



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 1, 2026 – 04:10 PM EDT

PDB ID : 9YU2 / pdb_00009yu2
Title : Crystal structure of evolved dirhodium artificial metalloenzyme 6A from Pfu POP for selective NH insertion
Authors : Saha, P.; Gonzalez-Gutierrez, G.; Lewis, J.C.
Deposited on : 2025-10-22
Resolution : 1.96 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

<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
Xtrriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.015 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

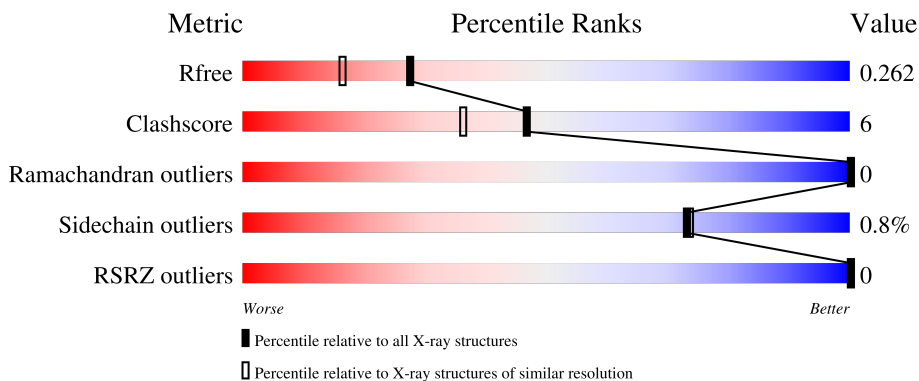
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.96 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}	180053	3494 (1.96-1.96)
Clashscore	190562	3612 (1.96-1.96)
Ramachandran outliers	187476	3587 (1.96-1.96)
Sidechain outliers	187428	3587 (1.96-1.96)
RSRZ outliers	180081	3495 (1.96-1.96)

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	624	
1	B	624	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 20474 atoms, of which 9951 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 Engineered Metalloenzyme 6A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	619	10057	3257	5021	847	920	12	0	0	0
1	B	612	9895	3213	4930	829	911	12	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	98	ARG	GLN	engineered mutation	UNP Q51714
A	99	PHE	GLY	engineered mutation	UNP Q51714
A	104	ALA	GLU	engineered mutation	UNP Q51714
A	119	ARG	ASP	engineered mutation	UNP Q51714
A	142	TYR	TRP	engineered mutation	UNP Q51714
A	146	ALA	PHE	engineered mutation	UNP Q51714
A	194	GLU	PHE	engineered mutation	UNP Q51714
A	199	ALA	LYS	engineered mutation	UNP Q51714
A	202	ALA	ASP	engineered mutation	UNP Q51714
A	209	VAL	THR	engineered mutation	UNP Q51714
A	214	TYR	TRP	engineered mutation	UNP Q51714
A	301	GLY	SER	engineered mutation	UNP Q51714
A	326	HIS	TYR	engineered mutation	UNP Q51714
A	328	HIS	LEU	engineered mutation	UNP Q51714
A	401	SER	TYR	engineered mutation	UNP Q51714
A	464	LEU	ARG	engineered mutation	UNP Q51714
A	477	PHE	SER	engineered mutation	UNP Q51714
A	478	ALA	ASN	engineered mutation	UNP Q51714
A	617	LEU	-	expression tag	UNP Q51714
A	618	GLU	-	expression tag	UNP Q51714
A	619	HIS	-	expression tag	UNP Q51714
A	620	HIS	-	expression tag	UNP Q51714
A	621	HIS	-	expression tag	UNP Q51714
A	622	HIS	-	expression tag	UNP Q51714
A	623	HIS	-	expression tag	UNP Q51714

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	624	HIS	-	expression tag	UNP Q51714
B	98	ARG	GLN	engineered mutation	UNP Q51714
B	99	PHE	GLY	engineered mutation	UNP Q51714
B	104	ALA	GLU	engineered mutation	UNP Q51714
B	119	ARG	ASP	engineered mutation	UNP Q51714
B	142	TYR	TRP	engineered mutation	UNP Q51714
B	146	ALA	PHE	engineered mutation	UNP Q51714
B	194	GLU	PHE	engineered mutation	UNP Q51714
B	199	ALA	LYS	engineered mutation	UNP Q51714
B	202	ALA	ASP	engineered mutation	UNP Q51714
B	209	VAL	THR	engineered mutation	UNP Q51714
B	214	TYR	TRP	engineered mutation	UNP Q51714
B	301	GLY	SER	engineered mutation	UNP Q51714
B	326	HIS	TYR	engineered mutation	UNP Q51714
B	328	HIS	LEU	engineered mutation	UNP Q51714
B	401	SER	TYR	engineered mutation	UNP Q51714
B	464	LEU	ARG	engineered mutation	UNP Q51714
B	477	PHE	SER	engineered mutation	UNP Q51714
B	478	ALA	ASN	engineered mutation	UNP Q51714
B	617	LEU	-	expression tag	UNP Q51714
B	618	GLU	-	expression tag	UNP Q51714
B	619	HIS	-	expression tag	UNP Q51714
B	620	HIS	-	expression tag	UNP Q51714
B	621	HIS	-	expression tag	UNP Q51714
B	622	HIS	-	expression tag	UNP Q51714
B	623	HIS	-	expression tag	UNP Q51714
B	624	HIS	-	expression tag	UNP Q51714

- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0
2	B	1	Total Cl 1 1	0	0

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

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

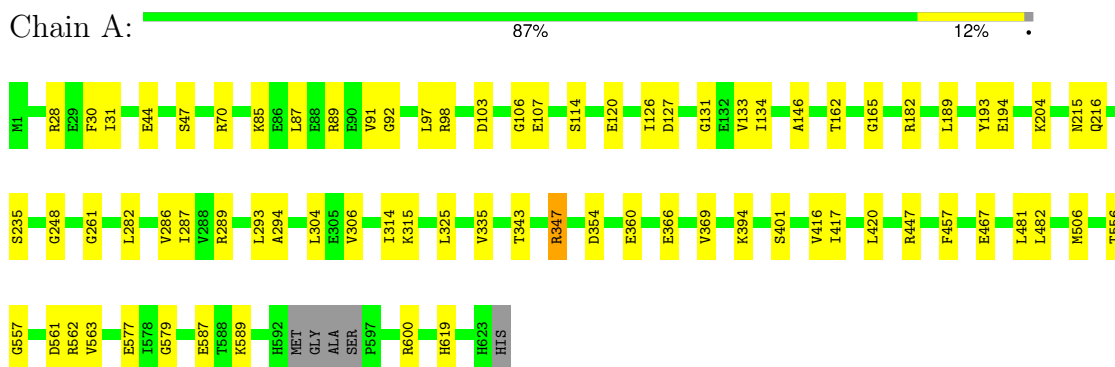
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	290	Total 290	O 290	0	0
4	B	229	Total 229	O 229	0	0

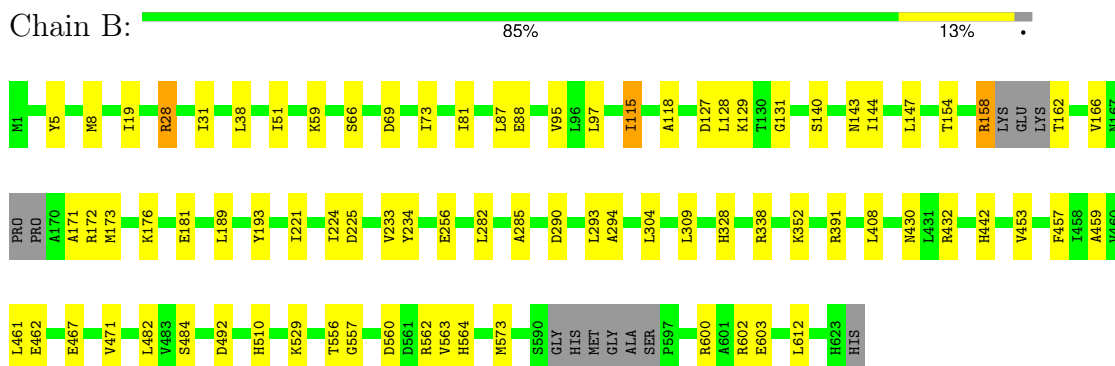
3 Residue-property plots

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: Engineered Metalloenzyme 6A



• Molecule 1: Engineered Metalloenzyme 6A



4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	111.52Å 67.88Å 187.47Å 90.00° 105.49° 90.00°	Depositor
Resolution (Å)	57.39 – 1.96 57.39 – 1.96	Depositor EDS
% Data completeness (in resolution range)	55.7 (57.39-1.96) 54.9 (57.39-1.96)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 1.97Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.222 , 0.260 0.225 , 0.262	Depositor DCC
R_{free} test set	2677 reflections (2.76%)	wwPDB-VP
Wilson B-factor (Å ²)	25.6	Xtrriage
Anisotropy	0.106	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 34.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.090 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	20474	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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, 0.333 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: NA, CL

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	0.27	1/5163 (0.0%)	0.43	0/6967
1	B	0.28	0/5087	0.48	0/6865
All	All	0.28	1/10250 (0.0%)	0.46	0/13832

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	2
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	417	ILE	CA-CB	6.12	1.58	1.53

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	28	ARG	Sidechain
1	A	347	ARG	Sidechain
1	A	70	ARG	Sidechain
1	B	158	ARG	Sidechain
1	B	28	ARG	Sidechain

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	5036	5021	5020	53	1
1	B	4965	4930	4929	62	1
2	A	1	0	0	1	0
2	B	1	0	0	1	0
3	A	1	0	0	0	0
4	A	290	0	0	28	3
4	B	229	0	0	26	2
All	All	10523	9951	9949	114	6

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 6.

The worst 5 of 114 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:215:ASN:ND2	4:A:802:HOH:O	2.07	0.87
1:A:44:GLU:OE2	4:A:801:HOH:O	2.06	0.73
1:B:73:ILE:HD11	1:B:97:LEU:HD21	1.72	0.71
1:B:564:HIS:O	4:B:801:HOH:O	2.10	0.70
1:B:285:ALA:O	4:B:802:HOH:O	2.10	0.70

The worst 5 of 6 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:5:TYR:OH	1:B:290:ASP:OD2[1_565]	2.12	0.08
4:B:892:HOH:O	4:B:998:HOH:O[4_455]	2.12	0.08
4:A:805:HOH:O	4:A:1022:HOH:O[1_545]	2.13	0.07
1:A:47:SER:O	1:A:89:ARG:NH2[2_556]	2.15	0.05
4:A:874:HOH:O	4:A:947:HOH:O[1_545]	2.15	0.05

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	615/624 (99%)	596 (97%)	19 (3%)	0	100	100
1	B	604/624 (97%)	575 (95%)	29 (5%)	0	100	100
All	All	1219/1248 (98%)	1171 (96%)	48 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	534/537 (99%)	531 (99%)	3 (1%)	78	79
1	B	525/537 (98%)	520 (99%)	5 (1%)	68	67
All	All	1059/1074 (99%)	1051 (99%)	8 (1%)	73	74

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	352	LYS
1	B	158	ARG
1	B	115	ILE
1	B	31	ILE
1	B	140	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	143	ASN
1	B	328	HIS
1	B	621	HIS
1	A	451	GLN
1	A	489	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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	619/624 (99%)	-1.29	0 100 100	13, 30, 57, 98	0
1	B	612/624 (98%)	-1.16	0 100 100	18, 37, 79, 126	0
All	All	1231/1248 (98%)	-1.23	0 100 100	13, 33, 73, 126	0

There are no RSRZ outliers to report.

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	CL	A	701	1/1	0.99	0.06	28,28,28,28	0
2	CL	B	701	1/1	0.99	0.10	36,36,36,36	0
3	NA	A	702	1/1	0.99	0.05	30,30,30,30	0

6.5 Other polymers [i](#)

There are no such residues in this entry.