Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 01:16 pm BST

PDB ID	:	4N5L
Title	:	Crystal structure of (R)-3-hydroxybutyryl-CoA dehydrogenase from Ralstonia eutropha
Authors	:	Kim, J.-E.; Kim, S.; Kim, K.-J.
Deposited on	:	2013-10-10
Resolution	:	1.65 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 ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

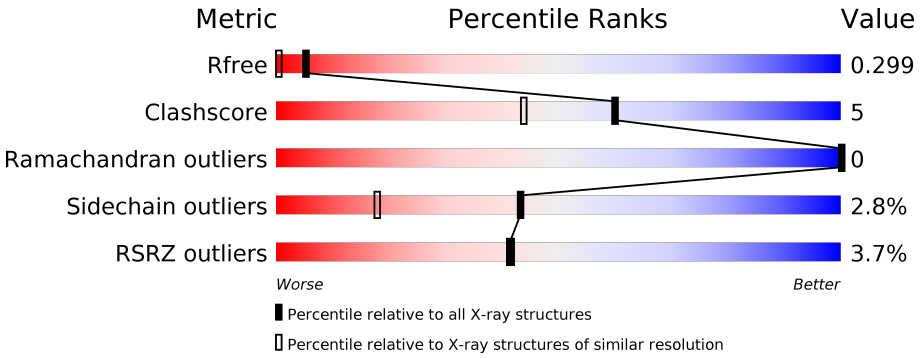
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.65 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	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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 on 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	270	4% 80% 9% • 10%
1	B	270	3% 73% 16% • 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3983 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 Acetoacetyl-CoA reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	244	Total	C	N	O	S	0	0	0
			1836	1149	324	354	9			
1	B	244	Total	C	N	O	S	0	0	0
			1836	1149	324	354	9			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	EXPRESSION TAG	UNP P14697
A	-22	SER	-	EXPRESSION TAG	UNP P14697
A	-21	TYR	-	EXPRESSION TAG	UNP P14697
A	-20	TYR	-	EXPRESSION TAG	UNP P14697
A	-19	HIS	-	EXPRESSION TAG	UNP P14697
A	-18	HIS	-	EXPRESSION TAG	UNP P14697
A	-17	HIS	-	EXPRESSION TAG	UNP P14697
A	-16	HIS	-	EXPRESSION TAG	UNP P14697
A	-15	HIS	-	EXPRESSION TAG	UNP P14697
A	-14	HIS	-	EXPRESSION TAG	UNP P14697
A	-13	ASP	-	EXPRESSION TAG	UNP P14697
A	-12	TYR	-	EXPRESSION TAG	UNP P14697
A	-11	ASP	-	EXPRESSION TAG	UNP P14697
A	-10	ILE	-	EXPRESSION TAG	UNP P14697
A	-9	PRO	-	EXPRESSION TAG	UNP P14697
A	-8	THR	-	EXPRESSION TAG	UNP P14697
A	-7	GLU	-	EXPRESSION TAG	UNP P14697
A	-6	ASN	-	EXPRESSION TAG	UNP P14697
A	-5	LEU	-	EXPRESSION TAG	UNP P14697
A	-4	TYR	-	EXPRESSION TAG	UNP P14697
A	-3	PHE	-	EXPRESSION TAG	UNP P14697
A	-2	GLN	-	EXPRESSION TAG	UNP P14697
A	-1	GLY	-	EXPRESSION TAG	UNP P14697
A	0	ALA	-	EXPRESSION TAG	UNP P14697
B	-23	MET	-	EXPRESSION TAG	UNP P14697

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-22	SER	-	EXPRESSION TAG	UNP P14697
B	-21	TYR	-	EXPRESSION TAG	UNP P14697
B	-20	TYR	-	EXPRESSION TAG	UNP P14697
B	-19	HIS	-	EXPRESSION TAG	UNP P14697
B	-18	HIS	-	EXPRESSION TAG	UNP P14697
B	-17	HIS	-	EXPRESSION TAG	UNP P14697
B	-16	HIS	-	EXPRESSION TAG	UNP P14697
B	-15	HIS	-	EXPRESSION TAG	UNP P14697
B	-14	HIS	-	EXPRESSION TAG	UNP P14697
B	-13	ASP	-	EXPRESSION TAG	UNP P14697
B	-12	TYR	-	EXPRESSION TAG	UNP P14697
B	-11	ASP	-	EXPRESSION TAG	UNP P14697
B	-10	ILE	-	EXPRESSION TAG	UNP P14697
B	-9	PRO	-	EXPRESSION TAG	UNP P14697
B	-8	THR	-	EXPRESSION TAG	UNP P14697
B	-7	GLU	-	EXPRESSION TAG	UNP P14697
B	-6	ASN	-	EXPRESSION TAG	UNP P14697
B	-5	LEU	-	EXPRESSION TAG	UNP P14697
B	-4	TYR	-	EXPRESSION TAG	UNP P14697
B	-3	PHE	-	EXPRESSION TAG	UNP P14697
B	-2	GLN	-	EXPRESSION TAG	UNP P14697
B	-1	GLY	-	EXPRESSION TAG	UNP P14697
B	0	ALA	-	EXPRESSION TAG	UNP P14697

- Molecule 2 is water.

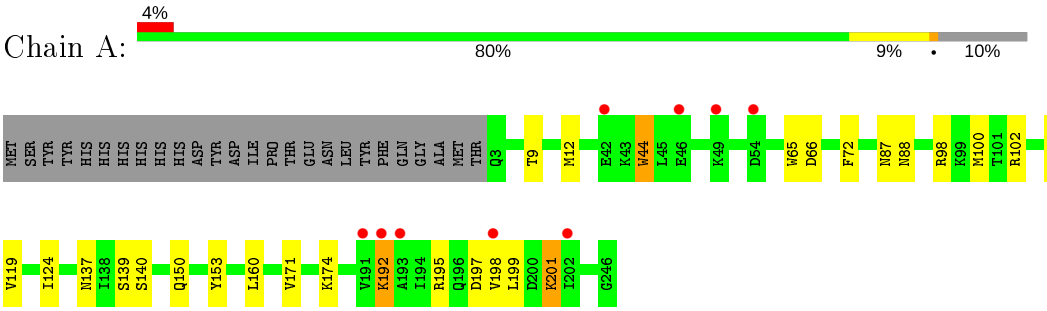
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	171	Total O 171 171	0	0
2	B	140	Total O 140 140	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

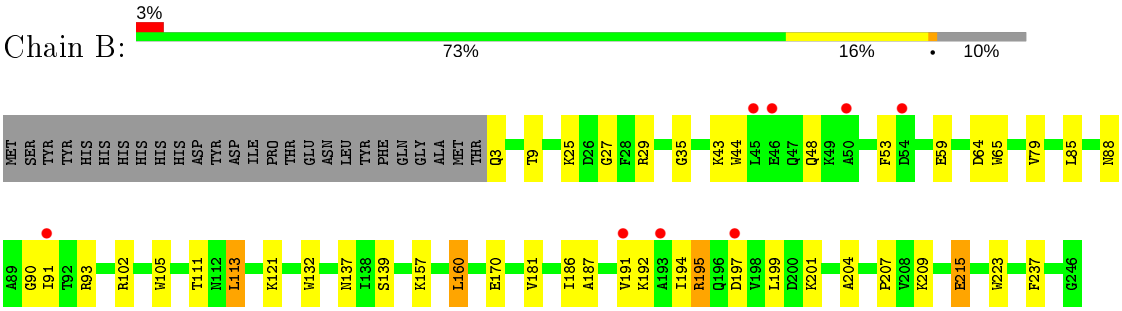
- Molecule 1: Acetoacetyl-CoA reductase

Chain A:



- Molecule 1: Acetoacetyl-CoA reductase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	87.91Å 94.34Å 137.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.18 – 1.65 32.16 – 1.65	Depositor EDS
% Data completeness (in resolution range)	87.2 (32.18-1.65) 87.2 (32.16-1.65)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	3.59 (at 1.65Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.245 , 0.292 0.256 , 0.299	Depositor DCC
R_{free} test set	3034 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	16.3	Xtriage
Anisotropy	0.705	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 35.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3983	wwPDB-VP
Average B, all atoms (Å^2)	23.0	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 63.16 % 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 1.0086e-05. 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	0.98	3/1865 (0.2%)	0.98	3/2520 (0.1%)
1	B	1.24	7/1865 (0.4%)	1.13	6/2520 (0.2%)
All	All	1.12	10/3730 (0.3%)	1.06	9/5040 (0.2%)

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

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	44	TRP	CD2-CE2	7.38	1.50	1.41
1	A	65	TRP	CD2-CE2	6.16	1.48	1.41
1	A	44	TRP	CD2-CE2	5.98	1.48	1.41
1	B	132	TRP	CD2-CE2	5.89	1.48	1.41
1	B	170	GLU	CD-OE2	-5.85	1.19	1.25
1	A	105	TRP	CD2-CE2	5.80	1.48	1.41
1	B	223	TRP	CD2-CE2	5.51	1.48	1.41
1	B	105	TRP	CD2-CE2	5.46	1.47	1.41
1	B	65	TRP	CD2-CE2	5.43	1.47	1.41
1	B	215	GLU	CD-OE2	5.34	1.31	1.25

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	102	ARG	NE-CZ-NH2	-9.80	115.40	120.30
1	A	102	ARG	NE-CZ-NH1	8.13	124.37	120.30
1	B	85	LEU	CB-CG-CD1	6.64	122.29	111.00
1	B	64	ASP	CB-CG-OD1	6.45	124.11	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	102	ARG	NE-CZ-NH2	-5.85	117.38	120.30
1	A	98	ARG	NE-CZ-NH2	-5.78	117.41	120.30
1	B	181	VAL	CG1-CB-CG2	5.73	120.06	110.90
1	B	113	LEU	CB-CG-CD1	5.47	120.29	111.00
1	B	160	LEU	CB-CG-CD1	-5.30	101.99	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	192	LYS	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	1836	0	1829	14	0
1	B	1836	0	1829	21	0
2	A	171	0	0	2	0
2	B	140	0	0	4	0
All	All	3983	0	3658	34	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 5.

All (34) 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:201:LYS:HE3	2:A:421:HOH:O	1.62	0.98
1:B:90:GLY:HA2	1:B:111:THR:HG22	1.66	0.78
1:B:201:LYS:HG3	2:B:430:HOH:O	1.84	0.76
1:B:195:ARG:HD3	1:B:197:ASP:OD1	1.92	0.69
1:B:93:ARG:HD3	2:B:349:HOH:O	1.91	0.68
1:B:43:LYS:HD2	2:B:379:HOH:O	1.92	0.68
1:A:139:SER:HB2	1:A:160:LEU:HD22	1.80	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:139:SER:HB2	1:B:160:LEU:HD22	1.84	0.59
1:B:35:GLY:O	1:B:59:GLU:HG2	2.02	0.59
1:A:124:ILE:HD11	1:A:171:VAL:HG11	1.87	0.56
1:A:192:LYS:HD3	1:A:199:LEU:HD22	1.88	0.55
1:A:174:LYS:HE2	2:A:385:HOH:O	2.07	0.54
1:B:9:THR:O	1:B:88:ASN:HB3	2.08	0.53
1:A:100:MET:O	1:B:121:LYS:HE2	2.08	0.53
1:B:139:SER:HA	1:B:157:LYS:HD2	1.92	0.52
1:A:9:THR:O	1:A:88:ASN:HB3	2.11	0.51
1:B:191:VAL:HG13	1:B:199:LEU:HD11	1.93	0.51
1:B:29:ARG:HE	1:B:79:VAL:HG13	1.75	0.50
1:B:137:ASN:CB	1:B:160:LEU:HD21	2.43	0.49
1:A:160:LEU:HD23	1:A:160:LEU:C	2.33	0.48
1:A:87:ASN:HB2	1:A:137:ASN:HD22	1.78	0.48
1:B:3:GLN:HG2	1:B:27:GLY:O	2.14	0.48
1:B:186:ILE:HG22	1:B:187:ALA:N	2.29	0.46
1:B:204:ALA:O	1:B:209:LYS:NZ	2.34	0.46
1:A:150:GLN:HB3	1:A:153:TYR:HB3	1.99	0.45
1:B:48:GLN:O	1:B:53:PHE:HB2	2.16	0.45
1:A:198:VAL:O	1:A:201:LYS:HB2	2.17	0.44
1:A:139:SER:OG	1:A:140:SER:N	2.51	0.43
1:A:12:MET:HG2	1:A:44:TRP:CD2	2.55	0.42
1:B:137:ASN:HB3	1:B:160:LEU:HD21	2.02	0.42
1:B:90:GLY:HA2	1:B:111:THR:CG2	2.44	0.42
1:A:72:PHE:CZ	1:A:119:VAL:HG13	2.55	0.41
1:B:237:PHE:CD1	1:B:237:PHE:N	2.88	0.41
1:B:91:ILE:HD12	2:B:349:HOH:O	2.20	0.41

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	242/270 (90%)	236 (98%)	6 (2%)	0	100	100
1	B	242/270 (90%)	235 (97%)	7 (3%)	0	100	100
All	All	484/540 (90%)	471 (97%)	13 (3%)	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	193/217 (89%)	189 (98%)	4 (2%)	53	29
1	B	193/217 (89%)	186 (96%)	7 (4%)	35	11
All	All	386/434 (89%)	375 (97%)	11 (3%)	43	18

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	ASP
1	A	195	ARG
1	A	197	ASP
1	A	201	LYS
1	B	25	LYS
1	B	113	LEU
1	B	192	LYS
1	B	194	ILE
1	B	195	ARG
1	B	207	PRO
1	B	215	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3	GLN
1	A	87	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	112	ASN
1	A	137	ASN
1	A	169	GLN
1	B	87	ASN
1	B	137	ASN

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 carbohydrates 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, 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	244/270 (90%)	0.10	10 (4%) 37 37	12, 21, 39, 56	0
1	B	244/270 (90%)	0.21	8 (3%) 46 47	12, 19, 35, 59	0
All	All	488/540 (90%)	0.16	18 (3%) 41 41	12, 20, 38, 59	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	193	ALA	5.9
1	B	193	ALA	3.7
1	A	198	VAL	3.3
1	B	54	ASP	3.2
1	A	49	LYS	3.1
1	A	191	VAL	3.0
1	B	197	ASP	2.8
1	A	54	ASP	2.8
1	B	46	GLU	2.6
1	A	192	LYS	2.4
1	A	202	ILE	2.4
1	B	45	LEU	2.2
1	B	191	VAL	2.2
1	B	50	ALA	2.2
1	A	46	GLU	2.1
1	A	113	LEU	2.1
1	A	42	GLU	2.1
1	B	91	ILE	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 carbohydrates 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.