



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 30, 2017 – 09:12 PM EST

PDB ID : 3RI6
Title : A Novel Mechanism of Sulfur Transfer Catalyzed by O-Acetylhomoserine
Sulfhydrylase in Methionine Biosynthetic Pathway of Wolinella succinogenes
Authors : Tran, T.H.; Krishnamoorthy, K.; Begley, T.P.; Ealick, S.E.
Deposited on : 2011-04-13
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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028442
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028442

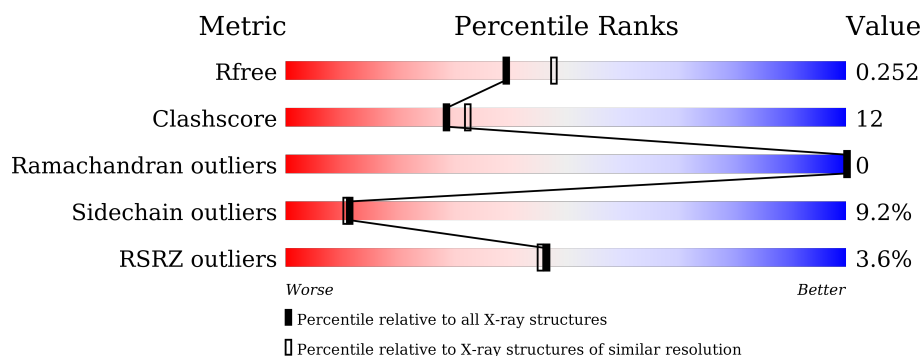
1 Overall quality at a glance

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}	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	430	<div> <div>2%</div> <div> <div></div> <div>65%</div> <div>13%</div> <div>• •</div> <div>18%</div> </div> </div>
1	B	430	<div> <div>4%</div> <div> <div></div> <div>62%</div> <div>13%</div> <div>•</div> <div>21%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5448 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 O-ACETYLHOMOSERINE SULFHYDRYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	353	Total	C	N	O	S	0	0	0
			2669	1708	449	500	12			
1	B	341	Total	C	N	O	S	0	0	0
			2596	1665	437	483	11			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	EXPRESSION TAG	UNP Q7M9C8
A	-21	GLY	-	EXPRESSION TAG	UNP Q7M9C8
A	-20	SER	-	EXPRESSION TAG	UNP Q7M9C8
A	-19	ASP	-	EXPRESSION TAG	UNP Q7M9C8
A	-18	LYS	-	EXPRESSION TAG	UNP Q7M9C8
A	-17	ILE	-	EXPRESSION TAG	UNP Q7M9C8
A	-16	HIS	-	EXPRESSION TAG	UNP Q7M9C8
A	-15	HIS	-	EXPRESSION TAG	UNP Q7M9C8
A	-14	HIS	-	EXPRESSION TAG	UNP Q7M9C8
A	-13	HIS	-	EXPRESSION TAG	UNP Q7M9C8
A	-12	HIS	-	EXPRESSION TAG	UNP Q7M9C8
A	-11	HIS	-	EXPRESSION TAG	UNP Q7M9C8
A	-10	SER	-	EXPRESSION TAG	UNP Q7M9C8
A	-9	SER	-	EXPRESSION TAG	UNP Q7M9C8
A	-8	GLY	-	EXPRESSION TAG	UNP Q7M9C8
A	-7	GLU	-	EXPRESSION TAG	UNP Q7M9C8
A	-6	ASN	-	EXPRESSION TAG	UNP Q7M9C8
A	-5	LEU	-	EXPRESSION TAG	UNP Q7M9C8
A	-4	TYR	-	EXPRESSION TAG	UNP Q7M9C8
A	-3	PHE	-	EXPRESSION TAG	UNP Q7M9C8
A	-2	GLN	-	EXPRESSION TAG	UNP Q7M9C8
A	-1	GLY	-	EXPRESSION TAG	UNP Q7M9C8
A	0	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-22	MET	-	EXPRESSION TAG	UNP Q7M9C8
B	-21	GLY	-	EXPRESSION TAG	UNP Q7M9C8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	SER	-	EXPRESSION TAG	UNP Q7M9C8
B	-19	ASP	-	EXPRESSION TAG	UNP Q7M9C8
B	-18	LYS	-	EXPRESSION TAG	UNP Q7M9C8
B	-17	ILE	-	EXPRESSION TAG	UNP Q7M9C8
B	-16	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-15	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-14	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-13	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-12	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-11	HIS	-	EXPRESSION TAG	UNP Q7M9C8
B	-10	SER	-	EXPRESSION TAG	UNP Q7M9C8
B	-9	SER	-	EXPRESSION TAG	UNP Q7M9C8
B	-8	GLY	-	EXPRESSION TAG	UNP Q7M9C8
B	-7	GLU	-	EXPRESSION TAG	UNP Q7M9C8
B	-6	ASN	-	EXPRESSION TAG	UNP Q7M9C8
B	-5	LEU	-	EXPRESSION TAG	UNP Q7M9C8
B	-4	TYR	-	EXPRESSION TAG	UNP Q7M9C8
B	-3	PHE	-	EXPRESSION TAG	UNP Q7M9C8
B	-2	GLN	-	EXPRESSION TAG	UNP Q7M9C8
B	-1	GLY	-	EXPRESSION TAG	UNP Q7M9C8
B	0	HIS	-	EXPRESSION TAG	UNP Q7M9C8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	115	Total	O	0	0
			115	115		
2	B	68	Total	O	0	0
			68	68		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Chain A:

Position	Amino Acid	Information Content (bits)
1	Met	0.25
2	Gly	0.15
3	Ser	0.10
4	Asp	0.05
5	Lys	0.05
6	Ile	0.05
7	His	0.05
8	His	0.05
9	His	0.05
10	His	0.15
11	His	0.05
12	Ser	0.05
13	Ser	0.05
14	Gly	0.05
15	Glu	0.05
16	Asn	0.05
17	Leu	0.05
18	Tyr	0.05
19	Phe	0.05
20	Gln	0.05
21	Gly	0.05
22	His	0.05
23	Met	0.05
24	R2	0.05
25	T5	0.15
26	T6	0.10
27	R7	0.05
28	A8	0.05
29	L9	0.05
30	P12	0.15
31	Lys	0.05
32	Ala	0.05
33	Lys	0.05
34	Arg	0.05
35	Asp	0.05
36	Val	0.05
37	His	0.05
38	Ala	0.05
39	Leu	0.05
40	Arg	0.05
41	Thr	0.05
42	Pro	0.05
43	Val	0.05
44	Ttr	0.05
45	Asp	0.05
46	Asn	0.05
47	Ala	0.05
48	Ala	0.05
49	Phe	0.05
50	Glu	0.05
51	Phe	0.05
52	Glu	0.05
53	Asn	0.05
54	Ser	0.05
55	Asp	0.05
56	Cut	0.05
57	Ile	0.05
58	Ala	0.05
59	Gln	0.05
60	Val	0.05
61	Ser	0.05
62	Lys	0.05
63	Leu	0.05
64	Gly	0.05
65	Arg	0.05
66	Ala	0.05
67	Leu	0.05
68	Gly	0.05
69	His	0.05
70	Val	0.05
71	Tyr	0.05
72	Ser	0.05
73	Gly	0.05
74	Arg	0.05
75	Ser	0.05
76	S57	0.15
77	L64	0.10
78	L63	0.05
79	L75	0.05
80	L78	0.05
81	S82	0.05
82	A86	0.05
83	L92	0.05
84	R96	0.05
85	T104	0.05
86	D105	0.05
87	R106	0.05
88	L119	0.05
89	D130	0.05
90	V131	0.05
91	M132	0.05
92	K146	0.05
93	L147	0.05
94	L148	0.05
95	F149	0.05
96	L150	0.05
97	E151	0.05
98	T152	0.05
99	N155	0.05
100	P156	0.05
101	Q157	0.05
102	L158	0.05
103	Q159	0.05
104	T162	0.05

Chain B:

4% 62% 13% 1% 21%

Amino Acid	Frequency (%)
MET	4%
GLY	62%
SER	13%
ASP	1%
LYS	21%
ILE	
HIS	
HIS	
HIS	
HIS	
HIS	
HIS	
HIS	
HIS	
GLY	
GLY	
ASN	
LEU	
TYR	
PHE	
GLN	
GLY	
HIS	
MET	
R2	
G3	
T6	
R7	
V11	
PRQ	
LVS	
ALA	
LVS	
ARG	
ASP	
VAL	
HIS	
GLY	
ALA	
LEU	
ARG	
THR	
PRQ	
VAL	
TYR	
ASP	
ASN	
ALA	
ALA	
PHE	
GLU	
PHE	
GLU	
ASN	
SER	
ASP	
GLU	
TYR	
ALA	
VAL	
SER	
L64	
R67	
L68	
T72	
L75	
L78	
A86	
T89	
L92	
R96	
V101	
V102	
T103	
F108	
F115	
L119	
D130	
V131	
M132	
K146	
L147	
L148	
F149	
T152	
I153	
A154	
M155	
P156	
L163	
V170	
I175	
P176	
V178	
V179	
D180	
T184	
P185	
P186	
V200	
L207	
SER	
GLY	
GLY	
G211	
V214	
L218	
K227	
A233	
P234	
A237	
M264	
L270	
R278	
Q287	
V292	
L293	
P297	
Q298	
V299	
K300	
G301	
V302	
N303	
H304	
P305	
S306	
L307	
F312	
I315	
A316	
K317	
F318	
Q319	
T328	
V338	
M341	
D342	
L346	
R349	
I360	
L361	
S362	
P363	
V384	
HIS	
VAL	
ILE	
TYR	
ALA	
LEU	
ASN	
SER	
HIS	
GLU	
GLU	
ARG	
LEU	
LYS	
LEU	
LEU	
GLU	
I381	
S382	
P383	
A384	
R387	
D387	
L398	
K399	
E400	
L403	
L406	
CYS	

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	161.80 Å 62.50 Å 91.60 Å 90.00° 120.50° 90.00°	Depositor
Resolution (Å)	30.00 – 2.20 29.69 – 2.19	Depositor EDS
% Data completeness (in resolution range)	99.0 (30.00-2.20) 98.9 (29.69-2.19)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.18 (at 2.20 Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.225 , 0.251 0.226 , 0.252	Depositor DCC
R_{free} test set	2015 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	23.3	Xtriage
Anisotropy	0.394	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 40.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5448	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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	1.05	3/2718 (0.1%)	0.95	10/3691 (0.3%)
1	B	1.07	5/2644 (0.2%)	0.99	11/3590 (0.3%)
All	All	1.06	8/5362 (0.1%)	0.97	21/7281 (0.3%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	101	VAL	CB-CG1	-8.35	1.35	1.52
1	B	178	VAL	CB-CG1	-8.29	1.35	1.52
1	B	103	THR	CB-CG2	-6.33	1.31	1.52
1	A	282	SER	CB-OG	-6.28	1.34	1.42
1	A	170	VAL	CB-CG2	-5.56	1.41	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	178	VAL	CG1-CB-CG2	-10.03	94.85	110.90
1	B	278	ARG	NE-CZ-NH1	7.94	124.27	120.30
1	B	278	ARG	NE-CZ-NH2	-7.23	116.69	120.30
1	A	96	ARG	NE-CZ-NH2	-6.88	116.86	120.30
1	B	89	THR	OG1-CB-CG2	-6.32	95.45	110.00

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	2669	0	2680	62	0
1	B	2596	0	2627	60	0
2	A	115	0	0	8	0
2	B	68	0	0	1	0
All	All	5448	0	5307	122	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 12.

The worst 5 of 122 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:132:MET:CE	1:B:312:PHE:CE1	2.30	1.14
1:A:104:THR:HG22	1:A:106:ARG:H	0.99	1.12
1:B:132:MET:CE	1:B:312:PHE:HE1	1.65	1.04
1:B:132:MET:HE3	1:B:312:PHE:HE1	1.19	1.02
1:A:132:MET:HE1	1:A:159:GLN:HB3	1.39	1.01

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	347/430 (81%)	339 (98%)	8 (2%)	0	100	100
1	B	333/430 (77%)	330 (99%)	3 (1%)	0	100	100
All	All	680/860 (79%)	669 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	286/365 (78%)	263 (92%)	23 (8%)	15	15
1	B	281/365 (77%)	252 (90%)	29 (10%)	9	8
All	All	567/730 (78%)	515 (91%)	52 (9%)	11	11

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	6	THR
1	B	92	LEU
1	B	317	LYS
1	B	57	SER
1	B	68	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	319	GLN
1	A	354	HIS
1	B	303	ASN
1	A	304	HIS
1	B	298	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates 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	353/430 (82%)	-0.25	9 (2%) 61 60	13, 22, 40, 54	0
1	B	341/430 (79%)	-0.11	16 (4%) 35 34	14, 22, 36, 45	0
All	All	694/860 (80%)	-0.18	25 (3%) 46 45	13, 22, 38, 54	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	301	CYS	4.5
1	B	57	SER	3.9
1	B	338	TYR	3.9
1	B	381	ILE	3.2
1	A	338	TYR	2.7

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 carbohydrates 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.