Full wwPDB X-ray Structure Validation Report (i)

Aug 20, 2023 – 08:13 AM EDT

PDB ID	:	2IIK
Title	:	Crystal Structure of human peroxisomal acetyl-CoA acyl transferase 1 (ACAA1)
Authors	:	Papagrigoriou, E.; Johansson, C.; Smee, C.; Kavanagh, K.; Pike, A.C.W.; Sundstrom, M.; Weigelt, J.; Edwards, A.; Arrowsmith, C.H.; Gileadi, O.; Gorrec, F.; Umeano, C.; von Delft, F.; Oppermann, U.; Structural Genomics Consortium (SGC)
Deposited on	:	2006-09-28
Resolution	:	2.55 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

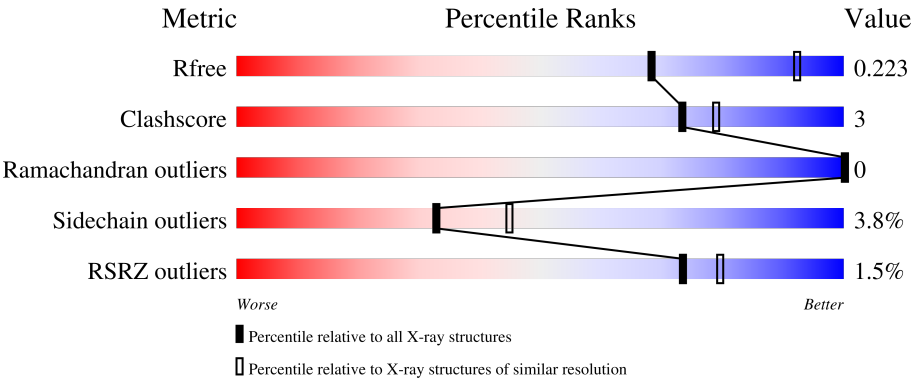
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

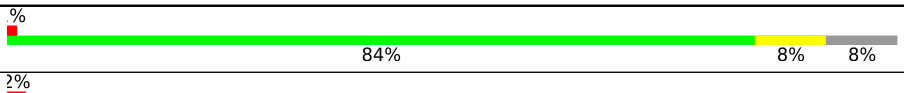
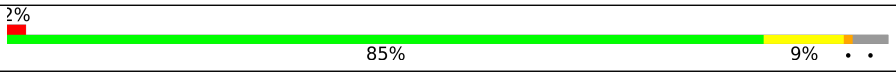
The reported resolution of this entry is 2.55 Å.

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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	418	
1	B	418	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5873 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 3-ketoacyl-CoA thiolase, peroxisomal.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	S	0	1	0
			2756	1725	476	536	19			
1	B	400	Total	C	N	O	S	0	1	0
			2888	1804	504	560	20			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	MET	-	cloning artifact	UNP P09110
A	8	HIS	-	expression tag	UNP P09110
A	9	HIS	-	expression tag	UNP P09110
A	10	HIS	-	expression tag	UNP P09110
A	11	HIS	-	expression tag	UNP P09110
A	12	HIS	-	expression tag	UNP P09110
A	13	HIS	-	expression tag	UNP P09110
A	14	SER	-	cloning artifact	UNP P09110
A	15	SER	-	cloning artifact	UNP P09110
A	16	GLY	-	cloning artifact	UNP P09110
A	17	VAL	-	cloning artifact	UNP P09110
A	18	ASP	-	cloning artifact	UNP P09110
A	19	LEU	-	cloning artifact	UNP P09110
A	20	GLY	-	cloning artifact	UNP P09110
A	21	THR	-	cloning artifact	UNP P09110
A	22	GLU	-	cloning artifact	UNP P09110
A	23	ASN	-	cloning artifact	UNP P09110
A	24	LEU	-	cloning artifact	UNP P09110
A	25	TYR	-	cloning artifact	UNP P09110
A	26	PHE	-	cloning artifact	UNP P09110
A	27	GLN	-	cloning artifact	UNP P09110
A	28	SER	-	cloning artifact	UNP P09110
A	29	MET	-	cloning artifact	UNP P09110
B	7	MET	-	cloning artifact	UNP P09110
B	8	HIS	-	expression tag	UNP P09110

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	HIS	-	expression tag	UNP P09110
B	10	HIS	-	expression tag	UNP P09110
B	11	HIS	-	expression tag	UNP P09110
B	12	HIS	-	expression tag	UNP P09110
B	13	HIS	-	expression tag	UNP P09110
B	14	SER	-	cloning artifact	UNP P09110
B	15	SER	-	cloning artifact	UNP P09110
B	16	GLY	-	cloning artifact	UNP P09110
B	17	VAL	-	cloning artifact	UNP P09110
B	18	ASP	-	cloning artifact	UNP P09110
B	19	LEU	-	cloning artifact	UNP P09110
B	20	GLY	-	cloning artifact	UNP P09110
B	21	THR	-	cloning artifact	UNP P09110
B	22	GLU	-	cloning artifact	UNP P09110
B	23	ASN	-	cloning artifact	UNP P09110
B	24	LEU	-	cloning artifact	UNP P09110
B	25	TYR	-	cloning artifact	UNP P09110
B	26	PHE	-	cloning artifact	UNP P09110
B	27	GLN	-	cloning artifact	UNP P09110
B	28	SER	-	cloning artifact	UNP P09110
B	29	MET	-	cloning artifact	UNP P09110

- Molecule 2 is water.

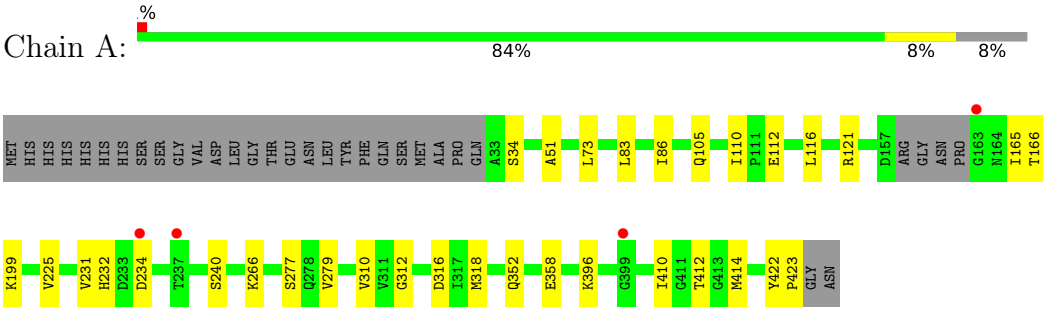
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	118	Total O 119 119	0	1
2	B	108	Total O 110 110	0	2

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.

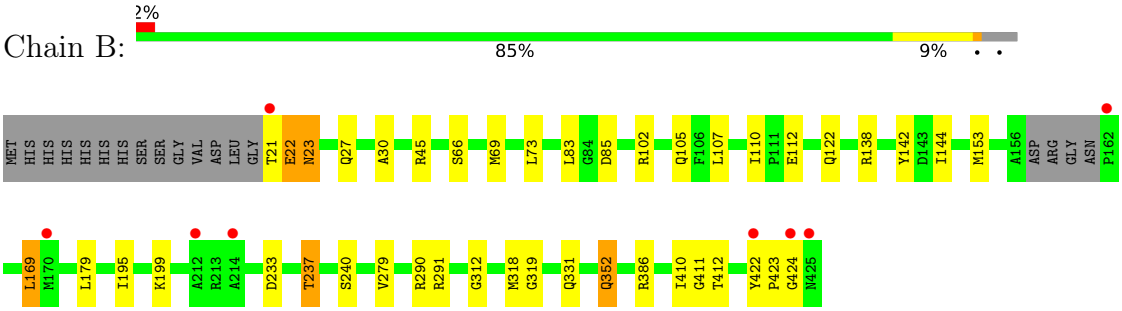
• Molecule 1: 3-ketoacyl-CoA thiolase, peroxisomal

Chain A:



• Molecule 1: 3-ketoacyl-CoA thiolase, peroxisomal

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.15Å 96.84Å 142.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	80.06 – 2.55 47.92 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.0 (80.06-2.55) 99.0 (47.92-2.55)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.76 (at 2.54Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.179 , 0.223 0.180 , 0.223	Depositor DCC
R_{free} test set	1274 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	34.2	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5873	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.86 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.2195e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.60	0/2796	0.70	1/3792 (0.0%)
1	B	0.58	0/2932	0.71	1/3972 (0.0%)
All	All	0.59	0/5728	0.70	2/7764 (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	B	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	165	ILE	CG1-CB-CG2	7.84	128.66	111.40
1	B	144	ILE	CB-CA-C	-6.32	98.97	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	424	GLY	Peptide

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	2756	0	2766	17	0
1	B	2888	0	2901	24	0
2	A	119	0	0	2	0
2	B	110	0	0	1	0
All	All	5873	0	5667	37	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 3.

All (37) 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:233:ASP:HB2	1:B:237:THR:HG22	1.68	0.74
1:B:169:LEU:HD11	1:B:179:LEU:HD11	1.77	0.67
1:A:51:ALA:HB3	1:A:279:VAL:HG12	1.77	0.67
1:B:153:MET:CE	1:B:279:VAL:HG13	2.31	0.61
1:B:169:LEU:HD11	1:B:179:LEU:CD1	2.34	0.56
1:A:191:GLU:OE2	1:A:266:LYS:NZ	2.36	0.56
1:A:422:TYR:CD1	1:A:423:PRO:HD2	2.41	0.56
1:B:195:ILE:HD12	1:B:195:ILE:N	2.22	0.55
1:A:105:GLN:HG3	1:A:110:ILE:HB	1.90	0.52
1:A:51:ALA:CB	1:A:279:VAL:HG12	2.40	0.51
1:B:45:ARG:HD2	1:B:386:ARG:HG2	1.94	0.50
1:B:21:THR:HG23	1:B:22:GLU:H	1.77	0.49
1:A:316:ASP:HB2	2:A:443:HOH:O	2.14	0.48
1:B:319:GLY:O	1:B:352:GLN:HG2	2.14	0.48
1:B:105:GLN:HG3	1:B:110:ILE:HB	1.96	0.47
1:A:199:LYS:HE3	1:A:358:GLU:OE2	2.15	0.47
1:A:192:ARG:NH2	2:A:441:HOH:O	2.47	0.47
1:B:66:SER:HA	1:B:69:MET:HE2	1.96	0.47
1:A:166:THR:HG21	1:B:107:LEU:HD11	1.98	0.46
1:A:412:THR:HB	1:B:102:ARG:HG2	1.97	0.46
1:B:352:GLN:NE2	2:B:448:HOH:O	2.49	0.46
1:B:411:GLY:O	1:B:412:THR:OG1	2.33	0.45
1:A:73:LEU:HD21	1:A:83:LEU:HD11	1.98	0.45
1:B:27:GLN:HB3	1:B:30:ALA:HB2	1.99	0.45
1:A:396:LYS:HA	1:A:422:TYR:CE1	2.52	0.44
1:B:122:GLN:HB2	1:B:410:ILE:HG23	2.00	0.43
1:A:121:ARG:NH1	1:A:310:VAL:HG11	2.34	0.43
1:B:153:MET:HE2	1:B:279:VAL:HG13	2.01	0.42
1:B:85:ASP:HB2	1:B:142:TYR:CE2	2.54	0.42
1:A:112:GLU:O	1:B:312:GLY:HA3	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:23:ASN:OD1	1:B:23:ASN:C	2.58	0.42
1:B:21:THR:HG23	1:B:22:GLU:N	2.35	0.42
1:A:410:ILE:HB	1:A:414:MET:HB2	2.02	0.41
1:B:422:TYR:CE2	1:B:423:PRO:O	2.74	0.41
1:A:312:GLY:HA3	1:B:112:GLU:O	2.21	0.41
1:A:86:ILE:O	1:A:116:LEU:HA	2.21	0.40
1:B:73:LEU:HD21	1:B:83:LEU:HD11	2.04	0.40

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	383/418 (92%)	375 (98%)	8 (2%)	0	100	100
1	B	397/418 (95%)	389 (98%)	8 (2%)	0	100	100
All	All	780/836 (93%)	764 (98%)	16 (2%)	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	281/325 (86%)	270 (96%)	11 (4%)	32	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	296/325 (91%)	283 (96%)	13 (4%)	28 38
All	All	577/650 (89%)	553 (96%)	24 (4%)	33 40

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	SER
1	A	192	ARG
1	A	225	VAL
1	A	231	VAL
1	A	232	HIS
1	A	234	ASP
1	A	240	SER
1	A	277[A]	SER
1	A	277[B]	SER
1	A	318	MET
1	A	352	GLN
1	B	22	GLU
1	B	23	ASN
1	B	138	ARG
1	B	169	LEU
1	B	199	LYS
1	B	237	THR
1	B	240	SER
1	B	290	ARG
1	B	291	ARG
1	B	318	MET
1	B	331[A]	GLN
1	B	331[B]	GLN
1	B	352	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	200	GLN

5.3.3 RNA

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

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, 95^{th} 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	386/418 (92%)	-0.27	4 (1%)	82	86	22, 26, 31, 40	0
1	B	400/418 (95%)	-0.25	8 (2%)	65	72	24, 26, 32, 44	0
All	All	786/836 (94%)	-0.26	12 (1%)	73	79	22, 26, 32, 44	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	422	TYR	3.6
1	B	21	THR	3.1
1	B	425	ASN	2.9
1	B	424	GLY	2.5
1	B	162	PRO	2.5
1	A	399	GLY	2.2
1	A	163	GLY	2.2
1	A	234	ASP	2.1
1	A	237	THR	2.1
1	B	212	ALA	2.1
1	B	214	ALA	2.1
1	B	170	MET	2.0

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.