



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2021 – 03:22 AM EDT

PDB ID : 2R4P  
Title : Crystal structure of the long-chain fatty acid transporter FadL mutant G212E  
Authors : Hearn, E.M.; Patel, D.R.; Lepore, B.W.; Indic, M.; van den Berg, B.  
Deposited on : 2007-08-31  
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](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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
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) : 2.23.2

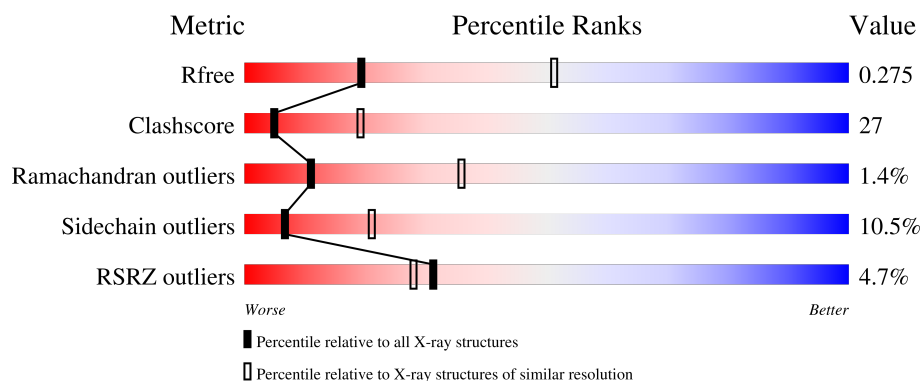
# 1 Overall quality at a glance

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))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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  $\geq 3$ , 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	427	<div> <div>2%</div> <div>53%</div> <div>39%</div> <div>5%</div> </div>
1	B	427	<div> <div>7%</div> <div>52%</div> <div>41%</div> <div>5%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6496 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.

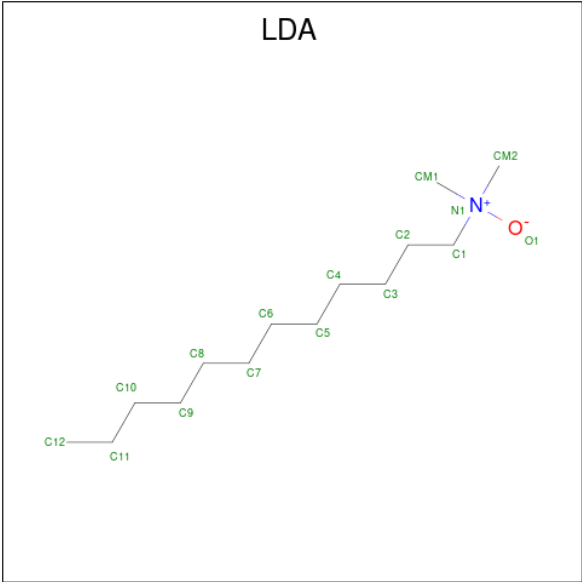
- Molecule 1 is a protein called Long-chain fatty acid transport protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	418	Total	C	N	O	S	0	0	0
			3232	2043	548	635	6			
1	B	418	Total	C	N	O	S	0	0	0
			3232	2043	548	635	6			

There are 16 discrepancies between the modelled and reference sequences:

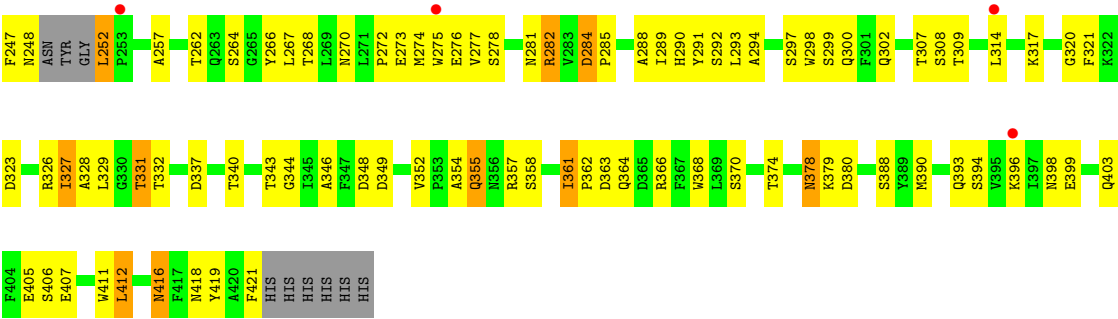
Chain	Residue	Modelled	Actual	Comment	Reference
A	212	GLU	GLY	engineered mutation	UNP P10384
A	197	THR	ILE	conflict	UNP P10384
A	422	HIS	-	expression tag	UNP P10384
A	423	HIS	-	expression tag	UNP P10384
A	424	HIS	-	expression tag	UNP P10384
A	425	HIS	-	expression tag	UNP P10384
A	426	HIS	-	expression tag	UNP P10384
A	427	HIS	-	expression tag	UNP P10384
B	212	GLU	GLY	engineered mutation	UNP P10384
B	197	THR	ILE	conflict	UNP P10384
B	422	HIS	-	expression tag	UNP P10384
B	423	HIS	-	expression tag	UNP P10384
B	424	HIS	-	expression tag	UNP P10384
B	425	HIS	-	expression tag	UNP P10384
B	426	HIS	-	expression tag	UNP P10384
B	427	HIS	-	expression tag	UNP P10384

- Molecule 2 is LAURYL DIMETHYLAMINE-N-OXIDE (three-letter code: LDA) (formula: C<sub>14</sub>H<sub>31</sub>NO).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			16	14	1	1		
2	B	1	Total	C	N	O	0	0
			16	14	1	1		





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	232.51Å 69.97Å 84.15Å 90.00° 100.84° 90.00°	Depositor
Resolution (Å)	10.00 – 2.90 38.25 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.4 (10.00-2.90) 98.4 (38.25-2.90)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.90Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.267 , 0.288 0.256 , 0.275	Depositor DCC
$R_{free}$ test set	2245 reflections (7.71%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.0	Xtriage
Anisotropy	0.315	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 59.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	6496	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.21% 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

Bond lengths and bond angles in the following residue types are not validated in this section: LDA

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	0.46	0/3319	0.66	0/4518
1	B	0.42	0/3319	0.65	0/4518
All	All	0.44	0/6638	0.65	0/9036

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

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	3232	0	3013	167	0
1	B	3232	0	3013	178	0
2	A	16	0	31	5	0
2	B	16	0	31	5	0
All	All	6496	0	6088	345	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 27.

The worst 5 of 345 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:308:SER:HB3	1:B:314:LEU:HD21	1.28	1.13
1:A:308:SER:HB3	1:A:314:LEU:HD21	1.29	1.08
1:B:58:ILE:HB	1:B:70:ALA:HB3	1.39	1.02
1:A:58:ILE:HB	1:A:70:ALA:HB3	1.41	0.98
1:B:32:ARG:HD2	1:B:228:ARG:HH22	1.29	0.97

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	414/427 (97%)	382 (92%)	26 (6%)	6 (1%)	11	36
1	B	414/427 (97%)	376 (91%)	32 (8%)	6 (1%)	11	36
All	All	828/854 (97%)	758 (92%)	58 (7%)	12 (1%)	11	36

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	247	PHE
1	A	378	ASN
1	B	247	PHE
1	B	378	ASN
1	A	2	GLY

### 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	330/338 (98%)	295 (89%)	35 (11%)	6	20
1	B	330/338 (98%)	296 (90%)	34 (10%)	7	22
All	All	660/676 (98%)	591 (90%)	69 (10%)	7	21

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	309	THR
1	B	327	ILE
1	B	361	ILE
1	A	327	ILE
1	A	317	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	137	ASN
1	B	300	GLN
1	B	138	ASN
1	B	270	ASN
1	B	303	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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

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	LDA	A	501	-	12,15,15	2.15	1 (8%)	14,17,17	1.72	4 (28%)
2	LDA	B	502	-	12,15,15	2.11	1 (8%)	14,17,17	1.52	2 (14%)

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. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	LDA	A	501	-	-	9/13/13/13	-
2	LDA	B	502	-	-	3/13/13/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	LDA	O1-N1	-7.12	1.25	1.42
2	B	502	LDA	O1-N1	-6.99	1.25	1.42

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	LDA	CM1-N1-C1	-3.99	101.85	110.23
2	B	502	LDA	CM1-N1-C1	-3.19	103.54	110.23
2	A	501	LDA	CM2-N1-C1	2.83	116.17	110.23
2	B	502	LDA	C9-C8-C7	-2.43	102.11	114.42
2	A	501	LDA	C9-C8-C7	-2.15	103.49	114.42

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	LDA	C2-C1-N1-O1
2	A	501	LDA	C2-C1-N1-CM1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	A	501	LDA	C2-C1-N1-CM2
2	A	501	LDA	C11-C10-C9-C8
2	B	502	LDA	C11-C10-C9-C8

There are no ring outliers.

2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	LDA	5	0
2	B	502	LDA	5	0

## 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	418/427 (97%)	0.12	9 (2%) 62 59	7, 31, 63, 92	0
1	B	418/427 (97%)	0.51	30 (7%) 15 11	16, 47, 90, 108	0
All	All	836/854 (97%)	0.32	39 (4%) 31 28	7, 38, 84, 108	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	175	ALA	7.1
1	B	180	GLN	5.2
1	B	181	GLY	5.2
1	B	174	PRO	4.4
1	B	184	LEU	4.4

### 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	LDA	B	502	16/16	0.80	0.41	70,71,77,78	0
2	LDA	A	501	16/16	0.86	0.36	67,70,76,76	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.