



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 02:06 pm GMT

PDB ID : 4R6L  
Title : Crystal structure of bacteriophytochrome RpBphP2 from photosynthetic bacterium *R. palustris*  
Authors : Yang, X.; Stojkovic, E.; Ozarowski, W.; Kuk, J.; Davydova, E.; Moffat, K.  
Deposited on : 2014-08-25  
Resolution : 3.40 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	:	1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix)	:	1.13
EDS	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

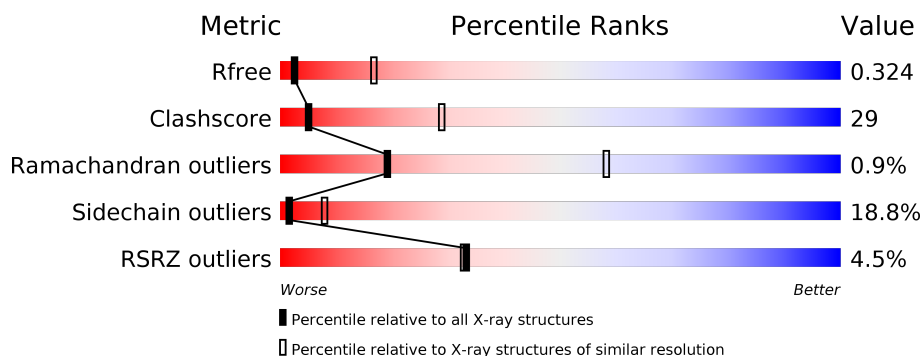
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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))
$R_{free}$	111664	1928 (3.50-3.30)
Clashscore	122126	2051 (3.50-3.30)
Ramachandran outliers	120053	2006 (3.50-3.30)
Sidechain outliers	120020	2006 (3.50-3.30)
RSRZ outliers	108989	1827 (3.50-3.30)

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	519	
1	B	519	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7198 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 Bacteriophytochrome (Light-regulated signal transduction histidine kinase), PhyB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	473	Total	C	N	O	S	0	0	0
			3639	2295	659	671	14			
1	B	448	Total	C	N	O	S	0	0	0
			3473	2191	627	641	14			

There are 26 discrepancies between the modelled and reference sequences:

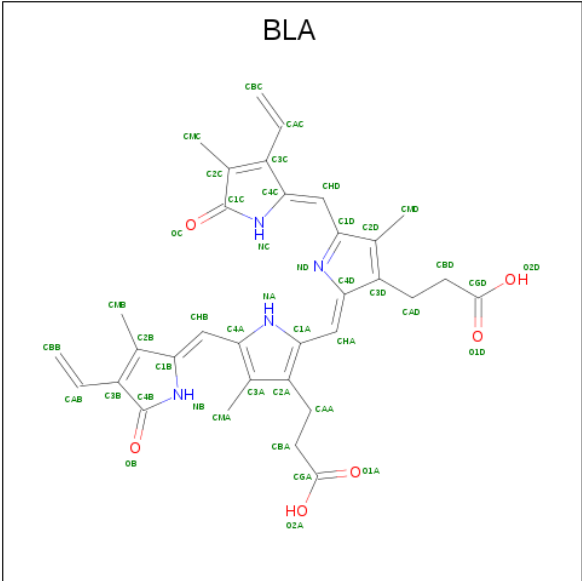
Chain	Residue	Modelled	Actual	Comment	Reference
A	507	LYS	-	EXPRESSION TAG	UNP Q6N5G3
A	508	LEU	-	EXPRESSION TAG	UNP Q6N5G3
A	509	ALA	-	EXPRESSION TAG	UNP Q6N5G3
A	510	ALA	-	EXPRESSION TAG	UNP Q6N5G3
A	511	ALA	-	EXPRESSION TAG	UNP Q6N5G3
A	512	LEU	-	EXPRESSION TAG	UNP Q6N5G3
A	513	GLU	-	EXPRESSION TAG	UNP Q6N5G3
A	514	HIS	-	EXPRESSION TAG	UNP Q6N5G3
A	515	HIS	-	EXPRESSION TAG	UNP Q6N5G3
A	516	HIS	-	EXPRESSION TAG	UNP Q6N5G3
A	517	HIS	-	EXPRESSION TAG	UNP Q6N5G3
A	518	HIS	-	EXPRESSION TAG	UNP Q6N5G3
A	519	HIS	-	EXPRESSION TAG	UNP Q6N5G3
B	507	LYS	-	EXPRESSION TAG	UNP Q6N5G3
B	508	LEU	-	EXPRESSION TAG	UNP Q6N5G3
B	509	ALA	-	EXPRESSION TAG	UNP Q6N5G3
B	510	ALA	-	EXPRESSION TAG	UNP Q6N5G3
B	511	ALA	-	EXPRESSION TAG	UNP Q6N5G3
B	512	LEU	-	EXPRESSION TAG	UNP Q6N5G3
B	513	GLU	-	EXPRESSION TAG	UNP Q6N5G3
B	514	HIS	-	EXPRESSION TAG	UNP Q6N5G3
B	515	HIS	-	EXPRESSION TAG	UNP Q6N5G3
B	516	HIS	-	EXPRESSION TAG	UNP Q6N5G3
B	517	HIS	-	EXPRESSION TAG	UNP Q6N5G3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	518	HIS	-	EXPRESSION TAG	UNP Q6N5G3
B	519	HIS	-	EXPRESSION TAG	UNP Q6N5G3

- Molecule 2 is BILIVERDINE IX ALPHA (three-letter code: BLA) (formula: C<sub>33</sub>H<sub>34</sub>N<sub>4</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			43	33	4	6		
2	B	1	Total	C	N	O	0	0
			43	33	4	6		



V488	V424	E352	L281	R208	ALA
A489	P425	A353		R209	F131
A490	L426	L354	H285	L210	F132
A491	S427	L355	R286	Y211	R133
E492	R428	E356	R287		R134
I493	T429	L357	K288	N214	T135
I494	P430	M358		P215	N136
R495	P431		Y291	V216	S137
D496	R432	S361		R217	A138
L497	R433	G362	V292	I218	I139
I498	V434	L363	D293	I219	R140
I499	M435	C364	L294	P220	R141
D500	M436	L365	D295	D221	L142
V501	M437	H366	G296	I222	Q143
I502	F438	R367	R297	M223	A144
LEU	R439	R368	Q298	Y224	A145
ARG	P440	E369		R225	E146
ASN	E441	G370	A308	P226	
THR	V442	V371	W309	V227	A151
LYS	M443		Q310		C152
LEU	Q444			P231	
ALA		Q376		D232	V159
ALA	A449	T377	V313	L233	R160
ALA	M450	P378	M314	N234	E161
LEU	M451	P379	E315	P235	I162
GLU	P452	I382	E316	V236	T163
HIS	D453	I383	A317	T237	G164
HIS	K454	D384	A318	G238	F165
HIS	S455	Q385	I319	R239	D166
HIS	VAL	Q388	T320	P240	R167
THR	THR	L389	R321	D241	
ALA	ALA	A390	Q322	I242	I170
GLU	GLU	G391	T323	L243	Y171
SER	SER	R392	L324	S244	R172
GLY	GLY		K325	F245	F173
ARG	ARG	S396	G326	A246	
LEU	LEU	E397	Q327	I247	D176
ARG	ARG	L398	I329		F177
PRO	PRO	F399	Q330	V254	S178
R466	R467	Q400	R331	H255	G179
S468	S468	T401	S332	L256	E180
P469	P469	D402	L333	E257	
		R403	I334	V258	C187
		L404	N335	R259	V190
		S405	D336	R260	E191
		T406	I337	N261	S192
		I407	E338	I262	Y193
		P409	Q339	G263	
		E410	L340	M264	
		A415	H341		L196
			D342	T267	H197
			H343	M268	F198
			R344	S269	P199
			A345	I270	
		A418	G346		D202
		S419	L347	R274	I203
		G420	A348	G275	P204
		V421	R349	E276	A205
		L422	N350	R277	Q206
		E423	L422	L278	A207

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.40 Å   174.40 Å   95.67 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	49.02 – 3.40 49.02 – 3.40	Depositor EDS
% Data completeness (in resolution range)	95.3 (49.02-3.40) 95.5 (49.02-3.40)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.28 (at 3.40 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, $R_{free}$	0.268 , 0.322 0.273 , 0.324	Depositor DCC
$R_{free}$ test set	1121 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	114.4	Xtriage
Anisotropy	0.435	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 91.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.105 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7198	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	156.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BLA

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	0.40	0/3716	0.83	3/5057 (0.1%)
1	B	0.43	0/3546	0.87	6/4825 (0.1%)
All	All	0.41	0/7262	0.85	9/9882 (0.1%)

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	0	1
1	B	0	3
All	All	0	4

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	355	LEU	CA-CB-CG	6.93	131.24	115.30
1	B	333	LEU	CB-CG-CD1	6.18	121.51	111.00
1	A	89	ALA	C-N-CD	-5.63	108.21	120.60
1	B	354	LEU	CA-CB-CG	5.58	128.13	115.30
1	A	273	LEU	CA-CB-CG	5.28	127.44	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	341	HIS	Peptide

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	B	38	ASP	Peptide
1	B	426	LEU	Peptide
1	B	68	PHE	Peptide

## 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	3639	0	3625	197	0
1	B	3473	0	3453	220	0
2	A	43	0	31	5	0
2	B	43	0	31	7	0
All	All	7198	0	7140	412	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 29.

The worst 5 of 412 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:B:202:ASP:OD1	1:B:466:ARG:NH2	1.95	1.00
1:B:274:ARG:NH1	1:B:308:ALA:O	1.99	0.94
1:A:75:ARG:NH2	1:A:92:ALA:O	2.01	0.94
1:B:330:GLN:NE2	1:B:493:GLU:OE2	2.01	0.93
1:B:451:ASN:OD1	1:B:453:ASP:N	2.05	0.90

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	463/519 (89%)	429 (93%)	30 (6%)	4 (1%)	19	58
1	B	436/519 (84%)	398 (91%)	34 (8%)	4 (1%)	19	58
All	All	899/1038 (87%)	827 (92%)	64 (7%)	8 (1%)	19	58

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	PRO
1	A	379	PRO
1	A	430	PRO
1	B	111	ASP
1	B	379	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	382/418 (91%)	313 (82%)	69 (18%)	2	9
1	B	368/418 (88%)	296 (80%)	72 (20%)	1	7
All	All	750/836 (90%)	609 (81%)	141 (19%)	1	8

5 of 141 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	472	TRP
1	B	91	ILE
1	B	451	ASN
1	A	479	ARG
1	B	35	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	305	GLN
1	A	330	GLN
1	B	29	HIS
1	B	143	GLN
1	B	255	HIS

### 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](#)

2 ligands are modelled in this entry.

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	BLA	A	900	1	36,46,46	2.98	17 (47%)	47,67,67	2.22	14 (29%)
2	BLA	B	900	1	36,46,46	3.11	14 (38%)	47,67,67	2.24	15 (31%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BLA	A	900	1	-	0/17/74/74	0/4/4/4
2	BLA	B	900	1	-	0/17/74/74	0/4/4/4

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	900	BLA	C1D-C2D	-4.01	1.37	1.45
2	B	900	BLA	C4D-C3D	-3.97	1.39	1.45
2	B	900	BLA	C3C-C4C	-3.73	1.39	1.45
2	A	900	BLA	C1C-C2C	-3.42	1.38	1.47
2	A	900	BLA	C3C-C4C	-3.38	1.40	1.45

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	900	BLA	C1A-CHA-C4D	-9.02	118.03	128.81
2	A	900	BLA	C1A-CHA-C4D	-8.74	118.37	128.81
2	A	900	BLA	C4C-CHD-C1D	-4.20	117.75	128.09
2	B	900	BLA	CHD-C4C-C3C	-3.85	117.86	128.01
2	A	900	BLA	CHA-C4D-C3D	-3.55	116.88	125.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	900	BLA	5	0
2	B	900	BLA	7	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	473/519 (91%)	0.13	9 (1%) 66 65	93, 131, 185, 270	0
1	B	448/519 (86%)	0.43	32 (7%) 16 17	119, 172, 245, 344	0
All	All	921/1038 (88%)	0.28	41 (4%) 33 33	93, 149, 232, 344	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	226	PRO	6.0
1	B	236	VAL	4.7
1	B	103	PHE	4.6
1	B	237	THR	4.6
1	B	191	GLU	3.7

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BLA	B	900	43/43	0.89	0.56	140,153,180,212	0
2	BLA	A	900	43/43	0.90	0.50	90,114,146,153	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.