



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 10, 2024 – 10:02 AM EST

PDB ID : 9BOU  
Title : Crystal structure of ATP-grasp ligase PruB from *Streptomyces coelicolor* A3(2)  
Authors : Patel, K.P.; Bruner, S.D.  
Deposited on : 2024-05-06  
Resolution : 2.91 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 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

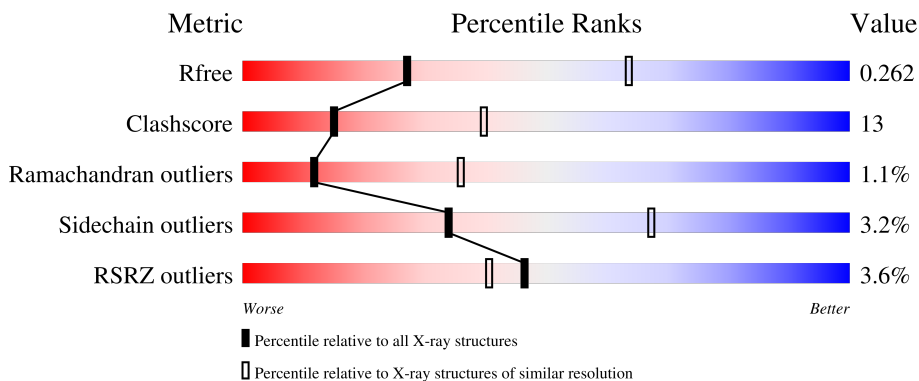
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.91 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))
$R_{free}$	164625	2797 (2.94-2.90)
Clashscore	180529	3049 (2.94-2.90)
Ramachandran outliers	177936	2981 (2.94-2.90)
Sidechain outliers	177891	2983 (2.94-2.90)
RSRZ outliers	164620	2799 (2.94-2.90)

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	339	
1	B	339	

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 8774 atoms, of which 4347 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 MvdD-like pre-ATP grasp domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	293	4428	1403	2202	413	405	5	0	0	0
1	B	284	4310	1366	2145	401	393	5	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	329	ALA	-	expression tag	UNP Q9L2K8
A	330	ALA	-	expression tag	UNP Q9L2K8
A	331	ALA	-	expression tag	UNP Q9L2K8
A	332	LEU	-	expression tag	UNP Q9L2K8
A	333	GLU	-	expression tag	UNP Q9L2K8
A	334	HIS	-	expression tag	UNP Q9L2K8
A	335	HIS	-	expression tag	UNP Q9L2K8
A	336	HIS	-	expression tag	UNP Q9L2K8
A	337	HIS	-	expression tag	UNP Q9L2K8
A	338	HIS	-	expression tag	UNP Q9L2K8
A	339	HIS	-	expression tag	UNP Q9L2K8
B	329	ALA	-	expression tag	UNP Q9L2K8
B	330	ALA	-	expression tag	UNP Q9L2K8
B	331	ALA	-	expression tag	UNP Q9L2K8
B	332	LEU	-	expression tag	UNP Q9L2K8
B	333	GLU	-	expression tag	UNP Q9L2K8
B	334	HIS	-	expression tag	UNP Q9L2K8
B	335	HIS	-	expression tag	UNP Q9L2K8
B	336	HIS	-	expression tag	UNP Q9L2K8
B	337	HIS	-	expression tag	UNP Q9L2K8
B	338	HIS	-	expression tag	UNP Q9L2K8
B	339	HIS	-	expression tag	UNP Q9L2K8

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

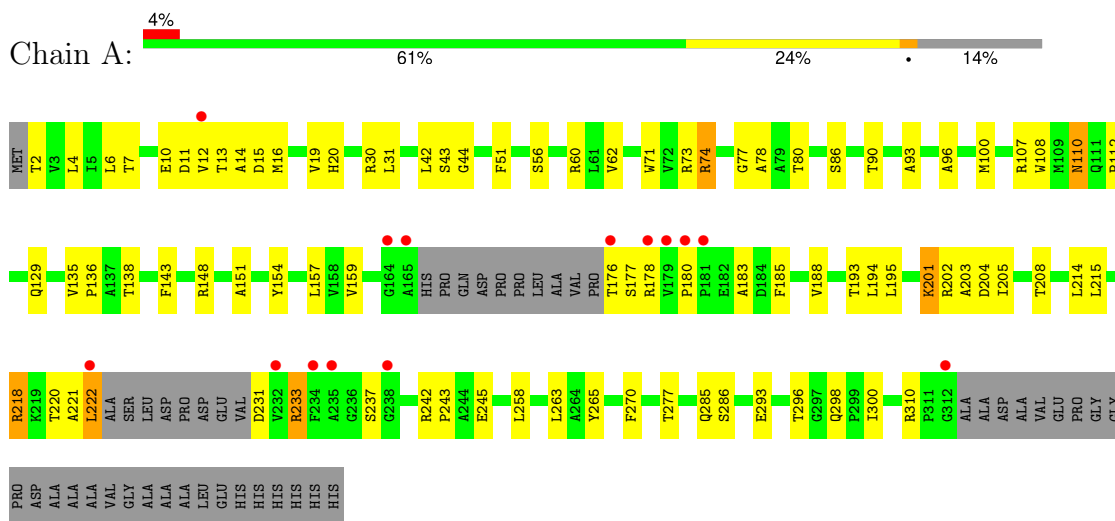
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	8	Total O 8 8	0	0
3	B	18	Total O 18 18	0	0

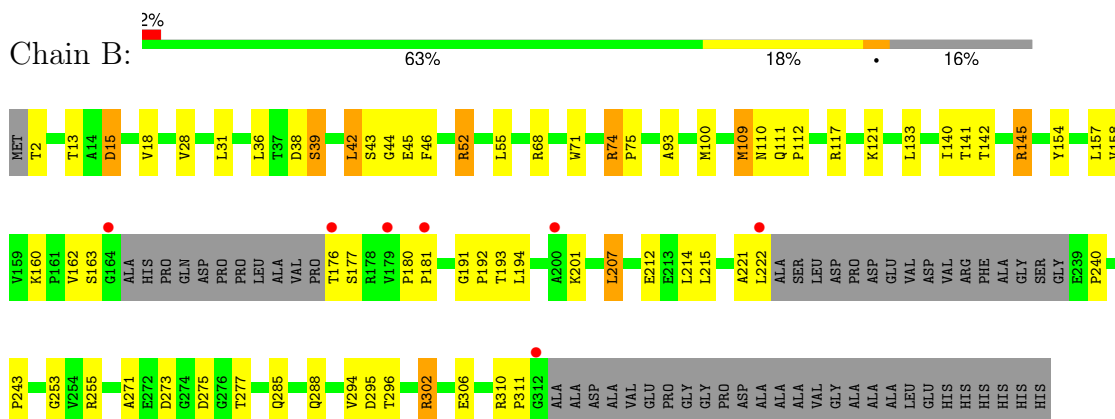
### 3 Residue-property plots [i](#)

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: MvdD-like pre-ATP grasp domain-containing protein



- Molecule 1: MvdD-like pre-ATP grasp domain-containing protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.04Å 222.04Å 87.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.65 – 2.91 49.65 – 2.91	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.65-2.91) 99.5 (49.65-2.91)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.52 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.207 , 0.258 0.209 , 0.262	Depositor DCC
$R_{free}$ test set	1210 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	74.4	Xtrriage
Anisotropy	0.062	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 53.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8774	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% 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: SO4

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	0.64	1/2279 (0.0%)	0.86	0/3103
1	B	0.66	0/2217	0.86	1/3020 (0.0%)
All	All	0.65	1/4496 (0.0%)	0.86	1/6123 (0.0%)

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	3
1	B	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	245	GLU	CG-CD	5.11	1.59	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	207	LEU	CB-CG-CD1	5.43	120.24	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	218	ARG	Sidechain

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	242	ARG	Sidechain
1	A	74	ARG	Sidechain
1	B	302	ARG	Sidechain

## 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	2226	2202	2201	71	0
1	B	2165	2145	2145	57	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
3	A	8	0	0	0	0
3	B	18	0	0	2	0
All	All	4427	4347	4346	118	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 13.

The worst 5 of 118 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:138:THR:HG23	1:A:194:LEU:HD21	1.67	0.77
1:A:215:LEU:N	1:A:215:LEU:HD12	2.03	0.74
1:A:222:LEU:HD12	1:A:222:LEU:O	1.87	0.73
1:A:2:THR:HG21	1:A:62:VAL:HG11	1.69	0.72
1:A:258:LEU:HD12	1:A:263:LEU:HB2	1.74	0.69

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	287/339 (85%)	265 (92%)	19 (7%)	3 (1%)	13	38
1	B	278/339 (82%)	260 (94%)	15 (5%)	3 (1%)	12	36
All	All	565/678 (83%)	525 (93%)	34 (6%)	6 (1%)	12	36

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	110	ASN
1	B	181	PRO
1	A	201	LYS
1	A	237	SER
1	B	240	PRO

### 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	221/253 (87%)	215 (97%)	6 (3%)	40	72
1	B	216/253 (85%)	208 (96%)	8 (4%)	29	62
All	All	437/506 (86%)	423 (97%)	14 (3%)	34	67

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	39	SER
1	B	42	LEU
1	B	158	VAL
1	B	109	MET
1	B	145	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

### 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](#)

2 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	SO4	A	401	-	4,4,4	0.35	0	6,6,6	0.41	0
2	SO4	B	401	-	4,4,4	0.38	0	6,6,6	0.36	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	SO4	1	0

## 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	293/339 (86%)	0.04	14 (4%) 36 31	51, 74, 113, 166	0
1	B	284/339 (83%)	-0.14	7 (2%) 58 52	49, 67, 106, 147	0
All	All	577/678 (85%)	-0.05	21 (3%) 46 40	49, 70, 109, 166	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	234	PHE	4.1
1	A	222	LEU	4.0
1	B	164	GLY	3.5
1	A	165	ALA	3.5
1	A	235	ALA	3.4

### 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( $\text{\AA}^2$ )	Q<0.9
2	SO4	A	401	5/5	0.98	0.06	49,59,65,67	0
2	SO4	B	401	5/5	0.98	0.07	51,63,67,88	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.