



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 11:39 pm BST

PDB ID : 1GXP  
Title : PhoB effector domain in complex with pho box DNA.  
Authors : Blanco, A.G.; Sola, M.; Gomis-Ruth, F.X.; Coll, M.  
Deposited on : 2002-04-08  
Resolution : 2.50 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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

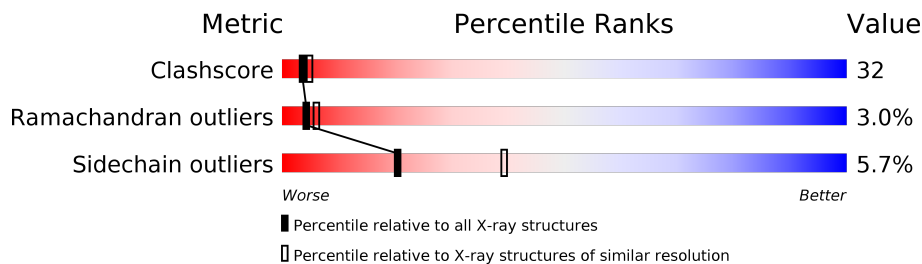
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	106	
1	B	106	
1	E	106	
1	F	106	
2	C	23	
2	G	23	
3	D	23	
3	H	23	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 5416 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 PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	103	Total 846	C 531	N 156	O 154	S 5	19	0	0
1	B	101	Total 834	C 523	N 154	O 152	S 5	20	0	0
1	E	102	Total 841	C 528	N 155	O 153	S 5	16	0	0
1	F	102	Total 841	C 528	N 155	O 153	S 5	22	0	0

- Molecule 2 is a DNA chain called 5'-D(\*GP\*AP\*GP\*CP\*TP\*GP\*TP\*CP\*AP\*TP\* AP\*AP\*AP\*GP\*TP\*TP\*GP\*TP\*CP\*AP\*CP\*GP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	23	Total 473	C 226	N 89	O 136	P 22	20	0	0
2	G	23	Total 473	C 226	N 89	O 136	P 22	20	0	0

- Molecule 3 is a DNA chain called 5'-D(\*CP\*CP\*CP\*GP\*TP\*GP\*AP\*CP\*AP\*AP\* CP\*TP\*TP\*TP\*AP\*TP\*GP\*AP\*CP\*AP\*GP\*CP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	23	Total 464	C 223	N 83	O 136	P 22	5	0	0
3	H	23	Total 464	C 223	N 83	O 136	P 22	8	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	38	Total 38	O 38	0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	B	31	Total O 31 31	0	0
4	C	8	Total O 8 8	0	0
4	D	17	Total O 17 17	0	0
4	E	41	Total O 41 41	0	0
4	F	30	Total O 30 30	0	0
4	G	6	Total O 6 6	0	0
4	H	9	Total O 9 9	0	0

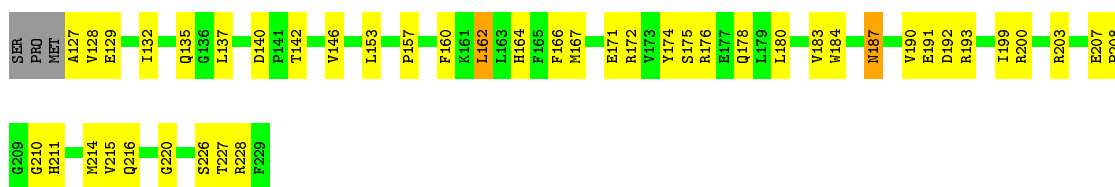
### 3 Residue-property plots [i](#)

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.

Note EDS was not executed.

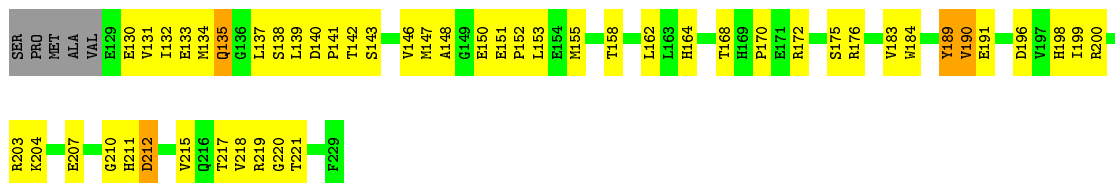
- Molecule 1: PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN

Chain A: 



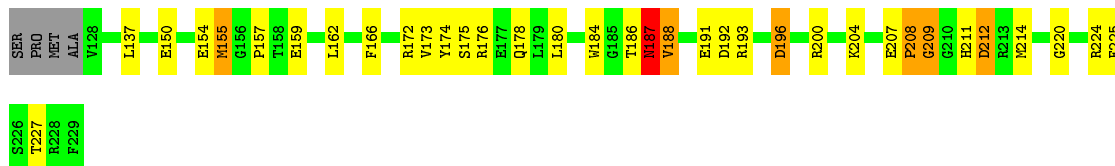
- Molecule 1: PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN

Chain B: 



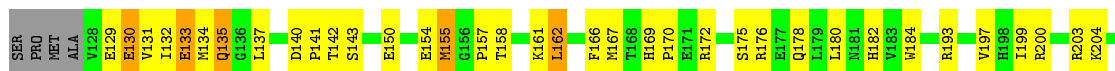
- Molecule 1: PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN

Chain E: 



- Molecule 1: PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN

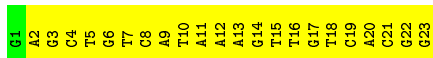
Chain F: 





- Molecule 2: 5'-D(\*GP\*AP\*GP\*CP\*TP\*GP\*TP\*CP\*AP\*TP\* AP\*AP\*AP\*GP\*TP\*TP\*GP\*TP\*CP\*AP\*CP\*GP\*G)-3'

Chain C: 96%



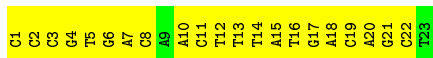
- Molecule 2: 5'-D(\*GP\*AP\*GP\*CP\*TP\*GP\*TP\*CP\*AP\*TP\* AP\*AP\*AP\*GP\*TP\*TP\*GP\*TP\*CP\*AP\*CP\*GP\*G)-3'

Chain G: 87%



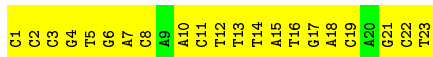
- Molecule 3: 5'-D(\*CP\*CP\*CP\*GP\*TP\*GP\*AP\*CP\*AP\*AP\* CP\*TP\*TP\*TP\*AP\*TP\*GP\*AP\*CP\*AP\*GP\*CP\*T)-3'

Chain D: 91%



- Molecule 3: 5'-D(\*CP\*CP\*CP\*GP\*TP\*GP\*AP\*CP\*AP\*AP\* CP\*TP\*TP\*TP\*AP\*TP\*GP\*AP\*CP\*AP\*GP\*CP\*T)-3'

Chain H: 91%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.11Å 74.11Å 289.68Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 2.50	Depositor
% Data completeness (in resolution range)	100.0 (40.00-2.50)	Depositor
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.244 , 0.289	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	5416	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	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	0.43	0/866	0.69	0/1168
1	B	0.45	0/854	0.69	0/1151
1	E	0.44	0/861	0.71	0/1161
1	F	0.43	0/861	0.74	1/1161 (0.1%)
2	C	0.41	0/531	0.77	0/819
2	G	0.46	0/531	0.73	0/819
3	D	0.37	0/519	0.71	0/798
3	H	0.45	0/519	0.75	0/798
All	All	0.43	0/5542	0.72	1/7875 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	130	GLU	N-CA-C	6.24	127.85	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	846	0	825	42	0
1	B	834	0	811	46	0
1	E	841	0	820	35	0
1	F	841	0	820	49	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	473	0	261	35	0
2	G	473	0	261	33	0
3	D	464	0	261	38	0
3	H	464	0	261	37	0
4	A	38	0	0	3	0
4	B	31	0	0	4	0
4	C	8	0	0	1	0
4	D	17	0	0	5	0
4	E	41	0	0	4	0
4	F	30	0	0	3	0
4	G	6	0	0	0	0
4	H	9	0	0	2	0
All	All	5416	0	4320	294	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 32.

The worst 5 of 294 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)
3:H:18:DA:H1'	3:H:19:DC:H5'	1.44	0.98
1:F:175:SER:H	1:F:178:GLN:NE2	1.61	0.97
2:C:21:DC:H1'	2:C:22:DG:H5'	1.45	0.94
1:F:175:SER:H	1:F:178:GLN:HE21	0.90	0.90
2:G:22:DG:H2''	2:G:23:DG:H5''	1.52	0.89

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	101/106 (95%)	89 (88%)	10 (10%)	2 (2%)	7	12
1	B	99/106 (93%)	90 (91%)	8 (8%)	1 (1%)	15	28
1	E	100/106 (94%)	86 (86%)	9 (9%)	5 (5%)	2	2
1	F	100/106 (94%)	87 (87%)	9 (9%)	4 (4%)	3	3
All	All	400/424 (94%)	352 (88%)	36 (9%)	12 (3%)	4	6

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	130	GLU
1	A	192	ASP
1	B	190	VAL
1	E	187	ASN
1	E	208	PRO

### 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	92/95 (97%)	89 (97%)	3 (3%)	38	64
1	B	91/95 (96%)	85 (93%)	6 (7%)	16	32
1	E	92/95 (97%)	86 (94%)	6 (6%)	17	33
1	F	92/95 (97%)	86 (94%)	6 (6%)	17	33
All	All	367/380 (97%)	346 (94%)	21 (6%)	20	39

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	159	GLU
1	E	187	ASN
1	F	162	LEU
1	B	212	ASP
1	F	172	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	178	GLN
1	E	164	HIS
1	F	135	GLN
1	B	164	HIS
1	E	178	GLN

### 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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.