wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 04:39 PM EDT

PDB ID	:	4IT5
Title	:	Chaperone HscB from Vibrio cholerae
Authors	:	Osipiuk, J.; Gu, M.; Papazisi, L.; Anderson, W.F.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2013-01-17
Resolution	:	2.15 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

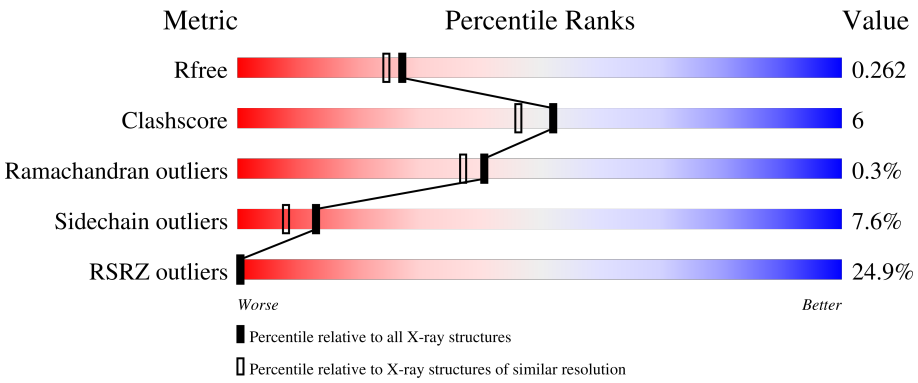
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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 of 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 <=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	174	20% 82% 11% • 6%
1	B	174	24% 75% 20% • •
1	C	174	17% 78% 14% • 6%
1	D	174	31% 74% 22% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5469 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 Co-chaperone protein HscB homolog.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	164	1326	838	228	252	1	7	0	0	0
1	B	169	1365	861	235	262	1	6	0	0	0
1	C	163	1317	833	226	250	1	7	0	0	0
1	D	169	1365	861	235	262	1	6	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q9KTX9
A	-1	ASN	-	EXPRESSION TAG	UNP Q9KTX9
A	0	ALA	-	EXPRESSION TAG	UNP Q9KTX9
B	-2	SER	-	EXPRESSION TAG	UNP Q9KTX9
B	-1	ASN	-	EXPRESSION TAG	UNP Q9KTX9
B	0	ALA	-	EXPRESSION TAG	UNP Q9KTX9
C	-2	SER	-	EXPRESSION TAG	UNP Q9KTX9
C	-1	ASN	-	EXPRESSION TAG	UNP Q9KTX9
C	0	ALA	-	EXPRESSION TAG	UNP Q9KTX9
D	-2	SER	-	EXPRESSION TAG	UNP Q9KTX9
D	-1	ASN	-	EXPRESSION TAG	UNP Q9KTX9
D	0	ALA	-	EXPRESSION TAG	UNP Q9KTX9

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	C	1	1	1	0	0
			Total	Ca		
2	D	2	2	2	0	0

- Molecule 3 is water.

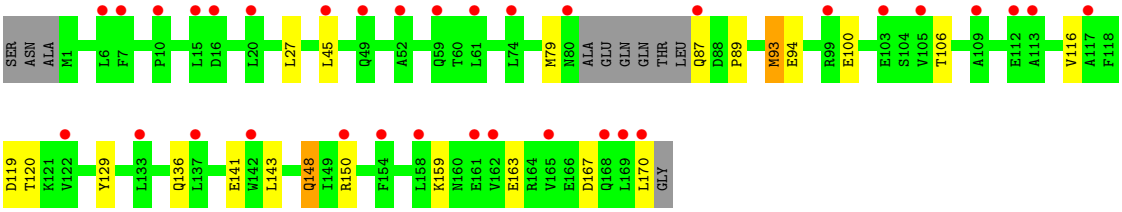
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	26	26	26	0	0
			Total	O		
3	B	15	15	15	0	0
			Total	O		
3	C	36	36	36	0	0
			Total	O		
3	D	14	14	14	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

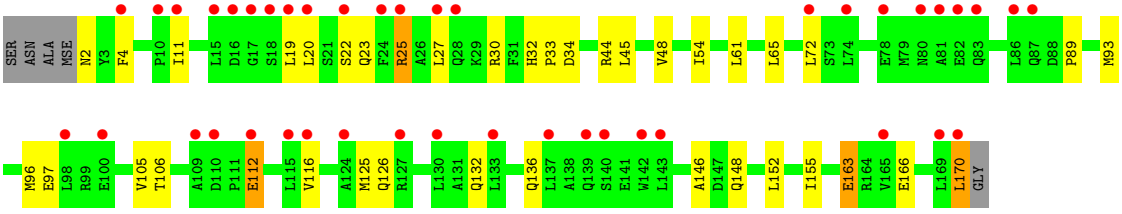
• Molecule 1: Co-chaperone protein HscB homolog

Chain A:



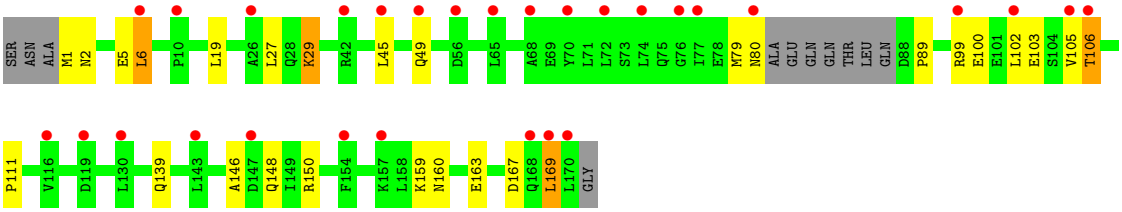
• Molecule 1: Co-chaperone protein HscB homolog

Chain B:



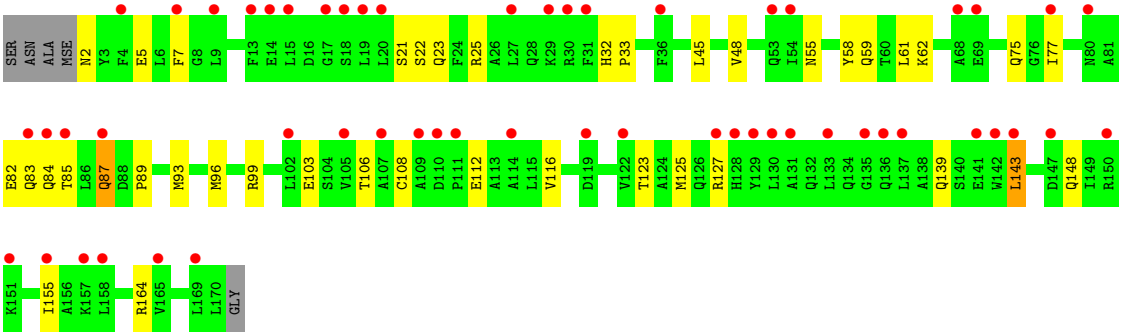
• Molecule 1: Co-chaperone protein HscB homolog

Chain C:



• Molecule 1: Co-chaperone protein HscB homolog

Chain D:


SER
ASN
ALA
MSE
N2
Y3
F4
E5
L6
F7
G8
L9
F13
E14
L15
D16
G17
S18
L19
L20
S21
S22
Q23
F24
R25
A26
L27
Q28
K29
R30
F31
H32
P33
F36
L45
V48
Q53
I54
N55
Y58
Q59
T60
L61
K62
A68
E69
Q75
G76
I77
N80
A81
E82
Q83
Q84
T85
L86
Q87
D88
P89
M93
M96
R99
L102
E103
S104
V105
T106
A107
C108
A109
D110
P111
E112
A113
A114
L115
V116
D119
V122
T123
A124
M125
Q126
R127
H128
Y129
L130
A131
Q132
L133
Q134
G135
Q136
L137
A138
Q139
S140
E141
W142
L143
D147
Q148
I149
R150
K151
I155
A156
K157
L158
R164
V165
L169
L170
GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.43Å 100.71Å 70.35Å 90.00° 90.47° 90.00°	Depositor
Resolution (Å)	48.56 – 2.15 48.56 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.6 (48.56-2.15) 96.9 (48.56-2.15)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.16 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.220 , 0.259 0.222 , 0.262	Depositor DCC
R_{free} test set	2099 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	43.1	Xtriage
Anisotropy	0.521	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.439 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5469	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.13% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: CA

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	0.26	0/1340	0.40	0/1794
1	B	0.25	0/1380	0.40	0/1851
1	C	0.26	0/1331	0.41	0/1782
1	D	0.25	0/1380	0.42	0/1851
All	All	0.26	0/5431	0.41	0/7278

There are no bond length outliers.

There are no bond angle outliers.

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	1326	0	1317	14	0
1	B	1365	0	1351	22	0
1	C	1317	0	1309	12	0
1	D	1365	0	1351	20	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	2	0	0	0	0
3	A	26	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	15	0	0	4	0
3	C	36	0	0	1	0
3	D	14	0	0	0	0
All	All	5469	0	5328	59	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 6.

The worst 5 of 59 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:170:LEU:HD11	1:B:163:GLU:HG3	1.62	0.80
1:B:163:GLU:OE2	3:B:315:HOH:O	2.00	0.77
1:C:1:MSE:HE3	1:C:6:LEU:HD12	1.72	0.71
1:A:136:GLN:HB2	1:A:141:GLU:HB2	1.73	0.69
1:B:45:LEU:HD22	1:D:89:PRO:HB2	1.77	0.67

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	160/174 (92%)	157 (98%)	3 (2%)	0	100 100
1	B	167/174 (96%)	163 (98%)	3 (2%)	1 (1%)	25 18
1	C	159/174 (91%)	153 (96%)	5 (3%)	1 (1%)	25 18
1	D	167/174 (96%)	163 (98%)	4 (2%)	0	100 100
All	All	653/696 (94%)	636 (97%)	15 (2%)	2 (0%)	41 37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	169	LEU
1	B	19	LEU

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	140/140 (100%)	132 (94%)	8 (6%)	20 16
1	B	144/140 (103%)	131 (91%)	13 (9%)	9 5
1	C	139/140 (99%)	127 (91%)	12 (9%)	10 6
1	D	144/140 (103%)	134 (93%)	10 (7%)	15 10
All	All	567/560 (101%)	524 (92%)	43 (8%)	13 8

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	100	GLU
1	D	61	LEU
1	C	106	THR
1	C	160	ASN
1	D	106	THR

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

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 5 are monoatomic - 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 (i)

6.1 Protein, DNA and RNA chains (i)

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	157/174 (90%)	1.35	34 (21%)	0	0	32, 52, 80, 97	0
1	B	163/174 (93%)	1.58	42 (25%)	0	0	37, 66, 105, 113	0
1	C	156/174 (89%)	1.30	29 (18%)	1	1	32, 52, 75, 94	0
1	D	163/174 (93%)	1.64	54 (33%)	0	0	37, 65, 100, 119	0
All	All	639/696 (91%)	1.47	159 (24%)	0	0	32, 58, 95, 119	0

The worst 5 of 159 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	4	PHE	8.2
1	D	15	LEU	7.0
1	D	13	PHE	6.4
1	B	82	GLU	6.2
1	B	170	LEU	5.7

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 monosaccharides 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	CA	B	201	1/1	0.68	0.13	60,60,60,60	0
2	CA	D	201	1/1	0.88	0.09	61,61,61,61	0
2	CA	D	202	1/1	0.92	0.12	55,55,55,55	0
2	CA	A	301	1/1	0.95	0.06	62,62,62,62	0
2	CA	C	201	1/1	0.97	0.10	59,59,59,59	0

6.5 Other polymers (i)

There are no such residues in this entry.