WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 22, 2024 – 02:05 AM EST

PDB ID	:	4RCN
Title	:	Structure and function of a single-chain, multi-domain long-chain acyl-coa carboxylase
Authors	:	Tran, T.H.; Tong, L.
Deposited on	:	2014-09-16
Resolution	:	3.01 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:

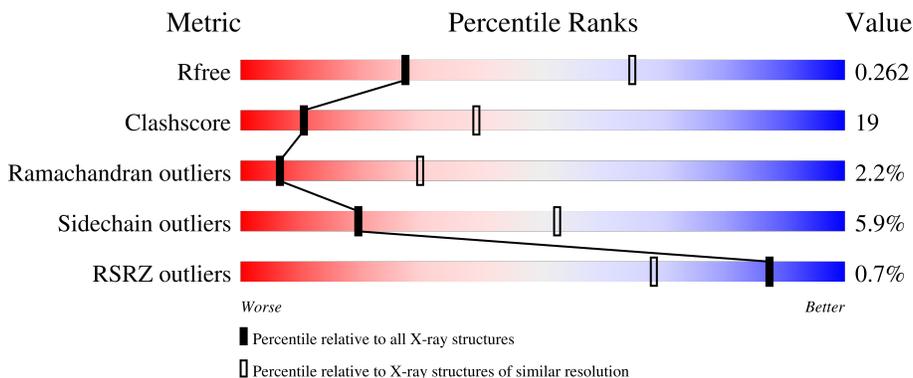|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.01 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	Percentile Ranks	Value
Rfree		0.262
Clashscore		19
Ramachandran outliers		2.2%
Sidechain outliers		5.9%
RSRZ outliers		0.7%

Worse *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2399 (3.04-3.00)
Clashscore	141614	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RSRZ outliers	127900	2287 (3.04-3.00)

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	1093	57% 30% · 11%
1	B	1093	.% 58% 31% · 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 14630 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 acyl-CoA carboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	972	Total 7211	C 4528	N 1319	O 1340	S 24	0	0	0
1	B	1000	Total 7419	C 4661	N 1354	O 1377	S 27	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1076	GLU	-	expression tag	UNP Q73VY8
A	1077	LEU	-	expression tag	UNP Q73VY8
A	1078	LEU	-	expression tag	UNP Q73VY8
A	1079	VAL	-	expression tag	UNP Q73VY8
A	1080	ASP	-	expression tag	UNP Q73VY8
A	1081	LYS	-	expression tag	UNP Q73VY8
A	1082	LEU	-	expression tag	UNP Q73VY8
A	1083	ALA	-	expression tag	UNP Q73VY8
A	1084	ALA	-	expression tag	UNP Q73VY8
A	1085	ALA	-	expression tag	UNP Q73VY8
A	1086	LEU	-	expression tag	UNP Q73VY8
A	1087	GLU	-	expression tag	UNP Q73VY8
A	1088	HIS	-	expression tag	UNP Q73VY8
A	1089	HIS	-	expression tag	UNP Q73VY8
A	1090	HIS	-	expression tag	UNP Q73VY8
A	1091	HIS	-	expression tag	UNP Q73VY8
A	1092	HIS	-	expression tag	UNP Q73VY8
A	1093	HIS	-	expression tag	UNP Q73VY8
B	1076	GLU	-	expression tag	UNP Q73VY8
B	1077	LEU	-	expression tag	UNP Q73VY8
B	1078	LEU	-	expression tag	UNP Q73VY8
B	1079	VAL	-	expression tag	UNP Q73VY8
B	1080	ASP	-	expression tag	UNP Q73VY8
B	1081	LYS	-	expression tag	UNP Q73VY8
B	1082	LEU	-	expression tag	UNP Q73VY8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1083	ALA	-	expression tag	UNP Q73VY8
B	1084	ALA	-	expression tag	UNP Q73VY8
B	1085	ALA	-	expression tag	UNP Q73VY8
B	1086	LEU	-	expression tag	UNP Q73VY8
B	1087	GLU	-	expression tag	UNP Q73VY8
B	1088	HIS	-	expression tag	UNP Q73VY8
B	1089	HIS	-	expression tag	UNP Q73VY8
B	1090	HIS	-	expression tag	UNP Q73VY8
B	1091	HIS	-	expression tag	UNP Q73VY8
B	1092	HIS	-	expression tag	UNP Q73VY8
B	1093	HIS	-	expression tag	UNP Q73VY8

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: long-chain acyl-CoA carboxylase

T1052 T1053 F1054 E1055 L1056 D1057 D1058 V1059 I1060 D1061 P1062 I1069 T1070 R1071 L1072 S1073 GLY GLY GLU LEU LEU VAL ASP LYS LEU ALA ALA LEU GLU HIS HIS HIS HIS HIS

• Molecule 1: long-chain acyl-CoA carboxylase

Chain B:

%

58%   31%   9%

M1 L5 R10 G11 E12 I13 A14 L15 R16 R19 T20 E23 L24 G25 M26 V31 Y32 A33 A34 D35 D36 P40 M49 P50 L51 P52 G53 S54 G55 P56 Y59 L60 A64 L65 V68 L76 Y81 G82 F83 E86 K92 T100 P104 D105

G106 R107 L111 N114 K115 T116 A117 A118 R119 A122 I123 V127 P128 T129 L130 T133 D134 G135 P136 S137 G138 V139 E140 D141 I142 E143 A144 F145 F146 A147 ALA HIS PRO GLY G152 I153 M154 I155 K156 ALA LEU ALA GLY GLY GLY ARG G165 M166 R167 K168 V169 H170 ASN ALA

ALA GLN ILE ALA GLY ALA TYR GLN ARG CYS ALA ALA GLU ALA ARG LEU GLY PHE GLY ASP PRO ALA L195 F196 A197 L200 L201 G202 D203 A204 R205 Q210 A216 H221 R228 D229 C230 S231 M232 Q233 R234 R235 Y236 Q237 E241 I242 Q246 G247 L248 L252 R253 R254

A255 L256 H257 Q258 V261 R262 L263 V267 R270 G271 L272 A273 T274 V275 E276 F277 L278 V279 S280 R283 L287 E288 V289 N290 P291 R292 I293 Q294 V295 E296 V299 Q312 I315 E319 L324 G325 L326 V337 E340 P341 A342 G346 Q350 L351 R352

A355 A360 S363 A373 P380 R383 V384 D385 T386 Y387 G388 L392 V393 Q397 Y398 D399 I406 V407 A408 V409 H410 G411 R415 R419 K420 L425 S426 E427 F428 E431 G432 V433 I437 R441 R442 Q447 R448 T451 G452 W453 V454 N455 V459

A468 A469 A470 LEU THR HIS GLN PRO GLY PRO GLN GLY ALA V482 G487 Q494 L495 T498 V499 V502 A503 P504 E505 A511 V516 V517 L518 E519 A520 M521 V526 P530 D531 A532 L533 V536 R537 V538 L539 V540 T541 P542 G543 Q544 V545 V546 G547 D550

T556 R557 T558 THR GLY GLY ASP ASP GLY GLY GLU SER GLU THR ALA ALA THR ASP LEU ASP LEU T576 D579 L580 R595 P596 D597 A598 V599 R602 R607 R608 T609 E612 N613 V618 D619 P620 F623 V624 E625 I631 A632 A633 Q634 R635 R638 S639 E640 E641 D642

L643 I644 A645 N646 T647 P648 A649 D650 G651 L652 A657 A671 S675 T679 A682 G683 T684 Q685 G686 M687 R688 N689 H690 A691 K692 T693 K703 R704 L705 P706 F710 A711 E712 G713 R717 P718 T721 ASP VAL GLY G725 V731 P732 T733 M736 R743 L746

S751 G752 R753 C754 F755 V764 C765 P772 G777 I784 V791 Y792 P793 P794 E795 A796 I797 G798 P799 Q803 R804 R805 N806 G810 A813 R814 D815 E816 A817 H818 A819 V820 S821 L822 A823 K824 Q825 Y826 L827 S828 Q831 W837 Q838 A839 P840 D841 P842 R843 V844

A845 R846 H847 V848 V849 P850 R853 L854 R855 A862 D868 S871 V872 L873 R876 Y879 G880 V884 T885 R889 V890 E891 G896 L897 I898 H904 L905 A908 A914 D915 K916 A917 G918 L921 E925 L934 M941 V942 G943 E947 T948 A951

F955 M958 L961 L965 L969 G970 M971 L974 R975 A986 R991 A992 P993 V997 A998 W999 P1000 G1002 E1010 R1014 Y1017 R1018 LYS GLU LEU ALA ALA VAL T1026 D1026 P1027 A1028 E1029 R1030 Q1031 N1032 L1033 R1036 L1037 V1038 A1039 T1052 D1057 D1058

V1059 I1060 D1061 P1062 R1066 I1069 T1070 R1071 L1072 S1073 GLY GLY GLU LEU LEU VAL ASP LYS LEU ALA ALA LEU GLU HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	220.88Å 220.88Å 220.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.39 – 3.01 49.39 – 3.01	Depositor EDS
% Data completeness (in resolution range)	91.0 (49.39-3.01) 91.1 (49.39-3.01)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 3.01Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.209 , 0.262 0.209 , 0.262	Depositor DCC
R_{free} test set	3516 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	63.4	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 54.4	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.026 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	14630	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.77% 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 ⓘ

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	0.37	0/7339	0.66	0/9980
1	B	0.39	0/7550	0.66	1/10259 (0.0%)
All	All	0.38	0/14889	0.66	1/20239 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	896	GLY	N-CA-C	-5.05	100.47	113.10

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	7211	0	7244	275	0
1	B	7419	0	7452	277	0
All	All	14630	0	14696	545	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 19.

The worst 5 of 545 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:632:ALA:H	1:A:647:THR:HG21	1.27	1.00
1:A:504:PRO:HD2	1:A:509:TYR:OH	1.65	0.96
1:A:210:GLN:HE21	1:A:228:ARG:HH12	1.03	0.94
1:B:753:ARG:HG2	1:B:753:ARG:HH11	1.34	0.92
1:B:92:LYS:HE2	1:B:107:ARG:NH1	1.84	0.92

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	958/1093 (88%)	870 (91%)	64 (7%)	24 (2%)	5	26
1	B	984/1093 (90%)	893 (91%)	73 (7%)	18 (2%)	8	35
All	All	1942/2186 (89%)	1763 (91%)	137 (7%)	42 (2%)	6	30

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	LEU
1	A	520	ALA
1	A	712	GLU
1	B	360	ALA
1	B	520	ALA

5.3.2 Protein sidechains ⓘ

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	718/796 (90%)	673 (94%)	45 (6%)	18	49
1	B	738/796 (93%)	697 (94%)	41 (6%)	21	54
All	All	1456/1592 (92%)	1370 (94%)	86 (6%)	19	52

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	283	ARG
1	B	539	LEU
1	B	293	ILE
1	B	397	GLN
1	B	743	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	206	HIS
1	B	904	HIS
1	B	233	GLN
1	B	613	ASN
1	B	219	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 ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	972/1093 (88%)	-0.40	5 (0%)	91	75	35, 63, 96, 139	0
1	B	1000/1093 (91%)	-0.38	8 (0%)	86	65	34, 61, 101, 138	0
All	All	1972/2186 (90%)	-0.39	13 (0%)	87	68	34, 62, 99, 139	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1025	THR	4.0
1	B	520	ALA	3.4
1	A	1026	ASP	3.2
1	B	1026	ASP	2.8
1	B	195	LEU	2.7

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK