



Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 10:01 pm BST

PDB ID : 4CTI
Title : Escherichia coli EnvZ histidine kinase catalytic part fused to Archaeoglobus fulgidus Af1503 HAMP domain
Authors : Ferris, H.U.; Coles, M.; Lupas, A.N.; Hartmann, M.D.
Deposited on : 2014-03-13
Resolution : 2.85 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
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

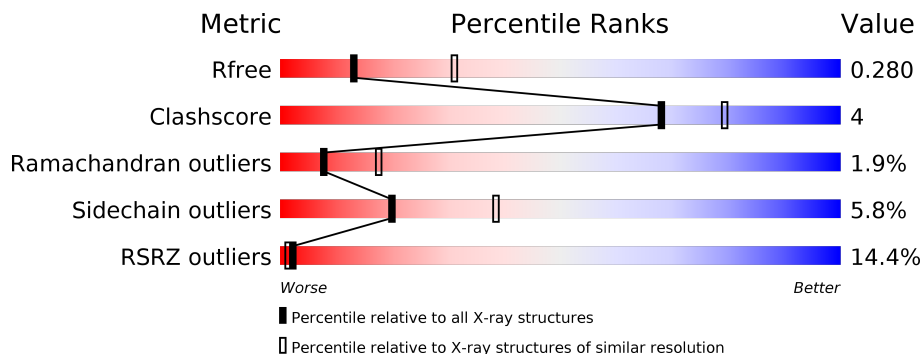
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.85 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	1031 (2.86-2.82)
Clashscore	141614	1078 (2.86-2.82)
Ramachandran outliers	138981	1050 (2.86-2.82)
Sidechain outliers	138945	1051 (2.86-2.82)
RSRZ outliers	127900	1019 (2.86-2.82)

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	273	
1	B	273	
1	C	273	
1	D	273	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6389 atoms, of which 0 are hydrogens and 0 are deuteriums.

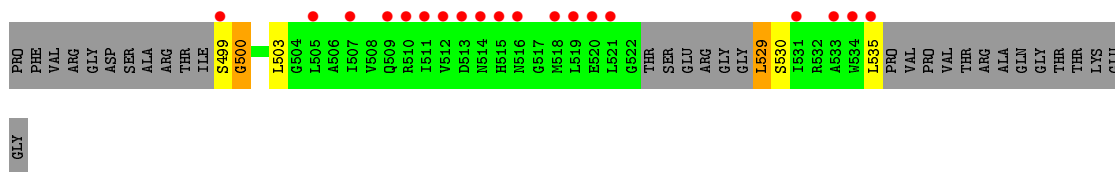
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called OSMOLARITY SENSOR PROTEIN ENVZ, AF1503.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	207	Total 1558	C 979	N 271	O 301	S 7	0	0	0
1	B	222	Total 1671	C 1050	N 288	O 325	S 8	0	0	0
1	C	203	Total 1508	C 943	N 264	O 294	S 7	0	0	0
1	D	223	Total 1652	C 1041	N 284	O 319	S 8	0	0	0

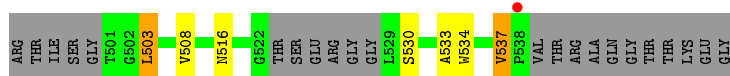
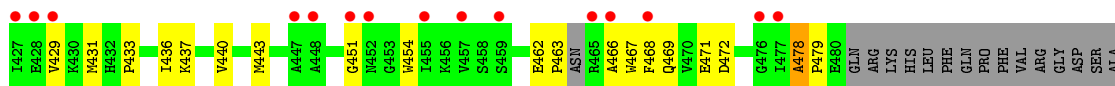
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	MET	-	expression tag	UNP O28769
A	291	PHE	ALA	engineered mutation	UNP O28769
B	277	MET	-	expression tag	UNP O28769
B	291	PHE	ALA	engineered mutation	UNP O28769
C	277	MET	-	expression tag	UNP O28769
C	291	PHE	ALA	engineered mutation	UNP O28769
D	277	MET	-	expression tag	UNP O28769
D	291	PHE	ALA	engineered mutation	UNP O28769



GLY

- Molecule 1: OSMOLARITY SENSOR PROTEIN ENVZ, AF1503



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.06Å 76.62Å 97.37Å 90.00° 107.03° 90.00°	Depositor
Resolution (Å)	38.83 – 2.85 38.83 – 2.85	Depositor EDS
% Data completeness (in resolution range)	99.2 (38.83-2.85) 99.3 (38.83-2.85)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 2.86Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.258 , 0.277 0.263 , 0.280	Depositor DCC
R_{free} test set	1170 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	68.6	Xtrriage
Anisotropy	0.384	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 66.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	6389	wwPDB-VP
Average B, all atoms (Å ²)	114.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% 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](#)

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	0.21	0/1574	0.39	0/2130
1	B	0.21	0/1692	0.42	0/2296
1	C	0.21	0/1519	0.41	1/2053 (0.0%)
1	D	0.21	0/1672	0.40	0/2270
All	All	0.21	0/6457	0.40	1/8749 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	529	LEU	CA-CB-CG	5.17	127.19	115.30

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	1558	0	1528	10	0
1	B	1671	0	1631	12	0
1	C	1508	0	1485	10	0
1	D	1652	0	1608	18	0
All	All	6389	0	6252	48	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 4.

All (48) 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:382:GLN:NE2	1:A:500:GLY:O	2.23	0.71
1:D:429:VAL:HG12	1:D:537:VAL:HG22	1.72	0.70
1:D:469:GLN:HE21	1:D:530:SER:HB2	1.60	0.66
1:C:382:GLN:NE2	1:C:500:GLY:O	2.28	0.64
1:B:462:GLU:HG3	1:B:463:PRO:HD2	1.84	0.59
1:C:457:VAL:HG22	1:C:470:VAL:HG22	1.85	0.58
1:D:451:GLY:HA3	1:D:472:ASP:HB2	1.86	0.57
1:B:308:ARG:NH2	1:B:310:ASP:OD2	2.40	0.55
1:A:423:TYR:CD1	1:A:460:GLY:HA2	2.44	0.53
1:B:305:HIS:HB3	1:B:308:ARG:HD3	1.92	0.51
1:A:467:TRP:HA	1:A:533:ALA:O	2.11	0.51
1:D:469:GLN:NE2	1:D:471:GLU:OE2	2.44	0.51
1:C:325:SER:OG	1:D:516:ASN:OD1	2.23	0.51
1:D:305:HIS:HB3	1:D:308:ARG:HD3	1.93	0.51
1:C:405:GLY:HA2	1:C:408:ILE:HD12	1.91	0.50
1:A:405:GLY:HA2	1:A:408:ILE:HD12	1.92	0.50
1:D:443:MET:HG3	1:D:508:VAL:HG21	1.94	0.50
1:C:466:ALA:HB3	1:C:535:LEU:HB2	1.94	0.49
1:D:462:GLU:HG3	1:D:463:PRO:HD2	1.94	0.49
1:B:389:THR:OG1	1:B:438:ARG:NH1	2.46	0.48
1:C:423:TYR:H	1:C:459:SER:HB2	1.78	0.48
1:D:467:TRP:HA	1:D:533:ALA:O	2.14	0.48
1:D:345:ARG:HH21	1:D:384:ILE:HD11	1.80	0.47
1:B:435:SER:HB3	1:B:511:ILE:HG23	1.98	0.46
1:A:280:ILE:HB	1:A:312:ILE:HD11	1.98	0.45
1:C:383:PHE:CG	1:D:340:VAL:HG21	2.51	0.45
1:B:308:ARG:HB3	1:B:308:ARG:HE	1.56	0.44
1:D:478:ALA:HA	1:D:479:PRO:HD3	1.88	0.44
1:A:435:SER:HB3	1:A:511:ILE:HG23	2.01	0.43
1:A:403:VAL:O	1:A:407:VAL:HG23	2.19	0.42
1:C:403:VAL:O	1:C:407:VAL:HG23	2.19	0.42
1:B:431:MET:SD	1:B:436:ILE:HG13	2.59	0.42
1:C:315:LEU:O	1:C:319:ILE:HG13	2.19	0.41
1:A:315:LEU:O	1:A:319:ILE:HG13	2.20	0.41
1:A:435:SER:O	1:A:511:ILE:HD13	2.20	0.41
1:B:386:TYR:OH	1:B:510:ARG:HG2	2.19	0.41
1:B:317:LYS:HB3	1:B:317:LYS:HE2	1.76	0.41
1:B:429:VAL:HG12	1:B:537:VAL:HG22	2.03	0.41
1:D:433:PRO:O	1:D:437:LYS:HB2	2.21	0.41
1:A:454:TRP:CD1	1:A:473:ASP:HB2	2.55	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:440:VAL:HG22	1:B:468:PHE:CE1	2.56	0.40
1:D:440:VAL:HG22	1:D:468:PHE:CZ	2.57	0.40
1:B:478:ALA:HA	1:B:479:PRO:HD3	1.96	0.40
1:D:383:PHE:HB2	1:D:503:LEU:HD11	2.02	0.40
1:D:466:ALA:O	1:D:534:TRP:HA	2.21	0.40
1:D:317:LYS:HB3	1:D:317:LYS:HE2	1.76	0.40
1:D:431:MET:SD	1:D:436:ILE:HG13	2.62	0.40
1:C:499:SER:OG	1:C:500:GLY:N	2.54	0.40

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	193/273 (71%)	185 (96%)	4 (2%)	4 (2%)	7	15
1	B	210/273 (77%)	197 (94%)	10 (5%)	3 (1%)	11	24
1	C	189/273 (69%)	182 (96%)	4 (2%)	3 (2%)	9	21
1	D	211/273 (77%)	196 (93%)	10 (5%)	5 (2%)	6	12
All	All	803/1092 (74%)	760 (95%)	28 (4%)	15 (2%)	8	17

All (15) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	394	PRO
1	A	467	TRP
1	C	467	TRP
1	A	462	GLU
1	C	462	GLU
1	A	453	GLY
1	B	480	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	537	VAL
1	D	478	ALA
1	D	537	VAL
1	B	478	ALA
1	D	393	MET
1	A	500	GLY
1	C	500	GLY
1	D	390	GLY

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	161/226 (71%)	150 (93%)	11 (7%)	16	31
1	B	173/226 (76%)	162 (94%)	11 (6%)	17	34
1	C	156/226 (69%)	149 (96%)	7 (4%)	27	51
1	D	168/226 (74%)	159 (95%)	9 (5%)	22	42
All	All	658/904 (73%)	620 (94%)	38 (6%)	20	38

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	327	LYS
1	A	333	ARG
1	A	335	LEU
1	A	337	MET
1	A	368	SER
1	A	454	TRP
1	A	467	TRP
1	A	473	ASP
1	A	503	LEU
1	A	529	LEU
1	A	530	SER
1	B	308	ARG
1	B	321	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	325	SER
1	B	336	LEU
1	B	365	LEU
1	B	374	GLU
1	B	423	TYR
1	B	428	GLU
1	B	432	HIS
1	B	454	TRP
1	B	503	LEU
1	C	327	LYS
1	C	333	ARG
1	C	335	LEU
1	C	368	SER
1	C	503	LEU
1	C	529	LEU
1	C	530	SER
1	D	308	ARG
1	D	321	ARG
1	D	325	SER
1	D	336	LEU
1	D	365	LEU
1	D	374	GLU
1	D	381	GLU
1	D	454	TRP
1	D	503	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	469	GLN

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

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	207/273 (75%)	0.59	23 (11%) 5 3	53, 91, 169, 202	0
1	B	222/273 (81%)	0.52	14 (6%) 20 14	55, 91, 163, 189	0
1	C	203/273 (74%)	1.77	61 (30%) 0 0	64, 104, 306, 412	0
1	D	223/273 (81%)	0.76	25 (11%) 5 3	60, 98, 167, 206	0
All	All	855/1092 (78%)	0.90	123 (14%) 2 1	53, 95, 251, 412	0

All (123) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	430	LYS	14.3
1	C	514	ASN	14.3
1	C	403	VAL	14.0
1	C	447	ALA	12.5
1	C	468	PHE	11.4
1	C	519	LEU	10.9
1	C	535	LEU	10.1
1	C	435	SER	9.0
1	C	515	HIS	8.6
1	C	459	SER	7.7
1	C	431	MET	7.3
1	C	407	VAL	7.2
1	C	422	LEU	6.7
1	C	513	ASP	6.5
1	C	533	ALA	6.4
1	C	400	LEU	6.3
1	C	432	HIS	6.1
1	D	427	ILE	5.7
1	C	450	TYR	5.6
1	A	417	GLU	5.5
1	C	518	MET	5.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	534	TRP	5.2
1	D	429	VAL	5.2
1	C	446	ASN	5.1
1	D	465	ARG	5.0
1	C	455	ILE	5.0
1	C	470	VAL	4.9
1	C	461	THR	4.7
1	A	431	MET	4.6
1	C	521	LEU	4.4
1	C	466	ALA	4.4
1	D	476	GLY	4.3
1	A	452	ASN	4.3
1	C	516	ASN	4.3
1	D	455	ILE	4.3
1	C	451	GLY	4.3
1	C	531	ILE	4.2
1	C	448	ALA	4.2
1	C	457	VAL	4.2
1	A	451	GLY	4.2
1	C	406	GLU	4.1
1	C	511	ILE	4.0
1	D	468	PHE	3.9
1	D	420	THR	3.7
1	A	475	PRO	3.6
1	C	443	MET	3.6
1	D	538	PRO	3.6
1	D	466	ALA	3.6
1	C	418	ILE	3.6
1	A	418	ILE	3.5
1	C	475	PRO	3.4
1	C	420	THR	3.4
1	C	460	GLY	3.4
1	B	447	ALA	3.3
1	A	471	GLU	3.3
1	D	393	MET	3.2
1	C	440	VAL	3.2
1	C	402	ALA	3.2
1	A	430	LYS	3.2
1	B	407	VAL	3.1
1	A	519	LEU	3.1
1	C	399	ASP	3.1
1	A	521	LEU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	473	ASP	3.0
1	C	458	SER	3.0
1	D	451	GLY	3.0
1	B	397	MET	3.0
1	B	451	GLY	2.9
1	A	401	ASN	2.9
1	C	471	GLU	2.9
1	D	477	ILE	2.8
1	A	307	ASN	2.8
1	D	448	ALA	2.8
1	D	459	SER	2.8
1	D	452	ASN	2.8
1	C	499	SER	2.8
1	C	417	GLU	2.8
1	D	423	TYR	2.8
1	C	473	ASP	2.7
1	A	466	ALA	2.7
1	C	510	ARG	2.7
1	D	428	GLU	2.7
1	D	298	ASN	2.7
1	C	507	ILE	2.7
1	B	418	ILE	2.6
1	C	401	ASN	2.6
1	D	421	ALA	2.6
1	B	329	LEU	2.5
1	A	436	ILE	2.5
1	D	447	ALA	2.5
1	A	450	TYR	2.5
1	D	400	LEU	2.5
1	C	439	ALA	2.4
1	D	403	VAL	2.4
1	A	400	LEU	2.4
1	C	509	GLN	2.4
1	C	434	LEU	2.4
1	A	454	TRP	2.4
1	B	476	GLY	2.4
1	C	361	GLN	2.4
1	C	449	ARG	2.4
1	A	408	ILE	2.4
1	B	427	ILE	2.4
1	C	520	GLU	2.4
1	B	428	GLU	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	512	VAL	2.3
1	C	421	ALA	2.3
1	D	457	VAL	2.3
1	B	465	ARG	2.3
1	C	505	LEU	2.2
1	C	452	ASN	2.2
1	A	403	VAL	2.2
1	D	398	ALA	2.2
1	A	407	VAL	2.2
1	A	529	LEU	2.2
1	B	429	VAL	2.1
1	B	455	ILE	2.1
1	B	464	ASN	2.1
1	A	468	PHE	2.1
1	C	309	ALA	2.1
1	C	474	GLY	2.1
1	D	299	LEU	2.1
1	A	518	MET	2.1

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.