



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 19, 2020 – 08:05 pm BST

PDB ID : 2PMU  
Title : Crystal structure of the DNA-binding domain of PhoP  
Authors : Wang, S.  
Deposited on : 2007-04-23  
Resolution : 1.78 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

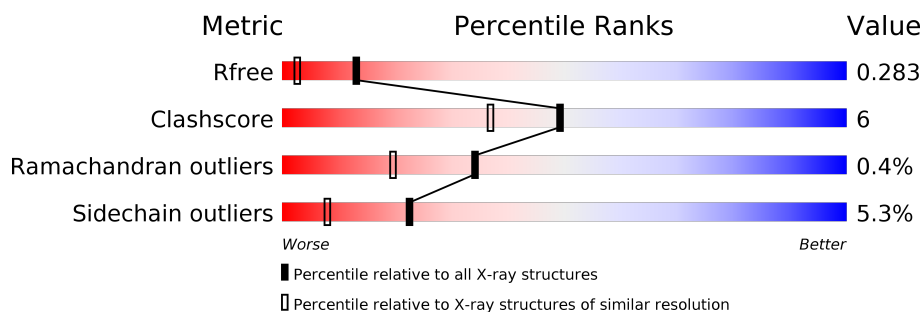
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.78 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	110	
1	B	110	
1	C	110	
1	D	110	
1	E	110	
1	F	110	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5358 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 response regulator PHOP.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	0	1	0
			805	520	138	147			
1	B	103	Total	C	N	O	0	1	0
			849	546	148	155			
1	C	92	Total	C	N	O	0	1	0
			760	491	133	136			
1	D	96	Total	C	N	O	0	1	0
			793	514	136	143			
1	E	103	Total	C	N	O	0	1	0
			849	546	148	155			
1	F	93	Total	C	N	O	0	1	0
			763	492	131	140			

There are 36 discrepancies between the modelled and reference sequences:

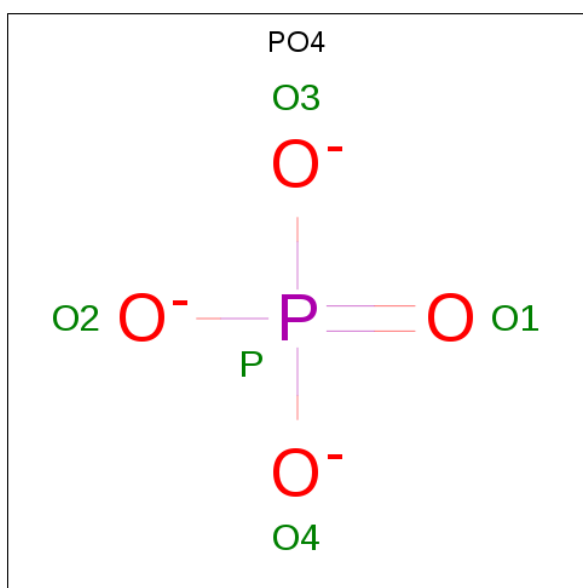
Chain	Residue	Modelled	Actual	Comment	Reference
A	140	GLY	-	EXPRESSION TAG	UNP P71814
A	141	THR	-	EXPRESSION TAG	UNP P71814
A	142	HIS	-	EXPRESSION TAG	UNP P71814
A	143	MET	-	EXPRESSION TAG	UNP P71814
A	248	LYS	-	EXPRESSION TAG	UNP P71814
A	249	LEU	-	EXPRESSION TAG	UNP P71814
B	140	GLY	-	EXPRESSION TAG	UNP P71814
B	141	THR	-	EXPRESSION TAG	UNP P71814
B	142	HIS	-	EXPRESSION TAG	UNP P71814
B	143	MET	-	EXPRESSION TAG	UNP P71814
B	248	LYS	-	EXPRESSION TAG	UNP P71814
B	249	LEU	-	EXPRESSION TAG	UNP P71814
C	140	GLY	-	EXPRESSION TAG	UNP P71814
C	141	THR	-	EXPRESSION TAG	UNP P71814
C	142	HIS	-	EXPRESSION TAG	UNP P71814
C	143	MET	-	EXPRESSION TAG	UNP P71814
C	248	LYS	-	EXPRESSION TAG	UNP P71814

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	249	LEU	-	EXPRESSION TAG	UNP P71814
D	140	GLY	-	EXPRESSION TAG	UNP P71814
D	141	THR	-	EXPRESSION TAG	UNP P71814
D	142	HIS	-	EXPRESSION TAG	UNP P71814
D	143	MET	-	EXPRESSION TAG	UNP P71814
D	248	LYS	-	EXPRESSION TAG	UNP P71814
D	249	LEU	-	EXPRESSION TAG	UNP P71814
E	140	GLY	-	EXPRESSION TAG	UNP P71814
E	141	THR	-	EXPRESSION TAG	UNP P71814
E	142	HIS	-	EXPRESSION TAG	UNP P71814
E	143	MET	-	EXPRESSION TAG	UNP P71814
E	248	LYS	-	EXPRESSION TAG	UNP P71814
E	249	LEU	-	EXPRESSION TAG	UNP P71814
F	140	GLY	-	EXPRESSION TAG	UNP P71814
F	141	THR	-	EXPRESSION TAG	UNP P71814
F	142	HIS	-	EXPRESSION TAG	UNP P71814
F	143	MET	-	EXPRESSION TAG	UNP P71814
F	248	LYS	-	EXPRESSION TAG	UNP P71814
F	249	LEU	-	EXPRESSION TAG	UNP P71814

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		
2	D	1	Total	O	P	0	1
			9	7	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	O	P	0	0
			5	4	1		
2	E	1	Total	O	P	0	0
			5	4	1		

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	K	0	0
			1	1		
3	D	1	Total	K	0	0
			1	1		
3	E	1	Total	K	0	0
			1	1		

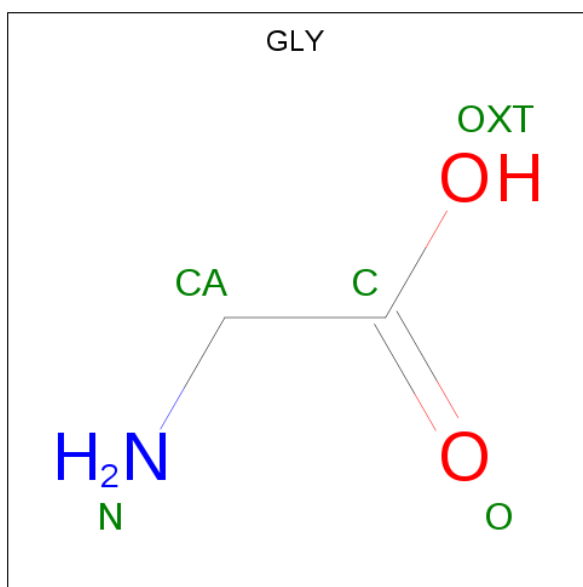
- Molecule 4 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	X	0	0
			1	1		
4	A	1	Total	X	0	0
			1	1		
4	D	1	Total	X	0	0
			1	1		
4	E	1	Total	X	0	0
			1	1		

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Cl	0	0
			1	1		
5	D	1	Total	Cl	0	0
			1	1		
5	E	2	Total	Cl	0	0
			2	2		

- Molecule 6 is GLYCINE (three-letter code: GLY) (formula: C<sub>2</sub>H<sub>5</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	O	0	0
			5	2	1	2		
6	B	1	Total	C	N	O	0	0
			5	2	1	2		
6	B	1	Total	C	N	O	0	0
			5	2	1	2		


- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	108	Total	O	0	0
			108	108		
7	B	72	Total	O	0	0
			72	72		
7	C	65	Total	O	0	0
			65	65		
7	D	107	Total	O	0	0
			107	107		
7	E	70	Total	O	0	0
			70	70		
7	F	67	Total	O	0	0
			67	67		

### 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. 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. 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: response regulator PHOP

Chain A: 



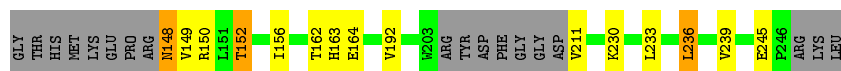
- Molecule 1: response regulator PHOP

Chain B: 




- Molecule 1: response regulator PHOP

Chain C: 




- Molecule 1: response regulator PHOP

Chain D: 



- Molecule 1: response regulator PHOP

Chain E: 



- Molecule 1: response regulator PHOP

Chain F: 

GLY	THR	HIS	MET	LYS	GLU	PRO	ARG	N148	V149	R150	L151	T152	I156	T162	H163	E164	W203	ARG	TYR	ASP	PHE	GLY	GLY	D210	V213	L233	L236	R244	E245	P246	ARG	LYS	LEU
-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	-----	-----	-----



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.88Å 101.13Å 86.93Å 90.00° 126.72° 90.00°	Depositor
Resolution (Å)	20.00 – 1.78 19.56 – 1.78	Depositor EDS
% Data completeness (in resolution range)	94.6 (20.00-1.78) 94.5 (19.56-1.78)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.196 , 0.238 0.253 , 0.283	Depositor DCC
$R_{free}$ test set	3310 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.1	Xtriage
Anisotropy	0.170	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 57.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.025 for -k,-h,-1/2*h+1/2*k-l 0.027 for k,h,-1/2*h-1/2*k-l 0.480 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5358	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: UNX, K, PO4, CL

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	0.82	0/827	0.81	0/1123
1	B	0.72	0/872	0.84	2/1183 (0.2%)
1	C	0.66	0/779	0.71	0/1057
1	D	0.81	0/814	0.78	0/1104
1	E	0.75	0/872	0.86	2/1183 (0.2%)
1	F	0.68	0/782	0.69	0/1062
All	All	0.74	0/4946	0.79	4/6712 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	E	223	ARG	NE-CZ-NH2	8.58	124.59	120.30
1	B	223	ARG	NE-CZ-NH1	8.52	124.56	120.30
1	B	223	ARG	NE-CZ-NH2	-8.22	116.19	120.30
1	E	223	ARG	NE-CZ-NH1	-7.15	116.72	120.30

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	805	0	809	12	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	849	0	854	12	0
1	C	760	0	777	13	0
1	D	793	0	801	4	0
1	E	849	0	854	10	0
1	F	763	0	773	8	0
2	A	5	0	0	0	0
2	D	14	0	0	0	0
2	E	5	0	0	0	0
3	A	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
5	B	1	0	0	1	0
5	D	1	0	0	1	0
5	E	2	0	0	1	0
6	B	15	0	6	0	0
7	A	108	0	0	0	0
7	B	72	0	0	0	0
7	C	65	0	0	4	0
7	D	107	0	0	0	0
7	E	70	0	0	0	0
7	F	67	0	0	1	0
All	All	5358	0	4874	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 6.

The worst 5 of 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)
1:F:210:ASP:O	1:F:213:VAL:HG12	1.76	0.84
1:D:237:ARG:NE	1:D:237:ARG:H	1.80	0.79
1:A:150:ARG:CG	1:A:150:ARG:HH11	1.94	0.79
1:B:204:ARG:HD3	7:C:306:HOH:O	1.83	0.78
1:C:149:VAL:HG13	1:C:150:ARG:H	1.52	0.74

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	97/110 (88%)	93 (96%)	4 (4%)	0	100	100
1	B	102/110 (93%)	93 (91%)	8 (8%)	1 (1%)	15	4
1	C	89/110 (81%)	86 (97%)	3 (3%)	0	100	100
1	D	93/110 (84%)	90 (97%)	3 (3%)	0	100	100
1	E	102/110 (93%)	95 (93%)	6 (6%)	1 (1%)	15	4
1	F	90/110 (82%)	89 (99%)	1 (1%)	0	100	100
All	All	573/660 (87%)	546 (95%)	25 (4%)	2 (0%)	34	25

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	145	GLU
1	E	145	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	89/99 (90%)	85 (96%)	4 (4%)	27	11
1	B	94/99 (95%)	91 (97%)	3 (3%)	39	22
1	C	85/99 (86%)	77 (91%)	8 (9%)	8	2
1	D	88/99 (89%)	86 (98%)	2 (2%)	50	34
1	E	94/99 (95%)	90 (96%)	4 (4%)	29	12

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	86/99 (87%)	79 (92%)	7 (8%)	11	3
All	All	536/594 (90%)	508 (95%)	28 (5%)	22	8

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	233	LEU
1	D	233	LEU
1	F	236	LEU
1	C	236	LEU
1	C	245	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	148	ASN
1	F	148	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

Of 19 ligands modelled in this entry, 4 are unknown and 7 are monoatomic - leaving 8 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	PO4	E	1003	-	4,4,4	0.89	0	6,6,6	0.65	0
2	PO4	A	1004	-	4,4,4	0.44	0	6,6,6	1.65	2 (33%)
2	PO4	D	1002	-	4,4,4	1.12	0	6,6,6	1.19	0
2	PO4	D	1001[A]	-	4,4,4	1.18	0	6,6,6	0.56	0
2	PO4	D	1001[B]	-	4,4,4	1.00	0	6,6,6	0.71	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	1004	PO4	O4-P-O2	3.24	118.37	107.97
2	A	1004	PO4	O4-P-O1	-2.20	102.84	110.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.