WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 26, 2014 – 04:46 PM GMT

PDB ID : 2HQ4

Title : Crystal Structure of ORF 1580 a hypothetical protein from Pyrococcus horikoshii

Authors : Li, Y.; Marshall, M.; Chang, J.; Zhao, M.; Zhang, M.; Xu, H.; Liu, Z.J.; Rose, J.P.; Wang, B.C.; Southeast Collaboratory for Structural Genomics; Southeast Collaboratory for Structural Genomics (SECSG)

Deposited on : 2006-07-18

Resolution : 1.99 Å(reported)

---

1 Overall quality at a glance ⓘ
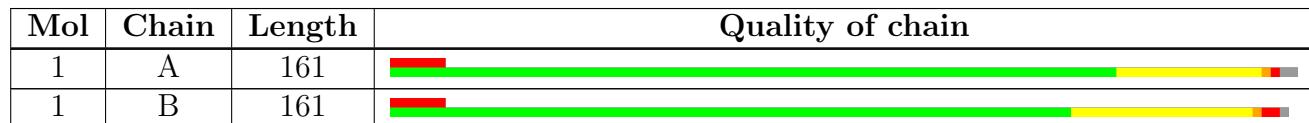
The reported resolution of this entry is 1.99 Å.

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	Percentile Ranks	Value
Clashscore		10
Ramachandran outliers		0.3%
Sidechain outliers		3.5%
RSRZ outliers		5.7%

Worse *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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	161	
1	B	161	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 2903 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 PH1570.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	158	Total	C	N	O	S	0	0	0
			1339	888	207	240	4			
1	B	159	Total	C	N	O	S	0	0	0
			1350	897	208	241	4			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	ASP	-	CLONING ARTIFACT	UNP O59278
A	-7	LEU	-	CLONING ARTIFACT	UNP O59278
A	-6	TYR	-	CLONING ARTIFACT	UNP O59278
A	-5	PHE	-	CLONING ARTIFACT	UNP O59278
A	-4	GLN	-	CLONING ARTIFACT	UNP O59278
A	-3	GLY	-	CLONING ARTIFACT	UNP O59278
A	-2	GLY	-	CLONING ARTIFACT	UNP O59278
A	-1	SER	-	CLONING ARTIFACT	UNP O59278
A	0	GLY	-	CLONING ARTIFACT	UNP O59278
B	-8	ASP	-	CLONING ARTIFACT	UNP O59278
B	-7	LEU	-	CLONING ARTIFACT	UNP O59278
B	-6	TYR	-	CLONING ARTIFACT	UNP O59278
B	-5	PHE	-	CLONING ARTIFACT	UNP O59278
B	-4	GLN	-	CLONING ARTIFACT	UNP O59278
B	-3	GLY	-	CLONING ARTIFACT	UNP O59278
B	-2	GLY	-	CLONING ARTIFACT	UNP O59278
B	-1	SER	-	CLONING ARTIFACT	UNP O59278
B	0	GLY	-	CLONING ARTIFACT	UNP O59278

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	103	Total	O	0	0
			103	103		

Continued on next page...
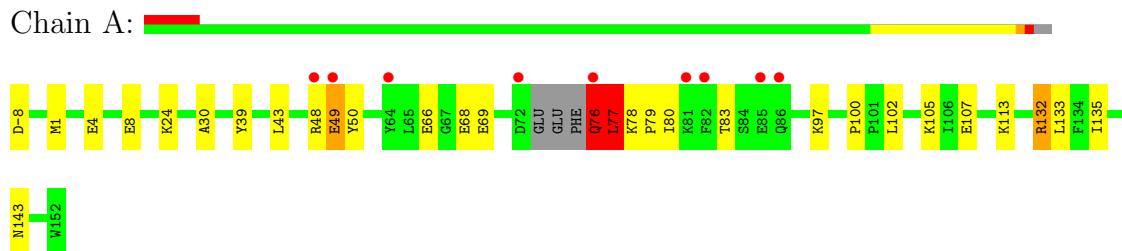
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	111	Total 111	O 111	0	0
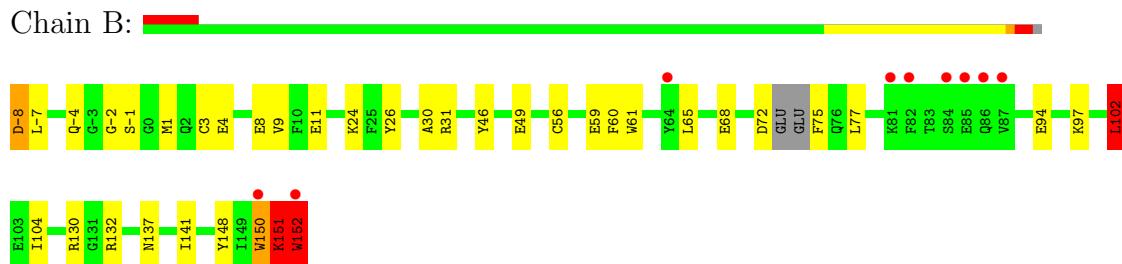
3 Residue-property plots ⓘ

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 PH1570

Chain A:

• Molecule 1: Hypothetical protein PH1570

Chain B:

4　Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	133.87Å　46.39Å　94.75Å 90.00°　132.42°　90.00°	Depositor
Resolution (Å)	20.00　–　1.99 47.33　–　1.99	Depositor EDS
% Data completeness (in resolution range)	98.9 (20.00-1.99) 97.4 (47.33-1.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.86 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.195　,　　0.232 0.196　,　(Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	26.0	Xtriage
Anisotropy	0.307	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 44.6	EDS
Estimated twinning fraction	0.032 for h+2*l,k,-h-l 0.016 for h,-k,-h-l 0.018 for -h-2*l,-k,l	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	0 of 56761 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2903	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.64% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]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	0.89	0/1373	0.81	4/1843 (0.2%)
1	B	0.91	2/1385 (0.1%)	0.84	4/1859 (0.2%)
All	All	0.90	2/2758 (0.1%)	0.83	8/3702 (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	A	1	2
1	B	0	3
All	All	1	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	59	GLU	C-N	-8.29	1.15	1.34
1	B	60	PHE	C-N	-7.90	1.15	1.34

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	152	TRP	N-CA-C	-7.68	90.27	111.00
1	B	1	MET	CG-SD-CE	-7.39	88.37	100.20
1	A	49	GLU	N-CA-C	7.09	130.14	111.00
1	A	50	TYR	N-CA-C	-6.02	94.75	111.00
1	A	132	ARG	NE-CZ-NH1	-5.63	117.48	120.30
1	A	77	LEU	N-CA-C	5.37	125.50	111.00
1	B	102	LEU	CA-CB-CG	5.28	127.45	115.30
1	B	-7	LEU	CA-CB-CG	5.10	127.03	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	49	GLU	CA

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	76	GLN	Peptide
1	A	77	LEU	Peptide
1	B	-8	ASP	Peptide
1	B	150	TRP	Peptide
1	B	151	LYS	Peptide

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	1339	0	1345	24	0
1	B	1350	0	1354	33	0
2	A	103	0	0	8	0
2	B	111	0	0	10	0
All	All	2903	0	2699	54	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 10.

All (54) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1:MET:HG2	2:A:249:HOH:O	1.27	1.27
1:A:48:ARG:NH1	2:A:174:HOH:O	1.84	1.09
1:A:66:GLU:HB3	1:A:68:GLU:OE1	1.52	1.08
1:A:1:MET:CG	2:A:249:HOH:O	1.94	1.01
1:A:1:MET:SD	2:A:249:HOH:O	2.29	0.91
1:B:65:LEU:HD11	1:B:152:TRP:CH2	2.10	0.86
1:A:-8:ASP:N	2:A:253:HOH:O	2.10	0.82
1:B:3:CYS:HG	1:B:56:CYS:HG	0.82	0.80
1:B:152:TRP:CE3	1:B:152:TRP:HA	2.18	0.78
1:B:-2:GLY:HA2	2:B:171:HOH:O	1.84	0.76

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:151:LYS:HB3	1:B:152:TRP:CE3	2.23	0.74
1:A:80:ILE:HD11	1:A:113:LYS:HB2	1.74	0.68
1:A:83:THR:HG22	2:A:254:HOH:O	1.93	0.68
1:B:152:TRP:HE3	1:B:152:TRP:HA	1.61	0.65
1:A:76:GLN:HG2	1:B:49:GLU:HB3	1.83	0.59
1:A:105:LYS:HD3	1:A:107:GLU:OE2	2.02	0.59
1:B:151:LYS:N	2:B:262:HOH:O	2.22	0.58
1:A:76:GLN:HA	1:A:77:LEU:HD12	1.87	0.56
1:B:31:ARG:NH1	2:B:257:HOH:O	2.24	0.56
1:B:-8:ASP:N	2:B:244:HOH:O	2.40	0.55
1:B:65:LEU:HD11	1:B:152:TRP:CZ2	2.42	0.55
1:A:68:GLU:H	1:A:68:GLU:CD	2.11	0.54
1:B:9:VAL:HG13	1:B:61:TRP:CZ2	2.43	0.54
1:B:75:PHE:HE1	1:B:150:TRP:HH2	1.55	0.54
1:B:151:LYS:HB3	1:B:152:TRP:CD2	2.43	0.53
1:B:11:GLU:HB2	2:B:187:HOH:O	2.09	0.53
1:B:150:TRP:NE1	2:B:259:HOH:O	2.03	0.53
1:B:137:ASN:OD1	1:B:141:ILE:CD1	2.59	0.51
1:B:30:ALA:HB2	1:B:132:ARG:HD2	1.92	0.51
1:A:97:LYS:HE3	2:A:192:HOH:O	2.10	0.51
1:B:68:GLU:HG3	2:B:260:HOH:O	2.13	0.49
1:A:8:GLU:OE2	2:A:244:HOH:O	2.19	0.49
1:B:150:TRP:HB3	2:B:262:HOH:O	2.13	0.49
1:A:43:LEU:HD11	1:A:48:ARG:HB3	1.94	0.48
1:A:39:TYR:OH	1:A:48:ARG:HB2	2.13	0.48
1:B:94:GLU:OE2	1:B:97:LYS:NZ	2.38	0.48
1:B:-8:ASP:OD2	2:B:261:HOH:O	2.20	0.47
1:A:76:GLN:HG2	1:B:49:GLU:CB	2.44	0.47
1:A:24:LYS:HE3	1:A:135:ILE:HD11	1.95	0.46
1:A:66:GLU:HB2	1:A:69:GLU:HG3	1.98	0.46
1:A:143:ASN:HD21	1:B:46:TYR:HD1	1.65	0.45
1:B:75:PHE:CE1	1:B:150:TRP:HH2	2.35	0.44
1:B:148:TYR:HB3	1:B:152:TRP:HD1	1.81	0.43
1:B:130:ARG:HH12	1:B:151:LYS:H	1.67	0.42
1:B:24:LYS:HE3	1:B:26:TYR:OH	2.19	0.42
1:A:143:ASN:H	1:A:143:ASN:HD22	1.66	0.42
1:B:8:GLU:OE2	2:B:171:HOH:O	2.22	0.42
1:B:102:LEU:CD1	1:B:104:ILE:HD11	2.49	0.42
1:B:150:TRP:O	1:B:151:LYS:HB2	2.20	0.42
1:B:151:LYS:HA	1:B:151:LYS:HD2	1.30	0.42
1:A:78:LYS:HA	1:A:79:PRO:HD2	1.86	0.42
1:A:30:ALA:HB2	1:A:132:ARG:HD2	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:100:PRO:HB3	1:A:133:LEU:CD2	2.51	0.40
1:B:-4:GLN:OE1	1:B:-1:SER:HB3	2.21	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	154/161 (96%)	152 (99%)	2 (1%)	0	100	100
1	B	155/161 (96%)	148 (96%)	6 (4%)	1 (1%)	33	24
All	All	309/322 (96%)	300 (97%)	8 (3%)	1 (0%)	50	44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	151	LYS

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	142/145 (98%)	138 (97%)	4 (3%)	56	54
1	B	143/145 (99%)	137 (96%)	6 (4%)	40	34
All	All	285/290 (98%)	275 (96%)	10 (4%)	48	43

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

Mol	Chain	Res	Type
1	A	4	GLU
1	A	49	GLU
1	A	76	GLN
1	A	102	LEU
1	B	4	GLU
1	B	72	ASP
1	B	77	LEU
1	B	102	LEU
1	B	151	LYS
1	B	152	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	ASN
1	A	143	ASN
1	B	28	ASN
1	B	143	ASN

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	158/161 (98%)	0.21	9 (5%)	23	22	15, 27, 56, 65	0
1	B	159/161 (98%)	0.39	9 (5%)	23	22	15, 27, 57, 65	0
All	All	317/322 (98%)	0.30	18 (5%)	23	22	15, 27, 56, 65	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	82	PHE	8.7
1	A	82	PHE	5.3
1	B	64	TYR	5.2
1	B	81	LYS	5.2
1	A	81	LYS	4.7
1	B	152	TRP	4.5
1	B	86	GLN	4.2
1	A	48	ARG	3.2
1	A	85	GLU	3.0
1	B	85	GLU	2.8
1	A	64	TYR	2.7
1	A	49	GLU	2.7
1	B	150	TRP	2.5
1	A	72	ASP	2.3
1	B	87	VAL	2.2
1	A	76	GLN	2.1
1	A	86	GLN	2.1
1	B	84	SER	2.0

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.