



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 05:40 pm GMT

PDB ID : 1R11  
Title : Structure Determination of the Dimeric Endonuclease in a Pseudo-face-centered P21 space group  
Authors : Li, H.; Zhang, Y.  
Deposited on : 2003-09-23  
Resolution : 2.70 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

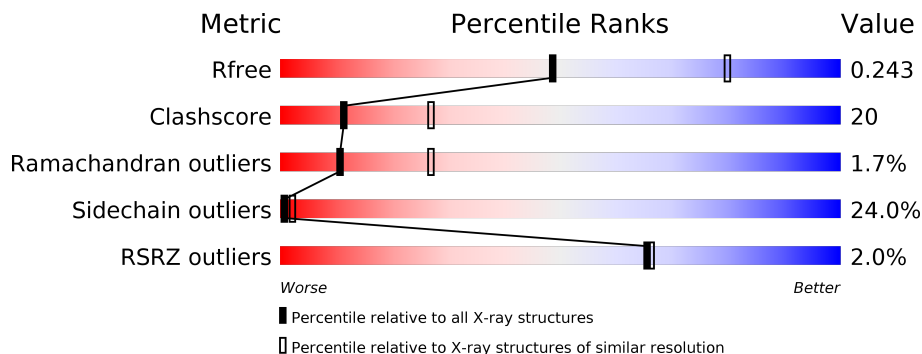
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	100719	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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. 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	305	<div> <div>2%</div> <div> <div></div> <div>39%</div> <div>39%</div> <div>16%</div> <div>5%</div> </div> </div>
1	B	305	<div> <div>2%</div> <div> <div></div> <div>49%</div> <div>34%</div> <div>14%</div> <div>..</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5042 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 tRNA-intron endonuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	11	0	0
			2521	1612	432	471	6			
1	B	303	Total	C	N	O	S	13	0	0
			2521	1612	432	471	6			

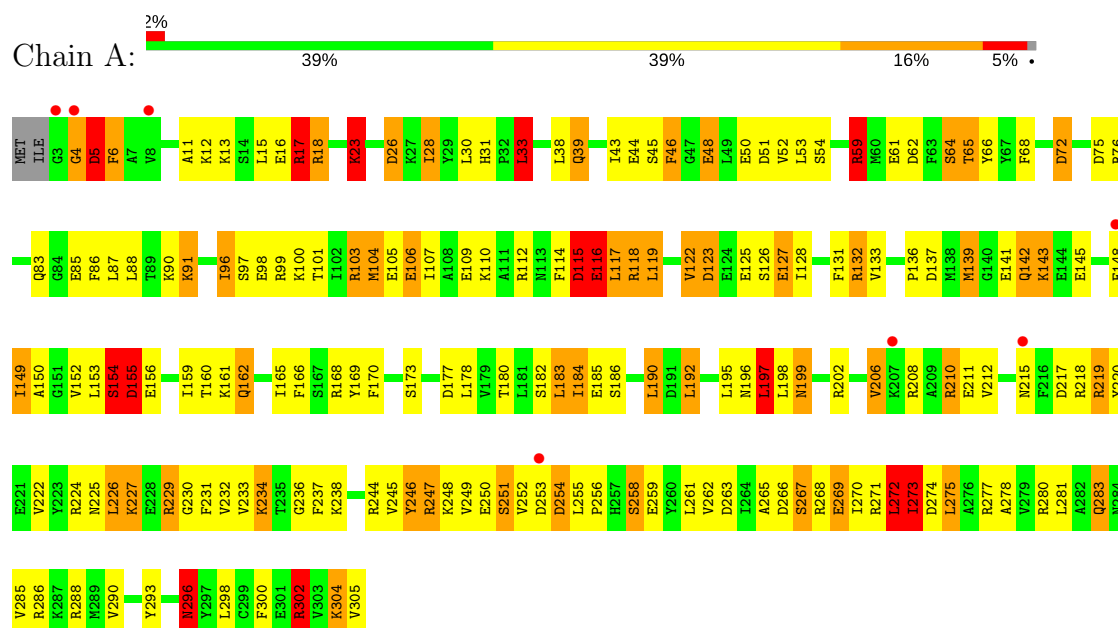
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	152	VAL	ILE	CONFLICT	UNP O29362
B	152	VAL	ILE	CONFLICT	UNP O29362

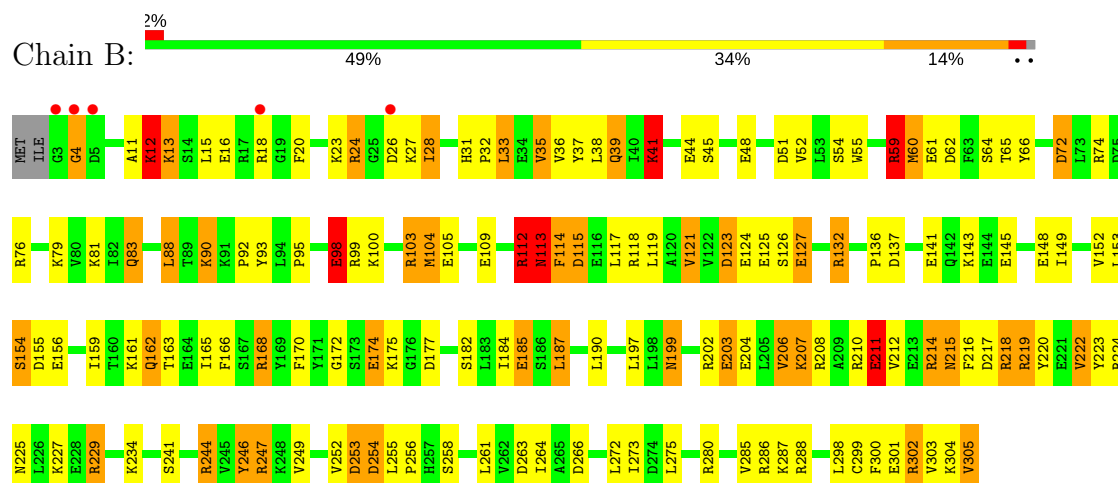
### 3 Residue-property plots

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: tRNA-intron endonuclease



#### • Molecule 1: tRNA-intron endonuclease



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.41Å 74.17Å 84.86Å 90.00° 106.13° 90.00°	Depositor
Resolution (Å)	26.66 – 2.70 26.65 – 2.69	Depositor EDS
% Data completeness (in resolution range)	(Not available) (26.66-2.70) 77.2 (26.65-2.69)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.00 (at 2.68Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.231 , 0.293 0.247 , 0.243	Depositor DCC
$R_{free}$ test set	633 reflections (4.75%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.5	Xtriage
Anisotropy	0.356	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 24.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.048 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	5042	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.44% 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

### 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	1.47	31/2565 (1.2%)	1.74	72/3440 (2.1%)
1	B	1.47	19/2565 (0.7%)	1.61	45/3440 (1.3%)
All	All	1.47	50/5130 (1.0%)	1.68	117/6880 (1.7%)

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	A	1	9
1	B	2	1
All	All	3	10

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	132	ARG	CZ-NH1	24.80	1.65	1.33
1	A	4	GLY	C-N	18.99	1.77	1.34
1	B	41	LYS	CE-NZ	-18.96	1.01	1.49
1	B	137	ASP	CB-CG	-13.52	1.23	1.51
1	A	115	ASP	C-N	12.02	1.61	1.34

The worst 5 of 117 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	132	ARG	NE-CZ-NH1	25.75	133.17	120.30
1	A	247	ARG	NE-CZ-NH1	25.48	133.04	120.30
1	A	127	GLU	OE1-CD-OE2	-20.03	99.27	123.30
1	B	41	LYS	CD-CE-NZ	19.40	156.31	111.70
1	B	286	ARG	NE-CZ-NH2	19.14	129.87	120.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	46	PHE	CA
1	B	113	ASN	CA
1	B	115	ASP	CA

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	115	ASP	Sidechain
1	A	116	GLU	Sidechain
1	A	117	LEU	Mainchain
1	A	119	LEU	Mainchain
1	A	5	ASP	Mainchain

## 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	2521	0	2529	112	3
1	B	2521	0	2534	98	6
All	All	5042	0	5063	199	6

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 20.

The worst 5 of 199 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:4:GLY:C	1:A:5:ASP:N	1.77	1.36
1:B:168:ARG:HG2	1:B:168:ARG:HH11	1.06	1.20
1:B:12:LYS:HE3	1:B:12:LYS:CA	1.75	1.10
1:B:12:LYS:HA	1:B:12:LYS:CE	1.79	1.08
1:A:48:GLU:HA	1:A:48:GLU:OE2	1.54	1.08

The worst 5 of 6 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:B:12:LYS:NZ	1:B:204:GLU:CD[1_554]	1.70	0.50
1:B:12:LYS:NZ	1:B:204:GLU:OE1[1_554]	1.75	0.45
1:A:43:ILE:O	1:B:103:ARG:NH1[2_656]	1.86	0.34
1:B:12:LYS:NZ	1:B:204:GLU:OE2[1_554]	2.05	0.15
1:A:192:LEU:O	1:B:24:ARG:NH1[1_455]	2.08	0.12

## 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	301/305 (99%)	272 (90%)	25 (8%)	4 (1%)	14	35
1	B	301/305 (99%)	267 (89%)	28 (9%)	6 (2%)	9	22
All	All	602/610 (99%)	539 (90%)	53 (9%)	10 (2%)	11	27

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	251	SER
1	A	273	ILE
1	B	113	ASN
1	A	104	MET
1	B	154	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	
1	A	271/273 (99%)	200 (74%)	71 (26%)	0	2
1	B	271/273 (99%)	212 (78%)	59 (22%)	1	3
All	All	542/546 (99%)	412 (76%)	130 (24%)	1	2

5 of 130 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	234	LYS
1	B	13	LYS
1	B	214	ARG
1	A	258	SER
1	A	286	ARG

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

Mol	Chain	Res	Type
1	A	215	ASN
1	B	284	ASN
1	B	83	GLN
1	A	162	GLN
1	B	39	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 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 ⓘ

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 [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	303/305 (99%)	0.07	7 (2%) 61 61	26, 27, 27, 28	9 (2%)
1	B	303/305 (99%)	-0.01	5 (1%) 70 72	26, 27, 27, 28	8 (2%)
All	All	606/610 (99%)	0.03	12 (1%) 65 66	26, 27, 27, 28	17 (2%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	GLY	4.3
1	B	4	GLY	3.5
1	A	4	GLY	2.9
1	B	5	ASP	2.8
1	B	26	ASP	2.6

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.