wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 08:58 pm BST

PDB ID	:	1UC8
Title	:	Crystal structure of a lysine biosynthesis enzyme, Lysx, from thermus thermophilus HB8
Authors	:	Sakai, H.; Vassylyeva, M.N.; Matsuura, T.; Sekine, S.; Nishiyama, M.; Terada, T.; Shirouzu, M.; Kuramitsu, S.; Vassylyev, D.G.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2003-04-09
Resolution	:	2.00 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

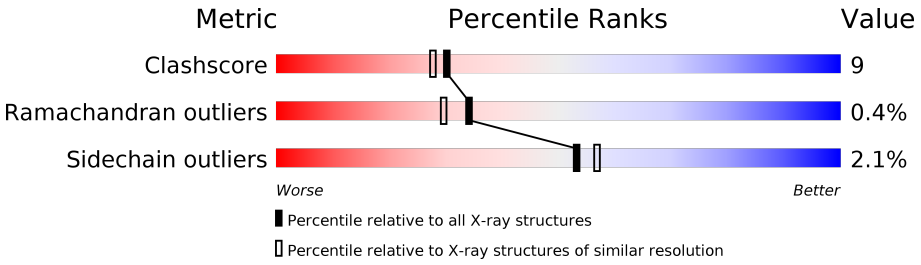
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
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 on 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	280	76% 14% • 9%
1	B	280	75% 14% • 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3951 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 lysine biosynthesis enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	254	Total	C	N	O	S	0	0	0
			1926	1231	334	353	8			
1	B	251	Total	C	N	O	S	0	0	0
			1912	1223	331	350	8			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	141	UNK	LYS	SEE REMARK 999	GB 29467705
A	142	UNK	VAL	SEE REMARK 999	GB 29467705
A	143	UNK	THR	SEE REMARK 999	GB 29467705
A	144	UNK	ASP	SEE REMARK 999	GB 29467705
A	145	UNK	ARG	SEE REMARK 999	GB 29467705
A	146	UNK	ALA	SEE REMARK 999	GB 29467705
A	147	UNK	ALA	SEE REMARK 999	GB 29467705
A	148	UNK	ALA	SEE REMARK 999	GB 29467705
A	149	UNK	GLU	SEE REMARK 999	GB 29467705
A	150	UNK	ALA	SEE REMARK 999	GB 29467705
A	151	UNK	LEU	SEE REMARK 999	GB 29467705
A	152	UNK	LEU	SEE REMARK 999	GB 29467705
A	153	UNK	GLU	SEE REMARK 999	GB 29467705
A	154	UNK	HIS	SEE REMARK 999	GB 29467705
B	141	UNK	LYS	SEE REMARK 999	GB 29467705
B	142	UNK	VAL	SEE REMARK 999	GB 29467705
B	143	UNK	THR	SEE REMARK 999	GB 29467705
B	144	UNK	ASP	SEE REMARK 999	GB 29467705
B	145	UNK	ARG	SEE REMARK 999	GB 29467705
B	146	UNK	ALA	SEE REMARK 999	GB 29467705
B	147	UNK	ALA	SEE REMARK 999	GB 29467705
B	148	UNK	ALA	SEE REMARK 999	GB 29467705
B	149	UNK	GLU	SEE REMARK 999	GB 29467705
B	150	UNK	ALA	SEE REMARK 999	GB 29467705
B	151	UNK	LEU	SEE REMARK 999	GB 29467705

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	152	UNK	LEU	SEE REMARK 999	GB 29467705
B	151	UNK	GLU	SEE REMARK 999	GB 29467705
B	152	UNK	HIS	SEE REMARK 999	GB 29467705

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	59	Total O 59 59	0	0
2	B	54	Total O 54 54	0	0

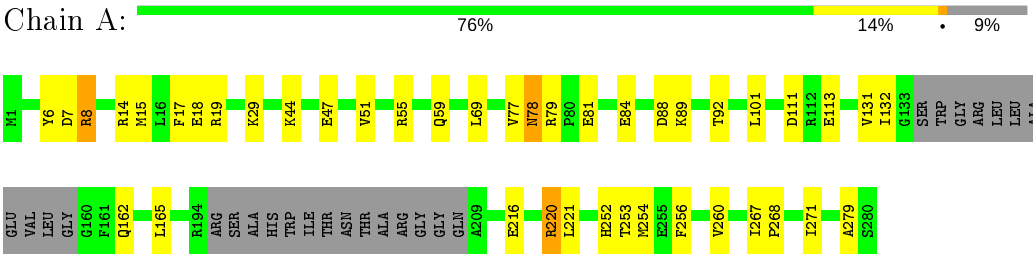
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

Note EDS was not executed.

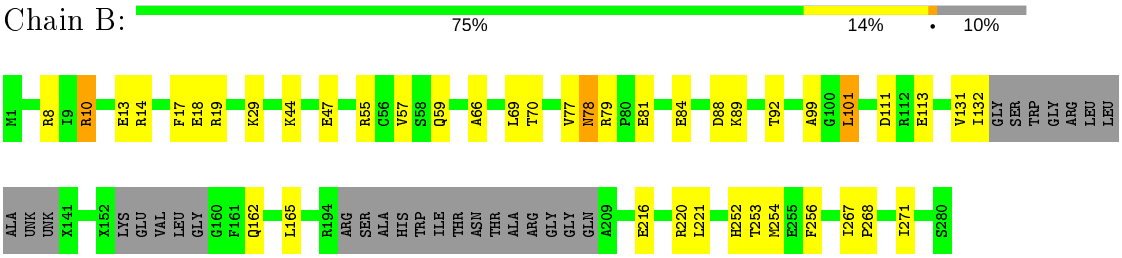
- Molecule 1: lysine biosynthesis enzyme

Chain A:



- Molecule 1: lysine biosynthesis enzyme

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	124.65Å 51.37Å 103.59Å 90.00° 122.81° 90.00°	Depositor
Resolution (Å)	43.53 – 2.00	Depositor
% Data completeness (in resolution range)	96.2 (43.53-2.00)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.240 , 0.269	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3951	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.35	0/1890	0.67	4/2561 (0.2%)
1	B	0.35	0/1886	0.67	3/2556 (0.1%)
All	All	0.35	0/3776	0.67	7/5117 (0.1%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	10	ARG	NE-CZ-NH1	-6.16	117.22	120.30
1	A	8	ARG	NE-CZ-NH2	-6.03	117.29	120.30
1	A	220	ARG	NE-CZ-NH2	-5.89	117.36	120.30
1	A	8	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	B	8	ARG	NE-CZ-NH1	-5.64	117.48	120.30

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	1926	0	1918	33	0
1	B	1912	0	1913	39	0
2	A	59	0	0	1	0
2	B	54	0	0	3	0
All	All	3951	0	3831	71	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 9.

The worst 5 of 71 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:10:ARG:HD3	1:B:13:GLU:OE1	1.73	0.89
1:B:132:ILE:HG13	1:B:162:GLN:HG2	1.72	0.71
1:A:89:LYS:HA	1:A:92:THR:CG2	2.22	0.69
1:B:88:ASP:HA	1:B:132:ILE:HD13	1.75	0.69
1:B:89:LYS:HA	1:B:92:THR:CG2	2.24	0.67

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	234/280 (84%)	228 (97%)	5 (2%)	1 (0%)	34	30
1	B	233/280 (83%)	227 (97%)	5 (2%)	1 (0%)	34	30
All	All	467/560 (83%)	455 (97%)	10 (2%)	2 (0%)	34	30

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	78	ASN
1	B	78	ASN

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	193/212 (91%)	189 (98%)	4 (2%)	53	57
1	B	193/212 (91%)	189 (98%)	4 (2%)	53	57
All	All	386/424 (91%)	378 (98%)	8 (2%)	53	57

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	221	LEU
1	B	221	LEU
1	B	69	LEU
1	A	101	LEU
1	B	55	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	252	HIS
1	B	252	HIS
1	B	103	GLN
1	A	251	ASN
1	B	251	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.