



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 03:48 AM EDT

PDB ID : 1M8W  
Title : CRYSTAL STRUCTURE OF THE PUMILIO-HOMOLOGY DOMAIN  
FROM HUMAN PUMILIO1 IN COMPLEX WITH NRE1-19 RNA  
Authors : Wang, X.; McLachlan, J.; Zamore, P.D.; Hall, T.M.T.  
Deposited on : 2002-07-26  
Resolution : 2.20 Å(reported)

This is a wwPDB X-ray Structure 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.20.1  
EDS : 2.36.2  
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) : 2.36.2

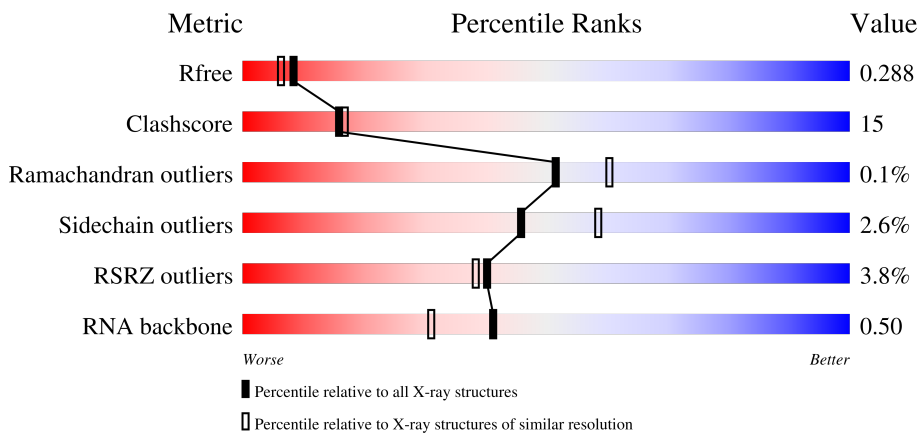
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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.



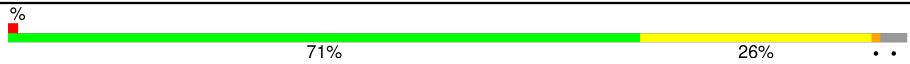
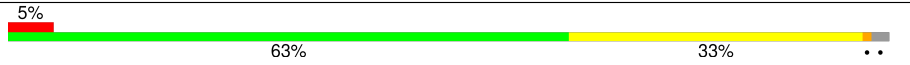
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)
RNA backbone	3102	1032 (2.60-1.80)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	C	8	
2	D	7	
3	E	8	
4	F	7	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	A	349	 <p>% 71% 26% ..</p>
5	B	349	 <p>5% 63% 33% ..</p>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6475 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 RNA chain called 5'-R(P\*UP\*UP\*GP\*UP\*AP\*UP\*AP\*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	8	167	75	25	59	8	0	8	0

- Molecule 2 is a RNA chain called 5'-R(P\*UP\*GP\*UP\*AP\*UP\*AP\*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	7	147	66	23	51	7	0	7	0

- Molecule 3 is a RNA chain called 5'-R(P\*UP\*UP\*GP\*UP\*CP\*CP\*AP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	8	168	75	27	58	8	0	8	0

- Molecule 4 is a RNA chain called 5'-R(P\*UP\*GP\*UP\*CP\*CP\*AP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	F	7	148	66	25	50	7	0	7	0

- Molecule 5 is a protein called Pumilio 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	A	340	2763	1749	500	497	17	0	0	0
5	B	341	2775	1758	501	499	17	0	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	5	Total O 5 5	0	0
6	E	6	Total O 6 6	0	0
6	F	6	Total O 6 6	0	0
6	A	205	Total O 205 205	0	0
6	B	85	Total O 85 85	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.

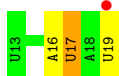
- Molecule 1: 5'-R(P\*UP\*UP\*GP\*UP\*AP\*UP\*AP\*U)-3'

Chain C: 



- Molecule 2: 5'-R(P\*UP\*GP\*UP\*AP\*UP\*AP\*U)-3'

Chain D: 



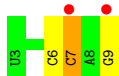
- Molecule 3: 5'-R(P\*UP\*UP\*GP\*UP\*CP\*CP\*AP\*G)-3'

Chain E: 



- Molecule 4: 5'-R(P\*UP\*GP\*UP\*CP\*CP\*AP\*G)-3'

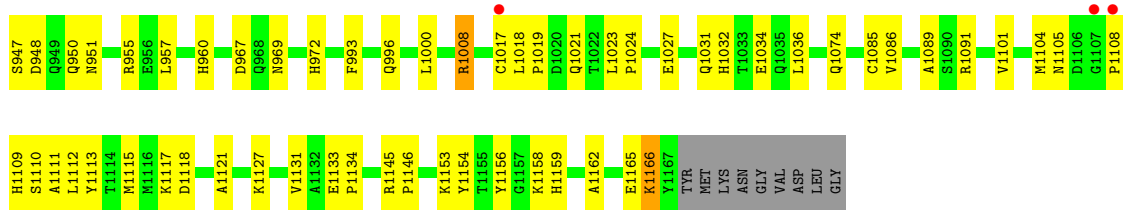
Chain F: 



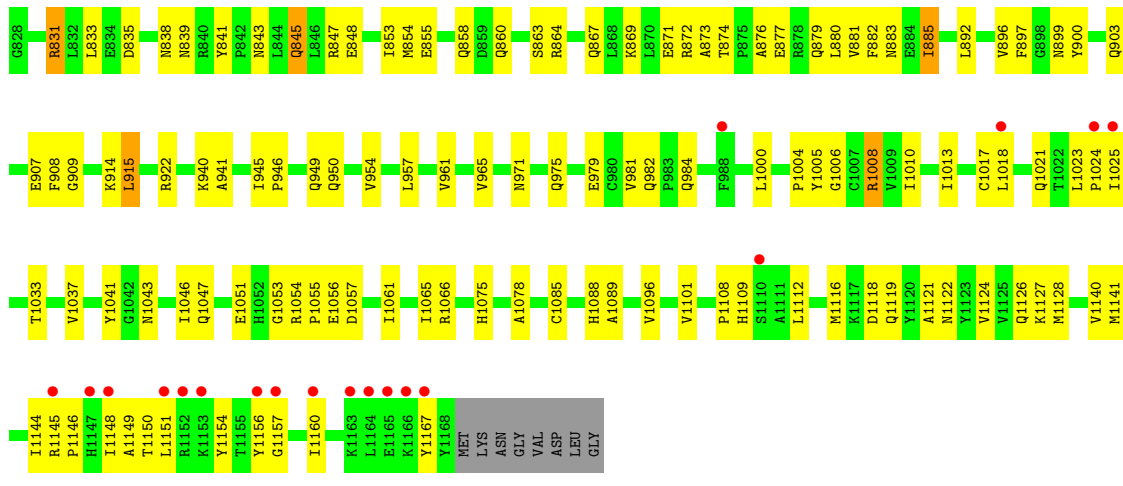
- Molecule 5: Pumilio 1

Chain A: 





• Molecule 5: Pumilio 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	262.73Å 37.73Å 82.52Å 90.00° 102.95° 90.00°	Depositor
Resolution (Å)	42.68 – 2.20 42.67 – 2.19	Depositor EDS
% Data completeness (in resolution range)	96.0 (42.68-2.20) 95.6 (42.67-2.19)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 2.20Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.224 , 0.289 0.224 , 0.288	Depositor DCC
$R_{free}$ test set	2447 reflections (5.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.9	Xtrriage
Anisotropy	0.519	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 48.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6475	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% 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	C	0.52	0/185	0.81	0/285
2	D	0.45	0/163	0.70	0/251
3	E	0.43	0/186	0.71	0/287
4	F	0.44	0/164	0.70	0/253
5	A	0.38	0/2816	0.55	0/3797
5	B	0.32	0/2829	0.52	0/3815
All	All	0.37	0/6343	0.56	0/8688

There are no bond length outliers.

There are no bond angle outliers.

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	C	167	0	74	4	0
2	D	147	0	66	3	0
3	E	168	0	73	7	0
4	F	148	0	66	3	0
5	A	2763	0	2781	71	0
5	B	2775	0	2790	93	0
6	A	205	0	0	9	0
6	B	85	0	0	3	0
6	C	5	0	0	0	0
6	E	6	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	F	6	0	0	0	0
All	All	6475	0	5850	176	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 15.

The worst 5 of 176 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)
5:A:1166:LYS:HD2	5:A:1166:LYS:H	1.23	0.98
5:A:879:GLN:HE21	5:A:883:ASN:HD21	1.22	0.88
5:A:1023:LEU:HB3	5:A:1024:PRO:HD3	1.60	0.84
5:A:1105:ASN:HA	5:A:1110:SER:HA	1.60	0.82
5:A:1101:VAL:HG21	5:A:1115:MET:SD	2.21	0.80

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	
5	A	338/349 (97%)	330 (98%)	8 (2%)	0	100	100
5	B	339/349 (97%)	325 (96%)	13 (4%)	1 (0%)	41	46
All	All	677/698 (97%)	655 (97%)	21 (3%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	B	1149	ALA

### 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	
5	A	303/310 (98%)	295 (97%)	8 (3%)	46	58
5	B	304/310 (98%)	296 (97%)	8 (3%)	46	58
All	All	607/620 (98%)	591 (97%)	16 (3%)	46	58

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	1008	ARG
5	B	957	LEU
5	B	831	ARG
5	B	915	LEU
5	A	1166	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	950	GLN
5	B	972	HIS
5	B	1119	GLN
5	B	968	GLN
5	B	984	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	7/8 (87%)	2 (28%)	0
2	D	6/7 (85%)	1 (16%)	0
3	E	7/8 (87%)	2 (28%)	0
4	F	6/7 (85%)	1 (16%)	0
All	All	26/30 (86%)	6 (23%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	C	13[A]	U
1	C	17[A]	U
2	D	17[A]	U
3	E	3[B]	U
3	E	7[B]	C

There are no RNA pucker outliers to report.

#### 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, 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	C	8/8 (100%)	0.05	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	32, 47, 70, 84	8 (100%)
2	D	7/7 (100%)	1.25	1 (14%) <span style="border: 1px solid red; padding: 2px;">2</span> <span style="border: 1px solid red; padding: 2px;">2</span>	51, 64, 78, 90	7 (100%)
3	E	8/8 (100%)	0.25	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	37, 53, 67, 77	8 (100%)
4	F	7/7 (100%)	1.93	2 (28%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	60, 66, 77, 82	7 (100%)
5	A	340/349 (97%)	-0.06	5 (1%) <span style="border: 1px solid blue; padding: 2px;">73</span> <span style="border: 1px solid blue; padding: 2px;">72</span>	21, 34, 68, 80	0
5	B	341/349 (97%)	0.24	19 (5%) <span style="border: 1px solid red; padding: 2px;">24</span> <span style="border: 1px solid red; padding: 2px;">23</span>	27, 48, 97, 105	0
All	All	711/728 (97%)	0.12	27 (3%) <span style="border: 1px solid red; padding: 2px;">40</span> <span style="border: 1px solid red; padding: 2px;">38</span>	21, 42, 77, 105	30 (4%)

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	F	9[B]	G	9.7
5	A	1107	GLY	6.8
5	B	1152	ARG	6.4
2	D	19[A]	U	5.6
5	B	1160	ILE	5.2

### 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.