



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 23, 2021 – 01:59 PM EDT

PDB ID : 2PUA
Title : CRYSTAL STRUCTURE OF THE LACI FAMILY MEMBER, PURR,
BOUND TO DNA: MINOR GROOVE BINDING BY ALPHA HELICES
Authors : Schumacher, R.G.; Choi, K.Y.; Zalkin, H.; Brennan, M.A.
Deposited on : 1997-10-04
Resolution : 2.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.2

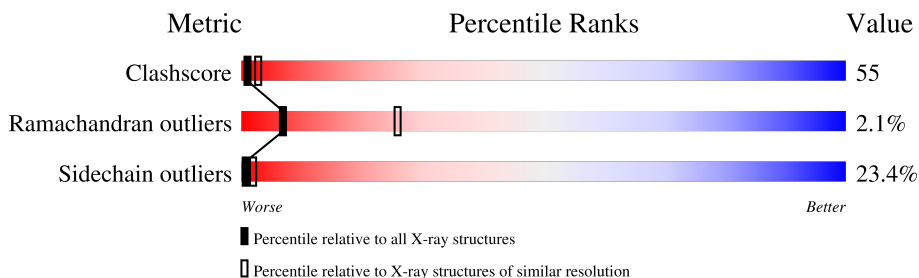
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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))
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	17	 24% 59% 18%
2	A	340	 31% 49% 18% .

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3065 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 DNA chain called DNA (5'-D(*TP*AP*CP*GP*CP*AP*AP*AP*CP*GP*TP*TP*TP*GP*CP*GP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	17	345	166	62	101	16	0	0	0

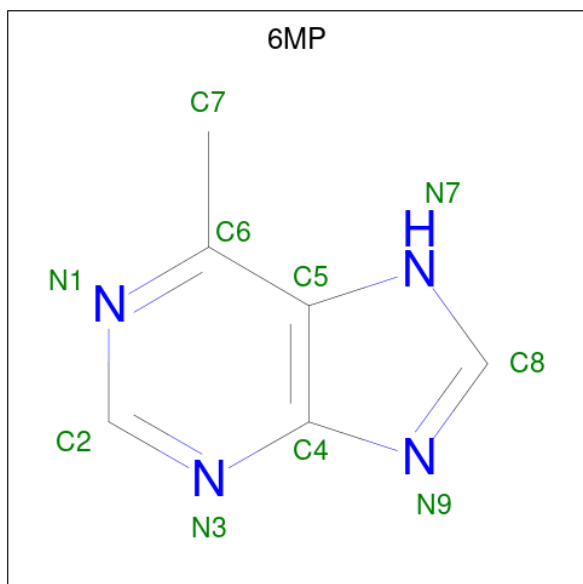
- Molecule 2 is a protein called PURINE REPRESSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	339	2651	1671	467	494	19	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	190	ALA	ARG	engineered mutation	UNP P0ACP7

- Molecule 3 is 6-METHYLPURINE (three-letter code: 6MP) (formula: C₆H₆N₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	N	0	0
			10	6	4		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	6	Total	O	0	0
			6	6		
4	A	53	Total	O	0	0
			53	53		

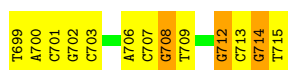
3 Residue-property plots [i](#)

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. 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. 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.

Note EDS was not executed.

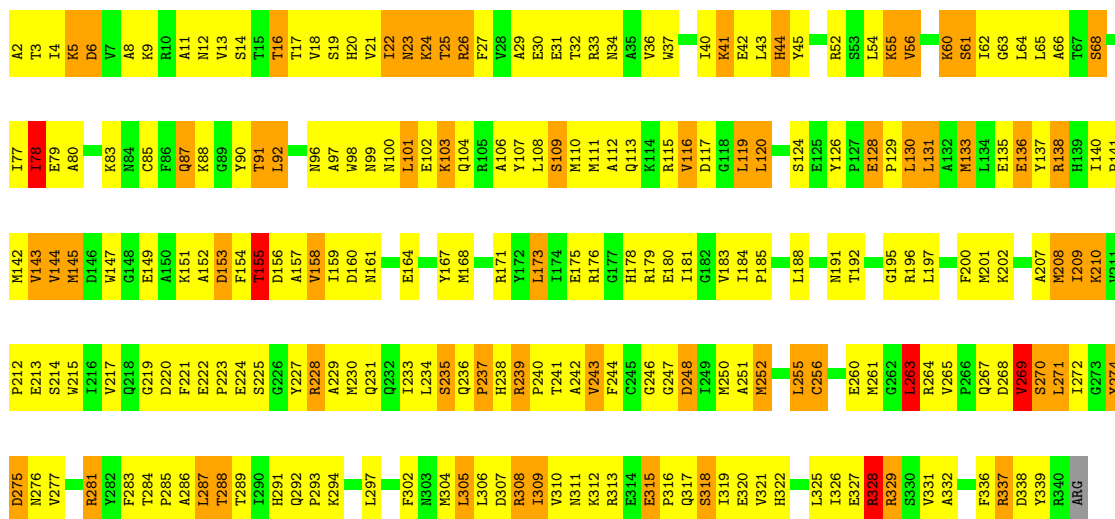
- Molecule 1: DNA (5'-D(*TP*AP*CP*GP*CP*AP*AP*AP*CP*GP*TP*TP*TP*GP*CP*GP*T)-3')

Chain B: 



- Molecule 2: PURINE REPRESSOR

Chain A: 



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	175.65Å 94.68Å 81.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.90	Depositor
% Data completeness (in resolution range)	99.0 (10.00-2.90)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	TNT	Depositor
R, R_{free}	0.166 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	3065	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

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: 6MP

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	B	1.11	4/386 (1.0%)	1.04	0/594
2	A	1.07	1/2705 (0.0%)	1.24	14/3660 (0.4%)
All	All	1.07	5/3091 (0.2%)	1.21	14/4254 (0.3%)

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
2	A	1	0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	714	DG	C8-N7	8.96	1.36	1.30
1	B	702	DG	C8-N7	8.29	1.35	1.30
1	B	712	DG	C8-N7	8.07	1.35	1.30
1	B	708	DG	C8-N7	7.98	1.35	1.30
2	A	78	ILE	CA-CB	-5.46	1.42	1.54

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	263	LEU	CA-CB-CG	-8.85	94.94	115.30
2	A	269	VAL	CB-CA-C	-7.18	97.76	111.40
2	A	248	ASP	CB-CG-OD2	6.69	124.32	118.30
2	A	287	LEU	CA-CB-CG	-6.57	100.19	115.30
2	A	173	LEU	CA-CB-CG	-6.40	100.59	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	328	ARG	CA

There are no planarity outliers.

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	B	345	0	194	20	0
2	A	2651	0	2633	306	0
3	A	10	0	6	0	0
4	A	53	0	0	7	0
4	B	6	0	0	0	0
All	All	3065	0	2833	323	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 55.

The worst 5 of 323 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)
2:A:159:ILE:HD11	2:A:320:GLU:HG2	1.22	1.21
1:B:712:DG:H2''	1:B:713:DC:H5''	1.27	1.09
2:A:236:GLN:HB2	2:A:237:PRO:HD2	1.32	1.06
1:B:713:DC:H2''	1:B:714:DG:H5''	1.12	1.04
2:A:224:GLU:HG3	2:A:228:ARG:HD2	1.36	1.02

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
2	A	337/340 (99%)	300 (89%)	30 (9%)	7 (2%)	7 26

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	136	GLU
2	A	246	GLY
2	A	275	ASP
2	A	312	LYS
2	A	124	SER

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
2	A	278/279 (100%)	213 (77%)	65 (23%)	1 2

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	281	ARG
2	A	308	ARG
2	A	116	VAL
2	A	109	SER
2	A	315	GLU

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

Mol	Chain	Res	Type
2	A	291	HIS
2	A	292	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	A	96	ASN
2	A	100	ASN
2	A	104	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

1 ligand 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
3	6MP	A	599	-	9,11,11	3.31	3 (33%)	5,15,15	2.39	2 (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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	6MP	A	599	-	-	-	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	599	6MP	C6-N1	8.08	1.39	1.32
3	A	599	6MP	C7-C6	4.74	1.54	1.50
3	A	599	6MP	C2-N1	2.11	1.37	1.33

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	599	6MP	C7-C6-C5	4.09	128.58	121.74
3	A	599	6MP	N3-C2-N1	-2.74	124.40	128.68

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.