



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 02:09 am GMT

PDB ID : 2WMP  
Title : Structure of the E. coli chaperone PapD in complex with the pilin domain of the PapGII adhesin  
Authors : Ford, B.A.; Verger, D.; Elam, J.S.; Dodson, K.W.; Pinkner, J.S.; Hultgren, S.J.  
Deposited on : 2009-07-02  
Resolution : 2.30 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

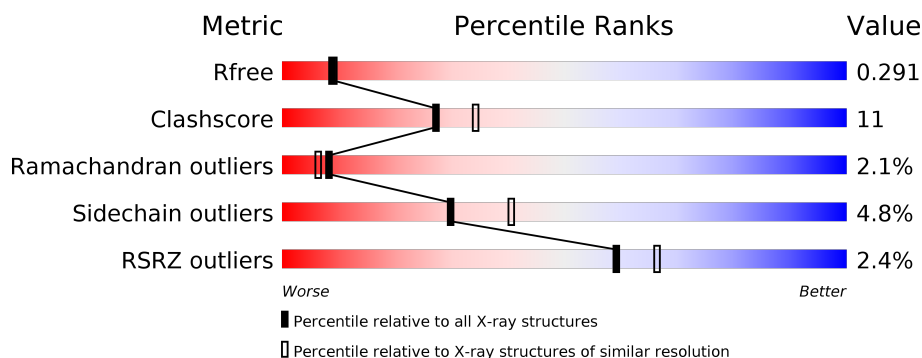
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (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  $\geq 3$ , 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	224	<div> <div>%</div> <div> <div></div> <div>77%</div> <div>18%</div> <div>• •</div> </div> </div>
2	B	123	<div> <div>5%</div> <div> <div></div> <div>72%</div> <div>25%</div> <div>• •</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2735 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 PAPD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	216	Total	C	N	O	S	0	0	1
			1671	1054	286	326	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	219	HIS	-	EXPRESSION TAG	UNP P15319
A	220	HIS	-	EXPRESSION TAG	UNP P15319
A	221	HIS	-	EXPRESSION TAG	UNP P15319
A	222	HIS	-	EXPRESSION TAG	UNP P15319
A	223	HIS	-	EXPRESSION TAG	UNP P15319
A	224	HIS	-	EXPRESSION TAG	UNP P15319

- Molecule 2 is a protein called PAPG PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	123	Total	C	N	O	S	0	0	0
			872	538	155	174	5			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	146	Total	O	0	0
			146	146		
3	B	46	Total	O	0	0
			46	46		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.56Å 90.47Å 64.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	200.00 – 2.30 64.49 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (200.00-2.30) 100.0 (64.49-2.30)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 2.29Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.227 , 0.289 0.227 , 0.291	Depositor DCC
$R_{free}$ test set	914 reflections (5.38%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.2	Xtriage
Anisotropy	0.381	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 39.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.027 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	2735	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.16% 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	0.35	0/1706	0.63	0/2325
2	B	0.33	0/889	0.62	0/1207
All	All	0.35	0/2595	0.62	0/3532

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	0

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	206	ILE	CB

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	1671	0	1635	30	0
2	B	872	0	809	30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	146	0	0	3	0
3	B	46	0	0	1	0
All	All	2735	0	2444	57	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 11.

All (57) 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)
2:B:208:HIS:CE1	2:B:293:LEU:HD21	2.13	0.84
2:B:252:LYS:HD3	2:B:264:ILE:HD12	1.62	0.80
2:B:219:HIS:HB3	2:B:293:LEU:HD23	1.62	0.78
2:B:293:LEU:HD22	2:B:293:LEU:N	2.04	0.72
1:A:37:ILE:HD13	1:A:50:VAL:HG12	1.75	0.67
2:B:219:HIS:HB3	2:B:293:LEU:CD2	2.27	0.64
1:A:32:LEU:HB2	1:A:93:ILE:HB	1.79	0.64
1:A:103:LEU:HD13	2:B:206:ILE:HD12	1.80	0.64
1:A:37:ILE:CD1	1:A:50:VAL:HG12	2.29	0.62
2:B:233:ALA:H	2:B:279:TRP:HB2	1.69	0.58
1:A:130:ASP:HA	1:A:214:VAL:HG21	1.88	0.56
1:A:19:THR:C	1:A:20:LEU:HD12	2.28	0.54
1:A:169:GLU:HG3	1:A:169:GLU:O	2.07	0.54
1:A:136:LYS:HE3	1:A:185:ASN:HD22	1.73	0.53
1:A:148:PRO:HD2	1:A:149:TYR:CE1	2.44	0.52
2:B:277:MET:HE2	2:B:285:GLN:HG2	1.91	0.52
2:B:219:HIS:CB	2:B:293:LEU:HD23	2.38	0.51
1:A:136:LYS:HD3	1:A:137:VAL:N	2.25	0.51
1:A:123:ARG:O	1:A:125:ASN:N	2.43	0.51
2:B:295:GLY:O	2:B:296:GLU:HB2	2.11	0.51
2:B:252:LYS:HD3	2:B:264:ILE:CD1	2.38	0.50
1:A:34:GLN:HG2	1:A:56:VAL:HG22	1.93	0.50
2:B:293:LEU:CD2	2:B:293:LEU:N	2.73	0.49
2:B:259:HIS:HB3	2:B:301:GLN:O	2.13	0.48
1:A:6:ARG:HA	2:B:195:GLY:H	1.79	0.48
2:B:217:ASN:N	2:B:217:ASN:HD22	2.11	0.48
1:A:176:ARG:HD2	3:A:2104:HOH:O	2.14	0.48
2:B:297:SER:OG	2:B:299:LYS:HD3	2.14	0.47
2:B:287:LEU:HD12	2:B:289:ILE:HD11	1.97	0.47
2:B:299:LYS:HD2	2:B:299:LYS:N	2.30	0.47
2:B:208:HIS:HB3	2:B:211:LEU:HD11	1.97	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:270:VAL:HG12	2:B:272:THR:H	1.79	0.46
2:B:277:MET:SD	2:B:287:LEU:HD23	2.56	0.46
1:A:8:ARG:NH2	2:B:316:PRO:OXT	2.48	0.46
2:B:301:GLN:N	2:B:302:PRO:CD	2.78	0.46
2:B:293:LEU:H	2:B:293:LEU:HD22	1.81	0.46
1:A:182:LYS:O	1:A:183:SER:HB3	2.16	0.46
1:A:37:ILE:HD12	1:A:51:ILE:HA	1.98	0.46
2:B:221:ALA:HA	3:B:2015:HOH:O	2.16	0.46
1:A:5:ASP:OD1	1:A:6:ARG:HD2	2.16	0.45
1:A:92:GLU:HG2	1:A:107:LEU:O	2.18	0.43
1:A:158:GLY:H	1:A:162:GLN:CD	2.22	0.43
1:A:37:ILE:HD13	1:A:50:VAL:CG1	2.47	0.42
3:A:2002:HOH:O	2:B:199:PRO:HG2	2.19	0.42
2:B:299:LYS:CD	2:B:299:LYS:N	2.83	0.42
1:A:173:LEU:CD2	1:A:173:LEU:N	2.82	0.41
2:B:209:GLY:O	2:B:211:LEU:HG	2.19	0.41
1:A:42:GLN:NE2	1:A:110:LYS:HZ2	2.18	0.41
1:A:155:GLY:HA3	1:A:163:ALA:O	2.20	0.41
1:A:42:GLN:NE2	1:A:110:LYS:NZ	2.68	0.41
1:A:19:THR:O	1:A:20:LEU:HD12	2.20	0.41
1:A:37:ILE:HD12	1:A:51:ILE:CA	2.51	0.41
1:A:191:LEU:C	1:A:191:LEU:HD23	2.42	0.41
1:A:195:ASN:HB2	3:A:2134:HOH:O	2.19	0.40
2:B:287:LEU:CD1	2:B:289:ILE:CD1	3.00	0.40
2:B:298:SER:C	2:B:299:LYS:HD2	2.40	0.40
1:A:29:LEU:C	1:A:61:PRO:HG3	2.42	0.40

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	214/224 (96%)	209 (98%)	4 (2%)	1 (0%)	32	39
2	B	121/123 (98%)	106 (88%)	9 (7%)	6 (5%)	2	1
All	All	335/347 (96%)	315 (94%)	13 (4%)	7 (2%)	8	6

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	300	ILE
1	A	124	PRO
2	B	294	TYR
2	B	296	GLU
2	B	302	PRO
2	B	303	GLY
2	B	304	VAL

### 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	183/200 (92%)	175 (96%)	8 (4%)	33	45
2	B	89/102 (87%)	84 (94%)	5 (6%)	25	33
All	All	272/302 (90%)	259 (95%)	13 (5%)	30	40

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	ARG
1	A	8	ARG
1	A	28	GLN
1	A	103	LEU
1	A	160	GLU
1	A	167	GLU
1	A	173	LEU
1	A	211	ARG
2	B	217	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	223	GLN
2	B	296	GLU
2	B	299	LYS
2	B	311	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	34	GLN
1	A	42	GLN
1	A	162	GLN
1	A	179	GLN
1	A	185	ASN
2	B	217	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 [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, 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	216/224 (96%)	-0.32	2 (0%) 84 87	9, 23, 43, 58	0
2	B	123/123 (100%)	0.20	6 (4%) 30 37	10, 32, 64, 72	0
All	All	339/347 (97%)	-0.13	8 (2%) 59 66	9, 26, 56, 72	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	297	SER	5.5
2	B	303	GLY	2.9
1	A	124	PRO	2.6
1	A	125	ASN	2.5
2	B	296	GLU	2.4
2	B	216	ALA	2.3
2	B	243	THR	2.2
2	B	302	PRO	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.