WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 03:13 PM EDT

PDB ID	:	7ULI
Title	:	Apo HMG-CoA Reductase from Arabidopsis thaliana (HMG1)
Authors	:	Haywood, J.; Bond, C.S.
Deposited on	:	2022-04-05
Resolution	:	1.90 Å(reported)

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

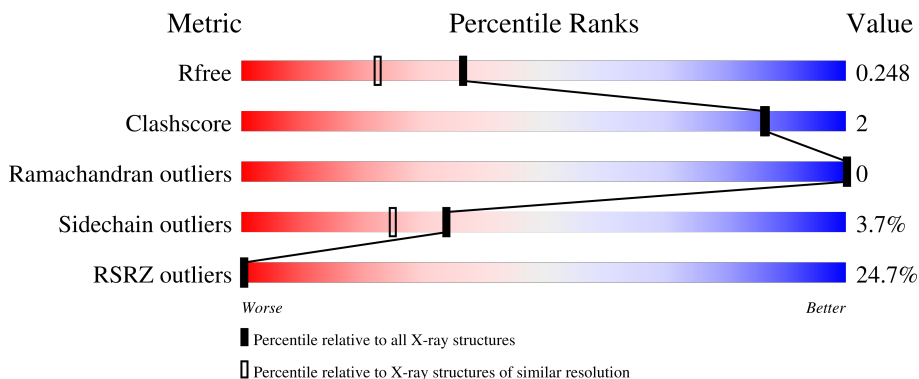|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.90 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	Percentile Ranks	Value
Rfree		0.248
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		3.7%
RSRZ outliers		24.7%

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 >=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	AAA	592	16% · 60% · · 36%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5398 atoms, of which 2644 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-hydroxy-3-methylglutaryl-coenzyme A reductase 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	377	Total 5307	C 1659	H 2644	N 462	O 518	S 24	91	1	1

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	91	Total 91	O 91	0	0

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: 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1
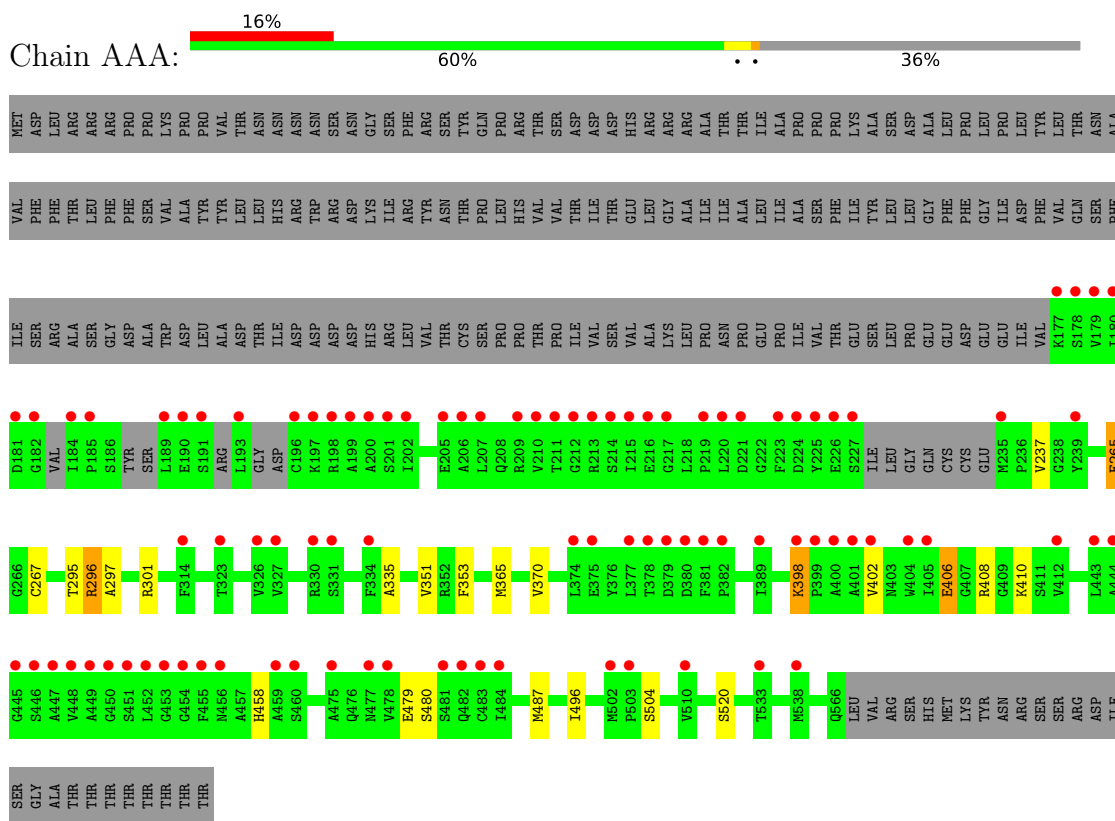
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, α, β, γ	85.58Å　85.58Å　266.65Å 90.00°　90.00°　90.00°	Depositor
Resolution (Å)	45.30　–　1.90 45.26　–　1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.30-1.90) 100.0 (45.26-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.207　,　0.243 0.218　,　0.248	Depositor DCC
R_{free} test set	2031 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	38.2	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 71.3	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5398	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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	AAA	0.75	0/2692	0.74	0/3632

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	AAA	2663	2644	2594	11	0
2	AAA	91	0	0	0	0
All	All	2754	2644	2594	11	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 2.

All (11) 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:AAA:408:ARG:O	1:AAA:504:SER:HA	2.07	0.54
1:AAA:265:GLU:HB3	1:AAA:267[A]:CYS:SG	2.51	0.50
1:AAA:402:VAL:O	1:AAA:406:GLU:HB2	2.14	0.48
1:AAA:353:PHE:CD1	1:AAA:370:VAL:HG22	2.50	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:335:ALA:HB2	1:AAA:365:MET:HE3	2.00	0.43
1:AAA:297:ALA:HA	1:AAA:351:VAL:O	2.18	0.43
1:AAA:296:ARG:HD3	1:AAA:296:ARG:HA	1.89	0.42
1:AAA:398:LYS:HB2	1:AAA:398:LYS:HE2	1.89	0.41
1:AAA:480:SER:HA	1:AAA:504:SER:O	2.20	0.41
1:AAA:295:THR:HA	1:AAA:353:PHE:O	2.21	0.41
1:AAA:487:MET:HB3	1:AAA:496:ILE:HD11	2.03	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	AAA	367/592 (62%)	356 (97%)	11 (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	AAA	273/490 (56%)	263 (96%)	10 (4%)	34	25

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	237	VAL
1	AAA	265	GLU
1	AAA	296	ARG
1	AAA	301	ARG
1	AAA	398	LYS
1	AAA	406	GLU
1	AAA	410	LYS
1	AAA	458	HIS
1	AAA	479	GLU
1	AAA	520	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	AAA	377/592 (63%)	1.13	93 (24%)	0	0	33, 47, 84, 116	0

All (93) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	451	SER	5.0
1	AAA	400	ALA	5.0
1	AAA	452	LEU	4.8
1	AAA	455	PHE	4.7
1	AAA	401	ALA	4.7
1	AAA	223	PHE	4.5
1	AAA	217	GLY	4.4
1	AAA	181	ASP	4.4
1	AAA	399	PRO	4.4
1	AAA	453	GLY	4.4
1	AAA	180	ILE	4.3
1	AAA	404	TRP	4.3
1	AAA	191	SER	4.2
1	AAA	331	SER	4.2
1	AAA	179	VAL	4.1
1	AAA	226	GLU	4.1
1	AAA	225	TYR	4.1
1	AAA	184	ILE	4.0
1	AAA	190	GLU	4.0
1	AAA	215	ILE	3.9
1	AAA	447	ALA	3.9
1	AAA	199	ALA	3.8
1	AAA	450	GLY	3.7
1	AAA	189	LEU	3.7
1	AAA	202	ILE	3.6
1	AAA	200	ALA	3.6
1	AAA	205	GLU	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	AAA	405	ILE	3.6
1	AAA	446	SER	3.6
1	AAA	483	CYS	3.6
1	AAA	454	GLY	3.5
1	AAA	402	VAL	3.5
1	AAA	448	VAL	3.5
1	AAA	206	ALA	3.5
1	AAA	196	CYS	3.5
1	AAA	227	SER	3.5
1	AAA	445	GLY	3.4
1	AAA	201	SER	3.3
1	AAA	185	PRO	3.3
1	AAA	216	GLU	3.3
1	AAA	207	LEU	3.3
1	AAA	221	ASP	3.3
1	AAA	375	GLU	3.2
1	AAA	220	LEU	3.2
1	AAA	533	THR	3.2
1	AAA	478	VAL	3.2
1	AAA	211	THR	3.1
1	AAA	182	GLY	3.0
1	AAA	389	ILE	3.0
1	AAA	398	LYS	3.0
1	AAA	477	ASN	3.0
1	AAA	449	ALA	3.0
1	AAA	177	LYS	2.9
1	AAA	212	GLY	2.9
1	AAA	443	LEU	2.9
1	AAA	459	ALA	2.9
1	AAA	198	ARG	2.9
1	AAA	235	MET	2.8
1	AAA	214	SER	2.8
1	AAA	209	ARG	2.8
1	AAA	484	ILE	2.8
1	AAA	334	PHE	2.8
1	AAA	327	VAL	2.8
1	AAA	379	ASP	2.7
1	AAA	197	LYS	2.7
1	AAA	326	VAL	2.7
1	AAA	377	LEU	2.7
1	AAA	210	VAL	2.6
1	AAA	510	VAL	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	AAA	444	ALA	2.6
1	AAA	224	ASP	2.6
1	AAA	538	MET	2.5
1	AAA	314	PHE	2.5
1	AAA	374	LEU	2.5
1	AAA	193	LEU	2.4
1	AAA	239	TYR	2.4
1	AAA	213	ARG	2.3
1	AAA	475	ALA	2.3
1	AAA	219	PRO	2.3
1	AAA	382	PRO	2.2
1	AAA	482	GLN	2.2
1	AAA	323	THR	2.2
1	AAA	378	THR	2.2
1	AAA	380	ASP	2.2
1	AAA	178	SER	2.2
1	AAA	460	SER	2.2
1	AAA	481	SER	2.1
1	AAA	412	VAL	2.1
1	AAA	330	ARG	2.1
1	AAA	503	PRO	2.1
1	AAA	502	MET	2.1
1	AAA	381	PHE	2.1
1	AAA	456	ASN	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.