wwPDB X-ray Structure Validation Summary Report (i)

Feb 28, 2014 – 11:08 PM GMT

PDB ID : 3I38
Title : Structure of a putative chaperone protein dnaj from klebsiella pneumoniae subsp. pneumoniae mgh 78578
Authors : Filippova, E.V.; Minasov, G.; Shuvalova, L.; Kiryukhina, O.; Bearden, J.; Joachimiak, A.; Anderson, W.F.; Midwest Center for Structural Genomics (MCSG)
Deposited on : 2009-06-30
Resolution : 2.30 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
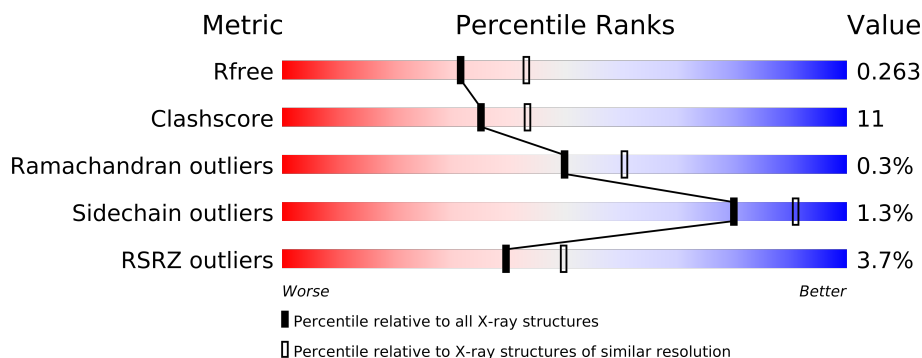We welcome your comments at validation@mail.wwpdb.org
A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.30 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
Rfree		0.263
Clashscore		11
Ramachandran outliers		0.3%
Sidechain outliers		1.3%
RSRZ outliers		3.7%

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))
R$_{free}$	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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	109	
1	B	109	
1	C	109	
1	D	109	
1	E	109	
1	F	109	
1	G	109	
1	H	109	
1	I	109	
1	J	109	
1	K	109	
1	L	109	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 9501 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 Putative chaperone DnaJ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	103	Total	C	N	O	Se	0	0	0
			795	517	138	139	1			
1	B	101	Total	C	N	O	Se	0	2	0
			793	516	139	137	1			
1	C	98	Total	C	N	O	Se	0	0	0
			753	492	128	132	1			
1	D	103	Total	C	N	O	Se	0	1	0
			804	522	140	141	1			
1	E	98	Total	C	N	O	Se	0	0	0
			753	492	128	132	1			
1	F	100	Total	C	N	O	Se	0	0	0
			773	504	134	134	1			
1	G	103	Total	C	N	O	Se	0	0	0
			795	517	138	139	1			
1	H	99	Total	C	N	O	Se	0	1	0
			771	501	132	137	1			
1	I	96	Total	C	N	O	Se	0	0	0
			737	481	126	129	1			
1	J	101	Total	C	N	O	Se	0	0	0
			780	508	135	136	1			
1	K	104	Total	C	N	O	Se	0	0	0
			799	520	139	139	1			
1	L	98	Total	C	N	O	Se	0	1	0
			762	497	130	134	1			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	197	SER	-	expression tag	UNP A6TH30
A	198	ASN	-	expression tag	UNP A6TH30
A	199	ALA	-	expression tag	UNP A6TH30
B	197	SER	-	expression tag	UNP A6TH30
B	198	ASN	-	expression tag	UNP A6TH30

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	199	ALA	-	expression tag	UNP A6TH30
C	197	SER	-	expression tag	UNP A6TH30
C	198	ASN	-	expression tag	UNP A6TH30
C	199	ALA	-	expression tag	UNP A6TH30
D	197	SER	-	expression tag	UNP A6TH30
D	198	ASN	-	expression tag	UNP A6TH30
D	199	ALA	-	expression tag	UNP A6TH30
E	197	SER	-	expression tag	UNP A6TH30
E	198	ASN	-	expression tag	UNP A6TH30
E	199	ALA	-	expression tag	UNP A6TH30
F	197	SER	-	expression tag	UNP A6TH30
F	198	ASN	-	expression tag	UNP A6TH30
F	199	ALA	-	expression tag	UNP A6TH30
G	197	SER	-	expression tag	UNP A6TH30
G	198	ASN	-	expression tag	UNP A6TH30
G	199	ALA	-	expression tag	UNP A6TH30
H	197	SER	-	expression tag	UNP A6TH30
H	198	ASN	-	expression tag	UNP A6TH30
H	199	ALA	-	expression tag	UNP A6TH30
I	197	SER	-	expression tag	UNP A6TH30
I	198	ASN	-	expression tag	UNP A6TH30
I	199	ALA	-	expression tag	UNP A6TH30
J	197	SER	-	expression tag	UNP A6TH30
J	198	ASN	-	expression tag	UNP A6TH30
J	199	ALA	-	expression tag	UNP A6TH30
K	-2	SER	-	expression tag	UNP A6TH30
K	-1	ASN	-	expression tag	UNP A6TH30
K	0	ALA	-	expression tag	UNP A6TH30
L	197	SER	-	expression tag	UNP A6TH30
L	198	ASN	-	expression tag	UNP A6TH30
L	199	ALA	-	expression tag	UNP A6TH30

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	21	Total 21	O 21	0	0
2	B	26	Total 26	O 26	0	0
2	C	9	Total 9	O 9	0	0
2	D	19	Total 19	O 19	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	3	Total 3	O 3	0	0
2	F	12	Total 12	O 12	0	0
2	G	27	Total 27	O 27	0	0
2	H	11	Total 11	O 11	0	0
2	I	3	Total 3	O 3	0	0
2	J	8	Total 8	O 8	0	0
2	K	20	Total 20	O 20	0	0
2	L	27	Total 27	O 27	0	0

3 Residue-property plots (i)

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.
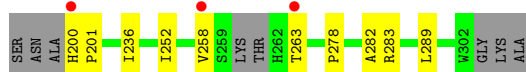
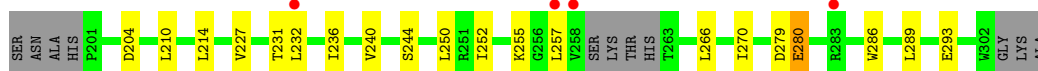- Molecule 1: Putative chaperone DnaJ

Chain A:

- Molecule 1: Putative chaperone DnaJ

Chain B:

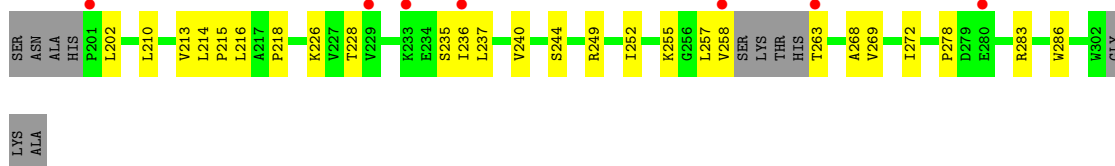- Molecule 1: Putative chaperone DnaJ

Chain C:

- Molecule 1: Putative chaperone DnaJ

Chain D:

- Molecule 1: Putative chaperone DnaJ

Chain E:

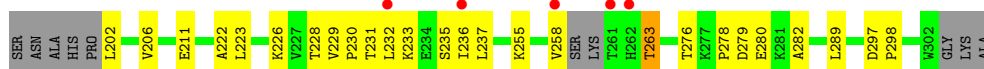- Molecule 1: Putative chaperone DnaJ

Chain F:

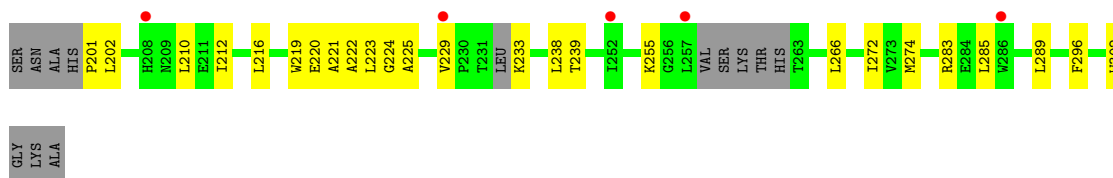• Molecule 1: Putative chaperone DnaJ

Chain G:

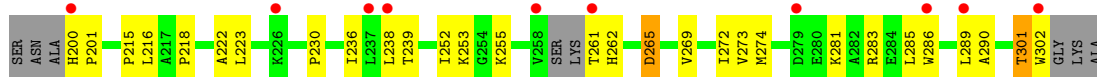• Molecule 1: Putative chaperone DnaJ

Chain H:

• Molecule 1: Putative chaperone DnaJ

Chain I:

• Molecule 1: Putative chaperone DnaJ

Chain J:

• Molecule 1: Putative chaperone DnaJ

Chain K:

• Molecule 1: Putative chaperone DnaJ

Chain L:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	61.80Å 61.63Å 92.91Å 81.75° 78.71° 83.79°	Depositor
Resolution (Å)	90.54 – 2.30 38.66 – 2.19	Depositor EDS
% Data completeness (in resolution range)	98.1 (90.54-2.30) 97.6 (38.66-2.19)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0051	Depositor
R, R_{free}	0.217 , 0.262 0.230 , 0.263	Depositor DCC
R_{free} test set	2942 reflections (5.36%)	DCC
Wilson B-factor (Å^2)	47.3	Xtriage
Anisotropy	0.163	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 33.4	EDS
Estimated twinning fraction	0.014 for -k,-h,-l	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	0 of 66296 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9501	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.74% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	0.68	0/814	0.81	0/1107
1	B	0.70	0/817	0.84	0/1109
1	C	0.57	0/769	0.79	0/1044
1	D	0.68	0/823	0.83	0/1119
1	E	0.48	0/769	0.78	0/1044
1	F	0.54	0/791	0.80	0/1075
1	G	0.67	0/814	0.84	1/1107 (0.1%)
1	H	0.52	0/787	0.76	0/1069
1	I	0.47	0/752	0.74	0/1019
1	J	0.57	0/798	0.77	0/1085
1	K	0.61	0/818	0.83	0/1113
1	L	0.67	0/778	0.89	1/1056 (0.1%)
All	All	0.60	0/9530	0.81	2/12947 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	279	ASP	CB-CG-OD2	5.85	123.56	118.30
1	L	224	GLY	N-CA-C	-5.21	100.06	113.10

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	795	0	834	17	0
1	B	793	0	834	12	0
1	C	753	0	795	24	0
1	D	804	0	841	10	0
1	E	753	0	795	26	0
1	F	773	0	808	24	0
1	G	795	0	834	12	0
1	H	771	0	804	19	0
1	I	737	0	771	30	0
1	J	780	0	815	37	0
1	K	799	0	839	19	0
1	L	762	0	802	23	0
2	A	21	0	0	1	0
2	B	26	0	0	0	0
2	C	9	0	0	0	0
2	D	19	0	0	1	0
2	E	3	0	0	0	0
2	F	12	0	0	0	0
2	G	27	0	0	0	0
2	H	11	0	0	0	0
2	I	3	0	0	0	0
2	J	8	0	0	1	0
2	K	20	0	0	2	0
2	L	27	0	0	3	0
All	All	9501	0	9772	202	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 11.

The worst 5 of 202 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:236:ILE:HG21	1:C:252:ILE:HD13	1.28	1.12
1:J:238:LEU:HD23	1:J:239:THR:N	1.74	1.01
1:I:201:PRO:HB2	1:I:212:ILE:HD11	1.43	0.96
1:A:231:THR:HG23	2:A:118:HOH:O	1.73	0.89
1:C:231:THR:HG22	1:C:232:LEU:H	1.37	0.89

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	101/109 (93%)	95 (94%)	6 (6%)	0	100	100
1	B	99/109 (91%)	95 (96%)	4 (4%)	0	100	100
1	C	94/109 (86%)	88 (94%)	5 (5%)	1 (1%)	21	21
1	D	102/109 (94%)	95 (93%)	7 (7%)	0	100	100
1	E	94/109 (86%)	84 (89%)	10 (11%)	0	100	100
1	F	96/109 (88%)	87 (91%)	9 (9%)	0	100	100
1	G	101/109 (93%)	94 (93%)	6 (6%)	1 (1%)	22	23
1	H	96/109 (88%)	91 (95%)	5 (5%)	0	100	100
1	I	90/109 (83%)	79 (88%)	11 (12%)	0	100	100
1	J	97/109 (89%)	88 (91%)	8 (8%)	1 (1%)	22	23
1	K	102/109 (94%)	94 (92%)	8 (8%)	0	100	100
1	L	95/109 (87%)	94 (99%)	1 (1%)	0	100	100
All	All	1167/1308 (89%)	1084 (93%)	80 (7%)	3 (0%)	50	60

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	301	THR
1	C	280	GLU
1	G	276	THR

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	85/87 (98%)	83 (98%)	2 (2%)	61	79
1	B	85/87 (98%)	85 (100%)	0	100	100
1	C	80/87 (92%)	79 (99%)	1 (1%)	80	91
1	D	86/87 (99%)	86 (100%)	0	100	100
1	E	80/87 (92%)	80 (100%)	0	100	100
1	F	82/87 (94%)	79 (96%)	3 (4%)	45	60
1	G	85/87 (98%)	85 (100%)	0	100	100
1	H	82/87 (94%)	80 (98%)	2 (2%)	61	79
1	I	77/87 (88%)	77 (100%)	0	100	100
1	J	83/87 (95%)	82 (99%)	1 (1%)	82	92
1	K	84/87 (97%)	84 (100%)	0	100	100
1	L	81/87 (93%)	77 (95%)	4 (5%)	35	45
All	All	990/1044 (95%)	977 (99%)	13 (1%)	80	91

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	279	ASP
1	H	263	THR
1	L	255	LYS
1	F	277	LYS
1	L	202	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	262	HIS
1	L	245	GLN
1	K	245	GLN
1	I	287	GLN
1	K	200	HIS

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	103/109 (94%)	0.07	2 (1%)	64	73	17, 27, 44, 61	0
1	B	101/109 (92%)	0.21	3 (2%)	48	58	15, 25, 42, 60	0
1	C	98/109 (89%)	0.34	4 (4%)	35	46	16, 28, 37, 69	0
1	D	103/109 (94%)	-0.03	0	100	100	15, 29, 39, 69	0
1	E	98/109 (89%)	0.63	7 (7%)	16	23	22, 32, 37, 68	0
1	F	100/109 (91%)	0.25	4 (4%)	36	47	24, 31, 38, 72	0
1	G	103/109 (94%)	-0.05	0	100	100	19, 28, 39, 66	0
1	H	99/109 (90%)	0.35	5 (5%)	27	37	20, 30, 37, 62	0
1	I	96/109 (88%)	0.60	5 (5%)	26	36	23, 31, 46, 79	0
1	J	101/109 (92%)	0.62	10 (9%)	8	13	23, 32, 44, 79	0
1	K	104/109 (95%)	0.10	3 (2%)	49	59	18, 28, 42, 70	0
1	L	98/109 (89%)	0.06	1 (1%)	79	87	18, 28, 37, 68	0
All	All	1204/1308 (92%)	0.26	44 (3%)	39	50	15, 30, 41, 79	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	258	VAL	4.5
1	F	296	PHE	4.4
1	J	237	LEU	4.3
1	H	261	THR	4.2
1	E	201	PRO	4.0

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.