



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 06:07 AM UTC

PDB ID : 4FR5 / pdb_00004fr5
Title : Crystal Structure of Shikimate Dehydrogenase (aroE) Y210S Mutant from *Helicobacter pylori* in Complex with Shikimate
Authors : Cheng, W.C.; Chen, T.J.; Wang, W.C.
Deposited on : 2012-06-26
Resolution : 2.20 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
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

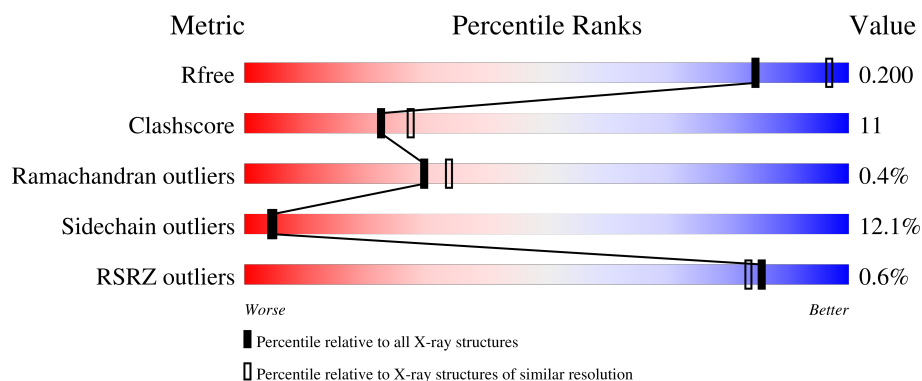
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	6164 (2.20-2.20)
Clashscore	190562	6851 (2.20-2.20)
Ramachandran outliers	187476	6768 (2.20-2.20)
Sidechain outliers	187428	6769 (2.20-2.20)
RSRZ outliers	180081	6166 (2.20-2.20)

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	271	% 72% 22% • •
1	B	271	72% 18% 7% • •

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4297 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.

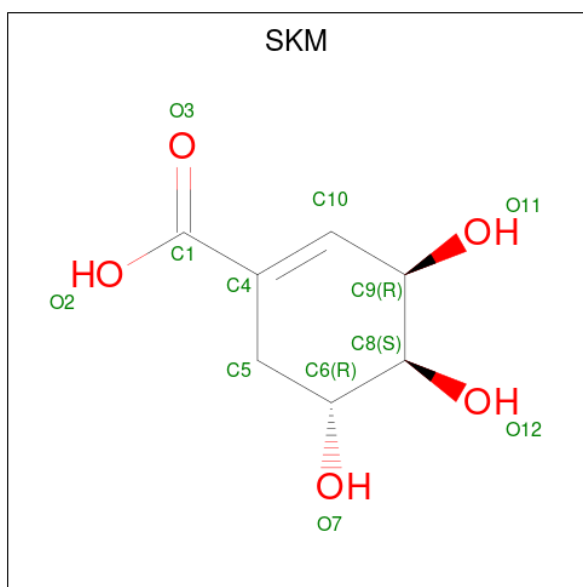
- Molecule 1 is a protein called Shikimate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	263	Total	C	N	O	S	0	0	0
			2058	1338	341	369	10			
1	B	263	Total	C	N	O	S	0	0	0
			2058	1338	341	369	10			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	210	SER	TYR	engineered mutation	UNP P56119
A	264	LEU	-	expression tag	UNP P56119
A	265	GLU	-	expression tag	UNP P56119
A	266	HIS	-	expression tag	UNP P56119
A	267	HIS	-	expression tag	UNP P56119
A	268	HIS	-	expression tag	UNP P56119
A	269	HIS	-	expression tag	UNP P56119
A	270	HIS	-	expression tag	UNP P56119
A	271	HIS	-	expression tag	UNP P56119
B	210	SER	TYR	engineered mutation	UNP P56119
B	264	LEU	-	expression tag	UNP P56119
B	265	GLU	-	expression tag	UNP P56119
B	266	HIS	-	expression tag	UNP P56119
B	267	HIS	-	expression tag	UNP P56119
B	268	HIS	-	expression tag	UNP P56119
B	269	HIS	-	expression tag	UNP P56119
B	270	HIS	-	expression tag	UNP P56119
B	271	HIS	-	expression tag	UNP P56119

- Molecule 2 is (3R,4S,5R)-3,4,5-TRIHIDROXYCYCLOHEX-1-ENE-1-CARBOXYLIC ACID (CCD ID: SKM) (formula: C₇H₁₀O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	7	5		
2	B	1	Total	C	O	0	0
			12	7	5		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	77	Total	O	0	0
			77	77		
3	B	80	Total	O	0	0
			80	80		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	46.51Å 88.52Å 118.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.20 30.00 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.3 (30.00-2.20) 98.2 (30.00-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.51 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.195 , 0.231 0.197 , 0.200	Depositor DCC
R_{free} test set	1299 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	29.7	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 51.5	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	0.94	EDS
Total number of atoms	4297	wwPDB-VP
Average B, all atoms (Å ²)	29.0	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 29.43 % 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 1.5524e-03. 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: SKM

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	1.04	1/2104 (0.0%)	0.90	3/2832 (0.1%)
1	B	1.03	1/2104 (0.0%)	0.90	2/2832 (0.1%)
All	All	1.03	2/4208 (0.0%)	0.90	5/5664 (0.1%)

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	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	82	ILE	CA-CB	5.28	1.62	1.54
1	A	250	ILE	CA-C	-5.26	1.48	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	95	GLU	N-CA-C	5.53	117.23	107.61
1	B	166	MET	N-CA-C	-5.36	106.62	114.39
1	A	225	THR	CB-CA-C	5.30	117.05	109.26
1	A	13	ILE	N-CA-C	5.27	120.29	109.34
1	A	14	LYS	N-CA-C	5.12	118.67	112.93

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	95	GLU	Peptide

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	2058	0	2099	33	0
1	B	2058	0	2099	55	0
2	A	12	0	9	0	0
2	B	12	0	9	0	0
3	A	77	0	0	2	0
3	B	80	0	0	0	0
All	All	4297	0	4216	88	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 11.

The worst 5 of 88 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:B:152:ARG:HD2	1:B:152:ARG:H	1.18	1.09
1:A:199:PHE:HB3	1:A:225:THR:HG21	1.58	0.85
1:B:254:LYS:CE	1:B:254:LYS:HA	2.08	0.83
1:B:152:ARG:HD2	1:B:152:ARG:N	1.91	0.83
1:B:196:LYS:HG2	1:B:223:LEU:HD21	1.59	0.82

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	261/271 (96%)	249 (95%)	11 (4%)	1 (0%)	30	34
1	B	261/271 (96%)	249 (95%)	11 (4%)	1 (0%)	30	34
All	All	522/542 (96%)	498 (95%)	22 (4%)	2 (0%)	30	34

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	114	LYS
1	A	150	SER

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	223/231 (96%)	197 (88%)	26 (12%)	5	5
1	B	223/231 (96%)	195 (87%)	28 (13%)	4	4
All	All	446/462 (96%)	392 (88%)	54 (12%)	5	4

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	27	LEU
1	B	82	ILE
1	B	231	LYS
1	B	28	THR
1	B	47	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	115	GLN
1	B	158	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	148	ASN
1	B	186	ASN
1	A	140	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](#)

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	SKM	A	500	-	12,12,12	2.11	5 (41%)	16,17,17	0.99	1 (6%)
2	SKM	B	500	-	12,12,12	1.99	5 (41%)	16,17,17	0.87	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SKM	A	500	-	-	0/4/20/20	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SKM	B	500	-	-	0/4/20/20	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	500	SKM	C9-C10	4.01	1.55	1.50
2	B	500	SKM	C5-C4	3.48	1.57	1.51
2	B	500	SKM	C10-C4	3.10	1.40	1.33
2	B	500	SKM	C1-C4	2.87	1.55	1.49
2	A	500	SKM	C10-C4	2.78	1.39	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	500	SKM	C6-C8-C9	2.78	114.11	110.38

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	263/271 (97%)	-0.28	2 (0%) 82 80	13, 26, 47, 63	0
1	B	263/271 (97%)	-0.19	1 (0%) 88 87	16, 27, 49, 61	0
All	All	526/542 (97%)	-0.23	3 (0%) 85 83	13, 27, 48, 63	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	152	ARG	2.9
1	A	211	GLY	2.3
1	B	152	ARG	2.1

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	SKM	A	500	12/12	0.80	0.12	47,55,56,57	0
2	SKM	B	500	12/12	0.90	0.08	32,34,36,39	0

6.5 Other polymers [i](#)

There are no such residues in this entry.