



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 07:19 AM GMT

PDB ID : 2HSQ  
Title : Human vinculin (head domain, Vh1, residues 1-258) in complex with Shigella's IpaA vinculin binding site 2 (residues 565-587)  
Authors : Izard, T.  
Deposited on : 2006-07-22  
Resolution : 3.97 Å(reported)

This is a wwPDB validation summary 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>

---

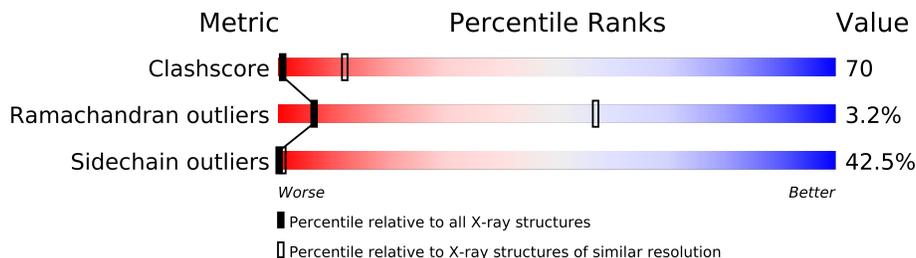
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.97 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	79885	1202 (4.46-3.50)
Ramachandran outliers	78287	1143 (4.46-3.50)
Sidechain outliers	78261	1130 (4.46-3.50)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	266	
2	B	23	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2239 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 Vinculin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	261	2065	1300	357	393	15	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	CLONING ARTIFACT	UNP P18206
A	-6	GLU	-	CLONING ARTIFACT	UNP P18206
A	-5	HIS	-	EXPRESSION TAG	UNP P18206
A	-4	HIS	-	EXPRESSION TAG	UNP P18206
A	-3	HIS	-	EXPRESSION TAG	UNP P18206
A	-2	HIS	-	EXPRESSION TAG	UNP P18206
A	-1	HIS	-	EXPRESSION TAG	UNP P18206
A	0	HIS	-	EXPRESSION TAG	UNP P18206
A	1	MET	-	INITIATING METHIONINE	UNP P18206

- Molecule 2 is a protein called Invasin ipaA.

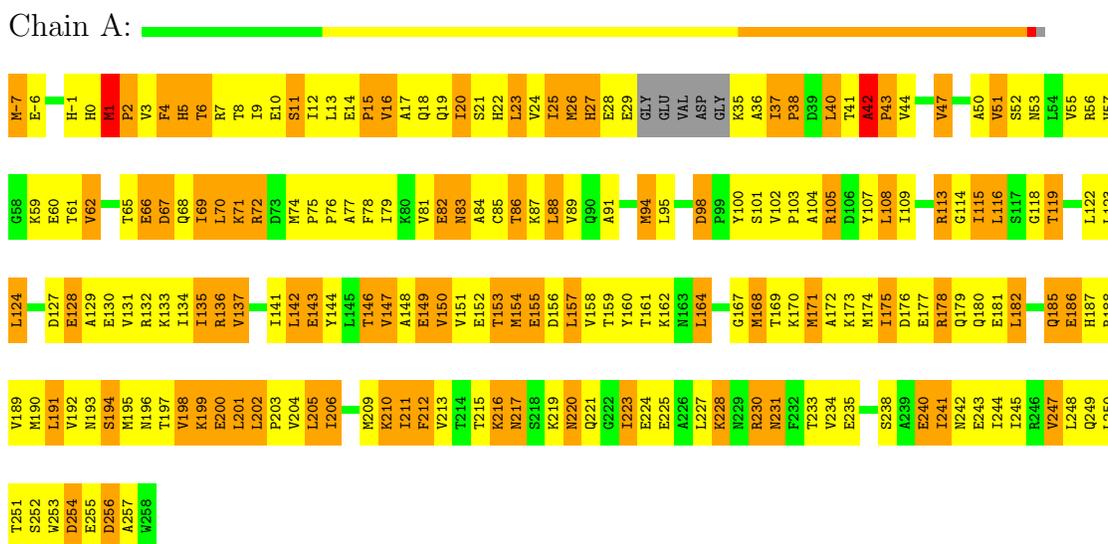
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	23	174	110	27	37	0	0	0

### 3 Residue-property plots

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.

Note EDS was not executed.

- Molecule 1: Vinculin



- Molecule 2: Invasin ipaA



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	I 4 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	203.17Å 203.17Å 203.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.00 – 3.97	Depositor
% Data completeness (in resolution range)	98.4 (58.00-3.97)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	BUSTER-TNT V. 1.3.2	Depositor
R, $R_{free}$	0.286 , 0.316	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2239	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	157.0	wwPDB-VP

## 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	0.22	0/2097	0.49	2/2839 (0.1%)
2	B	0.24	0/174	0.33	0/232
All	All	0.22	0/2271	0.48	2/3071 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	MET	C-N-CD	-14.76	88.14	120.60
1	A	42	ALA	C-N-CD	-6.09	107.19	120.60

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 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	2065	0	2105	302	0
2	B	174	0	184	30	0
All	All	2239	0	2289	317	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 70.

The worst 5 of 317 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:6:THR:HG23	1:A:9:ILE:HB	1.33	1.04
1:A:196:ASN:HA	1:A:199:LYS:HD2	1.36	1.04
1:A:13:LEU:HB3	1:A:116:LEU:HD11	1.45	0.98
1:A:224:GLU:HG2	1:A:228:LYS:HD3	1.45	0.98
1:A:25:ILE:HG22	1:A:26:MET:HE3	1.48	0.94

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 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	257/266 (97%)	211 (82%)	37 (14%)	9 (4%)	6	55
2	B	21/23 (91%)	16 (76%)	5 (24%)	0	100	100
All	All	278/289 (96%)	227 (82%)	42 (15%)	9 (3%)	6	57

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	PRO
1	A	220	ASN
1	A	67	ASP
1	A	43	PRO
1	A	149	GLU

### 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	234/237 (99%)	138 (59%)	96 (41%)	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	20/20 (100%)	8 (40%)	12 (60%)	0 0
All	All	254/257 (99%)	146 (58%)	108 (42%)	0 1

5 of 108 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	142	LEU
1	A	171	MET
2	B	573	VAL
1	A	146	THR
1	A	154	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	HIS
1	A	221	GLN
1	A	68	GLN
1	A	0	HIS
1	A	193	ASN

### 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

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers

EDS was not executed - this section will therefore be empty.