100 100	43, 74, 97, 131	0
All	All	866/866 (100%)	-0.26	1 (0%) 92 91	38, 74, 100, 131	1 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	205	GLY	2.2

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 oligosaccharides 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, 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	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	SO4	A	501	5/5	0.76	0.08	132,133,136,137	0
2	SO4	B	501	5/5	0.76	0.09	113,115,122,123	0
3	NA	B	502	1/1	0.97	0.04	54,54,54,54	0

6.5 Other polymers [i](#)

There are no such residues in this entry.