Full wwPDB X-ray Structure Validation Report (i)

Nov 16, 2023 – 03:05 AM JST

PDB ID	:	6KNI
Title	:	Crystal structure of SbnH in complex with the cofactor PLP, a PLP-dependent decarboxylase in Staphyloferrin B biothesynthesis
Authors	:	Tang, J.; Ju, Y.; Zhou, H.
Deposited on	:	2019-08-05
Resolution	:	1.97 Å(reported)

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

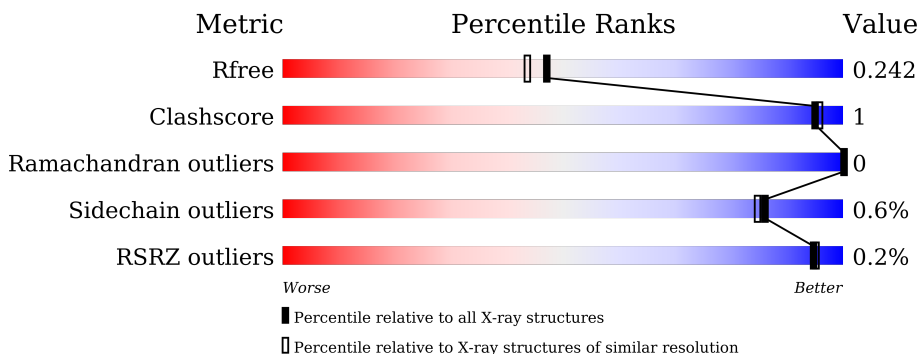|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

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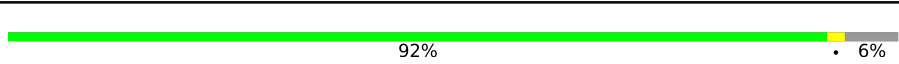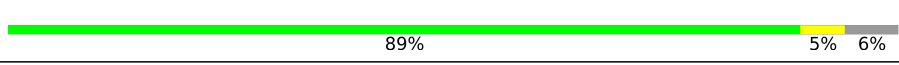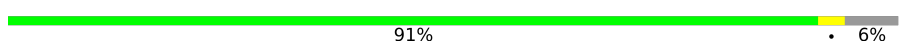*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.97 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		0.242
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		0.6%
RSRZ outliers		0.2%

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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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% 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	406	92% · 6%
1	B	406	89% 5% 6%
1	C	406	91% · 6%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9905 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.
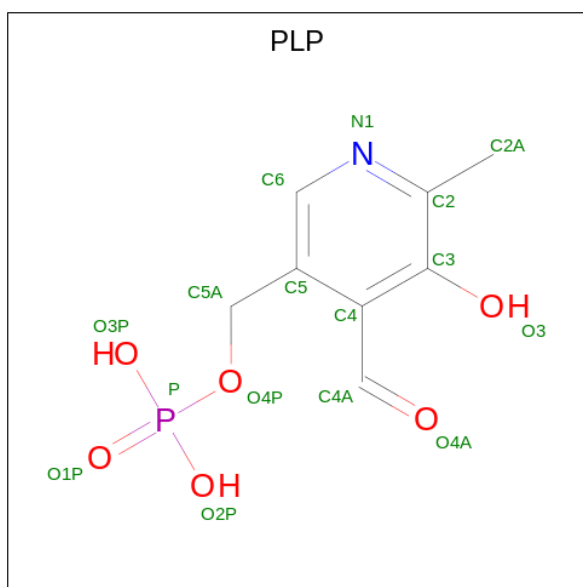
- Molecule 1 is a protein called Probable diaminopimelate decarboxylase protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	S	0	2	0
			3067	1971	517	570	9			
1	B	381	Total	C	N	O	S	0	6	0
			3115	1999	531	575	10			
1	C	383	Total	C	N	O	S	0	3	0
			3087	1982	521	575	9			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	401	HIS	-	expression tag	UNP A0A0H3JPF2
A	402	HIS	-	expression tag	UNP A0A0H3JPF2
A	403	HIS	-	expression tag	UNP A0A0H3JPF2
A	404	HIS	-	expression tag	UNP A0A0H3JPF2
A	405	HIS	-	expression tag	UNP A0A0H3JPF2
A	406	HIS	-	expression tag	UNP A0A0H3JPF2
B	401	HIS	-	expression tag	UNP A0A0H3JPF2
B	402	HIS	-	expression tag	UNP A0A0H3JPF2
B	403	HIS	-	expression tag	UNP A0A0H3JPF2
B	404	HIS	-	expression tag	UNP A0A0H3JPF2
B	405	HIS	-	expression tag	UNP A0A0H3JPF2
B	406	HIS	-	expression tag	UNP A0A0H3JPF2
C	401	HIS	-	expression tag	UNP A0A0H3JPF2
C	402	HIS	-	expression tag	UNP A0A0H3JPF2
C	403	HIS	-	expression tag	UNP A0A0H3JPF2
C	404	HIS	-	expression tag	UNP A0A0H3JPF2
C	405	HIS	-	expression tag	UNP A0A0H3JPF2
C	406	HIS	-	expression tag	UNP A0A0H3JPF2

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).

PLP

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P		0	0
			15	8	1	5	1			
2	B	1	Total	C	N	O	P		0	0
			15	8	1	5	1			
2	C	1	Total	C	N	O	P		0	0
			15	8	1	5	1			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	164	Total	O	0	0
			164	164		
3	B	208	Total	O	0	0
			208	208		
3	C	219	Total	O	0	0
			219	219		

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 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.
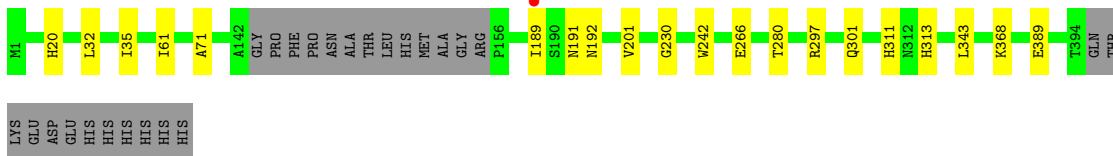
• Molecule 1: Probable diaminopimelate decarboxylase protein

Chain A:

• Molecule 1: Probable diaminopimelate decarboxylase protein

Chain B:

• Molecule 1: Probable diaminopimelate decarboxylase protein

Chain C:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	176.61Å 80.31Å 111.83Å 90.00° 102.15° 90.00°	Depositor
Resolution (Å)	50.00 – 1.97 109.32 – 1.97	Depositor EDS
% Data completeness (in resolution range)	97.5 (50.00-1.97) 97.4 (109.32-1.97)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.01 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.205 , 0.237 0.211 , 0.242	Depositor DCC
R_{free} test set	5228 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	23.9	Xtriage
Anisotropy	0.826	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 38.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9905	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.31% 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: PLP

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	0.36	0/3149	0.60	0/4280
1	B	0.37	0/3198	0.60	0/4341
1	C	0.37	0/3169	0.60	0/4306
All	All	0.37	0/9516	0.60	0/12927

There are no bond length outliers.

There are no bond angle outliers.

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	3067	0	2922	4	0
1	B	3115	0	2999	11	0
1	C	3087	0	2950	7	0
2	A	15	0	6	1	0
2	B	15	0	6	1	0
2	C	15	0	6	0	0
3	A	164	0	0	0	0
3	B	208	0	0	0	0
3	C	219	0	0	0	0
All	All	9905	0	8889	21	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 1.

All (21) 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:189:ILE:HD11	1:B:201:VAL:HG11	1.65	0.75
1:C:20:HIS:CD2	1:C:389:GLU:HB2	2.38	0.59
1:B:189:ILE:HG21	1:B:192:ASN:HB2	1.88	0.55
1:C:313:HIS:HE1	1:C:368:LYS:O	1.90	0.54
1:A:313:HIS:HE1	1:A:368:LYS:O	1.96	0.49
1:C:306:PRO:HG3	1:C:313:HIS:HB2	1.96	0.47
1:B:20:HIS:CD2	1:B:389:GLU:HB2	2.50	0.46
1:A:20:HIS:CD2	1:A:389:GLU:HB2	2.50	0.46
1:B:301:GLN:HB3	1:B:343:LEU:HD11	1.98	0.46
1:B:266:GLU:O	2:B:501:PLP:H6	2.16	0.45
1:A:383:LEU:HD22	1:B:71:ALA:HB3	1.99	0.45
1:B:313:HIS:HE1	1:B:368:LYS:O	2.00	0.44
1:C:305:LEU:HB3	1:C:306:PRO:HD3	2.00	0.43
1:B:191:ASN:HA	1:B:230:GLY:HA2	2.01	0.43
1:C:47:TYR:CE2	1:C:49:MET:HA	2.54	0.43
1:B:280:THR:OG1	1:B:297:ARG:O	2.29	0.42
1:A:266:GLU:O	2:A:501:PLP:H6	2.20	0.41
1:B:32:LEU:HD11	1:B:61:ILE:HG12	2.02	0.41
1:C:19:CYS:HB2	1:C:297:ARG:HD2	2.03	0.41
1:B:35:ILE:HG23	1:B:242:TRP:CZ2	2.55	0.41
1:C:313:HIS:CE1	1:C:368:LYS:O	2.72	0.40

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	379/406 (93%)	370 (98%)	9 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	383/406 (94%)	375 (98%)	8 (2%)	0	100	100
1	C	382/406 (94%)	374 (98%)	8 (2%)	0	100	100
All	All	1144/1218 (94%)	1119 (98%)	25 (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	328/358 (92%)	326 (99%)	2 (1%)	86	85
1	B	338/358 (94%)	337 (100%)	1 (0%)	92	92
1	C	332/358 (93%)	329 (99%)	3 (1%)	78	77
All	All	998/1074 (93%)	992 (99%)	6 (1%)	86	85

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ARG
1	A	311	HIS
1	B	311	HIS
1	C	270	PHE
1	C	311	HIS
1	C	345	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	311	HIS
1	A	313	HIS
1	B	313	HIS
1	C	302	GLN
1	C	313	HIS

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 ⓘ

3 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	PLP	C	501	1	15,15,16	3.17	3 (20%)	20,22,23	1.34	3 (15%)
2	PLP	B	501	1	15,15,16	3.12	3 (20%)	20,22,23	1.41	3 (15%)
2	PLP	A	501	1	15,15,16	3.19	3 (20%)	20,22,23	1.35	2 (10%)

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	PLP	C	501	1	-	0/6/6/8	0/1/1/1
2	PLP	B	501	1	-	0/6/6/8	0/1/1/1
2	PLP	A	501	1	-	0/6/6/8	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	PLP	C5-C4	8.26	1.49	1.40
2	C	501	PLP	C5-C4	8.24	1.49	1.40
2	B	501	PLP	C5-C4	7.91	1.49	1.40
2	A	501	PLP	C3-C2	7.81	1.48	1.40
2	B	501	PLP	C3-C2	7.80	1.48	1.40
2	C	501	PLP	C3-C2	7.59	1.48	1.40
2	C	501	PLP	C3-C4	4.18	1.48	1.40
2	B	501	PLP	C3-C4	4.16	1.48	1.40
2	A	501	PLP	C3-C4	4.04	1.48	1.40

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	PLP	C4A-C4-C5	3.12	124.14	120.94
2	C	501	PLP	C4A-C4-C5	2.96	123.98	120.94
2	A	501	PLP	C4A-C4-C5	2.87	123.89	120.94
2	C	501	PLP	C6-N1-C2	2.63	124.03	119.17
2	B	501	PLP	C6-N1-C2	2.61	124.00	119.17
2	A	501	PLP	C6-N1-C2	2.54	123.87	119.17
2	B	501	PLP	C3-C4-C5	-2.05	116.53	118.74
2	C	501	PLP	C3-C4-C5	-2.00	116.58	118.74

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	PLP	1	0
2	A	501	PLP	1	0

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 ⓘ

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	381/406 (93%)	-0.08	0	100	100	20, 28, 41, 53	0
1	B	381/406 (93%)	-0.10	1 (0%)	94	94	17, 25, 37, 53	0
1	C	383/406 (94%)	-0.13	1 (0%)	94	94	18, 24, 36, 52	0
All	All	1145/1218 (94%)	-0.10	2 (0%)	95	95	17, 25, 39, 53	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	189	ILE	2.4
1	C	309	TRP	2.2

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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	PLP	B	501	15/16	0.95	0.12	24,27,30,31	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	PLP	C	501	15/16	0.96	0.11	21,25,27,29	0
2	PLP	A	501	15/16	0.97	0.10	23,25,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.