



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:47 PM GMT

PDB ID : 1M41  
Title : Crystal structure of Escherichia coli alkanesulfonate monooxygenase SsuD at 2.3 Å resolution  
Authors : Eichhorn, E.; Davey, C.A.; Sargent, D.F.; Leisinger, T.; Richmond, T.J.  
Deposited on : 2002-07-02  
Resolution : 2.30 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](mailto: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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

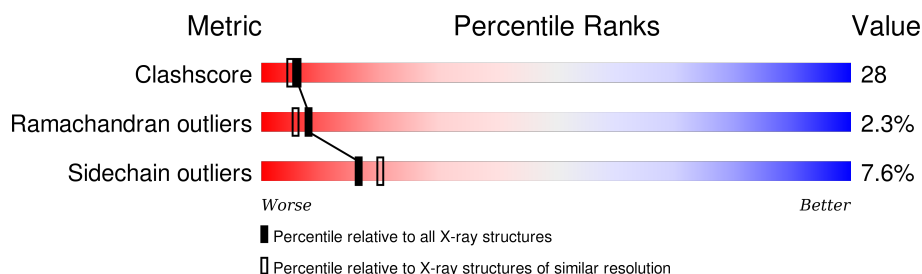
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)

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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	380	
1	B	380	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5408 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 FMNH2-dependent alkanesulfonate monooxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	0	0
			2549	1628	450	468	3			
1	B	328	Total	C	N	O	S	0	0	0
			2549	1628	450	468	3			

- Molecule 2 is water.

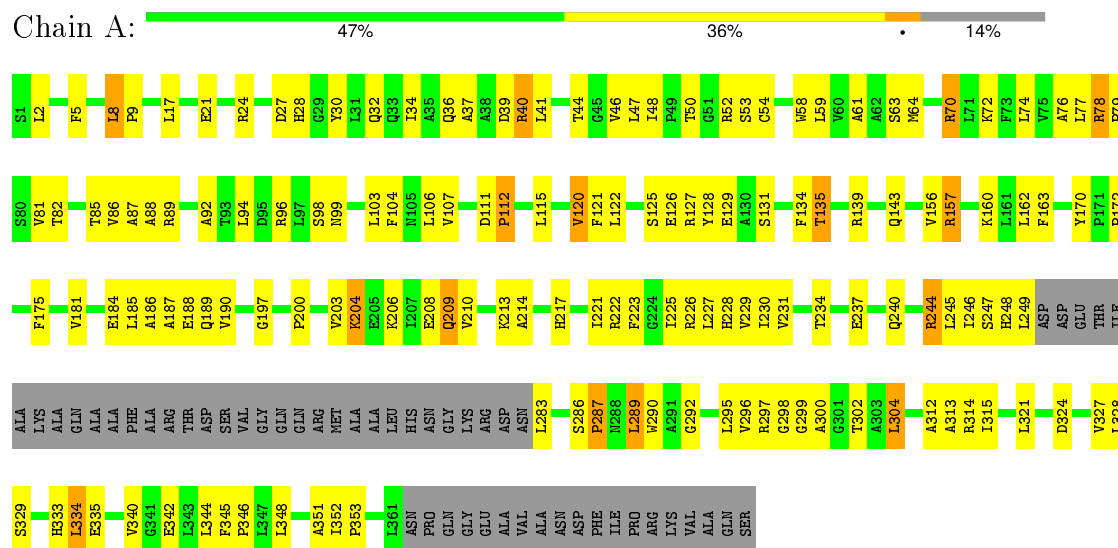
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	144	Total	O	0	0
			144	144		
2	B	166	Total	O	0	0
			166	166		

### 3 Residue-property plots

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.

Note EDS was not executed.

- Molecule 1: FMNH2-dependent alkanesulfonate monooxygenase



P1332	H1333	L1334	E1335	E1336	V1340	L1343	L1344	I1352	P1353	P1356	Q1359	P1360	L1361	ASN	PRO	GLN	GLY	GLU	ALA	VAL	ALA	ASN	ASP	PHE	ILE	PRO	ARG	LYS	VAL	ALA	GLN	SER
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.35Å 140.76Å 125.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.30	Depositor
% Data completeness (in resolution range)	94.4 (25.00-2.30)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.235 , 0.275	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5408	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

## 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	0.54	0/2612	0.79	2/3559 (0.1%)
1	B	0.54	0/2612	0.80	2/3559 (0.1%)
All	All	0.54	0/5224	0.80	4/7118 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1071	LEU	CA-CB-CG	7.75	133.13	115.30
1	A	334	LEU	CA-CB-CG	6.49	130.23	115.30
1	A	77	LEU	CA-CB-CG	5.85	128.75	115.30
1	B	1103	LEU	CA-CB-CG	5.10	127.03	115.30

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	2549	0	2526	165	0
1	B	2549	0	2523	132	0
2	A	144	0	0	9	0
2	B	166	0	0	11	0
All	All	5408	0	5049	286	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 28.

The worst 5 of 286 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:1246:ILE:HG13	1:B:1249:LEU:HD12	1.39	1.03
1:B:1340:VAL:HG13	1:B:1344:LEU:HD12	1.41	0.98
1:A:246:ILE:HG23	1:A:249:LEU:HD12	1.50	0.93
1:A:240:GLN:O	1:A:244:ARG:HG2	1.73	0.87
1:B:1222:ARG:HD2	1:B:1324:ASP:OD2	1.76	0.85

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	324/380 (85%)	290 (90%)	29 (9%)	5 (2%)	13	12
1	B	324/380 (85%)	292 (90%)	22 (7%)	10 (3%)	5	3
All	All	648/760 (85%)	582 (90%)	51 (8%)	15 (2%)	8	6

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	112	PRO
1	A	247	SER
1	B	1178	SER
1	B	1247	SER
1	B	1301	GLY



### 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	263/302 (87%)	242 (92%)	21 (8%)	15	18
1	B	263/302 (87%)	244 (93%)	19 (7%)	18	22
All	All	526/604 (87%)	486 (92%)	40 (8%)	16	20

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	304	LEU
1	B	1021	GLU
1	B	1310	THR
1	A	352	ILE
1	B	1024	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	316	ASN
1	B	1028	HIS
1	B	1209	GLN
1	A	248	HIS
1	B	1136	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.