



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:24 AM GMT

PDB ID : 2GYX
Title : Structure of aspartate semialdehyde dehydrogenase (ASADH) from *Streptococcus pneumoniae*
Authors : Faehnle, C.R.; Le Coq, J.; Liu, X.; Viola, R.E.
Deposited on : 2006-05-10
Resolution : 2.10 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

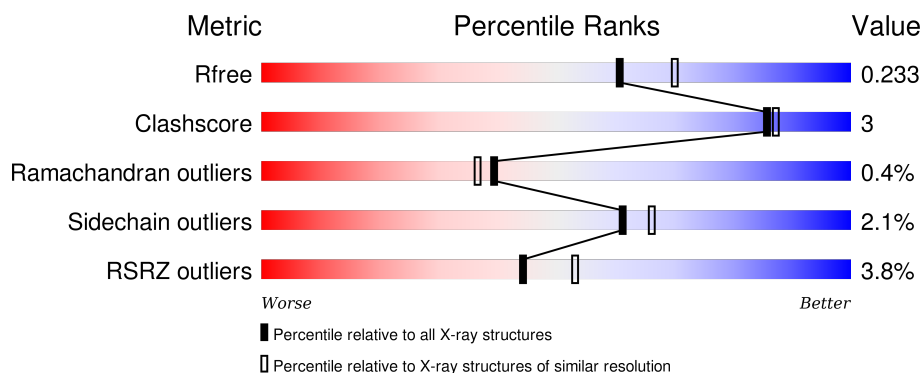
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	366	<div> <div>2%</div> <div>88%</div> <div>8%</div> <div>• •</div> </div>
1	B	366	<div> <div>3%</div> <div>84%</div> <div>9%</div> <div>7%</div> </div>
1	C	366	<div> <div>3%</div> <div>85%</div> <div>10%</div> <div>• •</div> </div>
1	D	366	<div> <div>6%</div> <div>82%</div> <div>11%</div> <div>• 7%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11565 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 Aspartate beta-semialdehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	352	Total	C	N	O	S	0	0	0
			2684	1699	450	522	13			
1	B	339	Total	C	N	O	S	0	0	0
			2594	1640	434	507	13			
1	C	350	Total	C	N	O	S	0	0	0
			2673	1693	448	519	13			
1	D	342	Total	C	N	O	S	0	0	0
			2611	1650	439	509	13			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	LEU	-	CLONING ARTIFACT	UNP Q8DQ00
A	360	GLU	-	CLONING ARTIFACT	UNP Q8DQ00
A	361	HIS	-	EXPRESSION TAG	UNP Q8DQ00
A	362	HIS	-	EXPRESSION TAG	UNP Q8DQ00
A	363	HIS	-	EXPRESSION TAG	UNP Q8DQ00
A	364	HIS	-	EXPRESSION TAG	UNP Q8DQ00
A	365	HIS	-	EXPRESSION TAG	UNP Q8DQ00
A	366	HIS	-	EXPRESSION TAG	UNP Q8DQ00
B	359	LEU	-	CLONING ARTIFACT	UNP Q8DQ00
B	360	GLU	-	CLONING ARTIFACT	UNP Q8DQ00
B	361	HIS	-	EXPRESSION TAG	UNP Q8DQ00
B	362	HIS	-	EXPRESSION TAG	UNP Q8DQ00
B	363	HIS	-	EXPRESSION TAG	UNP Q8DQ00
B	364	HIS	-	EXPRESSION TAG	UNP Q8DQ00
B	365	HIS	-	EXPRESSION TAG	UNP Q8DQ00
B	366	HIS	-	EXPRESSION TAG	UNP Q8DQ00
C	359	LEU	-	CLONING ARTIFACT	UNP Q8DQ00
C	360	GLU	-	CLONING ARTIFACT	UNP Q8DQ00
C	361	HIS	-	EXPRESSION TAG	UNP Q8DQ00
C	362	HIS	-	EXPRESSION TAG	UNP Q8DQ00
C	363	HIS	-	EXPRESSION TAG	UNP Q8DQ00

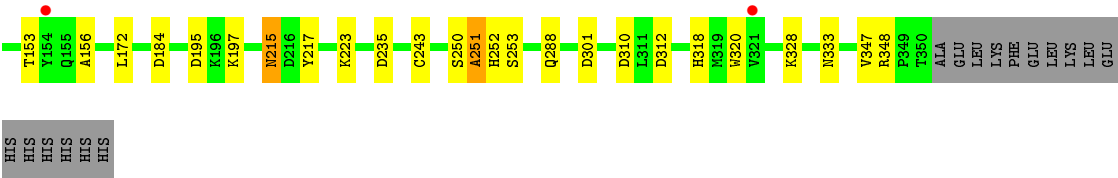
Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	364	HIS	-	EXPRESSION TAG	UNP Q8DQ00
C	365	HIS	-	EXPRESSION TAG	UNP Q8DQ00
C	366	HIS	-	EXPRESSION TAG	UNP Q8DQ00
D	359	LEU	-	CLONING ARTIFACT	UNP Q8DQ00
D	360	GLU	-	CLONING ARTIFACT	UNP Q8DQ00
D	361	HIS	-	EXPRESSION TAG	UNP Q8DQ00
D	362	HIS	-	EXPRESSION TAG	UNP Q8DQ00
D	363	HIS	-	EXPRESSION TAG	UNP Q8DQ00
D	364	HIS	-	EXPRESSION TAG	UNP Q8DQ00
D	365	HIS	-	EXPRESSION TAG	UNP Q8DQ00
D	366	HIS	-	EXPRESSION TAG	UNP Q8DQ00

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	296	Total	O	0	0
			296	296		
2	B	232	Total	O	0	0
			232	232		
2	C	242	Total	O	0	0
			242	242		
2	D	233	Total	O	0	0
			233	233		



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2 ₁	Depositor
Cell constants a, b, c, α , β , γ	72.48 Å 78.69 Å 243.41 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.11 – 2.10 29.07 – 2.10	Depositor EDS
% Data completeness (in resolution range)	94.6 (29.11-2.10) 94.6 (29.07-2.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.75 (at 2.10 Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.178 , 0.221 0.190 , 0.233	Depositor DCC
R_{free} test set	3948 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	27.8	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 57.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 77719 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11565	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.56% 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	0.38	0/2732	0.73	11/3711 (0.3%)
1	B	0.35	0/2640	0.70	7/3589 (0.2%)
1	C	0.37	0/2720	0.73	14/3694 (0.4%)
1	D	0.42	1/2658 (0.0%)	0.73	8/3613 (0.2%)
All	All	0.38	1/10750 (0.0%)	0.72	40/14607 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	87	ALA	C-O	5.27	1.33	1.23

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	195	ASP	CB-CG-OD2	6.57	124.21	118.30
1	A	195	ASP	CB-CG-OD2	6.27	123.94	118.30
1	A	216	ASP	CB-CG-OD2	6.26	123.94	118.30
1	C	30	ASP	CB-CG-OD2	6.24	123.91	118.30
1	D	301	ASP	CB-CG-OD2	6.11	123.80	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	2684	0	2685	12	0
1	B	2594	0	2580	13	0
1	C	2673	0	2673	19	0
1	D	2611	0	2605	19	0
2	A	296	0	0	0	0
2	B	232	0	0	1	0
2	C	242	0	0	4	0
2	D	233	0	0	1	0
All	All	11565	0	10543	58	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 58 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:C:288:GLN:HE22	1:D:197:LYS:H	1.36	0.73
1:A:197:LYS:H	1:B:288:GLN:HE22	1.38	0.71
1:A:288:GLN:HE22	1:B:197:LYS:H	1.40	0.70
1:B:156:ALA:H	1:B:252:HIS:HD2	1.40	0.69
1:C:197:LYS:H	1:D:288:GLN:HE22	1.40	0.69

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	348/366 (95%)	335 (96%)	11 (3%)	2 (1%)	30 24
1	B	333/366 (91%)	321 (96%)	11 (3%)	1 (0%)	46 45
1	C	344/366 (94%)	330 (96%)	13 (4%)	1 (0%)	46 45
1	D	338/366 (92%)	325 (96%)	12 (4%)	1 (0%)	46 45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1363/1464 (93%)	1311 (96%)	47 (3%)	5 (0%)	39 37

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	251	ALA
1	A	251	ALA
1	C	251	ALA
1	D	251	ALA
1	A	330	ALA

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	289/302 (96%)	286 (99%)	3 (1%)	82 87
1	B	280/302 (93%)	274 (98%)	6 (2%)	61 66
1	C	288/302 (95%)	282 (98%)	6 (2%)	61 66
1	D	282/302 (93%)	273 (97%)	9 (3%)	46 48
All	All	1139/1208 (94%)	1115 (98%)	24 (2%)	61 66

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	50	GLN
1	C	290	TYR
1	D	215	ASN
1	C	147	ASP
1	C	213	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	119	HIS
1	C	288	GLN
1	D	318	HIS
1	C	186	HIS
1	C	318	HIS

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	352/366 (96%)	-0.06	8 (2%) 64 70	18, 26, 38, 45	0
1	B	339/366 (92%)	0.13	12 (3%) 48 57	20, 32, 55, 86	0
1	C	350/366 (95%)	-0.01	11 (3%) 52 61	17, 29, 52, 78	0
1	D	342/366 (93%)	0.25	22 (6%) 23 30	17, 32, 73, 102	0
All	All	1383/1464 (94%)	0.08	53 (3%) 44 53	17, 29, 59, 102	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	60	THR	3.9
1	D	43	LYS	3.8
1	A	351	ALA	3.8
1	D	45	LEU	3.5
1	D	57	THR	3.4

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.