



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 01:48 AM GMT

PDB ID : 2B4W  
Title : Hypothetical protein from leishmania major  
Authors : Holmes, M.A.; Merritt, E.A.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)  
Deposited on : 2005-09-26  
Resolution : 1.98 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

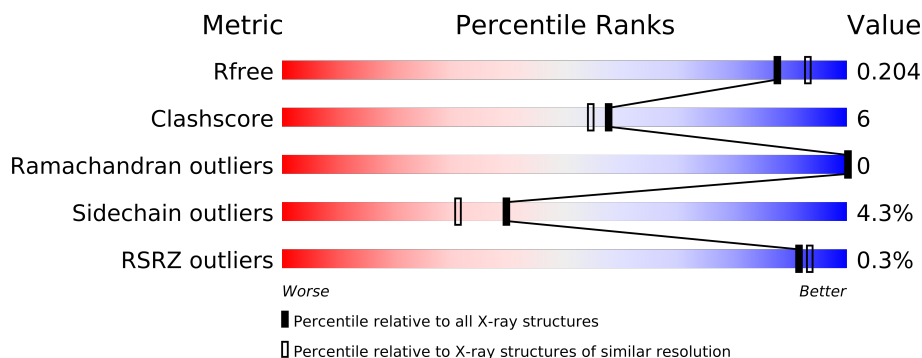
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance


The reported resolution of this entry is 1.98 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))
$R_{free}$	66092	6577 (2.00-1.96)
Clashscore	79885	8091 (2.00-1.96)
Ramachandran outliers	78287	7989 (2.00-1.96)
Sidechain outliers	78261	7987 (2.00-1.96)
RSRZ outliers	66119	6578 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	315	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2435 atoms, of which 0 are hydrogen and 0 are deuterium.

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 hypothetical protein, conserved.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	Se	0	0	0
			2250	1449	374	417	6	4			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q4QH86
A	93	MSE	MET	MODIFIED RESIDUE	UNP Q4QH86
A	132	MSE	MET	MODIFIED RESIDUE	UNP Q4QH86
A	271	MSE	MET	MODIFIED RESIDUE	UNP Q4QH86
A	311	MSE	MET	MODIFIED RESIDUE	UNP Q4QH86

- Molecule 2 is water.

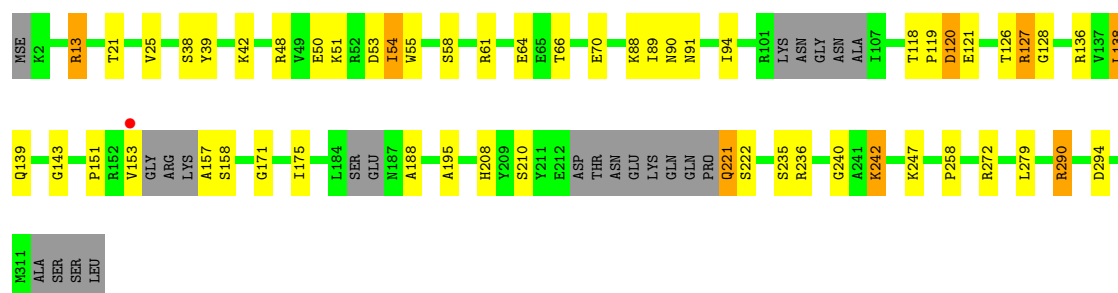
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	185	Total	O	0	0
			185	185		

### 3 Residue-property plots

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 ( $RSRZ > 2$ ). 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: hypothetical protein, conserved

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.40Å 97.40Å 113.60Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 1.98 30.70 – 1.98	Depositor EDS
% Data completeness (in resolution range)	94.5 (50.00-1.98) 99.4 (30.70-1.98)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.96 (at 1.98Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.132 , 0.211 0.154 , 0.204	Depositor DCC
$R_{free}$ test set	1420 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.7	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.8	EDS
Estimated twinning fraction	0.107 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.095 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.096 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l 0.095 for -h,2/3*h+1/3*k-2/3*l,-2/3*h-4/3*k-1/3*l 0.095 for 1/3*h+2/3*k+2/3*l,-k,4/3*h+2/3*k-1/3*l 0.106 for -1/3*h-2/3*k-2/3*l,-2/3*h-1/3*k+2/3*l,-2/3*h+2/3*k-1/3*l 0.359 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	0 of 28075 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2435	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.44	0/2301	1.26	16/3110 (0.5%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	48	ARG	NE-CZ-NH1	10.21	125.40	120.30
1	A	61	ARG	NE-CZ-NH1	8.82	124.71	120.30
1	A	48	ARG	CD-NE-CZ	8.22	135.10	123.60
1	A	61	ARG	CD-NE-CZ	7.88	134.63	123.60
1	A	127	ARG	NE-CZ-NH1	6.98	123.79	120.30
1	A	61	ARG	NE-CZ-NH2	-6.70	116.95	120.30
1	A	120	ASP	CB-CG-OD1	6.40	124.06	118.30
1	A	290	ARG	NE-CZ-NH1	6.33	123.46	120.30
1	A	48	ARG	NE-CZ-NH2	-5.96	117.32	120.30
1	A	138	LEU	CB-CG-CD1	5.71	120.72	111.00
1	A	13	ARG	CD-NE-CZ	5.67	131.53	123.60
1	A	272	ARG	NE-CZ-NH2	-5.66	117.47	120.30
1	A	138	LEU	CA-CB-CG	5.58	128.12	115.30
1	A	136	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	A	208	HIS	CA-CB-CG	5.54	123.02	113.60
1	A	272	ARG	NE-CZ-NH1	5.14	122.87	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2250	0	2172	28	0
2	A	185	0	0	5	0
All	All	2435	0	2172	28	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (28) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:188:ALA:HB1	1:A:210:SER:O	1.87	0.75
1:A:240:GLY:O	1:A:242:LYS:HE2	1.88	0.72
1:A:21:THR:HG23	2:A:415:HOH:O	1.88	0.72
1:A:51:LYS:HE3	1:A:53:ASP:OD1	2.03	0.59
1:A:139:GLN:HG2	2:A:501:HOH:O	2.01	0.59
1:A:88:LYS:HE2	2:A:446:HOH:O	2.08	0.53
1:A:151:PRO:O	1:A:157:ALA:HA	2.08	0.52
1:A:54:ILE:HG13	1:A:55:TRP:N	2.24	0.52
1:A:127:ARG:HD2	1:A:128:GLY:O	2.09	0.51
1:A:171:GLY:O	1:A:175:ILE:HG13	2.12	0.49
1:A:66:THR:OG1	1:A:70:GLU:OE2	2.30	0.49
1:A:221:GLN:HG3	2:A:556:HOH:O	2.12	0.48
1:A:13:ARG:HD2	1:A:294:ASP:OD2	2.14	0.48
1:A:89:ILE:HD13	1:A:94:ILE:CG1	2.44	0.48
1:A:25:VAL:O	1:A:25:VAL:HG23	2.14	0.48
1:A:138:LEU:HD23	1:A:195:ALA:HB1	1.97	0.46
1:A:42:LYS:HD2	1:A:64:GLU:OE2	2.16	0.46
1:A:118:THR:HG22	1:A:121:GLU:HB2	1.99	0.44
1:A:143:GLY:HA2	2:A:555:HOH:O	2.17	0.44
1:A:279:LEU:O	1:A:290:ARG:HA	2.17	0.44
1:A:138:LEU:HD23	1:A:195:ALA:CB	2.47	0.44
1:A:240:GLY:O	1:A:242:LYS:HG2	2.19	0.42
1:A:126:THR:HG23	1:A:127:ARG:N	2.35	0.42
1:A:39:TYR:CE1	1:A:119:PRO:HD2	2.56	0.41
1:A:50:GLU:OE2	1:A:58:SER:OG	2.38	0.41
1:A:235:SER:O	1:A:236:ARG:HB2	2.21	0.40
1:A:89:ILE:HD13	1:A:94:ILE:HG12	2.03	0.40
1:A:90:ASN:O	1:A:91:ASN:HB2	2.22	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	282/315 (90%)	273 (97%)	9 (3%)	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	234/258 (91%)	224 (96%)	10 (4%)	40	31

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	SER
1	A	54	ILE
1	A	120	ASP
1	A	153	VAL
1	A	158	SER
1	A	221	GLN
1	A	222	SER
1	A	242	LYS
1	A	247	LYS
1	A	258	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.



### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	292/315 (92%)	-0.21	1 (0%) 91 93	14, 27, 45, 66	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	153	VAL	2.8

### 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 carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

### 6.5 Other polymers ⓘ

There are no such residues in this entry.