



# Full wwPDB X-ray Structure Validation Report

Jun 10, 2014 – 02:00 PM EDT

PDB ID : 4L93  
Title : Crystal structure of Human Hsp90 with S36  
Authors : Li, J.; Ren, J.; Yang, M.; Xiong, B.; He, J.  
Deposited on : 2013-06-18  
Resolution : 1.84 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

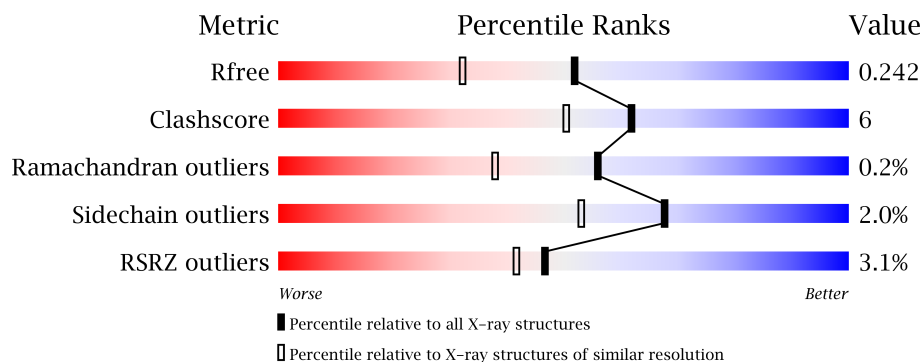
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : **FAILED**  
Xtriage (Phenix) : dev-1439  
EDS : stable23161  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23161

# 1 Overall quality at a glance

The reported resolution of this entry is 1.84 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))
$R_{free}$	66092	1857 (1.86-1.82)
Clashscore	79885	2149 (1.86-1.82)
Ramachandran outliers	78287	2124 (1.86-1.82)
Sidechain outliers	78261	2125 (1.86-1.82)
RSRZ outliers	66119	1857 (1.86-1.82)

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  $\geq 3$ , 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	228	
1	B	228	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	S36	A	301	-	X
2	S36	B	301	-	X

## 2 Entry composition i

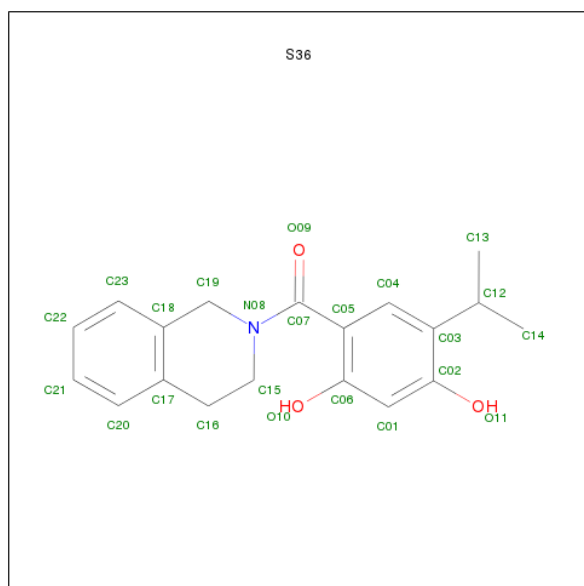
There are 3 unique types of molecules in this entry. The entry contains 3481 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Heat shock protein HSP 90-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	0	0	0
			1626	1033	268	320	5			
1	B	207	Total	C	N	O	S	0	0	0
			1626	1033	268	320	5			

- Molecule 2 is 3,4-DIHYDROISOQUINOLIN-2(1H)-YL[2,4-DIHYDROXY-5-(PROPAN-2-YL)PHENYL]METHANONE (three-letter code: S36) (formula: C<sub>19</sub>H<sub>21</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			23	19	1	3		
2	B	1	Total	C	N	O	0	0
			23	19	1	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	107	Total 107	O 107	0	0
3	B	76	Total 76	O 76	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.

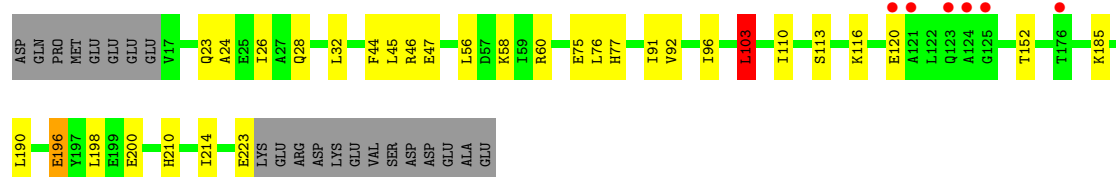
- Molecule 1: Heat shock protein HSP 90-alpha

Chain A:



- Molecule 1: Heat shock protein HSP 90-alpha

Chain B:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.32Å 88.83Å 98.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.56 – 1.84 40.47 – 1.85	Depositor EDS
% Data completeness (in resolution range)	91.8 (30.56-1.84) 90.6 (40.47-1.85)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.41 (at 1.84Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.2_432)	Depositor
R, $R_{free}$	0.223 , 0.260 0.215 , 0.242	Depositor DCC
$R_{free}$ test set	1820 reflections (4.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.2	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 41.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	3 of 48813 reflections (0.006%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3481	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 43.05 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.8662e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: S36

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	1.63	4/1652 (0.2%)	0.75	1/2228 (0.0%)
1	B	1.58	4/1652 (0.2%)	0.76	1/2228 (0.0%)
All	All	1.61	8/3304 (0.2%)	0.76	2/4456 (0.0%)

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	196	GLU	CB-CG	-8.83	1.35	1.52
1	A	196	GLU	CB-CG	-7.42	1.38	1.52
1	A	123	GLN	C-N	-5.83	1.20	1.34
1	A	38	TYR	CD1-CE1	-5.54	1.31	1.39
1	B	223	GLU	CD-OE2	-5.53	1.19	1.25
1	A	222	VAL	CB-CG2	-5.46	1.41	1.52
1	B	200	GLU	CD-OE1	-5.29	1.19	1.25
1	B	47	GLU	CD-OE2	-5.07	1.20	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	103	LEU	CB-CG-CD2	-6.58	99.81	111.00
1	A	32	LEU	CB-CG-CD2	-5.26	102.05	111.00

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	1626	0	1628	14	0
1	B	1626	0	1628	23	2
2	A	23	0	0	1	0
2	B	23	0	0	0	0
3	A	107	0	0	1	0
3	B	76	0	0	1	0
All	All	3481	0	3256	37	2

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 6.

All (37) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:24:ALA:O	1:B:28:GLN:HG3	1.59	1.03
1:B:103:LEU:HD21	1:B:152:THR:HG21	1.45	0.97
1:B:75:GLU:HG2	1:B:77:HIS:HD2	1.38	0.88
1:B:75:GLU:HG2	1:B:77:HIS:CD2	2.11	0.86
1:B:75:GLU:CG	1:B:77:HIS:HD2	1.95	0.79
1:A:196:GLU:HB2	3:A:490:HOH:O	1.85	0.76
1:B:196:GLU:HB2	3:B:446:HOH:O	1.87	0.74
1:B:75:GLU:CD	1:B:77:HIS:HD2	1.96	0.69
1:A:54:ASP:HB3	2:A:301:S36:C22	2.22	0.68
1:B:103:LEU:HD21	1:B:152:THR:CG2	2.23	0.64
1:B:26:ILE:N	1:B:26:ILE:HD12	2.14	0.62
1:B:75:GLU:CG	1:B:77:HIS:CD2	2.79	0.61
1:B:75:GLU:CD	1:B:77:HIS:CD2	2.75	0.60
1:B:26:ILE:H	1:B:26:ILE:HD12	1.69	0.56
1:B:23:GLN:HB2	1:B:26:ILE:HD13	1.90	0.52
1:B:58:LYS:HG3	1:B:96:ILE:HD13	1.92	0.52
1:B:92:VAL:HG22	1:B:185:LYS:HG3	1.93	0.50
1:A:75:GLU:HG2	1:A:77:HIS:CD2	2.47	0.49
1:A:43:ILE:C	1:A:43:ILE:HD12	2.34	0.48
1:A:113:SER:HB2	1:A:136:VAL:HG13	1.96	0.47
1:A:38:TYR:OH	1:A:129:SER:HB3	2.14	0.47
1:B:32:LEU:HD23	1:B:32:LEU:C	2.35	0.46
1:A:75:GLU:OE2	1:A:77:HIS:HD2	1.98	0.46
1:A:32:LEU:C	1:A:32:LEU:HD23	2.37	0.45
1:B:91:ILE:HD12	1:B:91:ILE:N	2.31	0.44

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:58:LYS:HG3	1:A:96:ILE:HD12	1.98	0.44
1:B:190:LEU:HD12	1:B:198:LEU:HD21	1.99	0.43
1:B:56:LEU:HD22	1:B:76:LEU:HB3	2.00	0.43
1:A:116:LYS:HB2	1:A:116:LYS:HE2	1.66	0.43
1:A:178:GLU:H	1:A:178:GLU:HG3	1.75	0.43
1:A:58:LYS:NZ	1:A:62:GLU:OE2	2.34	0.42
1:B:46:ARG:HD3	1:B:210:HIS:HB3	2.01	0.42
1:B:44:PHE:CG	1:B:45:LEU:N	2.89	0.41
1:A:26:ILE:HD12	1:A:26:ILE:HG23	1.86	0.40
1:B:75:GLU:OE2	1:B:77:HIS:CD2	2.74	0.40
1:A:177:GLY:O	1:A:178:GLU:C	2.59	0.40
1:B:214:ILE:HD13	1:B:214:ILE:HA	1.83	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:60:ARG:NH2	1:B:110:ILE:CD1[2_555]	1.54	0.66
1:B:60:ARG:CZ	1:B:113:SER:OG[2_555]	2.10	0.10

## 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	205/228 (90%)	196 (96%)	8 (4%)	1 (0%)	38	19
1	B	205/228 (90%)	199 (97%)	6 (3%)	0	100	100
All	All	410/456 (90%)	395 (96%)	14 (3%)	1 (0%)	56	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	178	GLU

### 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	178/198 (90%)	174 (98%)	4 (2%)	64	48
1	B	178/198 (90%)	175 (98%)	3 (2%)	73	59
All	All	356/396 (90%)	349 (98%)	7 (2%)	68	53

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	77	HIS
1	A	112	LYS
1	A	116	LYS
1	A	178	GLU
1	B	103	LEU
1	B	116	LYS
1	B	120	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	HIS
1	A	133	GLN
1	A	189	HIS
1	B	77	HIS
1	B	189	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	207/228 (90%)	-0.07	7 (3%) 43 38	17, 26, 48, 84	0
1	B	207/228 (90%)	0.06	6 (2%) 49 44	20, 29, 51, 88	0
All	All	414/456 (90%)	-0.01	13 (3%) 47 42	17, 28, 50, 88	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	GLY	5.2
1	B	123	GLN	5.2
1	B	124	ALA	4.2
1	A	111	ALA	3.5
1	B	176	THR	3.3
1	A	123	GLN	3.2
1	B	121	ALA	3.1
1	B	120	GLU	2.9
1	A	125	GLY	2.8
1	A	68	SER	2.2
1	A	65	THR	2.2
1	B	125	GLY	2.2
1	A	112	LYS	2.1

### 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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	S36	A	301	23/23	0.14	2.32	16,19,42,47	0
2	S36	B	301	23/23	0.18	2.24	19,25,60,68	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.