



wwPDB X-ray Structure Validation Summary Report

Mar 1, 2014 – 04:23 AM GMT

PDB ID : 3HGF
Title : Expression, purification, spectroscopical and crystallographical studies of segments of the nucleotide binding domain of the reticulocyte binding protein Py235 of Plasmodium yoelii
Authors : Gruber, A.; Manimekalai, M.S.S.; Balakrishna, A.M.; Hunke, C.; Jeyakanthan, J.; Preiser, P.R.; Gruber, G.
Deposited on : 2009-05-13
Resolution : 4.00 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

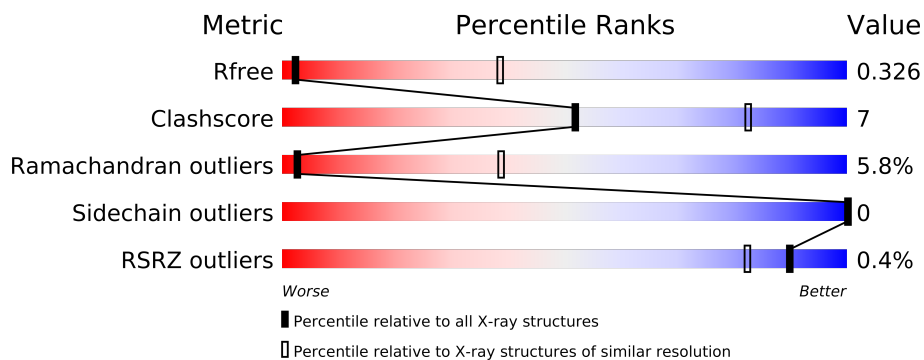
The following versions of software and data (see [references](#)) were used in the production of this report:

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 4.00 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	1035 (4.52-3.46)
Clashscore	79885	1235 (4.50-3.50)
Ramachandran outliers	78287	1170 (4.50-3.50)
Sidechain outliers	78261	1156 (4.50-3.50)
RSRZ outliers	66119	1035 (4.52-3.46)

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. 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	107	
1	B	107	
1	C	107	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1375 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 Rhoptyr protein fragment.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	0	0	0
			493	297	98	98			
1	B	89	Total	C	N	O	0	0	0
			446	268	89	89			
1	C	87	Total	C	N	O	0	0	0
			436	262	87	87			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	HIS	-	EXPRESSION TAG	UNP Q7RPU0
A	2	HIS	-	EXPRESSION TAG	UNP Q7RPU0
A	3	HIS	-	EXPRESSION TAG	UNP Q7RPU0
A	4	HIS	-	EXPRESSION TAG	UNP Q7RPU0
A	5	HIS	-	EXPRESSION TAG	UNP Q7RPU0
A	6	HIS	-	EXPRESSION TAG	UNP Q7RPU0
A	7	PRO	-	EXPRESSION TAG	UNP Q7RPU0
A	8	MET	-	EXPRESSION TAG	UNP Q7RPU0
A	9	VAL	-	EXPRESSION TAG	UNP Q7RPU0
B	1	HIS	-	EXPRESSION TAG	UNP Q7RPU0
B	2	HIS	-	EXPRESSION TAG	UNP Q7RPU0
B	3	HIS	-	EXPRESSION TAG	UNP Q7RPU0
B	4	HIS	-	EXPRESSION TAG	UNP Q7RPU0
B	5	HIS	-	EXPRESSION TAG	UNP Q7RPU0
B	6	HIS	-	EXPRESSION TAG	UNP Q7RPU0
B	7	PRO	-	EXPRESSION TAG	UNP Q7RPU0
B	8	MET	-	EXPRESSION TAG	UNP Q7RPU0
B	9	VAL	-	EXPRESSION TAG	UNP Q7RPU0
C	1	HIS	-	EXPRESSION TAG	UNP Q7RPU0
C	2	HIS	-	EXPRESSION TAG	UNP Q7RPU0
C	3	HIS	-	EXPRESSION TAG	UNP Q7RPU0
C	4	HIS	-	EXPRESSION TAG	UNP Q7RPU0
C	5	HIS	-	EXPRESSION TAG	UNP Q7RPU0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	6	HIS	-	EXPRESSION TAG	UNP Q7RPU0
C	7	PRO	-	EXPRESSION TAG	UNP Q7RPU0
C	8	MET	-	EXPRESSION TAG	UNP Q7RPU0
C	9	VAL	-	EXPRESSION TAG	UNP Q7RPU0

i

- Molecule 1: Rhoptry protein fragment

HIS	HIS	HIS	HIS	HIS	HIS	HIS	P7	M8	V9	K10	E11	I12	E13	V20	K27	Y28	I29	L36	L37	N38	I43	D46	K47	G61	K62	S63	I72	ASP	GLU	GLU	K76	K84	D92	E98	P101	E104	M105	N106	T107
-----	-----	-----	-----	-----	-----	-----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------

Chain B:

HIS	HIS	HIS	HIS	HIS	HIS	PRO	MET	VAL	LYS	GLU	ILE	ET3	A41	IY2	ASP	GLU	GLU	K76	ST9	ES0	N81	L93	D94	E95	I96	K97	E98	ETC4	MET	ASN	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	-----	-----	-----

Chain C:

HIS	HIS	HIS	HIS	HIS	PRO	MET	VAL	LYS	GLU	ILE	GLU	K14	Y30	ASP	ASN	MET	K34	L37	K66	I72	ASP	GLU	GLU	K76	K97	E98	Q99	S100	P101	K102	A103	E104	M105	N106	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α , β , γ	70.22Å 70.22Å 193.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.00 – 4.00 33.00 – 4.00	Depositor EDS
% Data completeness (in resolution range)	95.4 (33.00-4.00) 95.4 (33.00-4.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.09 (at 3.99Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.339 , 0.373 0.378 , 0.326	Depositor DCC
R_{free} test set	117 reflections (4.48%)	DCC
Wilson B-factor (Å ²)	133.2	Xtriage
Anisotropy	0.466	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	-0.12 , 98.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 2609 reflections (0.038%)	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	1375	wwPDB-VP
Average B, all atoms (Å ²)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.82% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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	0.27	0/493	0.45	0/687
1	B	0.27	0/445	0.45	0/620
1	C	0.27	0/434	0.45	0/603
All	All	0.27	0/1372	0.45	0/1910

There are no bond length outliers.

There are no bond angle outliers.

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	493	0	212	8	0
1	B	446	0	188	3	0
1	C	436	0	183	3	0
All	All	1375	0	583	14	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 7.

The worst 5 of 14 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:99:GLN:HA	1:C:103:ALA:HB2	1.83	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:9:VAL:C	1:A:11:GLU:H	2.12	0.53
1:A:36:LEU:C	1:A:38:ASN:H	2.14	0.51
1:B:96:ILE:C	1:B:98:GLU:H	2.15	0.49
1:B:79:SER:C	1:B:81:ASN:H	2.18	0.47

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	94/107 (88%)	70 (74%)	14 (15%)	10 (11%)	1	17
1	B	85/107 (79%)	67 (79%)	17 (20%)	1 (1%)	19	77
1	C	81/107 (76%)	63 (78%)	14 (17%)	4 (5%)	3	44
All	All	260/321 (81%)	200 (77%)	45 (17%)	15 (6%)	3	39

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	20	VAL
1	A	43	ILE
1	A	62	LYS
1	A	84	LYS
1	C	101	PRO

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	2/104 (2%)	2 (100%)	0	100	100
1	B	1/104 (1%)	1 (100%)	0	100	100
1	C	1/104 (1%)	1 (100%)	0	100	100
All	All	4/312 (1%)	4 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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, 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	98/107 (91%)	-0.43	0 100 100	18, 42, 56, 57	0
1	B	89/107 (83%)	-0.29	0 100 100	55, 63, 73, 73	0
1	C	87/107 (81%)	-0.08	1 (1%) 77 63	72, 101, 105, 105	0
All	All	274/321 (85%)	-0.27	1 (0%) 90 82	18, 62, 105, 105	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	37	LEU	2.4

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.