



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 03:30 pm GMT

PDB ID : 3CVR  
Title : Crystal structure of the full length IpaH3  
Authors : Zhu, Y.; Shao, F.  
Deposited on : 2008-04-19  
Resolution : 2.80 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

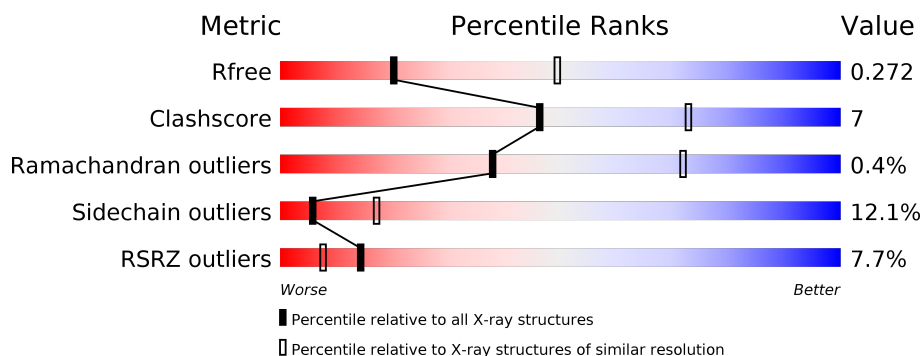
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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. 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  $\leq 5\%$ . 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	571	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3741 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 Invasion plasmid antigen.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	477	Total	C	N	O	S	Se	0	0	0
			3702	2341	632	718	5	6			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	344	ALA	THR	SEE REMARK 999	UNP Q83RJ4
A	500	LEU	GLN	SEE REMARK 999	UNP Q83RJ4
A	514	PRO	SER	SEE REMARK 999	UNP Q83RJ4
A	552	LEU	VAL	SEE REMARK 999	UNP Q83RJ4
A	562	PRO	SER	SEE REMARK 999	UNP Q83RJ4

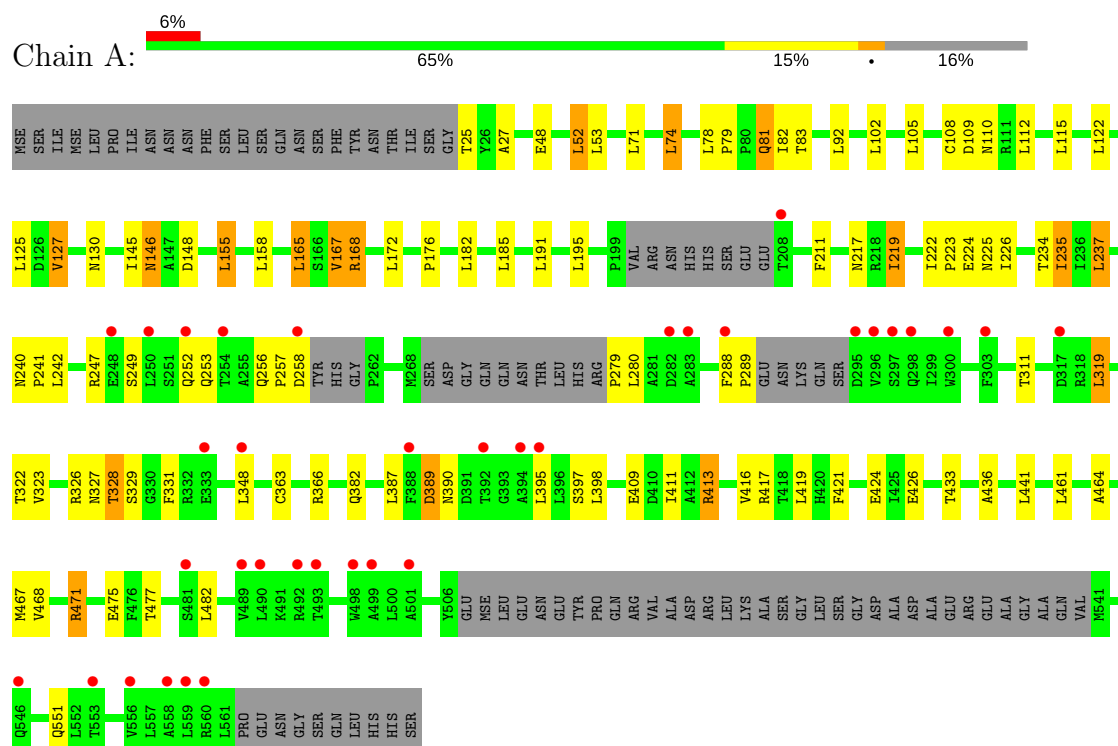
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	39	Total	O	0	0
			39	39		

### 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 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: Invasion plasmid antigen



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.19Å 154.19Å 85.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.76 – 2.80 48.76 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (48.76-2.80) 99.5 (48.76-2.80)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.76 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.251 , 0.277 0.248 , 0.272	Depositor DCC
$R_{free}$ test set	1318 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	71.3	Xtriage
Anisotropy	0.276	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 58.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3741	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.38	0/3767	0.58	1/5125 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	327	ASN	N-CA-C	5.38	125.53	111.00

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	3702	0	3516	52	0
2	A	39	0	0	0	0
All	All	3741	0	3516	52	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 7.

All (52) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance ( $\text{\AA}$ )	Clash overlap ( $\text{\AA}$ )
1:A:329:SER:HB3	1:A:331:PHE:H	1.15	1.09

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:288:PHE:HB3	1:A:289:PRO:HA	1.67	0.77
1:A:329:SER:HB3	1:A:331:PHE:N	1.98	0.76
1:A:319:LEU:O	1:A:322:THR:HG22	1.89	0.71
1:A:25:THR:HG22	1:A:27:ALA:H	1.58	0.69
1:A:411:ILE:HD11	1:A:467:MSE:HG2	1.75	0.68
1:A:366:ARG:CZ	1:A:413:ARG:HG3	2.24	0.66
1:A:224:GLU:HG3	1:A:419:LEU:HG	1.78	0.66
1:A:329:SER:CB	1:A:331:PHE:H	2.01	0.65
1:A:328:THR:HA	1:A:329:SER:C	2.23	0.59
1:A:110:ASN:HB2	1:A:130:ASN:HD21	1.68	0.58
1:A:146:ASN:ND2	1:A:148:ASP:H	2.03	0.56
1:A:81:GLN:NE2	1:A:81:GLN:H	2.06	0.54
1:A:155:LEU:HD13	1:A:172:LEU:HD21	1.89	0.53
1:A:471:ARG:NH1	1:A:475:GLU:OE1	2.42	0.53
1:A:112:LEU:H	1:A:130:ASN:HD22	1.57	0.52
1:A:222:ILE:HG13	1:A:421:PHE:HB3	1.93	0.51
1:A:235:ILE:HG22	1:A:237:LEU:HD13	1.91	0.51
1:A:219:ILE:H	1:A:240:ASN:HD22	1.57	0.51
1:A:79:PRO:HB2	1:A:82:ILE:HG23	1.92	0.51
1:A:222:ILE:HD12	1:A:226:ILE:HG21	1.94	0.50
1:A:48:GLU:O	1:A:52:LEU:HD22	2.12	0.49
1:A:115:LEU:HD11	1:A:127:VAL:HG21	1.95	0.48
1:A:319:LEU:HD11	1:A:331:PHE:HE2	1.78	0.48
1:A:366:ARG:NE	1:A:413:ARG:HG3	2.29	0.47
1:A:413:ARG:HG2	1:A:424:GLU:OE1	2.14	0.47
1:A:127:VAL:HG13	1:A:127:VAL:O	2.15	0.46
1:A:257:PRO:HA	1:A:258:ASP:HA	1.59	0.46
1:A:413:ARG:O	1:A:417:ARG:HB2	2.15	0.46
1:A:222:ILE:HG21	1:A:237:LEU:HD21	1.98	0.45
1:A:74:LEU:HD13	1:A:92:LEU:HD21	1.98	0.45
1:A:436:ALA:HA	1:A:441:LEU:HD12	1.99	0.44
1:A:92:LEU:H	1:A:110:ASN:HD22	1.64	0.44
1:A:279:PRO:HA	1:A:280:LEU:HA	1.69	0.44
1:A:146:ASN:C	1:A:146:ASN:HD22	2.21	0.43
1:A:288:PHE:HB3	1:A:289:PRO:CA	2.43	0.43
1:A:195:LEU:HD13	1:A:237:LEU:HD11	2.00	0.43
1:A:110:ASN:HB2	1:A:130:ASN:ND2	2.33	0.43
1:A:464:ALA:O	1:A:468:VAL:HG23	2.19	0.42
1:A:217:ASN:HB2	1:A:240:ASN:HD21	1.84	0.42
1:A:389:ASP:HA	1:A:390:ASN:HA	1.81	0.42
1:A:256:GLN:HA	1:A:257:PRO:HD3	1.93	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:249:SER:O	1:A:253:GLN:HG2	2.20	0.41
1:A:145:ILE:HB	1:A:165:LEU:HD23	2.03	0.41
1:A:223:PRO:HB2	1:A:225:ASN:OD1	2.21	0.41
1:A:158:LEU:HG	1:A:176:PRO:HG2	2.02	0.41
1:A:326:ARG:HH22	1:A:426:GLU:CD	2.24	0.41
1:A:382:GLN:HB3	1:A:387:LEU:HD22	2.02	0.41
1:A:240:ASN:C	1:A:242:LEU:H	2.25	0.41
1:A:148:ASP:CG	1:A:168:ARG:HG3	2.40	0.40
1:A:167:VAL:HG13	1:A:167:VAL:O	2.21	0.40
1:A:288:PHE:CB	1:A:289:PRO:HA	2.44	0.40

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	465/571 (81%)	434 (93%)	29 (6%)	2 (0%)	38 72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	328	THR
1	A	241	PRO

### 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	387/496 (78%)	340 (88%)	47 (12%)	6 17

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	52	LEU
1	A	53	LEU
1	A	71	LEU
1	A	74	LEU
1	A	78	LEU
1	A	81	GLN
1	A	83	THR
1	A	102	LEU
1	A	105	LEU
1	A	108	CYS
1	A	109	ASP
1	A	122	LEU
1	A	125	LEU
1	A	127	VAL
1	A	146	ASN
1	A	155	LEU
1	A	165	LEU
1	A	167	VAL
1	A	168	ARG
1	A	182	LEU
1	A	185	LEU
1	A	191	LEU
1	A	211	PHE
1	A	219	ILE
1	A	234	THR
1	A	235	ILE
1	A	237	LEU
1	A	247	ARG
1	A	252	GLN
1	A	311	THR
1	A	319	LEU
1	A	323	VAL
1	A	348	LEU
1	A	363	CYS
1	A	389	ASP
1	A	395	LEU
1	A	397	SER
1	A	398	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	409	GLU
1	A	413	ARG
1	A	416	VAL
1	A	433	THR
1	A	461	LEU
1	A	471	ARG
1	A	477	THR
1	A	482	LEU
1	A	551	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	81	GLN
1	A	110	ASN
1	A	130	ASN
1	A	146	ASN
1	A	169	ASN
1	A	240	ASN
1	A	327	ASN
1	A	372	ASN
1	A	373	ASN
1	A	474	ASN

### 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 carbohydrates 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

### 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	469/571 (82%)	0.53	36 (7%) <b>14</b> <b>8</b>	48, 64, 99, 103	6 (1%)

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	392	THR	5.8
1	A	499	ALA	4.9
1	A	559	LEU	4.5
1	A	296	VAL	4.0
1	A	295	ASP	4.0
1	A	493	THR	3.9
1	A	498	TRP	3.8
1	A	556	VAL	3.5
1	A	282	ASP	3.4
1	A	501	ALA	3.3
1	A	298	GLN	3.2
1	A	297	SER	3.2
1	A	300	TRP	3.0
1	A	490	LEU	3.0
1	A	481	SER	2.9
1	A	558	ALA	2.9
1	A	348	LEU	2.9
1	A	283	ALA	2.8
1	A	546	GLN	2.8
1	A	258	ASP	2.6
1	A	333	GLU	2.6
1	A	250	LEU	2.4
1	A	395	LEU	2.1
1	A	252	GLN	2.1
1	A	388	PHE	2.1
1	A	560	ARG	2.1
1	A	317	ASP	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	248	GLU	2.0
1	A	254	THR	2.0
1	A	553	THR	2.0
1	A	288	PHE	2.0
1	A	303	PHE	2.0
1	A	492	ARG	2.0
1	A	394	ALA	2.0
1	A	208	THR	2.0
1	A	489	VAL	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 carbohydrates 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.