



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 27, 2024 – 06:02 PM EST

PDB ID : 9BCF
Title : Chimeric protein of crocodile allergen Cro p 1.0101 and GFP
Authors : O'Malley, A.; Ruethers, T.; Lopata, A.L.; Chruszcz, M.
Deposited on : 2024-04-09
Resolution : 3.20 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

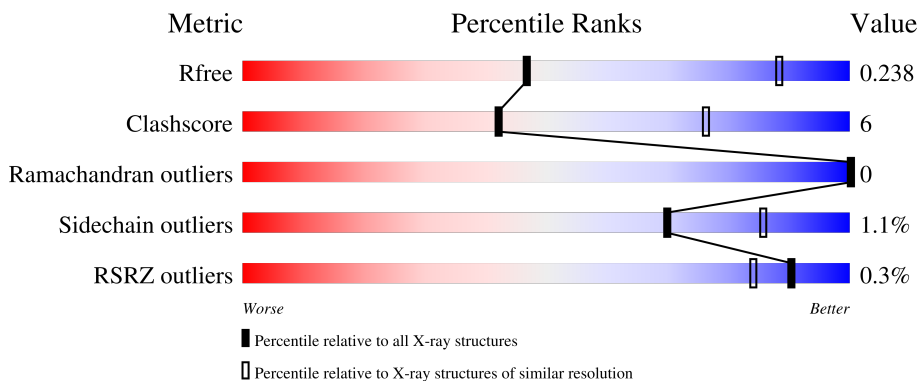
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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 $\leq 5\%$. 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	358	
1	B	358	
1	C	358	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6183 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 Parvalbumin, Green fluorescent protein chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	226	Total 1751	C 1109	N 293	O 344	S 5	0	0	0
1	B	341	Total 2644	C 1669	N 443	O 526	S 6	0	0	0
1	C	226	Total 1759	C 1115	N 296	O 343	S 5	0	0	0

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP A0A7M4EAX1
A	0	GLY	-	expression tag	UNP A0A7M4EAX1
A	109	GLY	-	linker	UNP A0A7M4EAX1
A	110	SER	-	linker	UNP A0A7M4EAX1
A	111	SER	-	linker	UNP A0A7M4EAX1
A	112	SER	-	linker	UNP A0A7M4EAX1
A	113	SER	-	linker	UNP A0A7M4EAX1
A	141	ARG	SER	conflict	UNP A0A059PIQ0
A	177	CRO	THR	chromophore	UNP A0A059PIQ0
A	?	-	TYR	chromophore	UNP A0A059PIQ0
A	?	-	GLY	chromophore	UNP A0A059PIQ0
A	180	LEU	GLN	conflict	UNP A0A059PIQ0
A	183	SER	ALA	conflict	UNP A0A059PIQ0
A	191	ARG	GLN	conflict	UNP A0A059PIQ0
A	275	TYR	ASN	conflict	UNP A0A059PIQ0
A	317	VAL	ALA	conflict	UNP A0A059PIQ0
A	334	ASP	PHE	conflict	UNP A0A059PIQ0
A	350	GLY	-	expression tag	UNP A0A059PIQ0
A	351	SER	-	expression tag	UNP A0A059PIQ0
A	352	GLY	-	expression tag	UNP A0A059PIQ0
A	353	HIS	-	expression tag	UNP A0A059PIQ0
A	354	HIS	-	expression tag	UNP A0A059PIQ0
A	355	HIS	-	expression tag	UNP A0A059PIQ0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	356	HIS	-	expression tag	UNP A0A059PIQ0
A	357	HIS	-	expression tag	UNP A0A059PIQ0
A	358	HIS	-	expression tag	UNP A0A059PIQ0
B	-1	MET	-	initiating methionine	UNP A0A7M4EAX1
B	0	GLY	-	expression tag	UNP A0A7M4EAX1
B	109	GLY	-	linker	UNP A0A7M4EAX1
B	110	SER	-	linker	UNP A0A7M4EAX1
B	111	SER	-	linker	UNP A0A7M4EAX1
B	112	SER	-	linker	UNP A0A7M4EAX1
B	113	SER	-	linker	UNP A0A7M4EAX1
B	141	ARG	SER	conflict	UNP A0A059PIQ0
B	177	CRO	THR	chromophore	UNP A0A059PIQ0
B	?	-	TYR	chromophore	UNP A0A059PIQ0
B	?	-	GLY	chromophore	UNP A0A059PIQ0
B	180	LEU	GLN	conflict	UNP A0A059PIQ0
B	183	SER	ALA	conflict	UNP A0A059PIQ0
B	191	ARG	GLN	conflict	UNP A0A059PIQ0
B	275	TYR	ASN	conflict	UNP A0A059PIQ0
B	317	VAL	ALA	conflict	UNP A0A059PIQ0
B	334	ASP	PHE	conflict	UNP A0A059PIQ0
B	350	GLY	-	expression tag	UNP A0A059PIQ0
B	351	SER	-	expression tag	UNP A0A059PIQ0
B	352	GLY	-	expression tag	UNP A0A059PIQ0
B	353	HIS	-	expression tag	UNP A0A059PIQ0
B	354	HIS	-	expression tag	UNP A0A059PIQ0
B	355	HIS	-	expression tag	UNP A0A059PIQ0
B	356	HIS	-	expression tag	UNP A0A059PIQ0
B	357	HIS	-	expression tag	UNP A0A059PIQ0
B	358	HIS	-	expression tag	UNP A0A059PIQ0
C	-1	MET	-	initiating methionine	UNP A0A7M4EAX1
C	0	GLY	-	expression tag	UNP A0A7M4EAX1
C	109	GLY	-	linker	UNP A0A7M4EAX1
C	110	SER	-	linker	UNP A0A7M4EAX1
C	111	SER	-	linker	UNP A0A7M4EAX1
C	112	SER	-	linker	UNP A0A7M4EAX1
C	113	SER	-	linker	UNP A0A7M4EAX1
C	141	ARG	SER	conflict	UNP A0A059PIQ0
C	177	CRO	THR	chromophore	UNP A0A059PIQ0
C	?	-	TYR	chromophore	UNP A0A059PIQ0
C	?	-	GLY	chromophore	UNP A0A059PIQ0
C	180	LEU	GLN	conflict	UNP A0A059PIQ0
C	183	SER	ALA	conflict	UNP A0A059PIQ0

Continued on next page...

Continued from previous page...

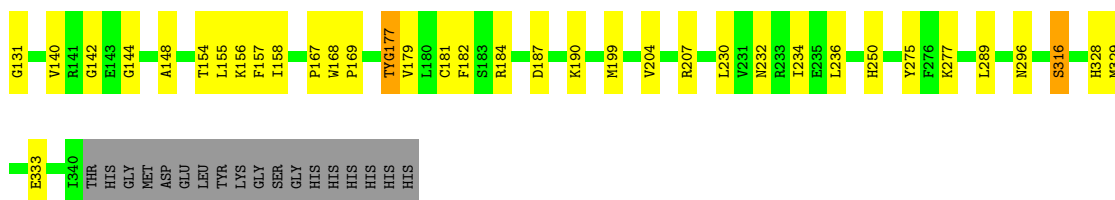
Chain	Residue	Modelled	Actual	Comment	Reference
C	191	ARG	GLN	conflict	UNP A0A059PIQ0
C	275	TYR	ASN	conflict	UNP A0A059PIQ0
C	317	VAL	ALA	conflict	UNP A0A059PIQ0
C	334	ASP	PHE	conflict	UNP A0A059PIQ0
C	350	GLY	-	expression tag	UNP A0A059PIQ0
C	351	SER	-	expression tag	UNP A0A059PIQ0
C	352	GLY	-	expression tag	UNP A0A059PIQ0
C	353	HIS	-	expression tag	UNP A0A059PIQ0
C	354	HIS	-	expression tag	UNP A0A059PIQ0
C	355	HIS	-	expression tag	UNP A0A059PIQ0
C	356	HIS	-	expression tag	UNP A0A059PIQ0
C	357	HIS	-	expression tag	UNP A0A059PIQ0
C	358	HIS	-	expression tag	UNP A0A059PIQ0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Ca 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	5	Total O 5 5	0	0
3	B	18	Total O 18 18	0	0
3	C	4	Total O 4 4	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	158.78Å 158.78Å 123.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.88 – 3.20 39.88 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.88-3.20) 99.9 (39.88-3.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 3.18Å)	Xtrriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.199 , 0.238 0.197 , 0.238	Depositor DCC
R_{free} test set	1280 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	96.6	Xtrriage
Anisotropy	0.315	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 98.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6183	wwPDB-VP
Average B, all atoms (Å ²)	116.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.36	0/1767	0.69	1/2402 (0.0%)
1	B	0.37	0/2671	0.72	0/3606
1	C	0.35	0/1775	0.67	0/2410
All	All	0.36	0/6213	0.70	1/8418 (0.0%)

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	B	0	1
1	C	0	2
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	175	LEU	CA-C-O	-5.50	108.56	120.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	207	ARG	Sidechain
1	C	184	ARG	Sidechain
1	C	207	ARG	Sidechain

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	1751	0	1641	18	0
1	B	2644	0	2564	34	0
1	C	1759	0	1665	29	0
2	B	2	0	0	0	0
3	A	5	0	0	0	0
3	B	18	0	0	0	0
3	C	4	0	0	0	0
All	All	6183	0	5870	74	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 6.

The worst 5 of 74 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:A:317:VAL:HG23	1:B:317:VAL:HG23	1.70	0.72
1:C:125:ILE:HD12	1:C:155:LEU:HD11	1.73	0.68
1:A:153:LEU:O	1:A:332:LEU:HD12	1.93	0.68
1:B:154:THR:C	1:B:155:LEU:HD12	2.14	0.67
1:B:78:THR:HG22	1:C:275:TYR:CE2	2.31	0.66

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	221/358 (62%)	217 (98%)	4 (2%)	0	100	100
1	B	336/358 (94%)	322 (96%)	14 (4%)	0	100	100
1	C	221/358 (62%)	217 (98%)	4 (2%)	0	100	100
All	All	778/1074 (72%)	756 (97%)	22 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	185/301 (62%)	183 (99%)	2 (1%)	70	86
1	B	284/301 (94%)	280 (99%)	4 (1%)	62	82
1	C	187/301 (62%)	186 (100%)	1 (0%)	86	93
All	All	656/903 (73%)	649 (99%)	7 (1%)	70	86

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	181	CYS
1	B	194	PHE
1	C	316	SER
1	B	256	PHE
1	B	85	PHE

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

3 non-standard protein/DNA/RNA residues 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
1	CRO	B	177	1	22,23,24	0.85	1 (4%)	30,32,34	1.37	4 (13%)
1	CRO	C	177	1	22,23,24	0.79	0	30,32,34	2.31	4 (13%)
1	CRO	A	177	1	22,23,24	0.85	1 (4%)	30,32,34	2.14	5 (16%)

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	CRO	B	177	1	-	4/12/31/32	0/2/2/2
1	CRO	C	177	1	-	4/12/31/32	0/2/2/2
1	CRO	A	177	1	-	4/12/31/32	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	177	CRO	C1-N2	2.52	1.35	1.32
1	A	177	CRO	C1-N2	2.01	1.35	1.32

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	177	CRO	CA1-C1-N3	-6.93	116.48	124.69
1	C	177	CRO	C1-CA1-N1	6.88	122.09	109.78
1	A	177	CRO	C1-CA1-N1	6.66	121.69	109.78
1	C	177	CRO	CA1-C1-N2	6.21	132.31	123.88
1	A	177	CRO	CA1-C1-N3	-5.79	117.83	124.69

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	177	CRO	N1-CA1-CB1-CG1
1	A	177	CRO	N1-CA1-CB1-OG1
1	B	177	CRO	N1-CA1-CB1-CG1
1	B	177	CRO	N1-CA1-CB1-OG1
1	B	177	CRO	C1-CA1-CB1-CG1

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	177	CRO	2	0
1	C	177	CRO	3	0
1	A	177	CRO	1	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	225/358 (62%)	-0.30	0 100 100	71, 125, 183, 211	0
1	B	340/358 (94%)	-0.42	2 (0%) 85 76	62, 94, 131, 185	0
1	C	225/358 (62%)	-0.32	0 100 100	92, 137, 179, 214	0
All	All	790/1074 (73%)	-0.36	2 (0%) 90 84	62, 112, 175, 214	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	343	GLY	2.8
1	B	143	GLU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [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, 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	RSCC	RSR	B-factors(Å ²)	Q<0.9
1	CRO	A	177	22/23	0.92	0.17	104,116,136,141	0
1	CRO	C	177	22/23	0.95	0.11	110,126,142,154	0
1	CRO	B	177	22/23	0.97	0.09	68,74,78,81	0

6.3 Carbohydrates [i](#)

There are no monosaccharides 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, 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	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	CA	B	401	1/1	0.99	0.02	130,130,130,130	0
2	CA	B	402	1/1	1.00	0.04	105,105,105,105	0

6.5 Other polymers [i](#)

There are no such residues in this entry.