



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 02:09 pm GMT

PDB ID : 4R70  
Title : Crystal structure of bacteriophytochrome RpBphP3 from photosynthetic bacterium *R. palustris*  
Authors : Yang, X.; Kuk, J.; Moffat, K.  
Deposited on : 2014-08-26  
Resolution : 2.85 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

<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  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
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

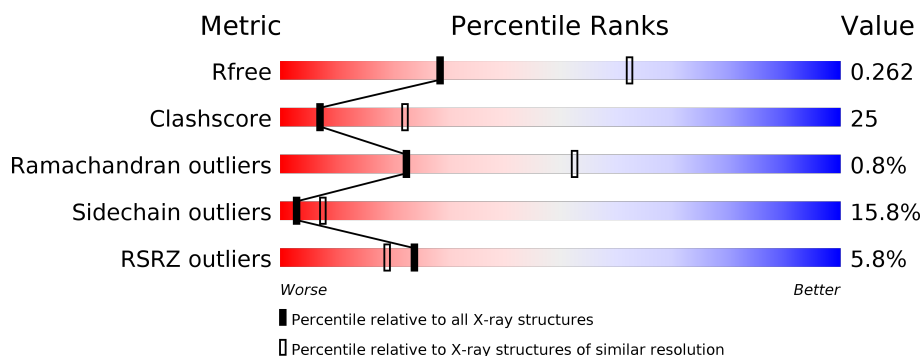
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.85 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	2469 (2.90-2.82)
Clashscore	112137	2749 (2.90-2.82)
Ramachandran outliers	110173	2687 (2.90-2.82)
Sidechain outliers	110143	2690 (2.90-2.82)
RSRZ outliers	101464	2487 (2.90-2.82)

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	529	<div> <div>4%</div> <div> <div></div> <div>46%</div> <div>39%</div> <div>9%</div> <div>6%</div> </div> </div>
1	B	529	<div> <div>6%</div> <div> <div></div> <div>43%</div> <div>42%</div> <div>7%</div> <div>8%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7830 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.

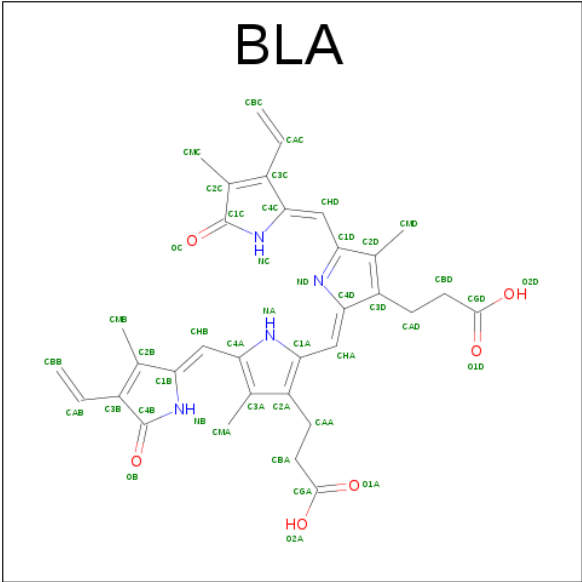
- Molecule 1 is a protein called Bacteriophytochrome (Light-regulated signal transduction histidine kinase), PhyB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	498	Total	C	N	O	S	0	2	0
			3915	2467	715	719	14			
1	B	486	Total	C	N	O	S	0	0	0
			3806	2403	689	700	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	522	LEU	-	EXPRESSION TAG	UNP Q6N5G2
A	523	GLU	-	EXPRESSION TAG	UNP Q6N5G2
A	524	HIS	-	EXPRESSION TAG	UNP Q6N5G2
A	525	HIS	-	EXPRESSION TAG	UNP Q6N5G2
A	526	HIS	-	EXPRESSION TAG	UNP Q6N5G2
A	527	HIS	-	EXPRESSION TAG	UNP Q6N5G2
A	528	HIS	-	EXPRESSION TAG	UNP Q6N5G2
A	529	HIS	-	EXPRESSION TAG	UNP Q6N5G2
B	522	LEU	-	EXPRESSION TAG	UNP Q6N5G2
B	523	GLU	-	EXPRESSION TAG	UNP Q6N5G2
B	524	HIS	-	EXPRESSION TAG	UNP Q6N5G2
B	525	HIS	-	EXPRESSION TAG	UNP Q6N5G2
B	526	HIS	-	EXPRESSION TAG	UNP Q6N5G2
B	527	HIS	-	EXPRESSION TAG	UNP Q6N5G2
B	528	HIS	-	EXPRESSION TAG	UNP Q6N5G2
B	529	HIS	-	EXPRESSION TAG	UNP Q6N5G2

- 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		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	16	Total	O	0	0
			16	16		
3	B	7	Total	O	0	0
			7	7		



D511	G441	G373	P218	ASP
L512	R442	A374	A219	GLU
V513	A443	S375	Q220	SER
V514	S444		S221	ARG
V516	T445	L379	R222	Y141
V517	T446	C380	I227	T142
L518	L447	S381	R228	I143
G519	M448	F382	P229	E144
R520	L449	D383	R229	F145
A521	W450	Q384	R231	V230
	F451	V385		F146
LEU	R452			R147
GLU		R390	I236	S148
HIS	T458	T391	G237	V149
HIS	W459	P392	Y238	R150
HIS	T460	S393	R239	V151
HIS	W461	D394	P240	R155
HIS	Q462	D395	S241	
HIS	G463	E396	P242	A165
HIS	D464	T397	L243	
	P465	Q398	V244	I168
	H466	A399	P245	A169
	R467	L400	D246	A170
	P468	A401	I247	S171
	V469	S402	N248	E172
	GLN	W403	P249	V173
	ILE	L404	R250	R174
	GLY	S405	L251	R175
	PRO	H406	G252	I176
	ARG	R407		
	GLY		I255	D180
	ARG	R410	D256	R181
	R477	G411		I182
	L478	T412	F259	K183
	Q479	F413		V184
	T480	Q414	L262	Y185
	R481	T415	R263	
	A482	Q416	S264	S192
	S483	Q417	V265	G193
	F484	L418	S266	
	E485	S419	P267	E198
		A420	T268	D199
	R488	S421	H269	R200
		F422	L270	D201
	V491	F423	E271	S202
	R492	E426	Y272	G203
	D493	Y427	M273	I204
	R494	Y428	V274	P205
		S429	N275	S206
	P497	D430	M276	L207
	W498	T431	G277	L208
	R499	A432	M278	D209
			S364	F210
	E502		E365	H211
	I503	L435	E366	A280
	V504	L436	E367	A281
		A437	M368	P212
	E508	V438	L369	P213
	I509	P439	A370	S214
	R510	L440	L371	S215
			S285	D216
			M372	I217

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	143.31Å 143.31Å 120.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.85 31.03 – 2.84	Depositor EDS
% Data completeness (in resolution range)	98.5 (20.00-2.85) 98.1 (31.03-2.84)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 2.85Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, $R_{free}$	0.185 , 0.260 0.187 , 0.262	Depositor DCC
$R_{free}$ test set	1690 reflections (5.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	94.6	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 88.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.029 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7830	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	124.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.20% 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	1.11	7/4011 (0.2%)	1.22	14/5461 (0.3%)
1	B	1.11	10/3893 (0.3%)	1.16	11/5301 (0.2%)
All	All	1.11	17/7904 (0.2%)	1.20	25/10762 (0.2%)

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	4
1	B	0	1
All	All	0	5

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	318	ALA	CA-CB	-7.80	1.36	1.52
1	A	298	CYS	CB-SG	-7.48	1.69	1.82
1	B	330	GLU	CD-OE2	7.15	1.33	1.25
1	A	184	VAL	CB-CG1	-7.12	1.38	1.52
1	B	327	VAL	CB-CG1	-6.66	1.38	1.52

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	397	ILE	CG1-CB-CG2	-9.61	90.26	111.40
1	A	207	LEU	CB-CG-CD2	-8.28	96.93	111.00
1	A	181	ARG	NE-CZ-NH1	-7.78	116.41	120.30
1	A	347	LEU	CB-CG-CD2	-7.70	97.92	111.00

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	166	CYS	CA-CB-SG	-6.89	101.60	114.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	137	ASP	Peptide
1	A	141	TYR	Peptide
1	A	441	GLY	Peptide
1	A	49	SER	Peptide
1	B	141	TYR	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	3915	0	3893	192	2
1	B	3806	0	3775	203	2
2	A	43	0	31	13	0
2	B	43	0	31	6	0
3	A	16	0	0	3	0
3	B	7	0	0	0	0
All	All	7830	0	7730	393	2

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 25.

The worst 5 of 393 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:97:GLY:O	1:B:100:ALA:N	2.00	0.95
1:B:217:ILE:HG22	1:B:222:ARG:HG2	1.46	0.94
1:A:149:VAL:HG21	1:B:338:VAL:HG11	1.52	0.92
1:B:265:VAL:HG13	1:B:270:LEU:HD11	1.53	0.90
1:A:351:LEU:HD11	1:A:357:LEU:HD12	1.55	0.87

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:442:ARG:NE	1:B:98:ASP:OD1[6_555]	1.99	0.21
1:A:390:ARG:NH2	1:B:458:THR:OG1[3_564]	2.06	0.14

## 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	498/529 (94%)	449 (90%)	45 (9%)	4 (1%)	22	52
1	B	480/529 (91%)	440 (92%)	36 (8%)	4 (1%)	22	52
All	All	978/1058 (92%)	889 (91%)	81 (8%)	8 (1%)	22	52

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	99	PRO
1	B	70	PRO
1	A	67	ARG
1	B	100	ALA
1	A	69	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	421/445 (95%)	350 (83%)	71 (17%)	2	6
1	B	409/445 (92%)	349 (85%)	60 (15%)	3	9
All	All	830/890 (93%)	699 (84%)	131 (16%)	3	7

5 of 131 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	438	VAL
1	B	65	LEU
1	B	438	VAL
1	A	447	LEU
1	A	471	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	269	HIS
1	B	501	HIS
1	B	398	GLN
1	B	42	HIS
1	B	344	GLN

### 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 ⓘ

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	3.08	17 (47%)	46,67,67	1.96	14 (30%)
2	BLA	B	900	1	36,46,46	3.09	17 (47%)	46,67,67	2.19	15 (32%)

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	-	0/22/74/74	0/4/4/4
2	BLA	B	900	1	-	0/22/74/74	0/4/4/4

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	900	BLA	C1C-C2C	-4.46	1.35	1.47
2	B	900	BLA	C1D-C2D	-4.13	1.36	1.45
2	A	900	BLA	C4D-C3D	-3.99	1.38	1.45
2	B	900	BLA	C1C-C2C	-3.54	1.37	1.47
2	A	900	BLA	C1B-C2B	-3.37	1.38	1.45

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	900	BLA	C1A-CHA-C4D	-5.18	122.49	128.77
2	B	900	BLA	C1A-CHA-C4D	-4.21	123.67	128.77
2	B	900	BLA	CAD-CBD-CGD	-3.90	106.00	112.66
2	A	900	BLA	CBC-CAC-C3C	-3.68	108.92	127.39
2	A	900	BLA	CHA-C4D-C3D	-3.57	116.92	125.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	900	BLA	13	0
2	B	900	BLA	6	0

## 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	498/529 (94%)	-0.17	23 (4%)	33 28	61, 107, 181, 236	0
1	B	486/529 (91%)	0.10	34 (6%)	17 12	65, 132, 200, 232	0
All	All	984/1058 (93%)	-0.04	57 (5%)	24 19	61, 119, 194, 236	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	477	ARG	6.4
1	B	355	GLN	5.3
1	A	114	GLY	5.1
1	B	353	ASP	4.9
1	A	476	ARG	4.8

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BLA	B	900	43/43	0.87	0.32	0.84	108,145,171,175	0
2	BLA	A	900	43/43	0.90	0.32	0.79	82,104,136,142	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.