WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Dec 14, 2023 – 04:31 pm GMT

PDB ID	:	2YOI
Title	:	Crystal Structure of Ancestral Thioredoxin Relative to Last Eukaryotes Common Ancestor (LECA) from the Precambrian Period
Authors	:	Gavira, J.A.; Ingles-Prieto, A.; Ibarra-Molero, B.; Sanchez-Ruiz, J.M.
Deposited on	:	2012-10-24
Resolution	:	1.30 Å(reported)

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

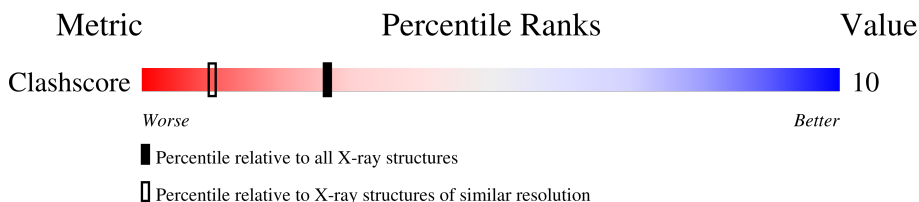|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (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 1.30 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
Clashscore		10

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))
Clashscore	141614	1101 (1.30-1.30)

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%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	106	82% 18%
1	B	106	82% 18%

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
5	ACT	A	1111	-	-	X	-

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 4357 atoms, of which 1995 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 LECA THIOREDOXIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	106	Total 1992	C 646	H 1015	N 139	O 181	S 11	0	22	0
1	B	106	Total 1925	C 626	H 977	N 137	O 176	S 9	0	17	0

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Cl 1	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total 2	Na 2	0	0
3	B	1	Total 1	Na 1	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total 2	Mg 2	0	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

ACT

OXT

O^-

C

CH3

O

O

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 7	C 2	H 3	O 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	221	Total 228	O 228	0	7
6	B	194	Total 199	O 199	0	5

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. 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. 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.

Note EDS failed to run properly.

• Molecule 1: LECA THIOREDOXIN

Chain A:
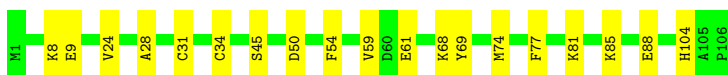
• Molecule 1: LECA THIOREDOXIN

Chain B:

4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	58.37Å　47.77Å　73.84Å 90.00°　98.49°　90.00°	Depositor
Resolution (Å)	25.58　–　1.30	Depositor
% Data completeness (in resolution range)	94.0 (25.58-1.30)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.77 (at 1.30Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.156　,　0.184	Depositor
Wilson B-factor (Å^2)	11.1	Xtriage
Anisotropy	0.006	Xtriage
L-test for twinning[2]	$< \vert L \vert > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4357	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.73% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \vert L \vert >$, $< L^2 >$ 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MG, CL, ACT, NA

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	0.54	0/1050	0.63	0/1407
1	B	0.54	0/1006	0.65	1/1351 (0.1%)
All	All	0.54	0/2056	0.64	1/2758 (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	25	ASP	CB-CG-OD1	5.38	123.14	118.30

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	977	1015	1035	15	0
1	B	948	977	992	21	0
2	A	1	0	0	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	2	0	0	0	0
5	A	4	3	3	5	0
6	A	228	0	0	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	199	0	0	8	0
All	All	2362	1995	2030	41	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 10.

All (41) 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)
5:A:1111:ACT:C	6:A:2144[A]:HOH:O	2.12	0.97
1:B:81:LYS:O	1:B:84[B]:LYS:HD2	1.71	0.89
5:A:1111:ACT:OXT	6:A:2028:HOH:O	1.95	0.83
1:A:61[A]:GLU:OE2	6:A:2146:HOH:O	1.95	0.81
5:A:1111:ACT:OXT	6:A:2144[A]:HOH:O	1.94	0.81
1:B:37[B]:ILE:CD1	1:B:94:GLN:HG2	2.10	0.80
1:B:35[B]:LYS:NZ	6:B:2091:HOH:O	2.15	0.78
1:A:61[B]:GLU:OE2	6:A:2141:HOH:O	2.02	0.78
5:A:1111:ACT:CH3	6:A:2144[A]:HOH:O	2.30	0.76
1:B:81:LYS:O	1:B:84[B]:LYS:CD	2.34	0.74
1:B:95[B]:GLU:OE2	6:B:2180:HOH:O	2.11	0.69
1:B:16:SER:OG	1:B:17[B]:GLU:OE1	2.08	0.69
1:A:61[A]:GLU:CD	6:A:2146:HOH:O	2.38	0.56
1:B:58:ASP:OD1	1:B:60[B]:ASP:OD1	2.23	0.55
1:B:36[B]:MET:HE3	1:B:36[B]:MET:HA	1.89	0.54
1:B:31[B]:CYS:SG	1:B:33:PRO:HD2	2.48	0.53
1:A:28:ALA:HB2	1:A:74[B]:MET:SD	2.49	0.53
5:A:1111:ACT:H3	6:A:2019:HOH:O	2.08	0.51
1:B:84[B]:LYS:NZ	6:B:2154:HOH:O	2.46	0.49
1:B:42[B]:GLU:OE1	6:B:2005:HOH:O	2.19	0.49
1:A:59:VAL:HG21	1:A:74[B]:MET:SD	2.53	0.48
1:B:37[B]:ILE:HD13	1:B:94:GLN:HG2	1.92	0.48
1:A:31[B]:CYS:HG	1:A:34[B]:CYS:HG	1.62	0.47
1:B:45:SER:HA	1:B:54:PHE:CE1	2.50	0.47
1:A:8[B]:LYS:HG2	1:A:9[B]:GLU:OE2	2.16	0.46
1:B:37[B]:ILE:HD12	1:B:94:GLN:HG2	1.94	0.46
1:B:17[B]:GLU:CD	6:B:2052:HOH:O	2.54	0.45
1:A:50[A]:ASP:OD1	1:A:50[A]:ASP:N	2.47	0.45
1:A:68[B]:LYS:HD3	1:A:69:TYR:CE2	2.52	0.44
1:B:42[B]:GLU:CD	6:B:2005:HOH:O	2.54	0.44
1:A:59:VAL:CG2	1:A:74[B]:MET:SD	3.05	0.44
1:B:84[B]:LYS:CE	6:B:2154:HOH:O	2.66	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:5:VAL:HG11	1:B:14:ILE:HD12	2.00	0.43
1:A:45:SER:HA	1:A:54:PHE:CE1	2.53	0.43
1:A:85:LYS:NZ	1:A:88[B]:GLU:OE1	2.42	0.43
1:B:84[B]:LYS:HE2	6:B:2154:HOH:O	2.19	0.42
1:A:8[B]:LYS:HG2	6:A:2028:HOH:O	2.20	0.41
1:A:24:VAL:O	1:A:77:PHE:HA	2.20	0.41
1:A:81[A]:LYS:CE	1:A:104:HIS:O	2.69	0.41
1:B:36[B]:MET:HE3	1:B:36[B]:MET:CA	2.51	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains ⓘ

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 ⓘ

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 for Mogul analysis.

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	$\#|Z| > 2$
5	ACT	A	1111	-	3,3,3	0.78	0	3,3,3	1.40	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1111	ACT	5	0

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 ⓘ

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

6.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.