



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 07:13 am BST

PDB ID : 5JBJ  
Title : Crystal structure of chicken LGP2 with 5'p 12-mer dsRNA at 3.6 Å resolution  
Authors : Cusack, S.; Uchikawa, E.; Lethier, M.  
Deposited on : 2016-04-13  
Resolution : 3.58 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<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  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

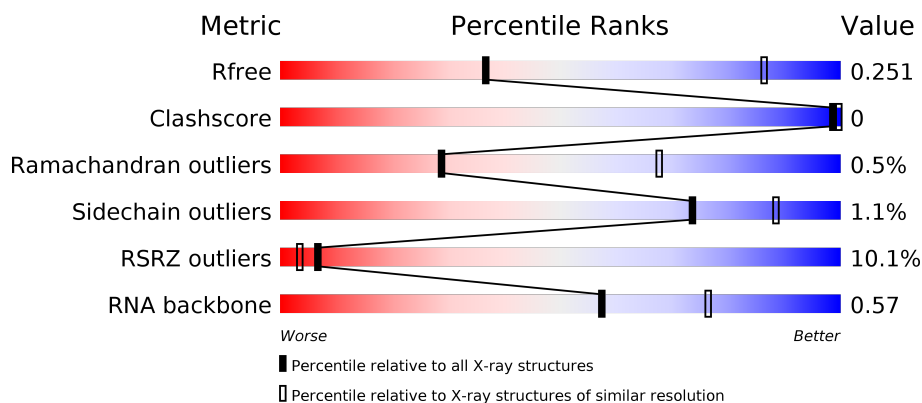
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.58 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))
$R_{free}$	130704	1094 (3.66-3.50)
Clashscore	141614	1181 (3.66-3.50)
Ramachandran outliers	138981	1143 (3.66-3.50)
Sidechain outliers	138945	1143 (3.66-3.50)
RSRZ outliers	127900	1012 (3.66-3.50)
RNA backbone	3102	1008 (4.10-3.00)

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 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	680	<div> <div>10%</div> <div> <div></div> <div>90%</div> <div>8%</div> </div> </div>
2	X	12	<div> <div>67%</div> <div>33%</div> </div>
2	Y	12	<div> <div>8%</div> <div>92%</div> <div>8%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5540 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 LGP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	627	Total	C	N	O	S	0	3	0
			5027	3146	917	931	33			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP G0YYQ5
A	-4	ALA	-	expression tag	UNP G0YYQ5
A	-3	MET	-	expression tag	UNP G0YYQ5
A	-2	GLY	-	expression tag	UNP G0YYQ5
A	-1	GLY	-	expression tag	UNP G0YYQ5
A	0	GLY	-	expression tag	UNP G0YYQ5
A	1	SER	-	expression tag	UNP G0YYQ5

- Molecule 2 is a RNA chain called RNA (5'-R(P\*GP\*GP\*UP\*AP\*GP\*CP\*GP\*CP\*UP\*AP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	12	Total	C	N	O	P	0	0	0
			256	114	46	84	12			
2	Y	12	Total	C	N	O	P	0	0	0
			256	114	46	84	12			

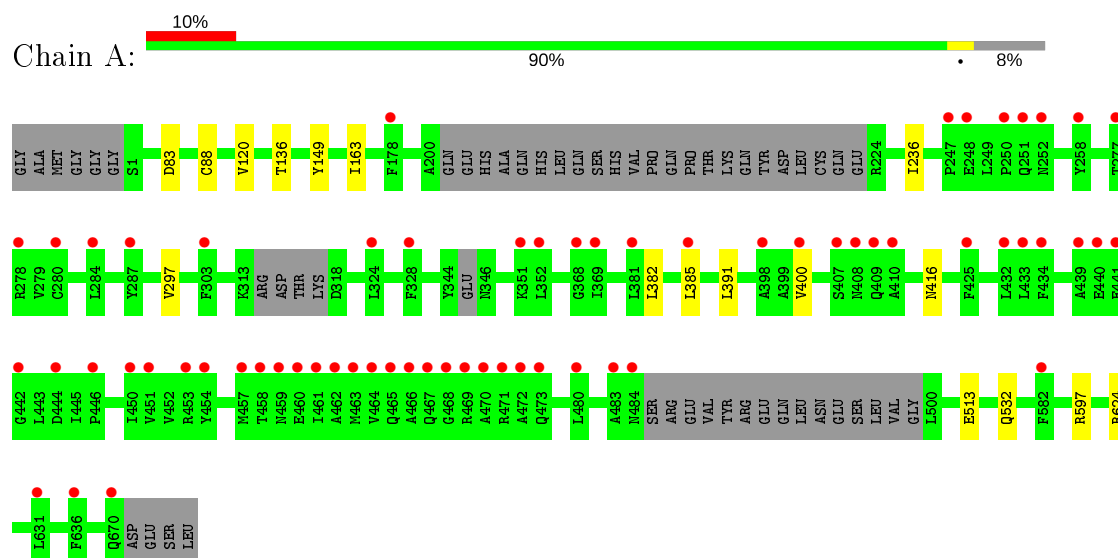
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

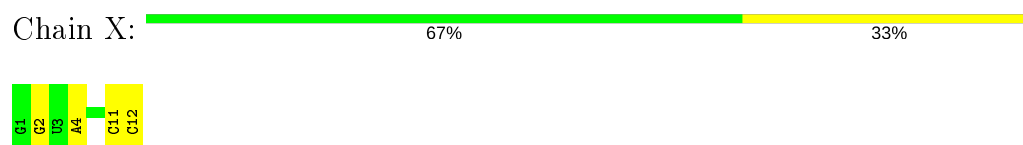
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

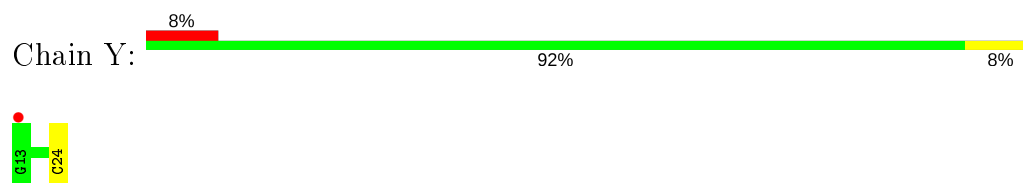
#### • Molecule 1: LGP2



#### • Molecule 2: RNA (5'-R(P\*GP\*GP\*UP\*AP\*GP\*CP\*GP\*CP\*UP\*AP\*CP\*C)-3')



#### • Molecule 2: RNA (5'-R(P\*GP\*GP\*UP\*AP\*GP\*CP\*GP\*CP\*UP\*AP\*CP\*C)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.09 Å 90.09 Å 196.73 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	78.02 – 3.58 45.05 – 3.65	Depositor EDS
% Data completeness (in resolution range)	99.8 (78.02-3.58) 99.9 (45.05-3.65)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.34 (at 3.66 Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.253 , 0.319 0.255 , 0.251	Depositor DCC
$R_{free}$ test set	509 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	165.8	Xtriage
Anisotropy	0.009	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 130.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.074 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5540	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	204.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical 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: ZN

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	0/5117	0.54	0/6903
2	X	0.21	0/285	0.67	0/442
2	Y	0.23	0/285	0.65	0/442
All	All	0.35	0/5687	0.55	0/7787

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

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	A	5027	0	4985	5	0
2	X	256	0	131	0	0
2	Y	256	0	131	1	0
3	A	1	0	0	0	0
All	All	5540	0	5247	5	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 0.

All (5) 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:385:LEU:HD12	1:A:391:LEU:HD13	1.89	0.54
1:A:236[A]:ILE:HD13	1:A:236[A]:ILE:N	2.26	0.49
1:A:382:LEU:HD22	1:A:400:VAL:HG23	1.95	0.48
1:A:149:TYR:CE2	1:A:163:ILE:HD11	2.52	0.44
1:A:597:ARG:NH2	2:Y:24:C:OP2	2.52	0.43

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	620/680 (91%)	581 (94%)	36 (6%)	3 (0%)	29	67

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	624	ARG
1	A	136	THR
1	A	297	VAL

### 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	545/589 (92%)	539 (99%)	6 (1%)	73	88

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	ASP
1	A	88	CYS
1	A	120	VAL
1	A	416	ASN
1	A	513	GLU
1	A	532	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	532	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	X	11/12 (91%)	4 (36%)	0
2	Y	11/12 (91%)	0	0
All	All	22/24 (91%)	4 (18%)	0

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	X	2	G
2	X	4	A
2	X	11	C
2	X	12	C

There are no RNA pucker outliers to report.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.



## 5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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

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	627/680 (92%)	0.39	65 (10%) 6 3	118, 199, 288, 346	0
2	X	12/12 (100%)	0.28	0 100 100	139, 158, 235, 247	0
2	Y	12/12 (100%)	0.39	1 (8%) 11 6	125, 162, 210, 218	0
All	All	651/704 (92%)	0.39	66 (10%) 7 3	118, 198, 287, 346	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	466	ALA	12.6
1	A	484	ASN	12.5
1	A	460	GLU	9.1
1	A	459	ASN	8.7
1	A	465	GLN	6.6
1	A	468	GLY	6.3
1	A	248	GLU	6.2
1	A	469	ARG	6.1
1	A	467	GLN	5.9
1	A	444	ASP	5.7
1	A	277	THR	5.4
1	A	432	LEU	5.3
1	A	473	GLN	5.2
1	A	328	PHE	4.6
1	A	463	MET	4.4
1	A	451	VAL	4.4
1	A	368	GLY	4.3
1	A	369	ILE	4.1
1	A	440	GLU	4.1
1	A	324	LEU	4.1
1	A	247	PRO	4.1
1	A	442	GLY	4.1
1	A	462	ALA	3.9

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	434	PHE	3.8
1	A	408	ASN	3.7
1	A	461	ILE	3.7
1	A	280	CYS	3.7
1	A	252	ASN	3.6
1	A	454	TYR	3.5
2	Y	13	G	3.5
1	A	433	LEU	3.4
1	A	439	ALA	3.4
1	A	251	GLN	3.3
1	A	464	VAL	3.2
1	A	381	LEU	3.2
1	A	470	ALA	3.1
1	A	441	GLU	3.1
1	A	471	ARG	3.0
1	A	483	ALA	2.9
1	A	458	THR	2.8
1	A	385	LEU	2.8
1	A	407	SER	2.7
1	A	409	GLN	2.7
1	A	631	LEU	2.7
1	A	425	PHE	2.6
1	A	287	TYR	2.6
1	A	284	LEU	2.6
1	A	446	PRO	2.5
1	A	450	ILE	2.5
1	A	670	GLN	2.5
1	A	480	LEU	2.4
1	A	303	PHE	2.4
1	A	351	LYS	2.4
1	A	453	ARG	2.3
1	A	250	PRO	2.3
1	A	410	ALA	2.2
1	A	582	PHE	2.2
1	A	472	ALA	2.2
1	A	278	ARG	2.2
1	A	178	PHE	2.2
1	A	400	VAL	2.1
1	A	457	MET	2.1
1	A	352	LEU	2.1
1	A	636	PHE	2.1
1	A	398	ALA	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	258	TYR	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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, 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	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	ZN	A	701	1/1	0.97	0.18	197,197,197,197	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.