W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 17, 2023 – 02:10 PM EST

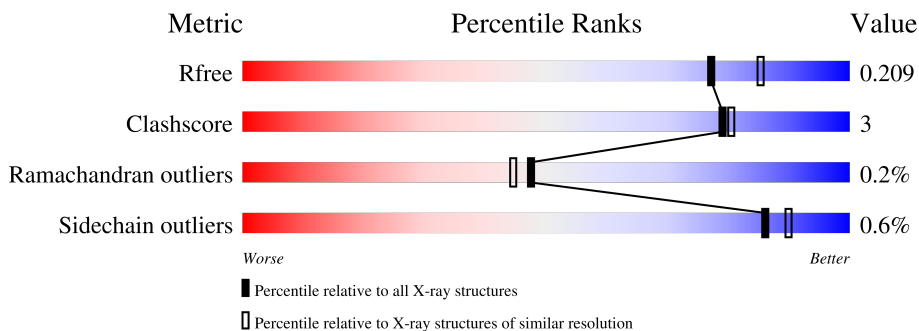|		
PDB ID	:	2P8J
Title	:	Crystal structure of S-adenosylmethionine-dependent methyltransferase (NP_349143.1) from Clostridium acetobutylicum at 2.00 A resolution
Authors	:	Joint Center for Structural Genomics (JCSG)
Deposited on	:	2007-03-22
Resolution	:	2.00 Å(reported)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
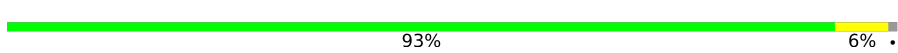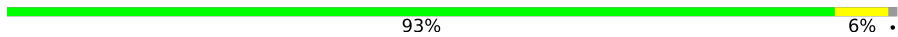*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	Percentile Ranks	Value
Rfree		0.209
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		0.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	209	93% 6% ·
1	B	209	93% 6% ·

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 3761 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 S-adenosylmethionine-dependent methyltransferase.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	206	Total	C	N	O	S	Se		0	3	0
			1651	1065	268	312	3	3				
1	B	207	Total	C	N	O	S	Se		0	7	0
			1686	1086	275	319	3	3				

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q97G40
A	1	MSE	MET	modified residue	UNP Q97G40
A	89	MSE	MET	modified residue	UNP Q97G40
A	101	MSE	MET	modified residue	UNP Q97G40
A	172	MSE	MET	modified residue	UNP Q97G40
B	0	GLY	-	expression tag	UNP Q97G40
B	1	MSE	MET	modified residue	UNP Q97G40
B	89	MSE	MET	modified residue	UNP Q97G40
B	101	MSE	MET	modified residue	UNP Q97G40
B	172	MSE	MET	modified residue	UNP Q97G40

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

wwPDB X-ray Structure Validation Summary Report

SO4

O1

O3 — S — O4

S

O2

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

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

C4

OH O4

O2

C3

C1

C2

HO

O1

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 7	C 4	O 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 7	C 4	O 3	0	0
3	B	1	Total 7	C 4	O 3	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
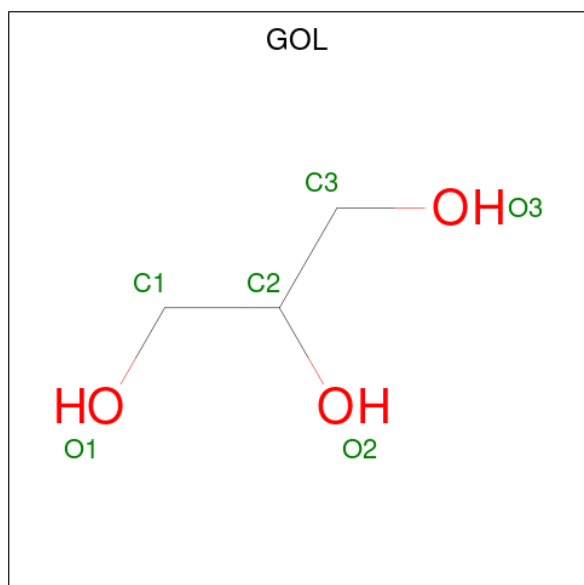
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 6	C 3	O 3	0	0
4	A	1	Total 6	C 3	O 3	0	0
4	B	1	Total 12	C 6	O 6	0	1
4	B	1	Total 6	C 3	O 3	0	0
4	B	1	Total 12	C 6	O 6	0	1
4	B	1	Total 12	C 6	O 6	0	1
4	B	1	Total 6	C 3	O 3	0	0

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).

PGE

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 10	C 6	O 4	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	130	Total 130	O 130	0	0
6	B	188	Total 188	O 188	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. 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. 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: S-adenosylmethionine-dependent methyltransferase

Chain A:
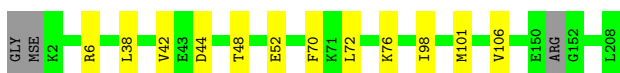
• Molecule 1: S-adenosylmethionine-dependent methyltransferase

Chain B:

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	83.86Å 83.86Å 159.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.16 – 2.00 29.97 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.5 (29.16-2.00) 99.6 (29.97-2.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.165 , 0.201 0.176 , 0.209	Depositor DCC
R_{free} test set	2229 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	34.9	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 60.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3761	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 4.46% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: PGE, GOL, PEG, 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	0.52	1/1685 (0.1%)	0.64	1/2256 (0.0%)
1	B	0.66	1/1732 (0.1%)	0.72	2/2318 (0.1%)
All	All	0.60	2/3417 (0.1%)	0.68	3/4574 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	101	MSE	SE-CE	-6.79	1.55	1.95
1	A	101	MSE	SE-CE	-6.50	1.57	1.95

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	70	PHE	CB-CA-C	-5.23	99.94	110.40
1	A	6	ARG	NE-CZ-NH2	-5.15	117.72	120.30
1	B	86	ASP	CB-CG-OD1	5.10	122.89	118.30

There are no chirality outliers.

There are no planarity outliers.

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	1651	0	1620	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1686	0	1680	12	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
3	A	7	0	10	0	0
3	B	14	0	20	0	0
4	A	12	0	16	2	0
4	B	48	0	64	3	0
5	B	10	0	14	4	0
6	A	130	0	0	2	0
6	B	188	0	0	2	0
All	All	3761	0	3424	19	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 3.

The worst 5 of 19 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:73:ASN:H	4:B:214:GOL:H2	1.52	0.74
1:A:42[A]:VAL:HG21	1:A:72:LEU:HD23	1.75	0.66
1:B:42[A]:VAL:HG21	1:B:72:LEU:HD23	1.78	0.65
1:B:7:GLN:HG2	6:B:270:HOH:O	1.97	0.64
1:B:173:LYS:NZ	5:B:210:PGE:H5	2.15	0.61

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	205/209 (98%)	200 (98%)	5 (2%)	0	100	100
1	B	212/209 (101%)	207 (98%)	4 (2%)	1 (0%)	29	23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	417/418 (100%)	407 (98%)	9 (2%)	1 (0%)	47 44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	2	LYS

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	172/183 (94%)	170 (99%)	2 (1%)	71 76
1	B	181/183 (99%)	181 (100%)	0	100 100
All	All	353/366 (96%)	351 (99%)	2 (1%)	86 90

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	ASP
1	A	70	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	57	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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

17 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	#$|Z| > 2$
4	GOL	A	213	-	5,5,5	0.32	0	5,5,5	0.55	0
2	SO4	B	209	-	4,4,4	0.26	0	6,6,6	0.48	0
2	SO4	A	210	-	4,4,4	0.34	0	6,6,6	0.46	0
4	GOL	B	215[B]	-	5,5,5	0.36	0	5,5,5	0.36	0
5	PGE	B	210	-	9,9,9	0.40	0	8,8,8	0.32	0
4	GOL	B	215[A]	-	5,5,5	0.41	0	5,5,5	0.27	0
2	SO4	A	209	-	4,4,4	0.10	0	6,6,6	0.58	0
4	GOL	B	217	-	5,5,5	0.45	0	5,5,5	0.70	0
4	GOL	B	213[B]	-	5,5,5	0.67	0	5,5,5	0.49	0
4	GOL	B	216[A]	-	5,5,5	0.36	0	5,5,5	0.39	0
3	PEG	B	211	-	6,6,6	0.56	0	5,5,5	0.35	0
3	PEG	A	211	-	6,6,6	0.49	0	5,5,5	0.32	0
3	PEG	B	212	-	6,6,6	0.75	0	5,5,5	0.66	0
4	GOL	B	216[B]	-	5,5,5	0.44	0	5,5,5	0.22	0
4	GOL	B	214	-	5,5,5	0.53	0	5,5,5	1.41	1 (20%)
4	GOL	A	212	-	5,5,5	0.43	0	5,5,5	0.23	0
4	GOL	B	213[A]	-	5,5,5	0.41	0	5,5,5	0.53	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
4	GOL	A	213	-	-	0/4/4/4	-
4	GOL	B	215[B]	-	-	2/4/4/4	-
5	PGE	B	210	-	-	3/7/7/7	-
4	GOL	B	215[A]	-	-	0/4/4/4	-
4	GOL	B	217	-	-	2/4/4/4	-
4	GOL	B	213[B]	-	-	4/4/4/4	-
4	GOL	B	216[A]	-	-	0/4/4/4	-
3	PEG	B	211	-	-	2/4/4/4	-
3	PEG	A	211	-	-	3/4/4/4	-
3	PEG	B	212	-	-	2/4/4/4	-
4	GOL	B	216[B]	-	-	2/4/4/4	-
4	GOL	B	214	-	-	2/4/4/4	-
4	GOL	A	212	-	-	3/4/4/4	-
4	GOL	B	213[A]	-	-	3/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	214	GOL	O2-C2-C1	2.05	118.17	109.12

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	213[A]	GOL	C1-C2-C3-O3
4	B	213[B]	GOL	O1-C1-C2-C3
4	B	215[B]	GOL	C1-C2-C3-O3
4	B	216[B]	GOL	O1-C1-C2-C3
4	B	217	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	213	GOL	2	0
5	B	210	PGE	4	0
4	B	214	GOL	3	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 ⓘ

6.1   Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3   Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4   Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5   Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.