



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 19, 2023 – 11:44 AM EDT

PDB ID : 5ES5  
Title : Crystal structure of the initiation module of LgrA in the "open" and "closed" adenylation states  
Authors : Reimer, J.M.; Aloise, M.N.; Schmeing, T.M.  
Deposited on : 2015-11-16  
Resolution : 2.80 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](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 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

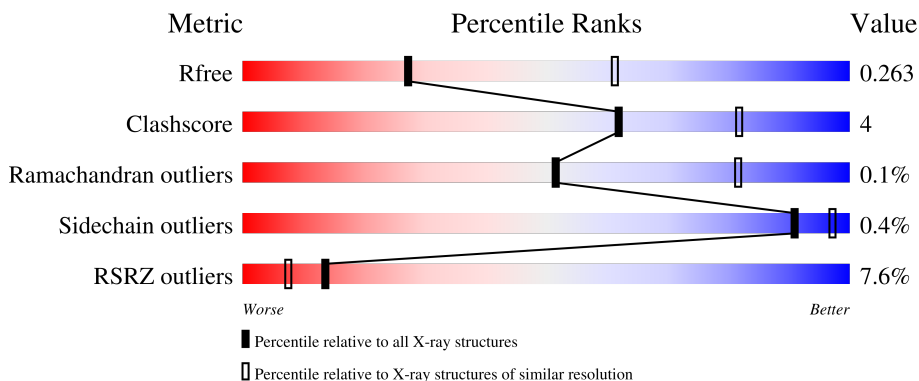
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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  $\geq 3$ , 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	776	 5% (poor fit), 81% (0-1 outliers), 7% (2 outliers), 12% (3+ outliers or not modelled)
1	B	776	 9% (poor fit), 78% (0-1 outliers), 10% (2 outliers), 12% (3+ outliers or not modelled)

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 21585 atoms, of which 10679 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 Linear gramicidin synthetase subunit A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	683	10764	3471	5324	930	1014	25	0	0	0
1	B	685	10802	3476	5355	930	1016	25	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP Q70LM7
B	2	GLY	-	expression tag	UNP Q70LM7
B	768	ALA	-	expression tag	UNP Q70LM7
B	769	ALA	-	expression tag	UNP Q70LM7
B	770	ALA	-	expression tag	UNP Q70LM7
B	771	GLU	-	expression tag	UNP Q70LM7
B	772	ASN	-	expression tag	UNP Q70LM7
B	773	LEU	-	expression tag	UNP Q70LM7
B	774	TYR	-	expression tag	UNP Q70LM7
B	775	PHE	-	expression tag	UNP Q70LM7
B	776	GLN	-	expression tag	UNP Q70LM7
A	1	MET	-	initiating methionine	UNP Q70LM7
A	2	GLY	-	expression tag	UNP Q70LM7
A	768	ALA	-	expression tag	UNP Q70LM7
A	769	ALA	-	expression tag	UNP Q70LM7
A	770	ALA	-	expression tag	UNP Q70LM7
A	771	GLU	-	expression tag	UNP Q70LM7
A	772	ASN	-	expression tag	UNP Q70LM7
A	773	LEU	-	expression tag	UNP Q70LM7
A	774	TYR	-	expression tag	UNP Q70LM7
A	775	PHE	-	expression tag	UNP Q70LM7
A	776	GLN	-	expression tag	UNP Q70LM7

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

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



VAL  
ALA  
PRO  
THR  
THR  
ASP  
GLU  
LEU  
GLU  
GLN  
GLN  
LEU  
LEU  
ALA  
GLN  
VAL  
TRP  
SER  
HIS  
SER  
VAL  
VAL  
LEU  
GLY  
LEU  
PRO  
GLN  
MET  
GLY  
ILE  
ASP  
ASP  
HIS  
PHE  
LEU  
LEU  
ARG  
GLY  
GLY  
ASP  
SER  
SER  
ILE  
LYS  
VAL  
MET  
GLN  
LEU  
ILE  
HIS  
GLN  
LEU  
LYS  
ASN  
ASN  
ILE  
GLY  
LEU  
SER  
LEU  
LEU  
ARG  
TYR  
ASP  
GLN  
LEU  
PHE

THR  
HIS  
PRO  
THR  
ILE  
ARG  
GLN  
LEU  
LEU  
ARG  
LEU  
LEU  
THR  
GLU  
GLN  
ALA  
SER  
ALA  
ALA  
GLU  
ASN  
LEU  
TYR  
PHE  
GLN

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	278.84Å 278.84Å 82.83Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.05 – 2.80 46.05 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.05-2.80) 92.3 (46.05-2.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.26 (at 2.81Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.227 , 0.263 0.227 , 0.263	Depositor DCC
$R_{free}$ test set	2961 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	77.0	Xtrriage
Anisotropy	0.313	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 86.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.014 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	21585	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	148.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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: SO4

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	0.25	0/5569	0.43	0/7564
1	B	0.26	0/5576	0.44	0/7574
All	All	0.26	0/11145	0.44	0/15138

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	5440	5324	5357	31	2
1	B	5447	5355	5362	64	1
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	A	5	0	0	0	0
3	B	4	0	0	0	0
All	All	10906	10679	10719	95	2

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 4.

The worst 5 of 95 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:B:611:GLU:C	1:B:612:LEU:HD23	1.64	1.17
1:A:199:LYS:O	1:A:510:ARG:NH1	2.03	0.91
1:A:512:ARG:NH2	1:A:561:GLU:OE2	2.14	0.80
1:B:611:GLU:O	1:B:612:LEU:HD23	1.85	0.75
1:B:231:TYR:O	1:B:235:ASN:ND2	2.23	0.72

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:637:ILE:O	1:B:682:GLN:NE2[3_765]	2.06	0.14
1:A:220:VAL:O	1:A:341:ARG:HH22[5_535]	1.51	0.09

## 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	681/776 (88%)	660 (97%)	20 (3%)	1 (0%)	51 81
1	B	683/776 (88%)	654 (96%)	28 (4%)	1 (0%)	51 81
All	All	1364/1552 (88%)	1314 (96%)	48 (4%)	2 (0%)	51 81

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	488	ILE
1	A	488	ILE

### 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	583/666 (88%)	583 (100%)	0	100	100
1	B	583/666 (88%)	578 (99%)	5 (1%)	78	94
All	All	1166/1332 (88%)	1161 (100%)	5 (0%)	91	97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	80	ASN
1	B	442	THR
1	B	472	ASN
1	B	606	ASN
1	B	612	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	235	ASN
1	B	450	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](#)

2 ligands 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	$\# Z  > 2$
2	SO4	A	801	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	B	801	-	4,4,4	0.13	0	6,6,6	0.07	0

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, 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	683/776 (88%)	0.48	35 (5%) 28 19	73, 117, 177, 194	0
1	B	685/776 (88%)	0.71	69 (10%) 7 4	86, 148, 192, 233	0
All	All	1368/1552 (88%)	0.59	104 (7%) 13 7	73, 139, 183, 233	0

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	ARG	6.3
1	B	655	ILE	6.1
1	B	646	LEU	5.9
1	B	589	LYS	5.8
1	A	37	ALA	5.3

### 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 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, 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
2	SO4	B	801	5/5	0.90	0.14	175,176,186,186	0
2	SO4	A	801	5/5	0.91	0.16	143,151,164,166	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.