



# wwPDB X-ray Structure Validation Summary Report i

Mar 1, 2014 – 04:05 AM GMT

PDB ID : 4FPD  
Title : Deprotonation of D96 in bacteriorhodopsin opens the proton uptake pathway  
Authors : Wang, T.; Sessions, A.O.; Lunde, C.S.; Rouani, S.; Glaeser, R.M.; Facciotti, M.T.; Duan, Y.  
Deposited on : 2012-06-22  
Resolution : 2.65 Å (reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

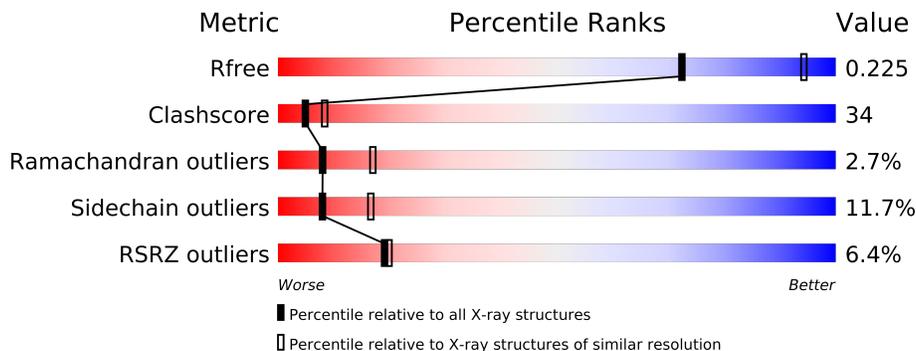
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.65 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))
$R_{free}$	66092	2232 (2.70-2.62)
Clashscore	79885	2700 (2.70-2.62)
Ramachandran outliers	78287	2657 (2.70-2.62)
Sidechain outliers	78261	2657 (2.70-2.62)
RSRZ outliers	66119	2234 (2.70-2.62)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	262	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	CL	A	302	-	X
4	LI1	A	303	-	X
4	LI1	A	304	-	X
4	LI1	A	305	-	X
4	LI1	A	306	X	X
4	LI1	A	307	-	X
4	LI1	A	308	-	X

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Geometry</b>	<b>Electron density</b>
4	LI1	A	309	-	X
4	LI1	A	310	-	X

## 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 1899 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

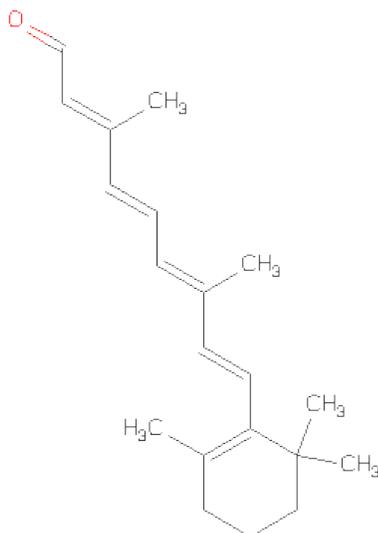
- Molecule 1 is a protein called Bacteriorhodopsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	225	1729	1158	266	295	10	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	96	GLY	ASP	ENGINEERED MUTATION	UNP P02945
A	171	CYS	PHE	ENGINEERED MUTATION	UNP P02945
A	219	LEU	PHE	ENGINEERED MUTATION	UNP P02945

- Molecule 2 is RETINAL (three-letter code: RET) (formula: C<sub>20</sub>H<sub>28</sub>O).

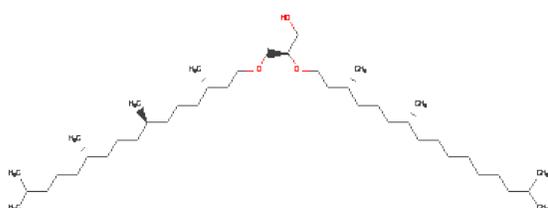


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	C	0	1
			40	40		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is 1-[2,6,10,14-TETRAMETHYL-HEXADECAN-16-YL]-2-[2,10,14-TRIMETHYLHEXADECAN-16-YL]GLYCEROL (three-letter code: LI1) (formula: C<sub>42</sub>H<sub>86</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C 11 11	0	0
4	A	1	Total C 18 18	0	0
4	A	1	Total C 8 8	0	0
4	A	1	Total C 5 5	0	0
4	A	1	Total C O 29 26 3	0	0
4	A	1	Total C 3 3	0	0
4	A	1	Total C 14 14	0	0
4	A	1	Total C 10 10	0	0

- Molecule 5 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	A	31	Total 31	O 31	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.93Å 60.93Å 107.70Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.95 – 2.65 47.39 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.95-2.65) 99.5 (47.39-2.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.65Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.223 , 0.240 0.221 , 0.225	Depositor DCC
$R_{free}$ test set	322 reflections (5.15%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.4	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 53.1	EDS
Estimated twinning fraction	0.088 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 6579 reflections (0.015%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	1899	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.14% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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: LI1, RET, 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	0.45	0/1774	0.64	0/2424

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1729	0	1797	128	0
2	A	40	0	53	5	0
3	A	1	0	0	1	0
4	A	98	0	135	9	0
5	A	31	0	0	5	0
All	All	1899	0	1985	131	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 34.

The worst 5 of 131 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:228:ALA:O	1:A:229:ILE:CG1	1.95	1.15
1:A:172:LYS:HD3	1:A:175:ARG:NH2	1.65	1.10
1:A:228:ALA:O	1:A:229:ILE:HG13	1.59	1.02
1:A:224:LEU:O	1:A:227:ARG:HD3	1.64	0.98
1:A:225:ARG:O	1:A:227:ARG:N	1.94	0.98

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	223/262 (85%)	204 (92%)	13 (6%)	6 (3%)	<b>8</b> <b>17</b>

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	162	SER
1	A	225	ARG
1	A	226	SER
1	A	228	ALA
1	A	171	CYS

### 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	180/205 (88%)	159 (88%)	21 (12%)	<b>8</b> <b>16</b>

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

Mol	Chain	Res	Type
1	A	163	MET
1	A	166	GLU
1	A	202	ASN
1	A	159	LYS
1	A	206	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	GLN
1	A	176	ASN
1	A	202	ASN

### 5.3.3 RNA [i](#)

There are no RNA chains 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](#)

Of 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 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	RET	A	301[A]	-	19,20,21	2.40	5 (26%)	26,27,28	1.69	4 (15%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	RET	A	301[B]	-	19,20,21	3.09	6 (31%)	26,27,28	3.89	11 (42%)
4	LI1	A	303	-	6,9,44	6.59	2 (33%)	3,7,51	2.42	1 (33%)
4	LI1	A	304	-	15,16,44	2.64	7 (46%)	10,16,51	3.65	6 (60%)
4	LI1	A	305	-	5,6,44	20.55	3 (60%)	2,4,51	5.79	1 (50%)
4	LI1	A	306	-	4,4,44	8.06	2 (50%)	3,3,51	4.25	3 (100%)
4	LI1	A	307	-	26,27,44	3.30	9 (34%)	19,26,51	3.90	8 (42%)
4	LI1	A	308	-	2,2,44	1.95	1 (50%)	0,1,51	0.00	-
4	LI1	A	309	-	13,13,44	16.40	2 (15%)	11,13,51	2.02	1 (9%)
4	LI1	A	310	-	9,9,44	3.42	5 (55%)	8,9,51	6.13	5 (62%)

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	RET	A	301[A]	-	-	0/13/30/31	0/1/1/1
2	RET	A	301[B]	-	-	0/13/30/31	0/1/1/1
4	LI1	A	303	-	-	0/4/5/49	0/0/0/0
4	LI1	A	304	-	1/1/2/8	0/12/13/49	0/0/0/0
4	LI1	A	305	-	-	0/1/2/49	0/0/0/0
4	LI1	A	306	-	-	0/2/2/49	0/0/0/0
4	LI1	A	307	-	-	0/23/25/49	0/0/0/0
4	LI1	A	308	-	-	0/0/0/49	0/0/0/0
4	LI1	A	309	-	-	0/11/11/49	0/0/0/0
4	LI1	A	310	-	-	0/7/8/49	0/0/0/0

The worst 5 of 42 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	309	LI1	C45-C46	-59.00	1.31	1.55
4	A	305	LI1	C52-C51	45.68	1.73	1.55
4	A	303	LI1	C19-C18	-15.53	1.21	1.49
4	A	306	LI1	C50-C51	12.38	1.60	1.55
4	A	307	LI1	C51-C52	-11.09	1.50	1.55

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	310	LI1	C25-C26-C27	14.51	143.43	114.68

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301[B]	RET	C20-C13-C12	13.39	139.75	118.09
4	A	307	LI1	C1-C2-C3	-10.42	88.39	111.88
4	A	307	LI1	O2-C41-C42	9.54	126.09	107.52
4	A	305	LI1	C52-C51-C50	-8.13	96.20	114.46

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	304	LI1	C18

There are no torsion outliers.

There are no ring outliers.

## 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

### 6.1 Protein, DNA and RNA chains

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	225/262 (85%)	0.04	13 (5%) <span style="border: 1px solid red; padding: 2px;">22</span> <span style="border: 1px solid red; padding: 2px;">23</span>	16, 29, 83, 93	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	160	ALA	4.8
1	A	228	ALA	4.6
1	A	229	ILE	3.8
1	A	161	GLU	3.4
1	A	167	VAL	3.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates

There are no carbohydrates in this entry.

### 6.4 Ligands

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	LI1	A	306	5/45	0.40	32.78	54,54,55,55	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	LI1	A	308	3/45	0.55	17.74	73,73,75,75	0
4	LI1	A	310	10/45	0.52	14.07	75,76,76,76	0
4	LI1	A	303	11/45	0.38	13.18	50,54,64,65	0
4	LI1	A	307	29/45	0.27	4.60	37,54,63,66	0
4	LI1	A	304	18/45	0.28	4.40	37,47,53,53	0
4	LI1	A	305	8/45	0.31	4.35	53,54,58,59	0
3	CL	A	302	1/1	0.23	4.08	38,38,38,38	0
4	LI1	A	309	14/45	0.26	2.50	45,48,53,53	0
2	RET	A	301[A]	20/21	0.16	1.13	19,20,27,29	20
2	RET	A	301[B]	20/21	0.16	1.13	7,9,15,17	20

## 6.5 Other polymers ⓘ

There are no such residues in this entry.