Full wwPDB X-ray Structure Validation Report (i)

Aug 19, 2023 – 09:27 PM EDT

PDB ID	:	2GQ0
Title	:	Crystal Structure of the Middle Domain of HtpG, the E. coli Hsp90
Authors	:	Harris, S.F.; Shiau, A.K.; Agard, D.A.
Deposited on	:	2006-04-19
Resolution	:	1.90 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 (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 (1)) were used in the production of this report:

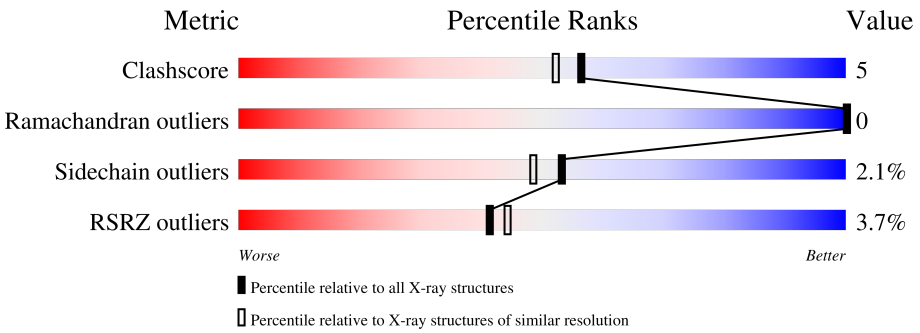
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.90 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))
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	303	3% 77% 7% • 16%
1	B	303	3% 78% 8% • 13%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4827 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 Chaperone protein htpG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	256	2156	1370	364	415	7	0	10	0
1	B	264	2190	1392	370	421	7	0	2	0

There are 74 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	229	MET	-	cloning artifact	UNP P0A6Z3
A	496	GLY	-	cloning artifact	UNP P0A6Z3
A	497	ASP	-	cloning artifact	UNP P0A6Z3
A	498	LEU	-	cloning artifact	UNP P0A6Z3
A	499	GLY	-	cloning artifact	UNP P0A6Z3
A	500	THR	-	cloning artifact	UNP P0A6Z3
A	501	LEU	-	cloning artifact	UNP P0A6Z3
A	502	VAL	-	cloning artifact	UNP P0A6Z3
A	503	PRO	-	cloning artifact	UNP P0A6Z3
A	504	ARG	-	cloning artifact	UNP P0A6Z3
A	505	GLY	-	cloning artifact	UNP P0A6Z3
A	506	SER	-	cloning artifact	UNP P0A6Z3
A	507	MET	-	cloning artifact	UNP P0A6Z3
A	508	ALA	-	cloning artifact	UNP P0A6Z3
A	509	ILE	-	cloning artifact	UNP P0A6Z3
A	510	SER	-	cloning artifact	UNP P0A6Z3
A	511	ASP	-	cloning artifact	UNP P0A6Z3
A	512	PRO	-	cloning artifact	UNP P0A6Z3
A	513	ASN	-	cloning artifact	UNP P0A6Z3
A	514	SER	-	cloning artifact	UNP P0A6Z3
A	515	SER	-	cloning artifact	UNP P0A6Z3
A	516	SER	-	cloning artifact	UNP P0A6Z3
A	517	VAL	-	cloning artifact	UNP P0A6Z3
A	518	ASP	-	cloning artifact	UNP P0A6Z3
A	519	LYS	-	cloning artifact	UNP P0A6Z3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	520	LEU	-	cloning artifact	UNP P0A6Z3
A	521	ALA	-	cloning artifact	UNP P0A6Z3
A	522	ALA	-	cloning artifact	UNP P0A6Z3
A	523	ALA	-	cloning artifact	UNP P0A6Z3
A	524	LEU	-	cloning artifact	UNP P0A6Z3
A	525	GLU	-	cloning artifact	UNP P0A6Z3
A	526	HIS	-	expression tag	UNP P0A6Z3
A	527	HIS	-	expression tag	UNP P0A6Z3
A	528	HIS	-	expression tag	UNP P0A6Z3
A	529	HIS	-	expression tag	UNP P0A6Z3
A	530	HIS	-	expression tag	UNP P0A6Z3
A	531	HIS	-	expression tag	UNP P0A6Z3
B	229	MET	-	cloning artifact	UNP P0A6Z3
B	496	GLY	-	cloning artifact	UNP P0A6Z3
B	497	ASP	-	cloning artifact	UNP P0A6Z3
B	498	LEU	-	cloning artifact	UNP P0A6Z3
B	499	GLY	-	cloning artifact	UNP P0A6Z3
B	500	THR	-	cloning artifact	UNP P0A6Z3
B	501	LEU	-	cloning artifact	UNP P0A6Z3
B	502	VAL	-	cloning artifact	UNP P0A6Z3
B	503	PRO	-	cloning artifact	UNP P0A6Z3
B	504	ARG	-	cloning artifact	UNP P0A6Z3
B	505	GLY	-	cloning artifact	UNP P0A6Z3
B	506	SER	-	cloning artifact	UNP P0A6Z3
B	507	MET	-	cloning artifact	UNP P0A6Z3
B	508	ALA	-	cloning artifact	UNP P0A6Z3
B	509	ILE	-	cloning artifact	UNP P0A6Z3
B	510	SER	-	cloning artifact	UNP P0A6Z3
B	511	ASP	-	cloning artifact	UNP P0A6Z3
B	512	PRO	-	cloning artifact	UNP P0A6Z3
B	513	ASN	-	cloning artifact	UNP P0A6Z3
B	514	SER	-	cloning artifact	UNP P0A6Z3
B	515	SER	-	cloning artifact	UNP P0A6Z3
B	516	SER	-	cloning artifact	UNP P0A6Z3
B	517	VAL	-	cloning artifact	UNP P0A6Z3
B	518	ASP	-	cloning artifact	UNP P0A6Z3
B	519	LYS	-	cloning artifact	UNP P0A6Z3
B	520	LEU	-	cloning artifact	UNP P0A6Z3
B	521	ALA	-	cloning artifact	UNP P0A6Z3
B	522	ALA	-	cloning artifact	UNP P0A6Z3
B	523	ALA	-	cloning artifact	UNP P0A6Z3
B	524	LEU	-	cloning artifact	UNP P0A6Z3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	525	GLU	-	cloning artifact	UNP P0A6Z3
B	526	HIS	-	expression tag	UNP P0A6Z3
B	527	HIS	-	expression tag	UNP P0A6Z3
B	528	HIS	-	expression tag	UNP P0A6Z3
B	529	HIS	-	expression tag	UNP P0A6Z3
B	530	HIS	-	expression tag	UNP P0A6Z3
B	531	HIS	-	expression tag	UNP P0A6Z3

- Molecule 2 is water.

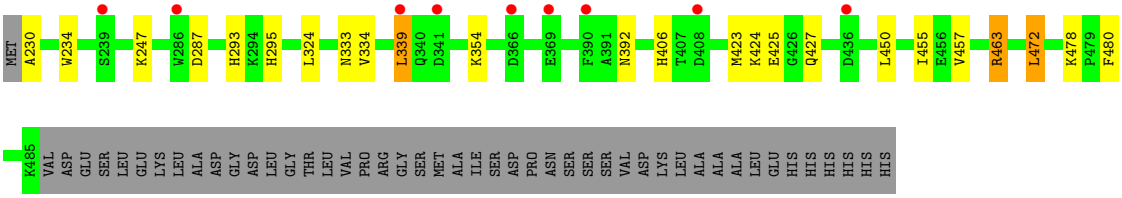
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	241	241	241	0	0
			Total	O		
2	B	240	240	240	0	0

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.

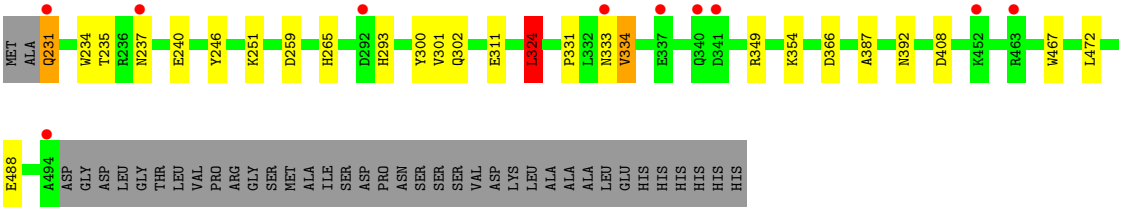
• Molecule 1: Chaperone protein htpG

Chain A:



• Molecule 1: Chaperone protein htpG

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	31.57Å 149.17Å 60.22Å 90.00° 99.01° 90.00°	Depositor
Resolution (Å)	19.83 – 1.90 19.83 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.7 (19.83-1.90) 96.7 (19.83-1.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.57 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.2.0005, CNS	Depositor
R, R_{free}	0.175 , 0.214 0.185 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	17.2	Xtriage
Anisotropy	0.167	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 68.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4827	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.35% 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)

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	0.62	0/2245	0.73	1/3034 (0.0%)
1	B	0.65	0/2247	0.72	2/3036 (0.1%)
All	All	0.64	0/4492	0.72	3/6070 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	324	LEU	CA-CB-CG	8.30	134.39	115.30
1	B	349	ARG	NE-CZ-NH1	5.69	123.15	120.30
1	A	463	ARG	NE-CZ-NH1	5.51	123.06	120.30

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	2156	0	2089	22	0
1	B	2190	0	2135	22	0
2	A	241	0	0	4	0
2	B	240	0	0	3	0
All	All	4827	0	4224	43	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 5.

All (43) 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:B:331:PRO:HG2	1:B:334:VAL:HG23	1.58	0.84
1:A:354:LYS:NZ	1:A:392:ASN:HD21	1.81	0.79
1:A:463:ARG:HG2	2:A:582:HOH:O	1.91	0.70
1:A:406:HIS:HE1	1:B:408:ASP:H	1.41	0.69
1:B:331:PRO:CG	1:B:334:VAL:HG23	2.28	0.64
1:A:354:LYS:HZ2	1:A:392:ASN:HD21	1.44	0.63
1:A:293:HIS:CE1	1:A:295:HIS:HB3	2.34	0.62
1:A:234:TRP:CE3	1:A:324:LEU:HD21	2.34	0.62
1:B:354:LYS:NZ	1:B:392:ASN:HD21	1.98	0.60
1:B:237:ASN:ND2	2:B:735:HOH:O	2.35	0.59
1:A:423:MET:SD	1:A:478:LYS:HD3	2.45	0.57
1:B:231:GLN:NE2	1:B:235:THR:OG1	2.36	0.57
1:A:472:LEU:HD22	1:A:480:PHE:CE1	2.40	0.57
1:B:234:TRP:CE3	1:B:324:LEU:HD21	2.40	0.56
1:B:293:HIS:NE2	2:B:585:HOH:O	2.33	0.53
1:A:424:LYS:HG3	1:A:455:ILE:HD13	1.91	0.53
1:A:354:LYS:HZ1	1:A:392:ASN:HD21	1.53	0.53
1:A:463:ARG:HG3	1:A:463:ARG:HH11	1.74	0.53
1:B:488:GLU:CD	1:B:488:GLU:H	2.12	0.53
1:A:334:VAL:CG2	1:A:339:LEU:HD13	2.39	0.52
1:B:237:ASN:O	1:B:240:GLU:HG2	2.09	0.52
1:B:300:TYR:CE2	1:B:324:LEU:HD22	2.45	0.52
1:B:246:TYR:OH	1:B:265:HIS:HD2	1.94	0.51
1:B:354:LYS:NZ	1:B:392:ASN:ND2	2.59	0.51
1:A:230:ALA:HB1	2:A:690:HOH:O	2.12	0.49
1:B:237:ASN:ND2	2:B:741:HOH:O	2.45	0.49
1:B:354:LYS:HZ1	1:B:392:ASN:HD21	1.61	0.48
1:A:354:LYS:HZ1	1:A:392:ASN:ND2	2.09	0.48
1:B:354:LYS:HZ1	1:B:392:ASN:ND2	2.11	0.48
1:A:287:ASP:O	1:A:293:HIS:HB2	2.14	0.48
1:A:354:LYS:NZ	1:A:392:ASN:ND2	2.58	0.47
1:A:247:LYS:NZ	2:A:756:HOH:O	2.42	0.47
1:B:331:PRO:CD	1:B:334:VAL:HG23	2.45	0.47
1:A:450:LEU:CD1	1:A:457:VAL:HG22	2.44	0.46
1:A:334:VAL:HG22	1:A:339:LEU:HD13	1.97	0.46
1:A:450:LEU:HD12	1:A:457:VAL:HG22	1.97	0.46
1:B:387:ALA:HB3	1:B:467:TRP:CH2	2.51	0.46
1:B:331:PRO:HD2	1:B:334:VAL:HG23	1.98	0.46
1:B:251:LYS:CE	1:B:259:ASP:OD1	2.64	0.46
1:B:301:VAL:HG12	1:B:302:GLN:HG3	1.99	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:427:GLN:NE2	1:A:455:ILE:HD12	2.33	0.44
1:A:333:ASN:ND2	2:A:713:HOH:O	2.52	0.41
1:B:331:PRO:HD2	1:B:334:VAL:CG2	2.50	0.41

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	264/303 (87%)	259 (98%)	5 (2%)	0	100	100
1	B	264/303 (87%)	260 (98%)	4 (2%)	0	100	100
All	All	528/606 (87%)	519 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	240/269 (89%)	236 (98%)	4 (2%)	60	57
1	B	240/269 (89%)	232 (97%)	8 (3%)	38	29
All	All	480/538 (89%)	468 (98%)	12 (2%)	53	41

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	339	LEU
1	A	425[A]	GLU
1	A	425[B]	GLU
1	A	472	LEU
1	B	231	GLN
1	B	311	GLU
1	B	324	LEU
1	B	333	ASN
1	B	334	VAL
1	B	366[A]	ASP
1	B	366[B]	ASP
1	B	472	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	295	HIS
1	A	392	ASN
1	A	406	HIS
1	A	427	GLN
1	B	231	GLN
1	B	237	ASN
1	B	265	HIS
1	B	272	GLN
1	B	333	ASN
1	B	376	GLN
1	B	392	ASN
1	B	406	HIS

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)

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 (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	256/303 (84%)	0.15	9 (3%)	44	47	10, 21, 34, 47	0
1	B	264/303 (87%)	0.11	10 (3%)	40	43	9, 19, 40, 49	0
All	All	520/606 (85%)	0.13	19 (3%)	41	44	9, 20, 37, 49	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	341	ASP	4.9
1	A	390	PHE	3.7
1	B	463	ARG	3.2
1	B	292	ASP	3.0
1	B	340	GLN	2.9
1	A	341	ASP	2.8
1	B	494	ALA	2.8
1	A	366	ASP	2.6
1	B	337	GLU	2.5
1	B	231	GLN	2.5
1	A	369	GLU	2.5
1	A	436	ASP	2.3
1	B	452	LYS	2.3
1	B	237	ASN	2.2
1	B	333	ASN	2.1
1	A	239	SER	2.1
1	A	408[A]	ASP	2.0
1	A	339	LEU	2.0
1	A	286	TRP	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.