



Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 08:46 pm BST

PDB ID : 1SVV
Title : Initial Structural Analysis of Leishmania major Threonine Aldolase
Authors : Hol, W.G.J.; Robien, M.A.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)
Deposited on : 2004-03-30
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.11
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) : 2.11

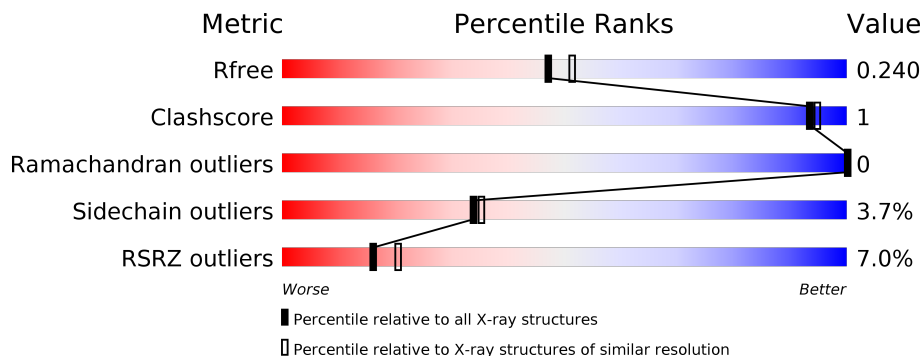
1 Overall quality at a glance

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}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 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 $\leq 5\%$. 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	359	 5% 89% 6% 5%
1	B	359	 8% 88% 7% 5%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5308 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 THREONINE ALDOLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	340	2593	1643	447	483	8	12	0	0	0
1	B	342	2593	1645	445	483	8	12	0	0	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	26	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	34	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	39	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	143	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	209	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	220	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	247	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	261	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	276	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	308	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	318	MSE	MET	MODIFIED RESIDUE	UNP O15839
A	331	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	1	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	26	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	34	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	39	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	143	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	209	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	220	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	247	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	261	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	276	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	308	MSE	MET	MODIFIED RESIDUE	UNP O15839
B	318	MSE	MET	MODIFIED RESIDUE	UNP O15839

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	331	MSE	MET	MODIFIED RESIDUE	UNP O15839

- Molecule 2 is UNKNOWN LIGAND (three-letter code: UNL) (formula:).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mn 1 1	0	0
2	A	1	Total Mn 1 1	0	0

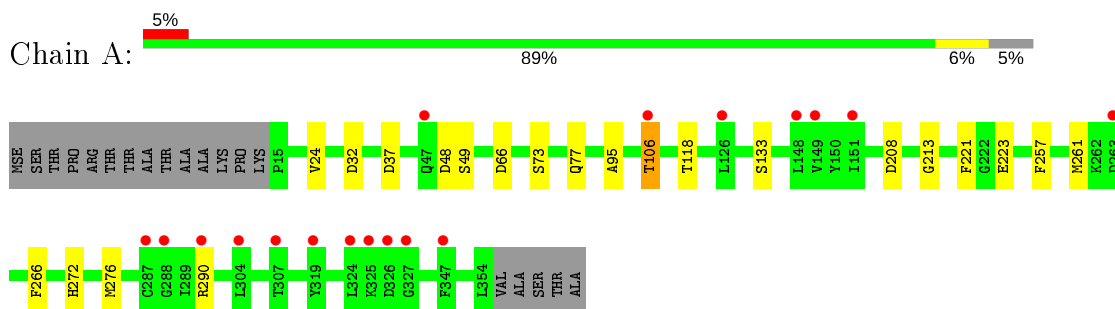
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	78	Total O 78 78	0	0
3	B	42	Total O 42 42	0	0

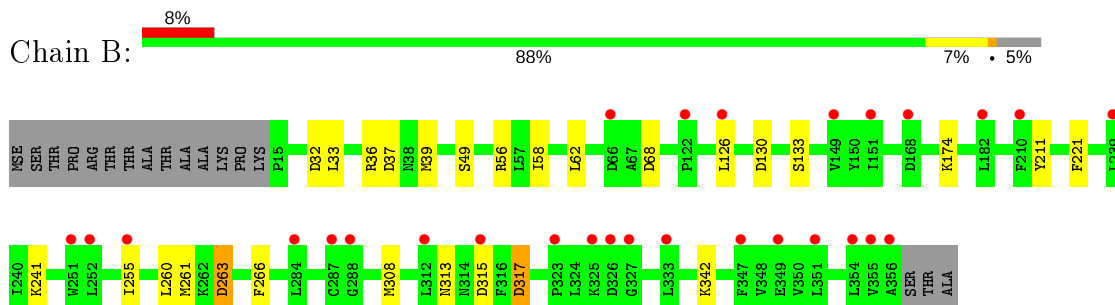
3 Residue-property plots [i](#)

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 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: THREONINE ALDOLASE



- Molecule 1: THREONINE ALDOLASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	119.09Å 119.09Å 129.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.17 – 2.10 19.04 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.17-2.10) 99.9 (19.04-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.41 (at 2.09Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.198 , 0.230 0.211 , 0.240	Depositor DCC
R_{free} test set	5285 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å ²)	48.0	Xtrriage
Anisotropy	0.024	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 48.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5308	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: UNL

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	0.60	0/2632	0.78	5/3547 (0.1%)
1	B	0.49	0/2632	0.73	7/3550 (0.2%)
All	All	0.54	0/5264	0.75	12/7097 (0.2%)

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	37	ASP	CB-CG-OD2	6.72	124.35	118.30
1	B	32	ASP	CB-CG-OD2	6.08	123.77	118.30
1	A	208	ASP	CB-CG-OD2	5.78	123.50	118.30
1	A	48	ASP	CB-CG-OD2	5.71	123.44	118.30
1	A	32	ASP	CB-CG-OD2	5.46	123.21	118.30
1	B	315	ASP	CB-CG-OD2	5.42	123.17	118.30
1	B	68	ASP	CB-CG-OD2	5.40	123.16	118.30
1	B	317	ASP	CB-CG-OD2	5.36	123.13	118.30
1	B	263	ASP	CB-CG-OD2	5.34	123.11	118.30
1	A	66	ASP	CB-CG-OD2	5.08	122.87	118.30
1	B	130	ASP	CB-CG-OD2	5.07	122.86	118.30
1	A	37	ASP	CB-CG-OD2	5.01	122.81	118.30

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	2593	0	2576	9	0
1	B	2593	0	2568	8	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	78	0	0	0	0
3	B	42	0	0	0	0
All	All	5308	0	5144	14	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 (14) 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:A:213:GLY:HA2	1:A:223:GLU:OE1	1.97	0.65
1:A:257:PHE:CD1	1:A:261:MSE:HE3	2.32	0.65
1:B:261:MSE:HE1	1:B:266:PHE:CD1	2.37	0.60
1:A:261:MSE:HE1	1:A:266:PHE:CD1	2.43	0.54
1:A:106:THR:CG2	1:B:241:LYS:CG	2.86	0.53
1:B:260:LEU:CB	1:B:261:MSE:HE2	2.41	0.51
1:A:106:THR:CG2	1:B:241:LYS:HG2	2.42	0.50
1:A:272:HIS:O	1:A:276:MSE:HG2	2.12	0.48
1:B:260:LEU:HB3	1:B:261:MSE:HE2	1.96	0.47
1:B:33:LEU:HD23	1:B:255:ILE:HG23	1.97	0.47
1:A:257:PHE:HD1	1:A:261:MSE:HE3	1.76	0.46
1:A:24:VAL:HG13	1:B:39:MSE:HA	1.98	0.45
1:B:58:ILE:O	1:B:62:LEU:HG	2.18	0.43
1:A:95:ALA:O	1:A:118:THR:HA	2.20	0.41

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	338/359 (94%)	327 (97%)	11 (3%)	0	100	100
1	B	340/359 (95%)	328 (96%)	12 (4%)	0	100	100
All	All	678/718 (94%)	655 (97%)	23 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	272/280 (97%)	265 (97%)	7 (3%)	46	50
1	B	270/280 (96%)	257 (95%)	13 (5%)	25	24
All	All	542/560 (97%)	522 (96%)	20 (4%)	34	35

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	SER
1	A	73	SER
1	A	77	GLN
1	A	106	THR
1	A	133	SER
1	A	221	PHE
1	A	290	ARG
1	B	36	ARG
1	B	49	SER
1	B	56	ARG
1	B	126	LEU
1	B	133	SER
1	B	174	LYS
1	B	211	TYR
1	B	221	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	263	ASP
1	B	308	MSE
1	B	313	ASN
1	B	317	ASP
1	B	342	LYS

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

Mol	Chain	Res	Type
1	A	114	HIS
1	A	264	ASN
1	A	306	ASN
1	A	313	ASN
1	A	314	ASN

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](#)

Of 2 ligands modelled in this entry, 2 are unknown - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

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	328/359 (91%)	0.26	18 (5%) 25 31	29, 37, 52, 67	0
1	B	330/359 (91%)	0.38	28 (8%) 10 13	31, 43, 59, 83	0
All	All	658/718 (91%)	0.32	46 (6%) 16 20	29, 41, 57, 83	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	288	GLY	4.7
1	B	347	PHE	4.3
1	B	327	GLY	3.9
1	B	326	ASP	3.8
1	A	149	VAL	3.5
1	B	151	ILE	3.5
1	A	288	GLY	3.4
1	A	307	THR	3.3
1	B	349	GLU	3.2
1	A	151	ILE	3.2
1	A	327	GLY	3.2
1	B	325	LYS	3.2
1	A	324	LEU	3.1
1	B	255	ILE	3.1
1	A	148	LEU	3.1
1	B	354	LEU	3.1
1	A	326	ASP	3.1
1	B	287	CYS	2.9
1	B	182	LEU	2.8
1	B	284	LEU	2.8
1	B	239	LEU	2.8
1	A	325	LYS	2.8
1	B	252	LEU	2.7
1	A	319	TYR	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	347	PHE	2.6
1	B	149	VAL	2.5
1	B	351	LEU	2.5
1	A	47	GLN	2.4
1	B	251	TRP	2.4
1	A	287	CYS	2.3
1	A	304	LEU	2.3
1	B	122	PRO	2.3
1	B	356	ALA	2.3
1	B	355	VAL	2.3
1	A	263	ASP	2.3
1	B	66	ASP	2.2
1	B	333	LEU	2.2
1	B	210	PHE	2.1
1	B	323	PRO	2.1
1	B	168	ASP	2.1
1	A	290	ARG	2.1
1	B	126	LEU	2.1
1	A	126	LEU	2.1
1	B	315	ASP	2.1
1	B	312	LEU	2.0
1	A	106	THR	2.0

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](#)

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	UNL	B	401	1/-	0.94	0.21	61,61,61,61	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	UNL	A	401	1/-	0.96	0.22	51,51,51,51	0

6.5 Other polymers [i](#)

There are no such residues in this entry.