



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 12:20 PM EDT

PDB ID : 2GGJ  
Title : The mutant Y218C of Deinococcus Radiodurans N-acylamino acid racemase  
Authors : Wang, W.C.; Chiu, W.C.  
Deposited on : 2006-03-24  
Resolution : 2.50 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

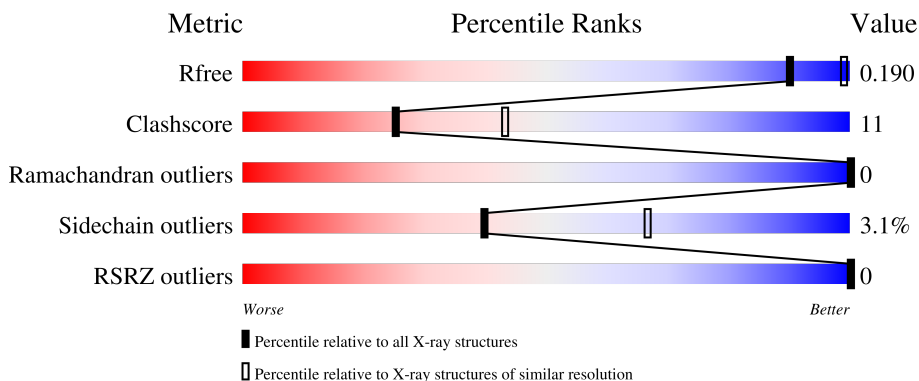
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	375	79% 14% . .
1	B	375	78% 17% . .
1	C	375	81% 14% . .
1	D	375	79% 16% . .

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 11801 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 N-acylamino acid racemase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	360	2760	1727	507	515	11	0	0	0
1	B	360	2760	1727	507	515	11	0	0	0
1	C	360	2760	1727	507	515	11	0	0	0
1	D	360	2760	1727	507	515	11	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	94	SER	ALA	SEE REMARK 999	UNP Q9RYA6
A	148	ASP	GLY	SEE REMARK 999	UNP Q9RYA6
A	158	ARG	LYS	SEE REMARK 999	UNP Q9RYA6
A	218	CYS	TYR	engineered mutation	UNP Q9RYA6
A	252	SER	ALA	SEE REMARK 999	UNP Q9RYA6
A	315	SER	PRO	SEE REMARK 999	UNP Q9RYA6
B	94	SER	ALA	SEE REMARK 999	UNP Q9RYA6
B	148	ASP	GLY	SEE REMARK 999	UNP Q9RYA6
B	158	ARG	LYS	SEE REMARK 999	UNP Q9RYA6
B	218	CYS	TYR	engineered mutation	UNP Q9RYA6
B	252	SER	ALA	SEE REMARK 999	UNP Q9RYA6
B	315	SER	PRO	SEE REMARK 999	UNP Q9RYA6
C	94	SER	ALA	SEE REMARK 999	UNP Q9RYA6
C	148	ASP	GLY	SEE REMARK 999	UNP Q9RYA6
C	158	ARG	LYS	SEE REMARK 999	UNP Q9RYA6
C	218	CYS	TYR	engineered mutation	UNP Q9RYA6
C	252	SER	ALA	SEE REMARK 999	UNP Q9RYA6
C	315	SER	PRO	SEE REMARK 999	UNP Q9RYA6
D	94	SER	ALA	SEE REMARK 999	UNP Q9RYA6
D	148	ASP	GLY	SEE REMARK 999	UNP Q9RYA6
D	158	ARG	LYS	SEE REMARK 999	UNP Q9RYA6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	218	CYS	TYR	engineered mutation	UNP Q9RYA6
D	252	SER	ALA	SEE REMARK 999	UNP Q9RYA6
D	315	SER	PRO	SEE REMARK 999	UNP Q9RYA6


- Molecule 2 is water.

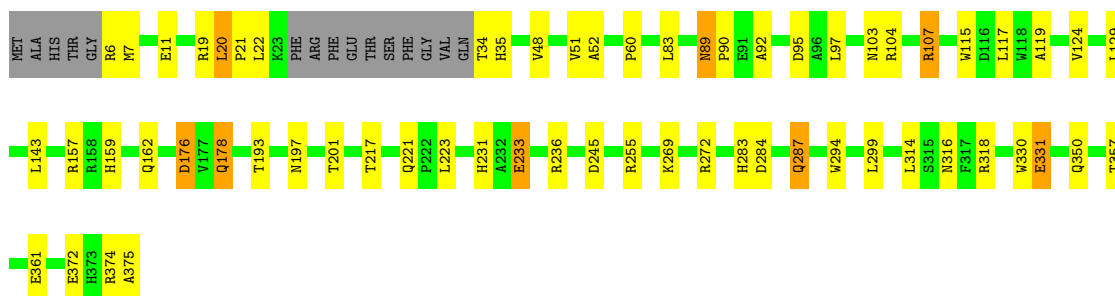
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	197	Total O 197 197	0	0
2	B	202	Total O 202 202	0	0
2	C	184	Total O 184 184	0	0
2	D	178	Total O 178 178	0	0

### 3 Residue-property plots


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.

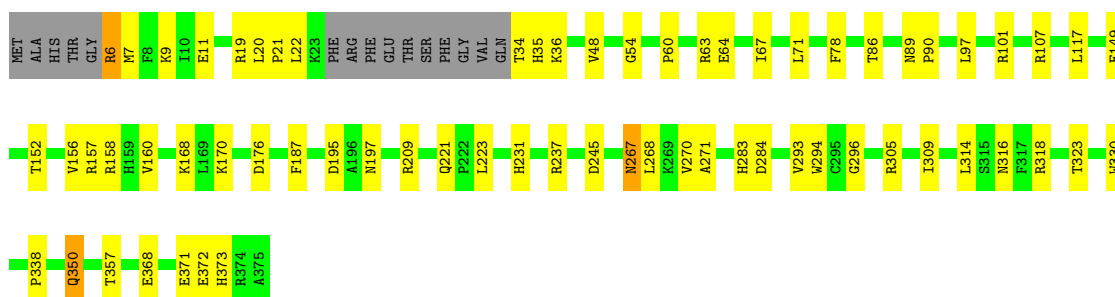
- Molecule 1: N-acylamino acid racemase

Chain A:  79% 14%




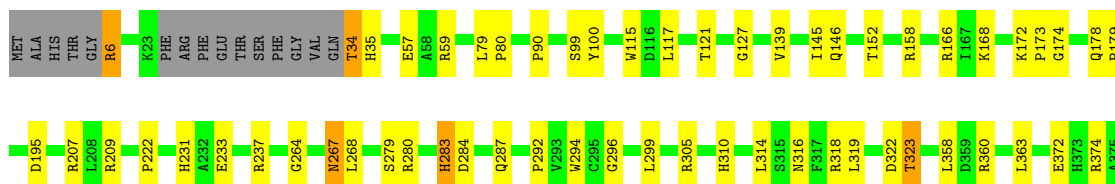
- Molecule 1: N-acylamino acid racemase

Chain B:  78% 17%




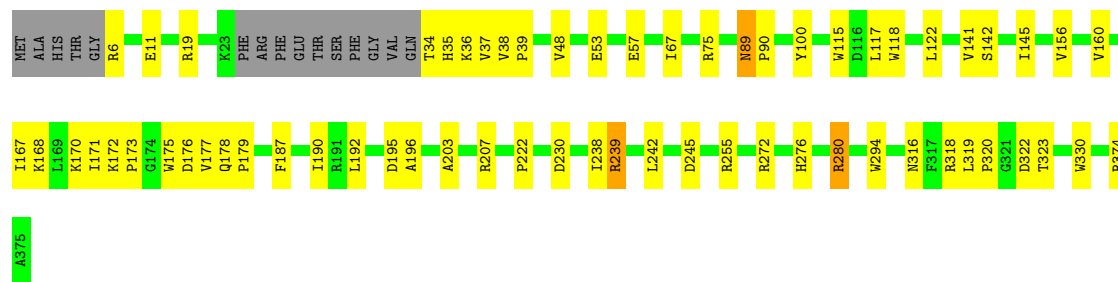
- Molecule 1: N-acylamino acid racemase

Chain C:  81% 14%



- Molecule 1: N-acylamino acid racemase

Chain D:  79% 16%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.51Å 116.51Å 120.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 29.13 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.50) 99.7 (29.13-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.87 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.162 , 0.208 0.163 , 0.190	Depositor DCC
$R_{free}$ test set	2812 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.8	Xtriage
Anisotropy	0.122	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 6.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.002 for -h,-l,-k 0.000 for -h,l,k 0.005 for l,-k,h 0.008 for -l,-k,-h 0.477 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11801	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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](#)

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	0.58	0/2810	0.70	0/3812
1	B	0.59	0/2810	0.71	1/3812 (0.0%)
1	C	0.57	0/2810	0.70	0/3812
1	D	0.59	0/2810	0.71	1/3812 (0.0%)
All	All	0.59	0/11240	0.71	2/15248 (0.0%)

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	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	101	ARG	NE-CZ-NH2	-6.19	117.20	120.30
1	D	280	ARG	NE-CZ-NH2	-5.45	117.58	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	97	LEU	Peptide

### 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	2760	0	2765	66	0
1	B	2760	0	2765	62	0
1	C	2760	0	2765	58	0
1	D	2760	0	2765	54	0
2	A	197	0	0	15	0
2	B	202	0	0	18	0
2	C	184	0	0	12	0
2	D	178	0	0	14	0
All	All	11801	0	11060	236	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 11.

The worst 5 of 236 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:294:TRP:CZ2	2:B:577:HOH:O	1.66	1.25
1:B:6:ARG:HA	2:B:418:HOH:O	1.46	1.15
1:B:11:GLU:HG2	2:B:492:HOH:O	1.50	1.11
1:C:34:THR:HB	1:C:35:HIS:HD2	1.15	1.07
1:C:323:THR:HG23	2:C:555:HOH:O	1.55	1.07

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	356/375 (95%)	337 (95%)	19 (5%)	0	100 100
1	B	356/375 (95%)	341 (96%)	15 (4%)	0	100 100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	356/375 (95%)	339 (95%)	17 (5%)	0	100	100
1	D	356/375 (95%)	345 (97%)	11 (3%)	0	100	100
All	All	1424/1500 (95%)	1362 (96%)	62 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	284/296 (96%)	270 (95%)	14 (5%)	25	47
1	B	284/296 (96%)	275 (97%)	9 (3%)	39	65
1	C	284/296 (96%)	276 (97%)	8 (3%)	43	70
1	D	284/296 (96%)	280 (99%)	4 (1%)	67	86
All	All	1136/1184 (96%)	1101 (97%)	35 (3%)	40	67

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	283	HIS
1	C	299	LEU
1	D	142	SER
1	A	357	THR
1	A	350	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	283	HIS
1	D	89	ASN
1	C	35	HIS
1	D	287	GLN
1	C	310	HIS

### 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](#)

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	360/375 (96%)	-1.12	0 100 100	4, 13, 31, 52	0
1	B	360/375 (96%)	-1.12	0 100 100	5, 13, 29, 47	0
1	C	360/375 (96%)	-1.02	0 100 100	4, 15, 34, 45	0
1	D	360/375 (96%)	-1.02	0 100 100	5, 15, 34, 46	0
All	All	1440/1500 (96%)	-1.07	0 100 100	4, 14, 33, 52	0

There are no RSRZ outliers to report.

### 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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.