



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:23 PM GMT

PDB ID : 4LDP
Title : Spinosyn Forosaminyltransferase SpnP
Authors : Isiorho, E.A.; Liu, H.-W.; Keatinge-Clay, A.K.
Deposited on : 2013-06-24
Resolution : 2.50 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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

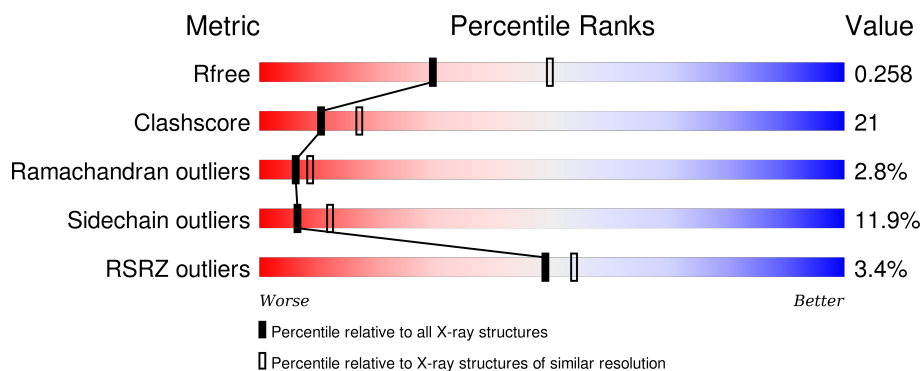
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

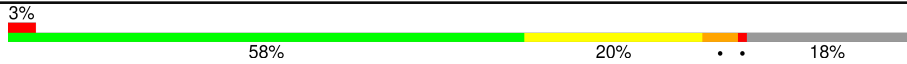

The reported resolution of this entry is 2.50 Å.

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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	455	
1	B	455	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5856 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 NDP-forosamyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	373	Total	C	N	O	S	0	0	0
			2899	1835	509	542	13			
1	B	369	Total	C	N	O	S	0	0	0
			2873	1821	504	535	13			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	HIS	-	EXPRESSION TAG	UNP Q9ALN7
A	-14	HIS	-	EXPRESSION TAG	UNP Q9ALN7
A	-13	HIS	-	EXPRESSION TAG	UNP Q9ALN7
A	-12	HIS	-	EXPRESSION TAG	UNP Q9ALN7
A	-11	HIS	-	EXPRESSION TAG	UNP Q9ALN7
A	-10	HIS	-	EXPRESSION TAG	UNP Q9ALN7
A	-9	SER	-	EXPRESSION TAG	UNP Q9ALN7
A	-8	SER	-	EXPRESSION TAG	UNP Q9ALN7
A	-7	GLY	-	EXPRESSION TAG	UNP Q9ALN7
A	-6	LEU	-	EXPRESSION TAG	UNP Q9ALN7
A	-5	VAL	-	EXPRESSION TAG	UNP Q9ALN7
A	-4	PRO	-	EXPRESSION TAG	UNP Q9ALN7
A	-3	ARG	-	EXPRESSION TAG	UNP Q9ALN7
A	-2	GLY	-	EXPRESSION TAG	UNP Q9ALN7
A	-1	SER	-	EXPRESSION TAG	UNP Q9ALN7
A	0	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-15	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-14	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-13	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-12	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-11	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-10	HIS	-	EXPRESSION TAG	UNP Q9ALN7
B	-9	SER	-	EXPRESSION TAG	UNP Q9ALN7
B	-8	SER	-	EXPRESSION TAG	UNP Q9ALN7
B	-7	GLY	-	EXPRESSION TAG	UNP Q9ALN7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	LEU	-	EXPRESSION TAG	UNP Q9ALN7
B	-5	VAL	-	EXPRESSION TAG	UNP Q9ALN7
B	-4	PRO	-	EXPRESSION TAG	UNP Q9ALN7
B	-3	ARG	-	EXPRESSION TAG	UNP Q9ALN7
B	-2	GLY	-	EXPRESSION TAG	UNP Q9ALN7
B	-1	SER	-	EXPRESSION TAG	UNP Q9ALN7
B	0	HIS	-	EXPRESSION TAG	UNP Q9ALN7

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	36	Total O 36 36	0	0
2	B	48	Total O 48 48	0	0

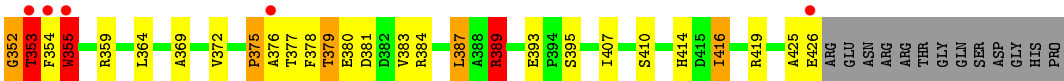
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 ($\text{RSRZ} > 2$). 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.

[illegible]

Chain B:

3% 51% 22% 6% 19%

HIS HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY S-1 A10 S11 S12 H13 F14 L17 V18 P19 A23 R32 C36 P37 N38 M39 V40 S41 M42 V43 E56 L57 E58 L59 ILE SER LEU ALA LYS ASN GLU LEU VAL LEU GLY SER GLY VAL SER ASP GLU LYS GLY ARG HIS PRO GLU LEU PHE ASP GLU LEU LEU SER ILE ASN SER GLY ARG ASP THR D99 A100 V101 E102 Q103 L104 H105 L106 D107 N108 D109 M110 M116 G117 K121 V122 W129 D130 R141 A142 H147 V148 R149 V152 D155 V156 S157 F164 Y167



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	162.42Å 162.42Å 81.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.93 – 2.50 46.89 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.93-2.50) 99.7 (46.89-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.09 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.200 , 0.257 0.200 , 0.258	Depositor DCC
R_{free} test set	1909 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	57.2	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 50.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 38152 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5856	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

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	0.76	0/2958	0.94	4/4030 (0.1%)
1	B	0.82	0/2937	1.04	12/4005 (0.3%)
All	All	0.79	0/5895	0.99	16/8035 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	149	ARG	NE-CZ-NH2	-9.24	115.68	120.30
1	A	25	ARG	NE-CZ-NH2	-8.10	116.25	120.30
1	B	149	ARG	NE-CZ-NH1	7.46	124.03	120.30
1	B	389	ARG	NE-CZ-NH1	6.77	123.69	120.30
1	B	280	LEU	CA-CB-CG	6.71	130.74	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	299	GLU	Peptide

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	2899	0	2909	101	0
1	B	2873	0	2870	149	0
2	A	36	0	0	5	0
2	B	48	0	0	3	1
All	All	5856	0	5779	247	1

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 21.

The worst 5 of 247 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:287:LEU:HD21	1:B:380:GLU:O	1.21	1.29
1:B:287:LEU:CD2	1:B:380:GLU:O	1.78	1.28
1:B:301:ARG:NE	1:B:301:ARG:HA	1.38	1.25
1:B:287:LEU:HD22	1:B:380:GLU:CA	1.66	1.24
1:B:287:LEU:CD2	1:B:383:VAL:HB	1.73	1.18

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:532:HOH:O	2:B:532:HOH:O[7_556]	1.21	0.99

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	363/455 (80%)	338 (93%)	17 (5%)	8 (2%)	8	13
1	B	363/455 (80%)	337 (93%)	14 (4%)	12 (3%)	5	6
All	All	726/910 (80%)	675 (93%)	31 (4%)	20 (3%)	6	9

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	259	LEU
1	A	274	GLN
1	B	100	ALA
1	B	253	GLY
1	B	254	LEU

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	317/386 (82%)	284 (90%)	33 (10%)	9	16
1	B	312/386 (81%)	270 (86%)	42 (14%)	5	9
All	All	629/772 (82%)	554 (88%)	75 (12%)	6	12

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	12	SER
1	B	110	ARG
1	B	379	THR
1	B	17	LEU
1	B	106	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	343	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	0	HIS
1	B	105	HIS
1	A	230	ASN
1	B	103	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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	373/455 (81%)	0.08	13 (3%) 48 53	34, 59, 96, 125	0
1	B	369/455 (81%)	0.04	12 (3%) 50 55	33, 53, 102, 135	0
All	All	742/910 (81%)	0.06	25 (3%) 49 54	33, 56, 101, 135	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	354	PHE	6.3
1	A	60	ILE	5.5
1	A	259	LEU	4.5
1	B	355	TRP	3.7
1	A	104	LEU	3.5

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.