wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2023 – 01:45 PM EDT

PDB ID	:	2R4N
Title	:	Crystal structure of the long-chain fatty acid transporter FadL mutant N33A
Authors	:	Hearn, E.M.; Patel, D.R.; Lepore, B.W.; Indic, M.; van den Berg, B.
Deposited on	:	2007-08-31
Resolution	:	3.20 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

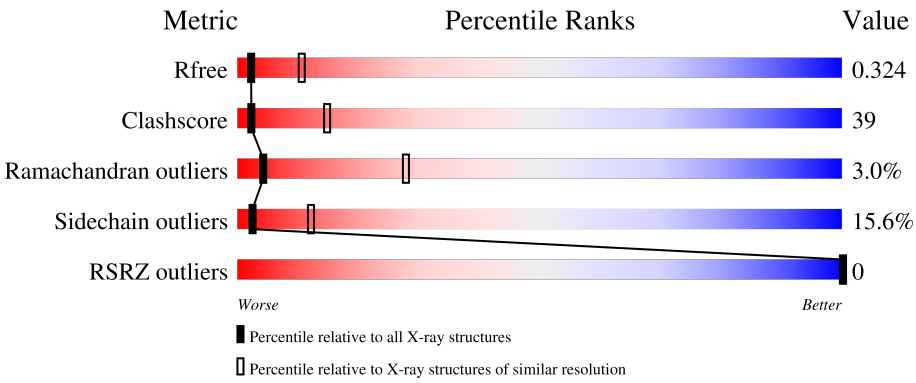
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 >=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 <=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	40% 49% 9% •
1	B	427	39% 49% 11% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	LDA	B	502	-	-	X	-

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6528 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.

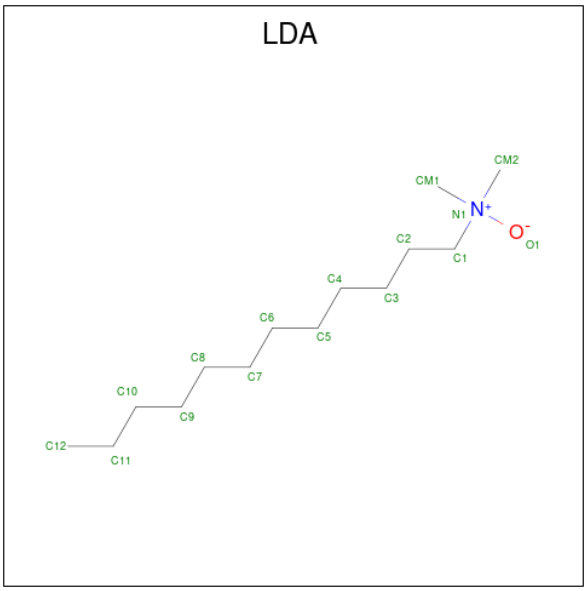
- Molecule 1 is a protein called Long-chain fatty acid transport protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	421	3248	2054	551	637	6	0	0	0
1	B	421	3248	2054	551	637	6	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	ALA	ASN	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	33	ALA	ASN	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: $C_{14}H_{31}NO$).



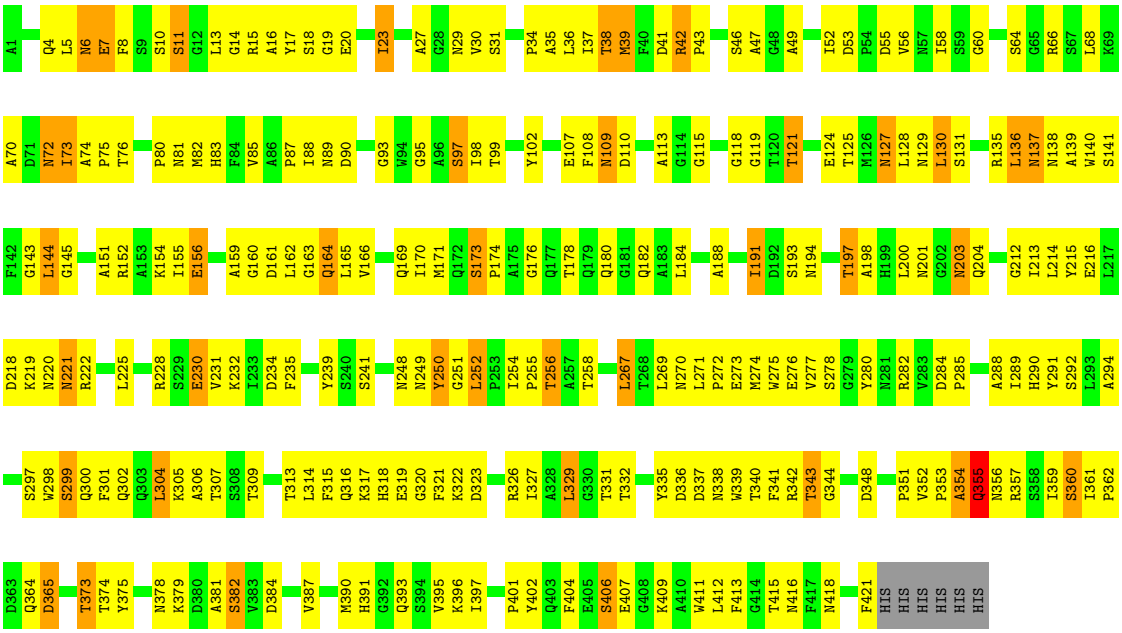
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	16	14	1	1	0	0
			Total	C	N	O		
2	B	1	16	14	1	1	0	0

3 Residue-property plots (i)

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.

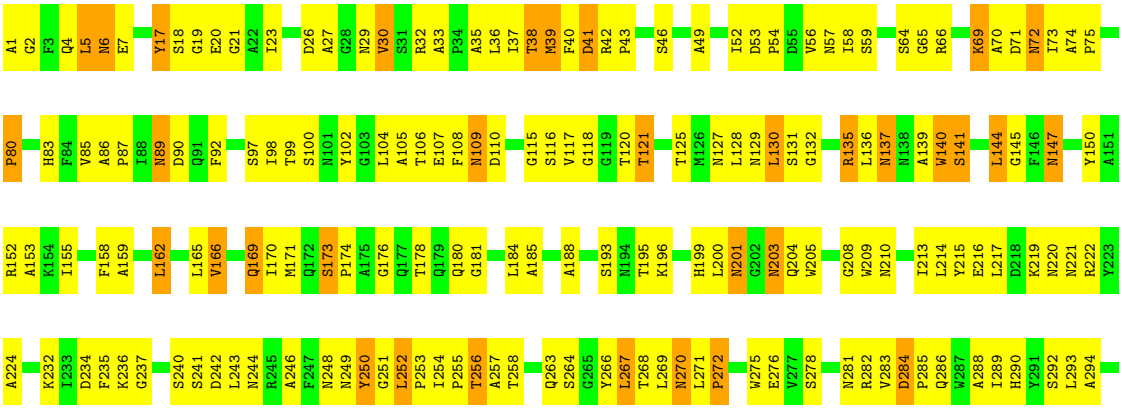
• Molecule 1: Long-chain fatty acid transport protein

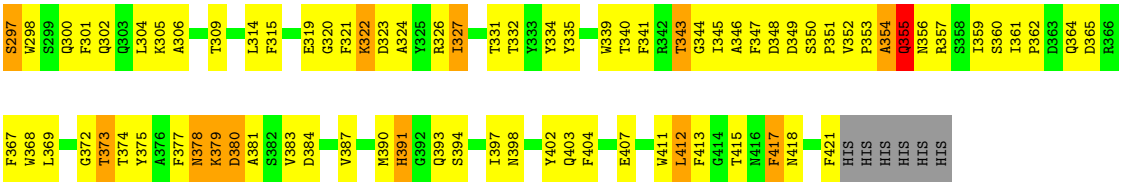
Chain A: 40% 49% 9%



• Molecule 1: Long-chain fatty acid transport protein

Chain B: 39% 49% 11%




S297 W298 S299 Q300 F301 Q302 Q303 L304 K305 A306
T309
L314 F315
E319 G320 F321 K322 D323 A324 Y325 R326 I327
T331 T332 Y333 Y334 Y335
W339 T340 F341 R342 T343 G344 I345 A346 F347 D348 D349 S350 P351 V352 P353 A354 Q355 N356 R357 S358 I359 S360 I361 P362 D363 Q364 D365 R366
F367 W368 L369
G372 T373 T374 Y375 A376 F377 N378 K379 D380 A381 S382 V383 D384
V387
M390 H391 G392 Q393 S394
I397 N398
Y402 Q403 F404
E407
W411 L412 F413 G414 T415 N416 F417 N418
F421 HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.21Å 39.99Å 175.00Å 90.00° 89.98° 90.00°	Depositor
Resolution (Å)	10.00 – 3.20 37.12 – 3.11	Depositor EDS
% Data completeness (in resolution range)	85.2 (10.00-3.20) 84.6 (37.12-3.11)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.75 (at 3.12Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.235 , 0.332 0.234 , 0.324	Depositor DCC
R_{free} test set	1170 reflections (7.65%)	wwPDB-VP
Wilson B-factor (Å^2)	32.0	Xtriage
Anisotropy	0.943	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 6.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.460 for h,-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6528	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.92% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: 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	0.47	0/3337	0.74	0/4544
1	B	0.48	0/3337	0.76	0/4544
All	All	0.48	0/6674	0.75	0/9088

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	335	TYR	Sidechain

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	3248	0	3028	254	0
1	B	3248	0	3028	237	0
2	A	16	0	31	5	0
2	B	16	0	31	10	0
All	All	6528	0	6118	489	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 39.

The worst 5 of 489 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:178:THR:HG22	1:B:180:GLN:H	1.10	1.12
1:A:178:THR:HG22	1:A:180:GLN:H	1.21	0.99
1:A:58:ILE:HB	1:A:70:ALA:HB3	1.45	0.97
1:B:152:ARG:HG2	1:B:203:ASN:HB3	1.43	0.97
1:A:164:GLN:HE21	1:A:164:GLN:N	1.66	0.94

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	419/427 (98%)	353 (84%)	55 (13%)	11 (3%)	5 31
1	B	419/427 (98%)	355 (85%)	50 (12%)	14 (3%)	4 25
All	All	838/854 (98%)	708 (84%)	105 (12%)	25 (3%)	4 28

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	38	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	354	ALA
1	B	355	GLN
1	A	39	MET
1	A	119	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	330/336 (98%)	280 (85%)	50 (15%)	3 13
1	B	330/336 (98%)	277 (84%)	53 (16%)	2 11
All	All	660/672 (98%)	557 (84%)	103 (16%)	2 12

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	69	LYS
1	B	144	LEU
1	B	391	HIS
1	B	80	PRO
1	B	130	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	180	GLN
1	B	281	ASN
1	B	189	ASN
1	B	248	ASN
1	B	300	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	LDA	A	501	-	12,15,15	2.05	1 (8%)	14,17,17	1.26	1 (7%)
2	LDA	B	502	-	12,15,15	2.14	1 (8%)	14,17,17	1.40	3 (21%)

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	-	-	4/13/13/13	-
2	LDA	B	502	-	-	5/13/13/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	502	LDA	O1-N1	-7.02	1.25	1.42
2	A	501	LDA	O1-N1	-6.91	1.26	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	502	LDA	C6-C5-C4	-2.72	100.64	114.42
2	B	502	LDA	CM2-N1-C1	2.20	114.86	110.23
2	B	502	LDA	CM1-N1-C1	-2.13	105.77	110.23
2	A	501	LDA	CM1-N1-C1	-2.09	105.85	110.23

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	502	LDA	C2-C1-N1-O1
2	B	502	LDA	C2-C1-N1-CM1
2	B	502	LDA	C7-C8-C9-C10
2	A	501	LDA	C6-C7-C8-C9
2	B	502	LDA	C2-C3-C4-C5

There are no ring outliers.

2 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	LDA	5	0
2	B	502	LDA	10	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, 95th 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^2)	Q<0.9
1	A	421/427 (98%)	-0.53	0	100	100	4, 20, 44, 84	0
1	B	421/427 (98%)	-0.47	0	100	100	4, 19, 44, 82	0
All	All	842/854 (98%)	-0.50	0	100	100	4, 20, 44, 84	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)

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, 95th 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(Å^2)	Q<0.9
2	LDA	A	501	16/16	0.92	0.42	21,21,21,21	0
2	LDA	B	502	16/16	0.93	0.33	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.