



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 09:30 PM UTC

PDB ID : 3KHZ / pdb_00003khz
Title : Crystal Structure of R350A mutant of Staphylococcus aureus metallopeptidase (Sapep/DapE) in the apo-form
Authors : Girish, T.S.; Gopal, B.
Deposited on : 2009-10-31
Resolution : 2.50 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
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

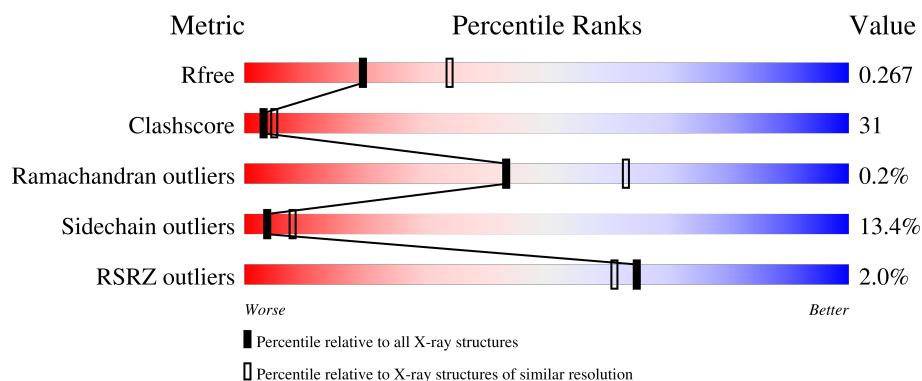
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	5829 (2.50-2.50)
Clashscore	190562	6492 (2.50-2.50)
Ramachandran outliers	187476	6378 (2.50-2.50)
Sidechain outliers	187428	6380 (2.50-2.50)
RSRZ outliers	180081	5833 (2.50-2.50)

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	492	2% 53% 28% 6% • 12%
1	B	492	2% 50% 29% 8% • 12%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 6880 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 dipeptidase SACOL1801.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	433	Total	C	N	O	S	17	0	0
			3382	2154	558	655	15			
1	B	435	Total	C	N	O	S	2	0	0
			3385	2153	552	664	16			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	expression tag	UNP Q5HF23
A	-21	GLY	-	expression tag	UNP Q5HF23
A	-20	SER	-	expression tag	UNP Q5HF23
A	-19	SER	-	expression tag	UNP Q5HF23
A	-18	HIS	-	expression tag	UNP Q5HF23
A	-17	HIS	-	expression tag	UNP Q5HF23
A	-16	HIS	-	expression tag	UNP Q5HF23
A	-15	HIS	-	expression tag	UNP Q5HF23
A	-14	HIS	-	expression tag	UNP Q5HF23
A	-13	HIS	-	expression tag	UNP Q5HF23
A	-12	SER	-	expression tag	UNP Q5HF23
A	-11	SER	-	expression tag	UNP Q5HF23
A	-10	GLY	-	expression tag	UNP Q5HF23
A	-9	LEU	-	expression tag	UNP Q5HF23
A	-8	VAL	-	expression tag	UNP Q5HF23
A	-7	PRO	-	expression tag	UNP Q5HF23
A	-6	ARG	-	expression tag	UNP Q5HF23
A	-5	GLY	-	expression tag	UNP Q5HF23
A	-4	SER	-	expression tag	UNP Q5HF23
A	-3	HIS	-	expression tag	UNP Q5HF23
A	-2	MET	-	expression tag	UNP Q5HF23
A	-1	ALA	-	expression tag	UNP Q5HF23
A	0	SER	-	expression tag	UNP Q5HF23
A	350	ALA	ARG	engineered mutation	UNP Q5HF23
B	-22	MET	-	expression tag	UNP Q5HF23

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-21	GLY	-	expression tag	UNP Q5HF23
B	-20	SER	-	expression tag	UNP Q5HF23
B	-19	SER	-	expression tag	UNP Q5HF23
B	-18	HIS	-	expression tag	UNP Q5HF23
B	-17	HIS	-	expression tag	UNP Q5HF23
B	-16	HIS	-	expression tag	UNP Q5HF23
B	-15	HIS	-	expression tag	UNP Q5HF23
B	-14	HIS	-	expression tag	UNP Q5HF23
B	-13	HIS	-	expression tag	UNP Q5HF23
B	-12	SER	-	expression tag	UNP Q5HF23
B	-11	SER	-	expression tag	UNP Q5HF23
B	-10	GLY	-	expression tag	UNP Q5HF23
B	-9	LEU	-	expression tag	UNP Q5HF23
B	-8	VAL	-	expression tag	UNP Q5HF23
B	-7	PRO	-	expression tag	UNP Q5HF23
B	-6	ARG	-	expression tag	UNP Q5HF23
B	-5	GLY	-	expression tag	UNP Q5HF23
B	-4	SER	-	expression tag	UNP Q5HF23
B	-3	HIS	-	expression tag	UNP Q5HF23
B	-2	MET	-	expression tag	UNP Q5HF23
B	-1	ALA	-	expression tag	UNP Q5HF23
B	0	SER	-	expression tag	UNP Q5HF23
B	350	ALA	ARG	engineered mutation	UNP Q5HF23

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	69	Total O 69 69	0	0
2	B	44	Total O 44 44	0	0

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.

[illegible]

Chain B:

Sequence logo for Chain B. The y-axis represents information content in bits (0.00 to 0.15). The x-axis shows positions 1 to 255. A color scale at the top indicates conservation levels: 2% (red), 50% (green), 29% (yellow), 8% (orange), and 12% (grey). Amino acid labels are provided for each position.

Position	Amino Acid	Information Content (bits)
1	Met	0.00
2	Gly	0.00
3	Ser	0.00
4	Ser	0.00
5	His	0.00
6	His	0.00
7	His	0.00
8	His	0.00
9	His	0.00
10	His	0.00
11	His	0.00
12	Ser	0.00
13	Ser	0.00
14	Gly	0.00
15	Leu	0.00
16	Val	0.00
17	Pro	0.00
18	Arg	0.00
19	Gly	0.00
20	Ser	0.00
21	His	0.00
22	His	0.00
23	Met	0.00
24	A-1	0.00
25	S0	0.00
26	M1	0.00
27	W2	0.00
28	K3	0.00
29	E4	0.00
30	K5	0.00
31	V6	0.00
32	Q7	0.00
33	Q8	0.00
34	Y9	0.00
35	E10	0.00
36	D11	0.00
37	Q12	0.00
38	I13	0.00
39	I14	0.00
40	K18	0.00
41	G19	0.00
42	L20	0.00
43	L21	0.00
44	E24	0.00
45	D28	0.00
46	A32	0.00
47	A36	0.00
48	P37	0.00
49	V38	0.00
50	P42	0.00
51	Y50	0.00
52	R55	0.00
53	D54	0.00
54	G68	0.00
55	G73	0.00
56	D77	0.00
57	V78	0.00
58	L79	0.00
59	L82	0.00
60	C83	0.00
61	H84	0.00
62	V85	0.00
63	D86	0.00
64	V87	0.00
65	V88	0.00
66	D92	0.00
67	G93	0.00
68	W94	0.00
69	D95	0.00
70	S96	0.00
71	N97	0.00
72	P98	0.00
73	F99	0.00
74	V103	0.00
75	T104	0.00
76	A107	0.00
77	I108	0.00
78	I109	0.00
79	A110	0.00
80	R111	0.00
81	G112	0.00
82	T113	0.00
83	L114	0.00
84	D115	0.00
85	D116	0.00
86	K117	0.00
87	T120	0.00
88	K128	0.00
89	I129	0.00
90	M133	0.00
91	N134	0.00
92	V135	0.00
93	D136	0.00
94	P137	0.00
95	M143	0.00
96	T147	0.00
97	D148	0.00
98	E149	0.00
99	R158	0.00
100	T167	0.00
101	L168	0.00
102	G169	0.00
103	D173	0.00
104	H180	0.00
105	G181	0.00
106	E182	0.00
107	K183	0.00
108	I184	0.00
109	I185	0.00
110	T186	0.00
111	L190	0.00
112	V191	0.00
113	Q192	0.00
114	M193	0.00
115	K194	0.00
116	L195	0.00
117	T196	0.00
118	E197	0.00
119	D198	0.00
120	P202	0.00
121	D203	0.00
122	Y204	0.00
123	E205	0.00
124	L206	0.00
125	T207	0.00
126	T208	0.00
127	F209	0.00
128	E213	0.00
129	V218	0.00
130	V228	0.00
131	K229	0.00
132	E230	0.00
133	N231	0.00
134	M232	0.00
135	V235	0.00
136	T236	0.00
137	F239	0.00
138	E240	0.00
139	L243	0.00
140	N246	0



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.75Å 133.52Å 67.71Å 90.00° 95.59° 90.00°	Depositor
Resolution (Å)	49.02 – 2.50 49.02 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.02-2.50) 100.0 (49.02-2.50)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.24 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.208 , 0.269 0.215 , 0.267	Depositor DCC
R_{free} test set	1984 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	49.0	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 44.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.029 for l,-k,h	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6880	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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](#)

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	0.85	1/3457 (0.0%)	1.02	13/4691 (0.3%)
1	B	0.85	2/3460 (0.1%)	1.00	11/4700 (0.2%)
All	All	0.85	3/6917 (0.0%)	1.01	24/9391 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	268	VAL	CA-CB	-8.20	1.49	1.55
1	B	87	VAL	CA-CB	-5.71	1.46	1.55
1	B	228	VAL	C-O	-5.01	1.18	1.24

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	355	PHE	N-CA-C	-7.79	98.95	110.48
1	A	230	GLU	CB-CA-C	-7.61	101.11	113.37
1	B	235	VAL	CB-CA-C	-7.59	102.25	111.97
1	A	39	GLY	CA-C-N	7.11	127.42	119.32
1	A	39	GLY	C-N-CA	7.11	127.42	119.32

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	3382	0	3220	198	0
1	B	3385	0	3195	212	0
2	A	69	0	0	5	0
2	B	44	0	0	6	0
All	All	6880	0	6415	410	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 31.

The worst 5 of 410 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:108:ILE:HD12	1:A:109:ILE:N	1.45	1.31
1:A:108:ILE:HD12	1:A:108:ILE:C	1.49	1.29
1:B:93:GLY:N	1:B:442:LYS:HD2	1.44	1.28
1:A:271:MET:HE3	1:A:271:MET:HA	1.17	1.16
1:A:357:PHE:CE1	1:A:361:MET:CE	2.32	1.12

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	429/492 (87%)	405 (94%)	23 (5%)	1 (0%)	43	63
1	B	431/492 (88%)	407 (94%)	23 (5%)	1 (0%)	43	63
All	All	860/984 (87%)	812 (94%)	46 (5%)	2 (0%)	43	63

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	441	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	274	SER

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	354/417 (85%)	309 (87%)	45 (13%)	4	9
1	B	354/417 (85%)	304 (86%)	50 (14%)	3	7
All	All	708/834 (85%)	613 (87%)	95 (13%)	4	8

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	135	VAL
1	B	253	THR
1	B	167	THR
1	B	195	LEU
1	B	307	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	384	HIS
1	B	389	ASN
1	B	453	ASN
1	A	384	HIS
1	A	381	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](#)

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	433/492 (88%)	0.07	9 (2%) 63 59	25, 48, 73, 97	5 (1%)
1	B	435/492 (88%)	0.06	8 (1%) 67 64	27, 48, 79, 97	1 (0%)
All	All	868/984 (88%)	0.07	17 (1%) 65 61	25, 48, 76, 97	6 (0%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	-1	ALA	4.6
1	A	404	THR	3.9
1	A	230	GLU	3.8
1	A	440	HIS	3.3
1	B	36	ALA	3.3

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 oligosaccharides 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.