



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 23, 2021 – 03:19 PM EST

PDB ID : 1TEL
Title : Crystal structure of a RubisCO-like protein from *Chlorobium tepidum*
Authors : Fedorov, A.A.; Fedorov, E.V.; Imker, H.J.; Gerlt, J.A.; Almo, S.C.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2004-05-25
Resolution : 2.70 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 ⓘ 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 : 2.16
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) : 2.16

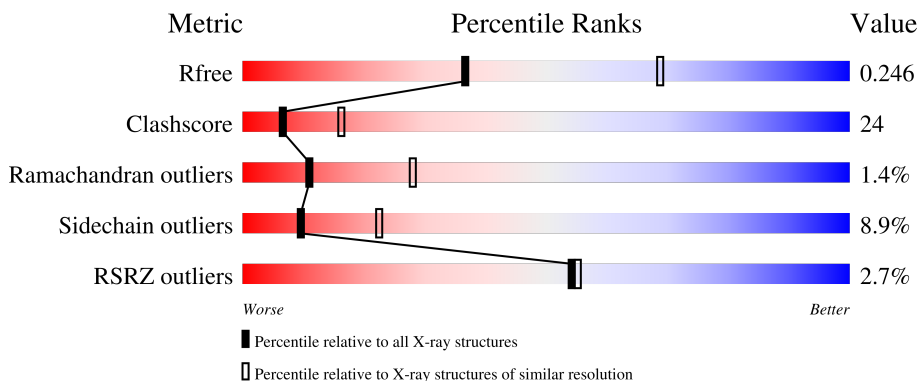
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 $\leq 5\%$. 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	435	 3% 63% 27% 7% •
1	B	435	 3% 64% 28% 6% •

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6651 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 ribulose biphosphate carboxylase, large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	424	3286	2098	552	616	20	0	0	0
1	B	428	3318	2116	558	624	20	0	0	0

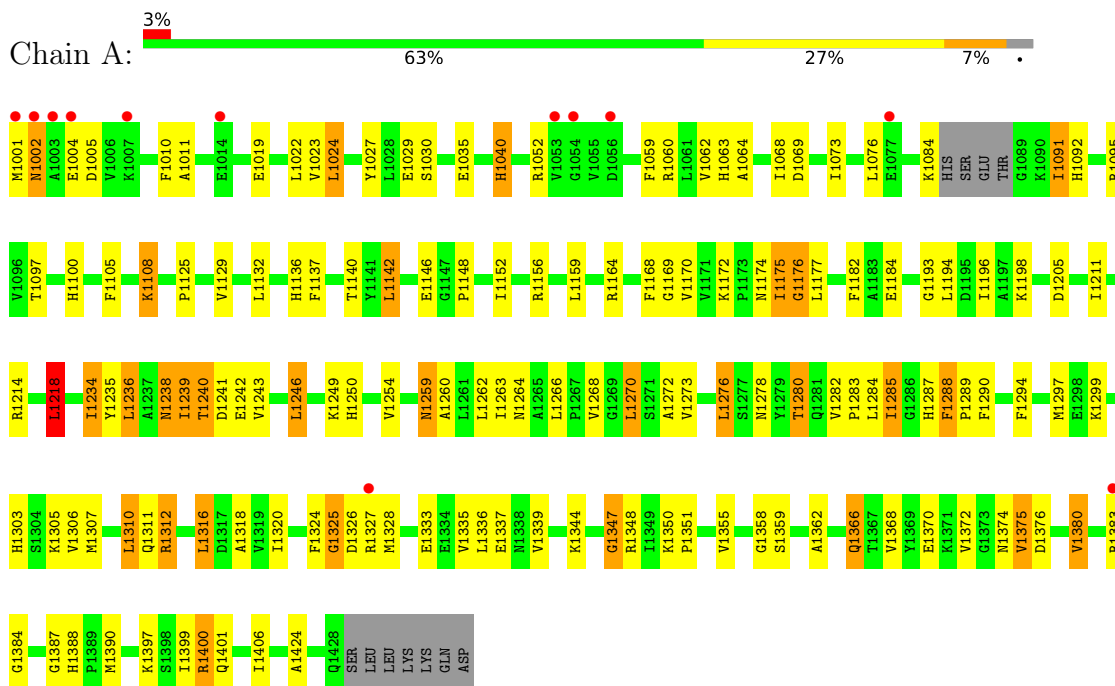
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	24	Total 24	O 24	0	0
2	B	23	Total 23	O 23	0	0

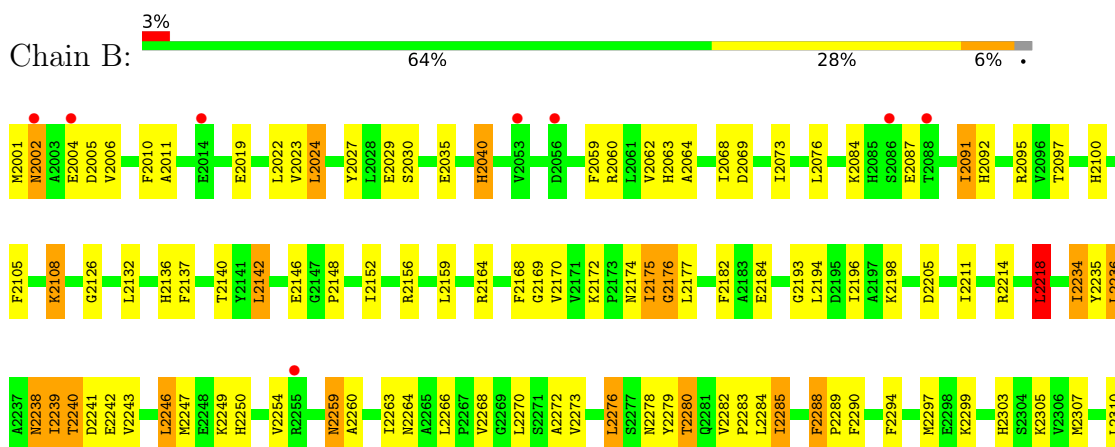
3 Residue-property plots

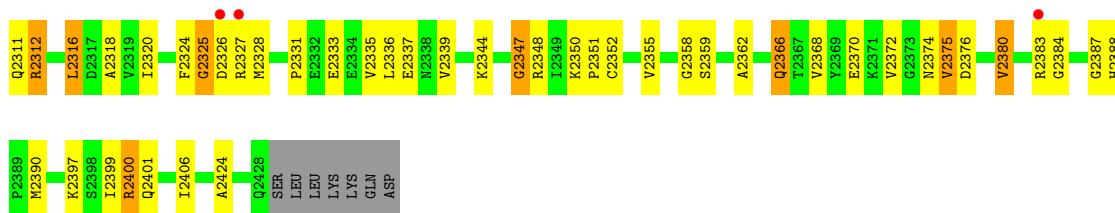
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: ribulose biphosphate carboxylase, large subunit



- Molecule 1: ribulose biphosphate carboxylase, large subunit





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	67.11Å 77.71Å 88.76Å 90.00° 99.57° 90.00°	Depositor
Resolution (Å)	25.00 – 2.70 19.91 – 2.70	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.70) 99.6 (19.91-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.01 (at 2.71Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.209 , 0.248 0.208 , 0.246	Depositor DCC
R_{free} test set	1199 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	21.3	Xtrriage
Anisotropy	0.167	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 39.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	6651	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.70% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.43	0/3363	0.69	1/4559 (0.0%)
1	B	0.43	0/3397	0.69	1/4607 (0.0%)
All	All	0.43	0/6760	0.69	2/9166 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	2218	LEU	CA-CB-CG	-5.22	103.30	115.30
1	A	1218	LEU	CA-CB-CG	-5.10	103.58	115.30

There are no chirality outliers.

There are no planarity outliers.

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	3286	0	3242	161	0
1	B	3318	0	3268	162	0
2	A	24	0	0	1	0
2	B	23	0	0	0	0
All	All	6651	0	6510	309	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 24.

The worst 5 of 309 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:2238:ASN:HD21	1:B:2240:THR:HG23	1.10	1.08
1:A:1238:ASN:HD21	1:A:1240:THR:HG23	1.07	1.05
1:B:2238:ASN:HD21	1:B:2240:THR:CG2	1.72	1.01
1:A:1076:LEU:HD12	1:A:1091:ILE:HD11	1.45	0.99
1:A:1238:ASN:HD21	1:A:1240:THR:CG2	1.74	0.98

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	420/435 (97%)	383 (91%)	31 (7%)	6 (1%)	11	28
1	B	426/435 (98%)	388 (91%)	32 (8%)	6 (1%)	11	28
All	All	846/870 (97%)	771 (91%)	63 (7%)	12 (1%)	11	28

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1175	ILE
1	A	1176	GLY
1	A	1326	ASP
1	B	2175	ILE
1	B	2176	GLY

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	350/361 (97%)	318 (91%)	32 (9%)	9	21
1	B	354/361 (98%)	323 (91%)	31 (9%)	10	23
All	All	704/722 (98%)	641 (91%)	63 (9%)	9	22

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1366	GLN
1	B	2040	HIS
1	B	2336	LEU
1	A	1375	VAL
1	A	1400	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1388	HIS
1	B	2049	GLN
1	B	2374	ASN
1	B	2002	ASN
1	B	2078	GLN

5.3.3 RNA [i](#)

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	424/435 (97%)	-0.25	12 (2%) 53 54	5, 16, 34, 55	0
1	B	428/435 (98%)	-0.23	11 (2%) 56 57	4, 16, 35, 53	0
All	All	852/870 (97%)	-0.24	23 (2%) 54 55	4, 16, 35, 55	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2053	VAL	4.1
1	A	1002	ASN	3.8
1	A	1054	GLY	3.8
1	B	2327	ARG	3.5
1	B	2004	GLU	3.4

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.