wwPDB X-ray Structure Validation Summary Report (i)

Oct 26, 2026 – 11:11 PM JST

PDB ID	:	9XIE / pdb_00009xie
Title	:	Crystal structure of Macrobrachium rosenbergii lipocalin (apo form)
Authors	:	Guan, L.; Zhao, G.
Deposited on	:	2025-11-02
Resolution	:	2.93 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 (i) 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 (i)) were used in the production of this report:

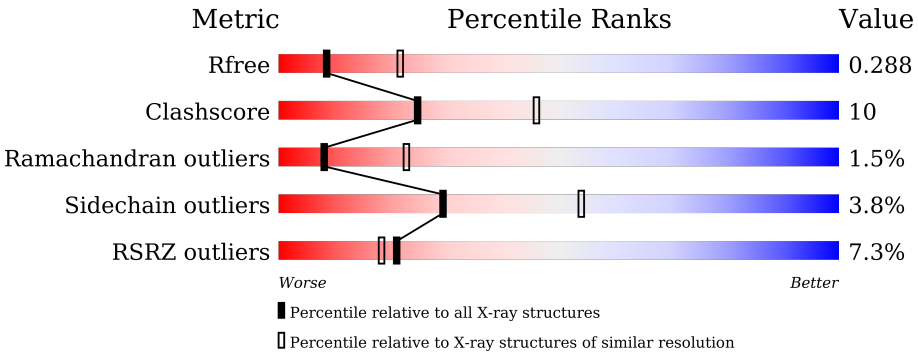
MolProbity	:	4-5-2 with Phenix2.2
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.93 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}	180053	1159 (2.96-2.92)
Clashscore	190562	1184 (2.96-2.92)
Ramachandran outliers	187476	1131 (2.96-2.92)
Sidechain outliers	187428	1131 (2.96-2.92)
RSRZ outliers	180081	1159 (2.96-2.92)

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 <=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	207	2% 67% 21% 11%
1	B	207	10% 65% 22% 12%
1	C	207	7% 68% 19% 12%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4436 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 Crustacyanin-like lipocalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	185	Total	C	N	O	S	0	0	0
			1489	951	250	277	11			
1	B	183	Total	C	N	O	S	0	0	0
			1473	942	248	273	10			
1	C	183	Total	C	N	O	S	0	0	0
			1474	943	248	273	10			

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A4Z4V4
A	-18	GLY	-	expression tag	UNP A4Z4V4
A	-17	SER	-	expression tag	UNP A4Z4V4
A	-16	SER	-	expression tag	UNP A4Z4V4
A	-15	HIS	-	expression tag	UNP A4Z4V4
A	-14	HIS	-	expression tag	UNP A4Z4V4
A	-13	HIS	-	expression tag	UNP A4Z4V4
A	-12	HIS	-	expression tag	UNP A4Z4V4
A	-11	HIS	-	expression tag	UNP A4Z4V4
A	-10	HIS	-	expression tag	UNP A4Z4V4
A	-9	SER	-	expression tag	UNP A4Z4V4
A	-8	SER	-	expression tag	UNP A4Z4V4
A	-7	GLY	-	expression tag	UNP A4Z4V4
A	-6	LEU	-	expression tag	UNP A4Z4V4
A	-5	VAL	-	expression tag	UNP A4Z4V4
A	-4	PRO	-	expression tag	UNP A4Z4V4
A	-3	ARG	-	expression tag	UNP A4Z4V4
A	-2	GLY	-	expression tag	UNP A4Z4V4
A	-1	SER	-	expression tag	UNP A4Z4V4
A	0	HIS	-	expression tag	UNP A4Z4V4
A	1	MET	-	expression tag	UNP A4Z4V4
B	-19	MET	-	initiating methionine	UNP A4Z4V4
B	-18	GLY	-	expression tag	UNP A4Z4V4

Continued on next page...

Continued from previous page...

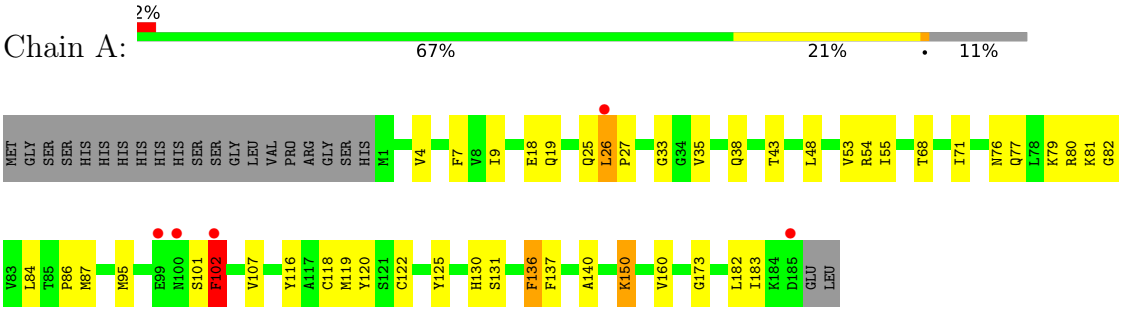
Chain	Residue	Modelled	Actual	Comment	Reference
B	-17	SER	-	expression tag	UNP A4Z4V4
B	-16	SER	-	expression tag	UNP A4Z4V4
B	-15	HIS	-	expression tag	UNP A4Z4V4
B	-14	HIS	-	expression tag	UNP A4Z4V4
B	-13	HIS	-	expression tag	UNP A4Z4V4
B	-12	HIS	-	expression tag	UNP A4Z4V4
B	-11	HIS	-	expression tag	UNP A4Z4V4
B	-10	HIS	-	expression tag	UNP A4Z4V4
B	-9	SER	-	expression tag	UNP A4Z4V4
B	-8	SER	-	expression tag	UNP A4Z4V4
B	-7	GLY	-	expression tag	UNP A4Z4V4
B	-6	LEU	-	expression tag	UNP A4Z4V4
B	-5	VAL	-	expression tag	UNP A4Z4V4
B	-4	PRO	-	expression tag	UNP A4Z4V4
B	-3	ARG	-	expression tag	UNP A4Z4V4
B	-2	GLY	-	expression tag	UNP A4Z4V4
B	-1	SER	-	expression tag	UNP A4Z4V4
B	0	HIS	-	expression tag	UNP A4Z4V4
B	1	MET	-	expression tag	UNP A4Z4V4
C	-19	MET	-	initiating methionine	UNP A4Z4V4
C	-18	GLY	-	expression tag	UNP A4Z4V4
C	-17	SER	-	expression tag	UNP A4Z4V4
C	-16	SER	-	expression tag	UNP A4Z4V4
C	-15	HIS	-	expression tag	UNP A4Z4V4
C	-14	HIS	-	expression tag	UNP A4Z4V4
C	-13	HIS	-	expression tag	UNP A4Z4V4
C	-12	HIS	-	expression tag	UNP A4Z4V4
C	-11	HIS	-	expression tag	UNP A4Z4V4
C	-10	HIS	-	expression tag	UNP A4Z4V4
C	-9	SER	-	expression tag	UNP A4Z4V4
C	-8	SER	-	expression tag	UNP A4Z4V4
C	-7	GLY	-	expression tag	UNP A4Z4V4
C	-6	LEU	-	expression tag	UNP A4Z4V4
C	-5	VAL	-	expression tag	UNP A4Z4V4
C	-4	PRO	-	expression tag	UNP A4Z4V4
C	-3	ARG	-	expression tag	UNP A4Z4V4
C	-2	GLY	-	expression tag	UNP A4Z4V4
C	-1	SER	-	expression tag	UNP A4Z4V4
C	0	HIS	-	expression tag	UNP A4Z4V4
C	1	MET	-	expression tag	UNP A4Z4V4

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

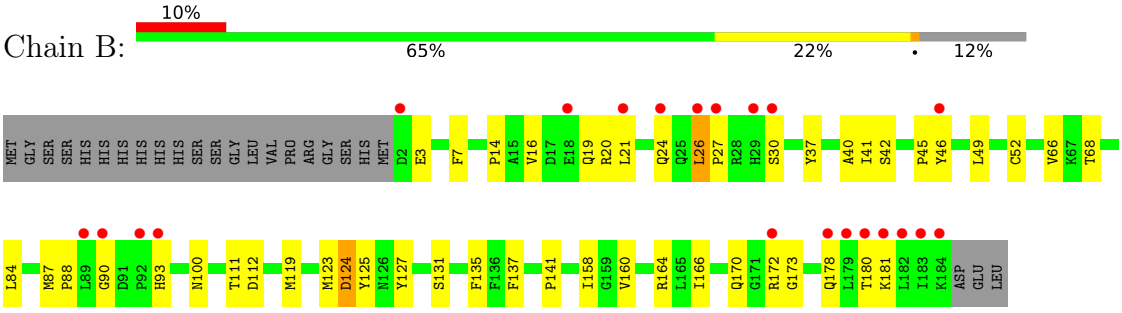
• Molecule 1: Crustacyanin-like lipocalin

Chain A:



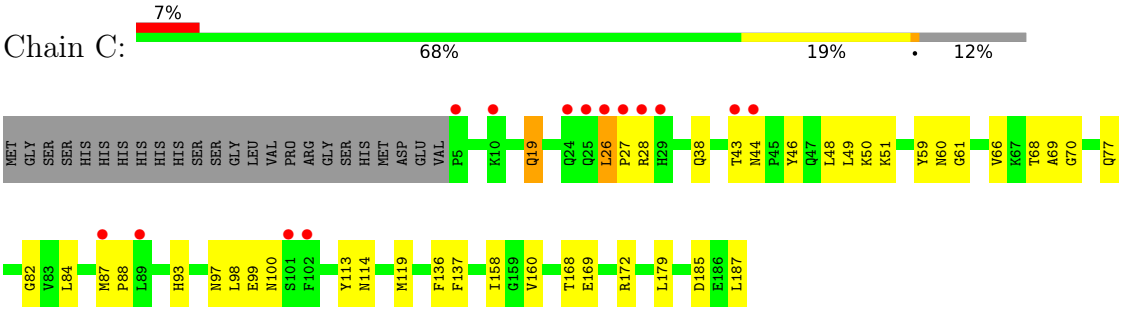
• Molecule 1: Crustacyanin-like lipocalin

Chain B:



• Molecule 1: Crustacyanin-like lipocalin

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	153.10Å 266.04Å 52.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.24 – 2.93 36.24 – 2.93	Depositor EDS
% Data completeness (in resolution range)	99.1 (36.24-2.93) 87.7 (36.24-2.93)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.93 (at 2.95Å)	Xtriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, R_{free}	0.225 , 0.287 0.226 , 0.288	Depositor DCC
R_{free} test set	2020 reflections (8.55%)	wwPDB-VP
Wilson B-factor (Å^2)	48.7	Xtriage
Anisotropy	0.464	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 51.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.032 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.044 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	4436	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.15% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	0.44	0/1528	0.66	0/2068
1	B	0.42	0/1512	0.62	0/2047
1	C	0.38	0/1513	0.60	0/2047
All	All	0.41	0/4553	0.63	0/6162

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	124	ASP	Peptide

5.2 Too-close contacts

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	1489	0	1437	32	0
1	B	1473	0	1421	34	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1474	0	1424	26	0
All	All	4436	0	4282	85	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 10.

The worst 5 of 85 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:C:87:MET:HG3	1:C:88:PRO:HD2	1.60	0.84
1:B:119:MET:HE1	1:B:137:PHE:CE2	2.30	0.66
1:A:4:VAL:HG11	1:B:87:MET:HE1	1.77	0.66
1:A:26:LEU:HB3	1:A:27:PRO:HD3	1.78	0.64
1:A:7:PHE:HD1	1:A:131:SER:HB2	1.62	0.63

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	183/207 (88%)	166 (91%)	13 (7%)	4 (2%)	5 15
1	B	181/207 (87%)	166 (92%)	13 (7%)	2 (1%)	11 29
1	C	181/207 (87%)	157 (87%)	22 (12%)	2 (1%)	11 29
All	All	545/621 (88%)	489 (90%)	48 (9%)	8 (2%)	8 23

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	26	LEU
1	B	26	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	100	ASN
1	A	33	GLY
1	A	43	THR

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	160/179 (89%)	154 (96%)	6 (4%)	29	54
1	B	158/179 (88%)	150 (95%)	8 (5%)	21	45
1	C	158/179 (88%)	154 (98%)	4 (2%)	42	66
All	All	476/537 (89%)	458 (96%)	18 (4%)	29	54

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	19	GLN
1	C	187	LEU
1	C	185	ASP
1	B	21	LEU
1	B	180	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	97	ASN
1	C	47	GLN
1	C	114	ASN
1	A	115	ASN
1	A	126	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

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 (i)

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^{th} 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	185/207 (89%)	0.00	5 (2%) 56 47	30, 45, 68, 95	0
1	B	183/207 (88%)	0.58	21 (11%) 9 9	42, 56, 101, 128	0
1	C	183/207 (88%)	0.58	14 (7%) 19 17	42, 62, 88, 128	0
All	All	551/621 (88%)	0.38	40 (7%) 21 18	30, 55, 89, 128	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	26	LEU	5.4
1	B	21	LEU	4.9
1	B	26	LEU	4.5
1	B	183	ILE	4.5
1	B	27	PRO	4.5

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 oligosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.