WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 01:40 AM GMT

PDB ID : 3SS0

Title : Engineered high-affinity halide-binding protein derived from YFP: fluoride complex

Authors : Wang, W.; Grimley, J.S.; Beese, L.S.; Hellinga, H.W.

Deposited on : 2011-07-07

Resolution : 1.49 Å(reported)

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 1.49 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	Percentile Ranks	Value
Rfree		0.188
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0.5%
RSRZ outliers		4.3%

Worse   *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution
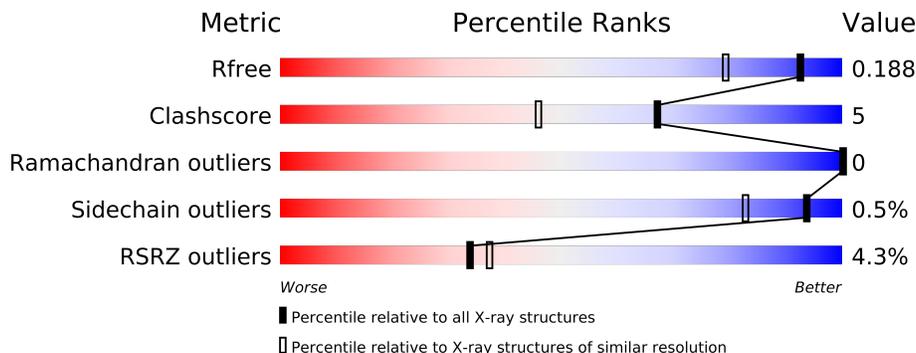
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1513 (1.50-1.50)
Clashscore	79885	1768 (1.50-1.50)
Ramachandran outliers	78287	1720 (1.50-1.50)
Sidechain outliers	78261	1718 (1.50-1.50)
RSRZ outliers	66119	1514 (1.50-1.50)

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 >=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	258	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 4174 atoms, of which 1895 are hydrogens 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 Green fluorescent protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	232	3843	1238	1895	330	373	7	0	11	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	VAL	-	INSERTION	UNP P42212
A	66	CR2	SER	CHROMOPHORE	UNP P42212
A	66	CR2	TYR	CHROMOPHORE	UNP P42212
A	66	CR2	GLY	CHROMOPHORE	UNP P42212
A	72	ALA	SER	ENGINEERED MUTATION	UNP P42212
A	79	ARG	LYS	ENGINEERED MUTATION	UNP P42212
A	183	ALA	GLN	ENGINEERED MUTATION	UNP P42212
A	203	TYR	THR	ENGINEERED MUTATION	UNP P42212
A	231	LEU	HIS	ENGINEERED MUTATION	UNP P42212
A	239	GLY	-	EXPRESSION TAG	UNP P42212
A	240	GLY	-	EXPRESSION TAG	UNP P42212
A	241	SER	-	EXPRESSION TAG	UNP P42212
A	242	ASN	-	EXPRESSION TAG	UNP P42212
A	243	ASP	-	EXPRESSION TAG	UNP P42212
A	244	TYR	-	EXPRESSION TAG	UNP P42212
A	245	LYS	-	EXPRESSION TAG	UNP P42212
A	246	ASP	-	EXPRESSION TAG	UNP P42212
A	247	ASP	-	EXPRESSION TAG	UNP P42212
A	248	ASP	-	EXPRESSION TAG	UNP P42212
A	249	ASP	-	EXPRESSION TAG	UNP P42212
A	250	LYS	-	EXPRESSION TAG	UNP P42212
A	251	GLY	-	EXPRESSION TAG	UNP P42212
A	252	GLY	-	EXPRESSION TAG	UNP P42212
A	253	SER	-	EXPRESSION TAG	UNP P42212
A	254	HIS	-	EXPRESSION TAG	UNP P42212
A	255	HIS	-	EXPRESSION TAG	UNP P42212
A	256	HIS	-	EXPRESSION TAG	UNP P42212

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	257	HIS	-	EXPRESSION TAG	UNP P42212
A	258	HIS	-	EXPRESSION TAG	UNP P42212
A	259	HIS	-	EXPRESSION TAG	UNP P42212

• Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	331	Total 331	O 331	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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: Green fluorescent protein

Chain A:
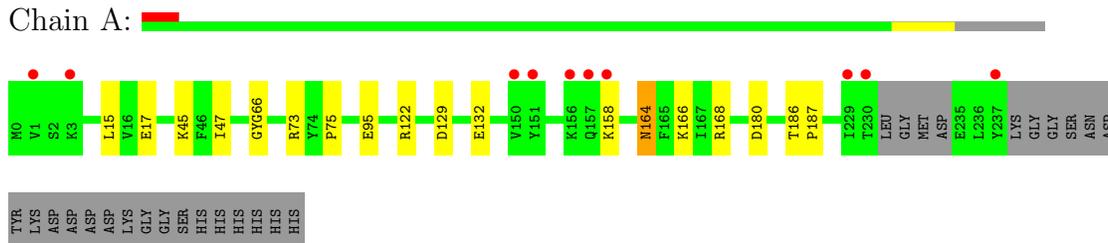
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.06Å 62.98Å 69.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.17 – 1.49 41.17 – 1.49	Depositor EDS
% Data completeness (in resolution range)	96.7 (41.17-1.49) 96.7 (41.17-1.49)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.40 (at 1.49Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.157 , 0.192 0.156 , 0.188	Depositor DCC
R_{free} test set	1866 reflections (5.19%)	DCC
Wilson B-factor (Å^2)	12.9	Xtriage
Anisotropy	0.589	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 39.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	1 of 35985 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4174	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.01% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]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: CR2

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	0.42	0/1971	0.60	0/2662

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	1948	1895	2	19	0
2	A	331	0	0	14	1
All	All	2279	1895	2	19	1

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 5.

All (19) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:73[B]:ARG:NH1	2:A:576:HOH:O	2.11	0.83
1:A:158:LYS:NZ	2:A:326:HOH:O	2.14	0.79

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:73[B]:ARG:CZ	2:A:576:HOH:O	2.29	0.79
1:A:75:PRO:HA	2:A:576:HOH:O	1.84	0.77
1:A:129:ASP:OD2	2:A:532:HOH:O	2.09	0.69
1:A:95:GLU:CD	2:A:263:HOH:O	2.37	0.63
1:A:17[A]:GLU:OE1	2:A:408:HOH:O	2.15	0.60
1:A:15:LEU:HD11	1:A:17[A]:GLU:OE2	2.09	0.52
1:A:168:ARG:NE	2:A:449:HOH:O	2.32	0.50
1:A:166:LYS:HG2	1:A:180[B]:ASP:OD1	2.15	0.47
1:A:164:ASN:ND2	2:A:535:HOH:O	2.50	0.45
1:A:66:CR2:HD1	1:A:66:CR2:N2	2.32	0.44
1:A:45:LYS:HE2	1:A:47:ILE:HD11	2.00	0.44
1:A:95:GLU:CG	2:A:263:HOH:O	2.65	0.44
1:A:73[B]:ARG:NH2	2:A:576:HOH:O	2.48	0.43
1:A:132:GLU:HB3	2:A:414:HOH:O	2.18	0.43
1:A:186[B]:THR:HG22	2:A:386:HOH:O	2.19	0.42
1:A:186[B]:THR:OG1	1:A:187:PRO:HD2	2.21	0.40
1:A:122:ARG:NH1	2:A:425:HOH:O	2.52	0.40

All (1) 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	Distance(Å)	Clash(Å)
2:A:270:HOH:O	2:A:287:HOH:O[4_545]	1.93	0.27

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	236/258 (92%)	233 (99%)	3 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	211/222 (95%)	210 (100%)	1 (0%)	94 83

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	164	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	#$|Z| > 2$
1	CR2	A	66	1	20,20,21	6.98	9 (45%)	25,27,29	5.40	10 (40%)

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
1	CR2	A	66	1	-	0/8/25/26	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CR2	O3-C3	21.43	1.26	1.11
1	A	66	CR2	CA2-C2	-13.50	1.33	1.48
1	A	66	CR2	C1-N2	11.62	1.56	1.33
1	A	66	CR2	C1-N3	9.34	1.51	1.37
1	A	66	CR2	CB2-CA2	6.77	1.39	1.35
1	A	66	CR2	CA3-N3	-4.64	1.43	1.47
1	A	66	CR2	C2-N3	-4.45	1.30	1.39
1	A	66	CR2	CA1-C1	-3.58	1.44	1.50
1	A	66	CR2	CG2-CB2	2.57	1.52	1.46

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	CR2	CA2-C2-N3	18.63	114.05	103.44
1	A	66	CR2	CA3-N3-C2	11.47	130.60	123.46
1	A	66	CR2	C2-CA2-N2	7.44	114.52	108.91
1	A	66	CR2	CA2-N2-C1	-6.87	100.28	105.81
1	A	66	CR2	N3-C1-N2	-6.82	105.48	111.56
1	A	66	CR2	O2-C2-CA2	-5.67	127.68	130.96
1	A	66	CR2	CB2-CA2-N2	-4.31	120.64	128.59
1	A	66	CR2	CA1-C1-N2	3.69	128.20	123.82
1	A	66	CR2	C1-CA1-N1	-3.49	104.81	112.84
1	A	66	CR2	O2-C2-N3	-2.99	118.26	124.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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

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, 95th 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^2)	Q<0.9
1	A	232/258 (89%)	-0.06	10 (4%)	34	37	8, 14, 37, 78	2 (0%)

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	VAL	5.7
1	A	237	TYR	4.9
1	A	157	GLN	4.1
1	A	151	TYR	3.9
1	A	230	THR	3.6
1	A	229	ILE	2.6
1	A	158	LYS	2.3
1	A	156	LYS	2.2
1	A	3	LYS	2.1
1	A	150	VAL	2.1

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

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, 95th 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^2)	Q<0.9
1	CR2	A	66	19/20	0.08	0.49	7,10,12,14	0

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.