



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 14, 2020 – 04:44 am BST

PDB ID : 1MJG  
Title : CRYSTAL STRUCTURE OF BIFUNCTIONAL CARBON MONOXIDE DEHYDROGENASE / ACETYL-COA SYNTHASE(CODH/ACS) FROM MOORELLA THERMOACETICA (F. CLOSTRIDIUM THERMOACETICUM)  
Authors : Doukov, T.I.; Iverson, T.M.; Seravalli, J.; Ragsdale, S.W.; Drennan, C.L.  
Deposited on : 2002-08-27  
Resolution : 2.20 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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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) : 2.11

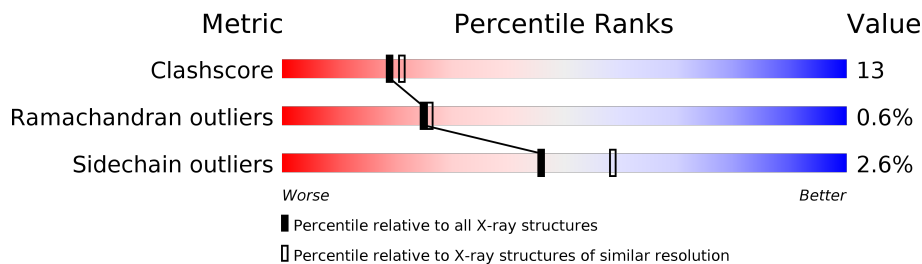
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain	
1	A	674	78%	20%
1	B	674	78%	20%
1	C	674	78%	20%
1	D	674	78%	21%
2	M	729	77%	21%
2	N	729	77%	21%
2	O	729	72%	26%
2	P	729	75%	24%

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
4	XCC	A	800	-	-	X	-
4	XCC	C	800	-	-	X	-
4	XCC	D	800	-	-	X	-
7	ACT	M	950	-	-	X	-
7	ACT	N	950	-	-	X	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 44653 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 CARBON MONOXIDE DEHYDROGENASE BETA SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	672	5087	3197	890	958	42	0	0	0
1	B	672	5087	3197	890	958	42	0	0	0
1	C	672	5087	3197	890	958	42	0	0	0
1	D	672	5087	3197	890	958	42	0	0	0

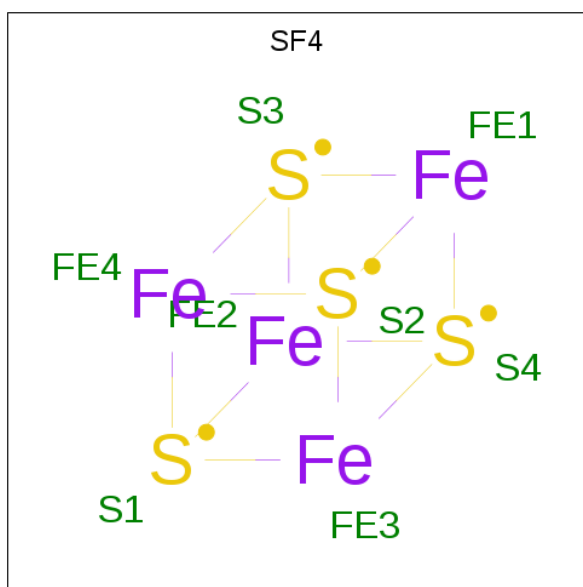
- Molecule 2 is a protein called Carbon monoxide dehydrogenase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	M	728	5735	3678	953	1069	35	0	0	0
2	N	728	5735	3678	953	1069	35	0	0	0
2	O	728	5735	3678	953	1069	35	0	0	0
2	P	728	5735	3678	953	1069	35	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

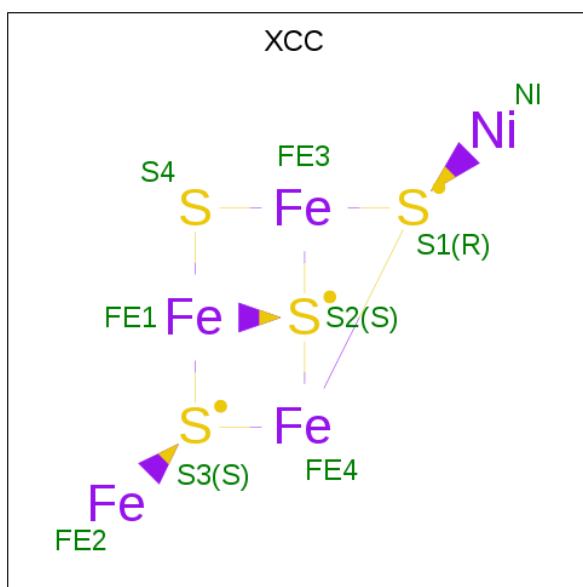
Chain	Residue	Modelled	Actual	Comment	Reference
M	685	SER	ARG	SEE REMARK 999	UNP P27988
N	685	SER	ARG	SEE REMARK 999	UNP P27988
O	685	SER	ARG	SEE REMARK 999	UNP P27988
P	685	SER	ARG	SEE REMARK 999	UNP P27988

- Molecule 3 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	Fe	S	0	0
			8	4	4		
3	A	1	Total	Fe	S	0	0
			8	4	4		
3	B	1	Total	Fe	S	0	0
			8	4	4		
3	C	1	Total	Fe	S	0	0
			8	4	4		
3	C	1	Total	Fe	S	0	0
			8	4	4		
3	D	1	Total	Fe	S	0	0
			8	4	4		
3	M	1	Total	Fe	S	0	0
			8	4	4		
3	N	1	Total	Fe	S	0	0
			8	4	4		
3	O	1	Total	Fe	S	0	0
			8	4	4		
3	P	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 4 is FE(4)-NI(1)-S(4) CLUSTER (three-letter code: XCC) (formula: Fe<sub>4</sub>NiS<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	Fe	Ni	S		
4	A	1	9	4	1	4	0	0
4	B	1	9	4	1	4	0	0
4	C	1	9	4	1	4	0	0
4	D	1	9	4	1	4	0	0

- Molecule 5 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
5	P	1	1	1	0	0
5	O	1	1	1	0	0
5	N	1	1	1	0	0
5	M	1	1	1	0	0

- Molecule 6 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

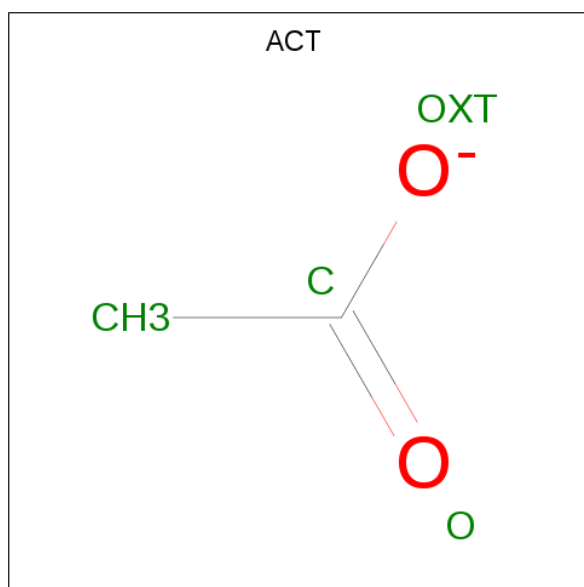
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ni		
6	P	1	1	1	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	O	1	Total Ni 1 1	0	0
6	N	1	Total Ni 1 1	0	0
6	M	1	Total Ni 1 1	0	0

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	M	1	Total C O 3 2 1	0	0
7	N	1	Total C O 3 2 1	0	0
7	O	1	Total C O 3 2 1	0	0
7	P	1	Total C O 3 2 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	179	Total O 179 179	0	0
8	B	217	Total O 217 217	0	0

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
8	C	148	Total 148	O 148	0	0
8	D	131	Total 131	O 131	0	0
8	M	206	Total 206	O 206	0	0
8	N	221	Total 221	O 221	0	0
8	O	26	Total 26	O 26	0	0
8	P	101	Total 101	O 101	0	0

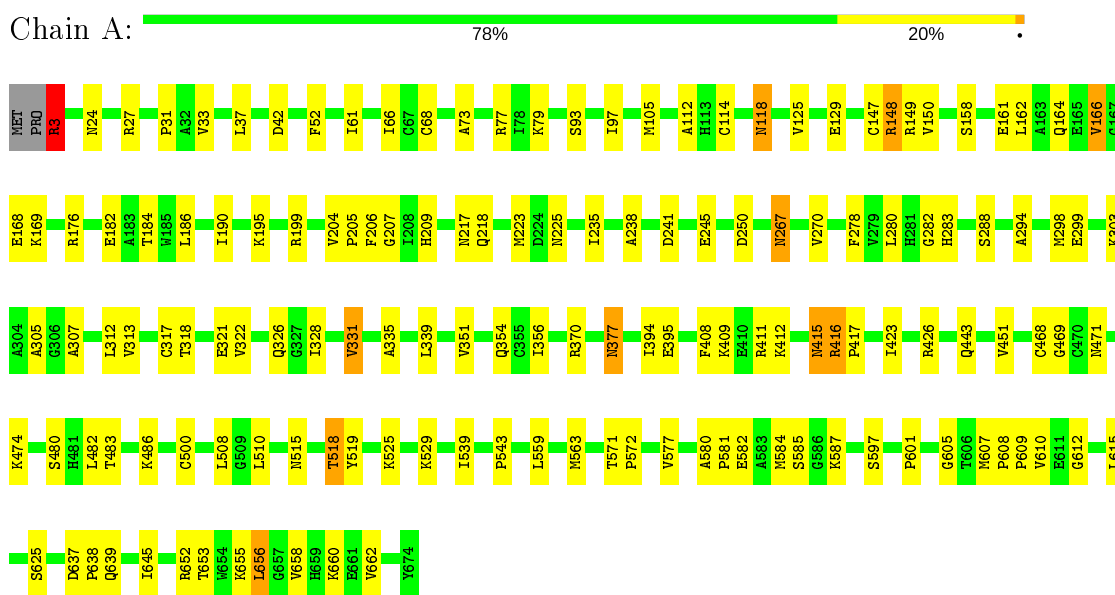


### 3 Residue-property plots [i](#)

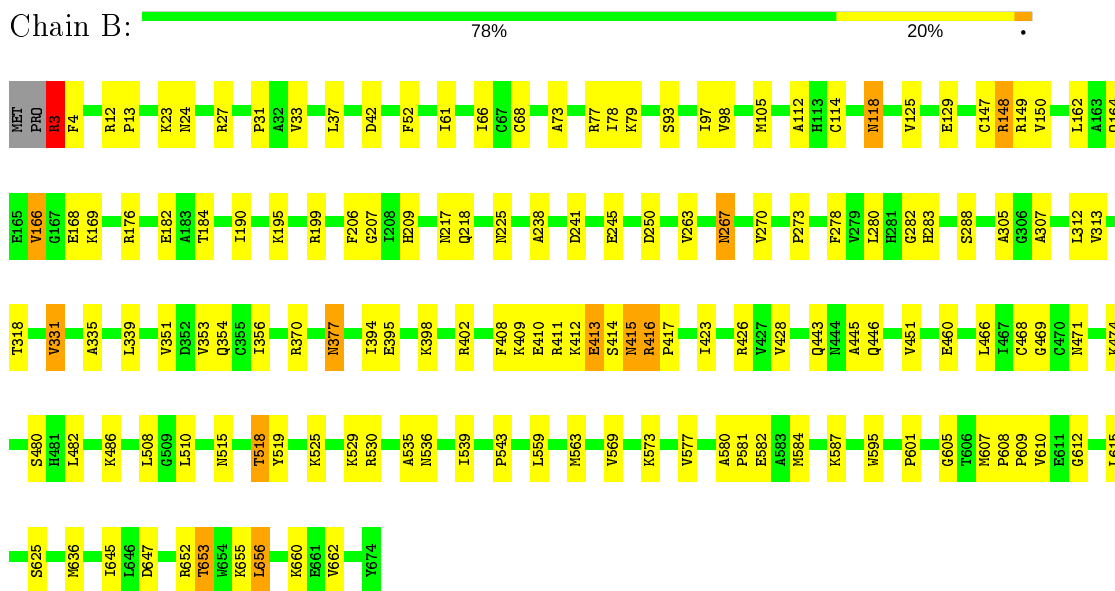
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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 was not executed.

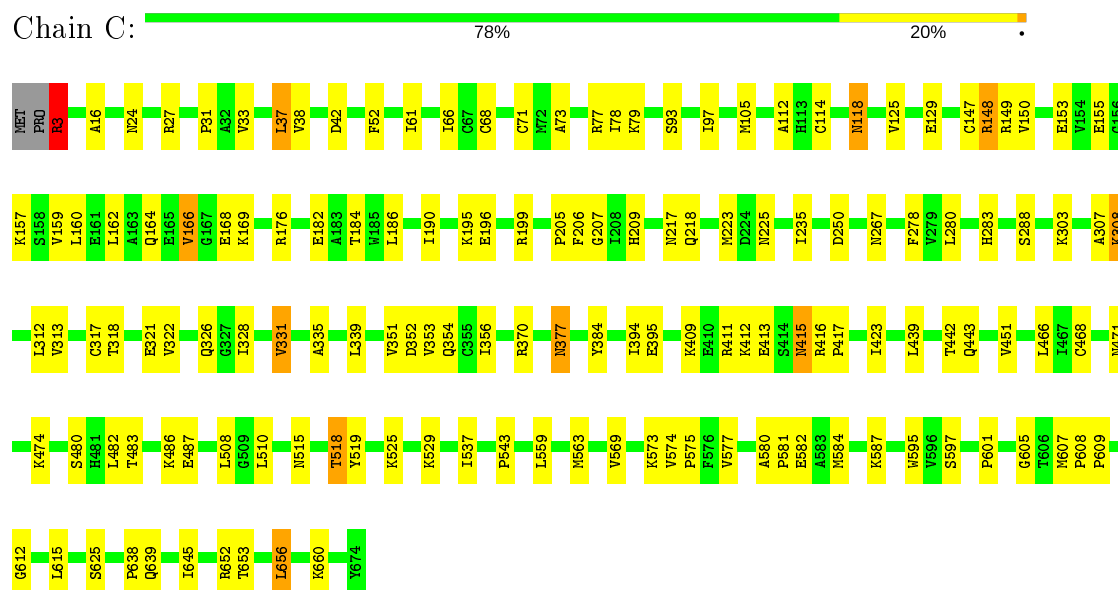
- Molecule 1: CARBON MONOXIDE DEHYDROGENASE BETA SUBUNIT



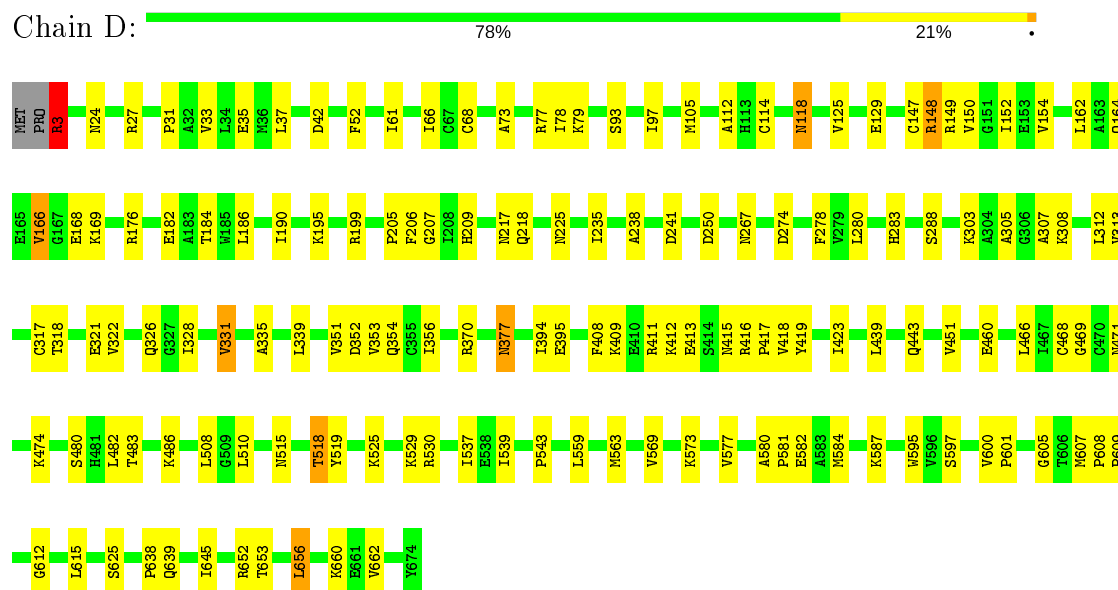
- Molecule 1: CARBON MONOXIDE DEHYDROGENASE BETA SUBUNIT



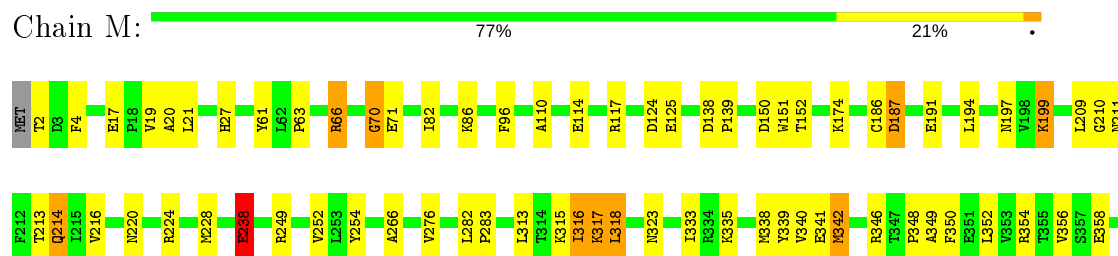
- Molecule 1: CARBON MONOXIDE DEHYDROGENASE BETA SUBUNIT

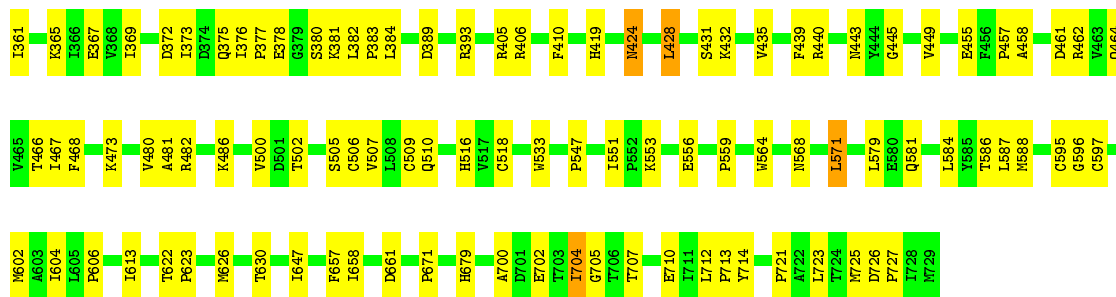


- Molecule 1: CARBON MONOXIDE DEHYDROGENASE BETA SUBUNIT

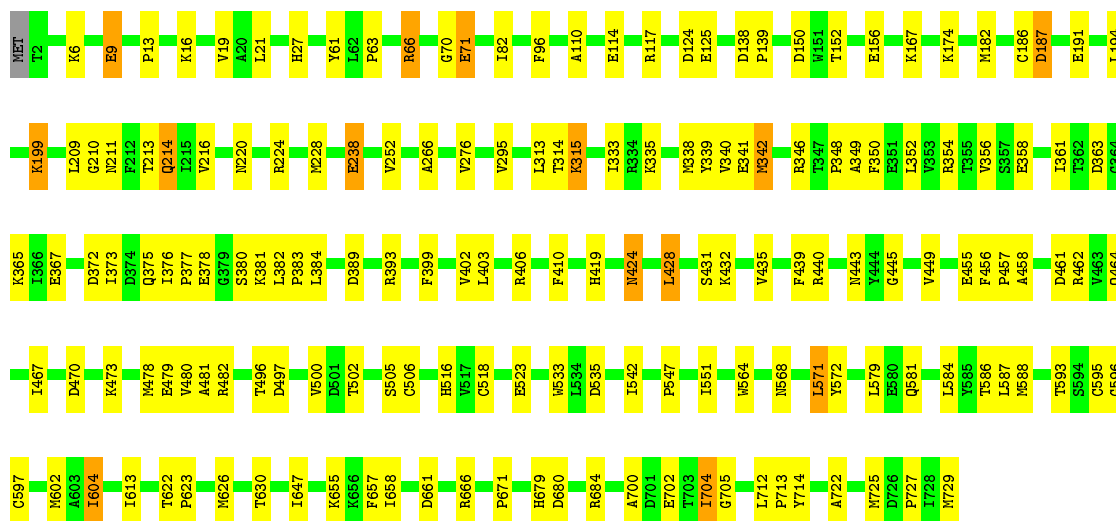
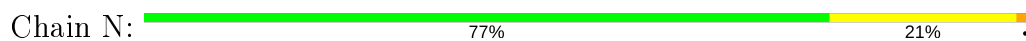


- Molecule 2: Carbon monoxide dehydrogenase alpha subunit

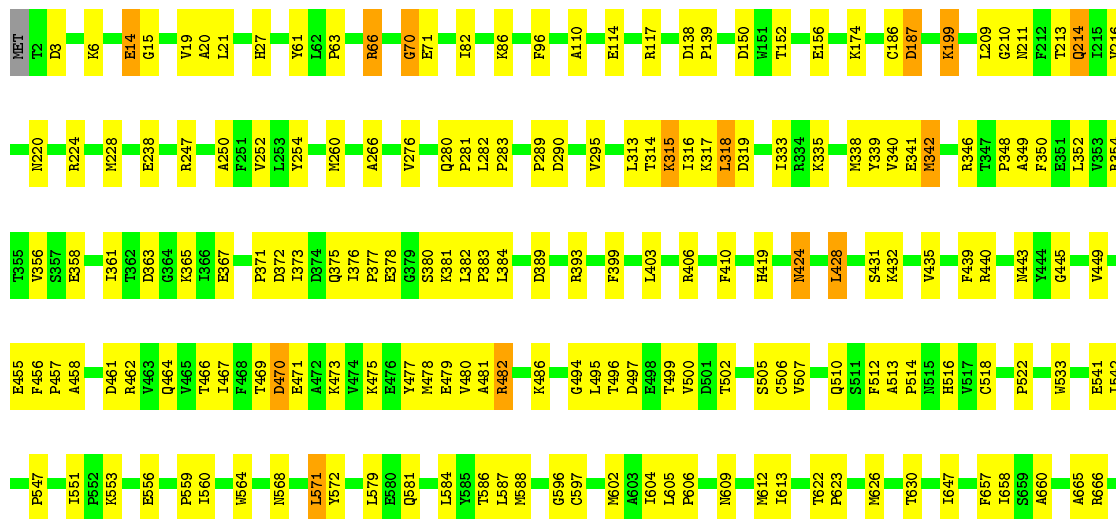




• Molecule 2: Carbon monoxide dehydrogenase alpha subunit



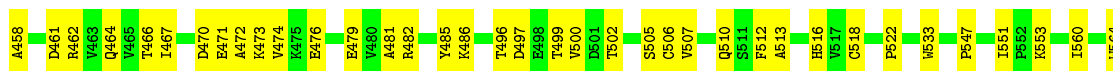
• Molecule 2: Carbon monoxide dehydrogenase alpha subunit





- Molecule 2: Carbon monoxide dehydrogenase alpha subunit

Chain P: 75% 24%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.75Å 136.97Å 141.53Å 101.45° 109.05° 103.94°	Depositor
Resolution (Å)	20.00 – 2.20	Depositor
% Data completeness (in resolution range)	90.6 (20.00-2.20)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.215 , 0.241	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	44653	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: XCC, SF4, ACT, CU1, NI

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	0.38	0/5179	0.61	1/7017 (0.0%)
1	B	0.39	0/5179	0.62	1/7017 (0.0%)
1	C	0.36	0/5179	0.61	2/7017 (0.0%)
1	D	0.34	0/5179	0.60	1/7017 (0.0%)
2	M	0.36	0/5869	0.61	3/7948 (0.0%)
2	N	0.37	0/5869	0.61	0/7948
2	O	0.30	0/5869	0.58	1/7948 (0.0%)
2	P	0.32	0/5869	0.59	0/7948
All	All	0.35	0/44192	0.60	9/59860 (0.0%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	3	ARG	NE-CZ-NH2	7.28	123.94	120.30
1	A	3	ARG	NE-CZ-NH1	7.19	123.90	120.30
1	D	3	ARG	NE-CZ-NH1	6.67	123.64	120.30
1	C	3	ARG	NE-CZ-NH2	6.54	123.57	120.30
2	M	238	GLU	CA-CB-CG	6.13	126.88	113.40

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	5087	0	5089	126	0
1	B	5087	0	5089	139	0
1	C	5087	0	5089	128	0
1	D	5087	0	5089	118	0
2	M	5735	0	5693	155	0
2	N	5735	0	5693	148	0
2	O	5735	0	5693	190	0
2	P	5735	0	5693	166	0
3	A	16	0	0	0	0
3	B	8	0	0	0	0
3	C	16	0	0	0	0
3	D	8	0	0	0	0
3	M	8	0	0	0	0
3	N	8	0	0	0	0
3	O	8	0	0	0	0
3	P	8	0	0	0	0
4	A	9	0	0	3	0
4	B	9	0	0	1	0
4	C	9	0	0	2	0
4	D	9	0	0	2	0
5	M	1	0	0	0	0
5	N	1	0	0	0	0
5	O	1	0	0	0	0
5	P	1	0	0	0	0
6	M	1	0	0	0	0
6	N	1	0	0	0	0
6	O	1	0	0	0	0
6	P	1	0	0	0	0
7	M	3	0	3	2	0
7	N	3	0	3	2	0
7	O	3	0	3	1	0
7	P	3	0	3	1	0
8	A	179	0	0	3	0
8	B	217	0	0	6	0
8	C	148	0	0	3	0
8	D	131	0	0	0	0
8	M	206	0	0	2	0
8	N	221	0	0	7	0
8	O	26	0	0	0	0
8	P	101	0	0	1	0
All	All	44653	0	43140	1127	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 13.

The worst 5 of 1127 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)
2:O:335:LYS:HD2	2:O:335:LYS:H	1.21	1.06
2:M:335:LYS:H	2:M:335:LYS:HD2	1.20	1.03
2:P:335:LYS:HD2	2:P:335:LYS:H	1.21	1.03
1:B:446:GLN:HE22	1:C:38:VAL:HG11	1.22	1.02
2:N:335:LYS:HD2	2:N:335:LYS:H	1.20	1.02

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	670/674 (99%)	646 (96%)	21 (3%)	3 (0%)	34	37
1	B	670/674 (99%)	644 (96%)	22 (3%)	4 (1%)	25	26
1	C	670/674 (99%)	647 (97%)	21 (3%)	2 (0%)	41	46
1	D	670/674 (99%)	646 (96%)	22 (3%)	2 (0%)	41	46
2	M	726/729 (100%)	696 (96%)	26 (4%)	4 (1%)	25	26
2	N	726/729 (100%)	696 (96%)	26 (4%)	4 (1%)	25	26
2	O	726/729 (100%)	682 (94%)	34 (5%)	10 (1%)	11	8
2	P	726/729 (100%)	691 (95%)	29 (4%)	6 (1%)	19	19
All	All	5584/5612 (100%)	5348 (96%)	201 (4%)	35 (1%)	25	26

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	M	316	ILE
2	O	315	LYS
1	B	415	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	267	ASN
1	B	267	ASN

### 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	541/543 (100%)	528 (98%)	13 (2%)	49	62
1	B	541/543 (100%)	528 (98%)	13 (2%)	49	62
1	C	541/543 (100%)	526 (97%)	15 (3%)	43	56
1	D	541/543 (100%)	528 (98%)	13 (2%)	49	62
2	M	610/611 (100%)	593 (97%)	17 (3%)	43	56
2	N	610/611 (100%)	593 (97%)	17 (3%)	43	56
2	O	610/611 (100%)	593 (97%)	17 (3%)	43	56
2	P	610/611 (100%)	597 (98%)	13 (2%)	53	67
All	All	4604/4616 (100%)	4486 (97%)	118 (3%)	46	58

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	656	LEU
2	M	424	ASN
2	P	209	LEU
2	M	17	GLU
2	M	209	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 83 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	217	ASN
2	M	323	ASN
2	P	240	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	225	ASN
1	D	639	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

Of 26 ligands modelled in this entry, 8 are monoatomic - leaving 18 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	$\# Z  > 2$
3	SF4	M	900	2	0,12,12	0.00	-	-		
3	SF4	A	750	1	0,12,12	0.00	-	-		
7	ACT	O	950	5	1,2,3	1.68	0	1,1,3	0.56	0
3	SF4	N	900	2	0,12,12	0.00	-	-		
4	XCC	D	800	1	0,11,11	0.00	-	-		
7	ACT	M	950	-	1,2,3	1.74	0	1,1,3	0.73	0
3	SF4	C	700	1	0,12,12	0.00	-	-		
4	XCC	B	800	1	0,11,11	0.00	-	-		
4	XCC	A	800	1	0,11,11	0.00	-	-		
4	XCC	C	800	1	0,11,11	0.00	-	-		
3	SF4	D	750	1	0,12,12	0.00	-	-		
3	SF4	B	750	1	0,12,12	0.00	-	-		

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	ACT	P	950	5	1,2,3	1.84	0	1,1,3	0.58	0
3	SF4	C	750	1	0,12,12	0.00	-	-		
7	ACT	N	950	-	1,2,3	2.11	1 (100%)	1,1,3	0.64	0
3	SF4	P	900	2	0,12,12	0.00	-	-		
3	SF4	O	900	2	0,12,12	0.00	-	-		
3	SF4	A	700	1	0,12,12	0.00	-	-		

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SF4	M	900	2	-	-	0/6/5/5
3	SF4	A	750	1	-	-	0/6/5/5
3	SF4	N	900	2	-	-	0/6/5/5
4	XCC	D	800	1	-	-	0/3/3/3
3	SF4	C	700	1	-	-	0/6/5/5
4	XCC	B	800	1	-	-	0/3/3/3
4	XCC	A	800	1	-	-	0/3/3/3
4	XCC	C	800	1	-	-	0/3/3/3
3	SF4	D	750	1	-	-	0/6/5/5
3	SF4	B	750	1	-	-	0/6/5/5
3	SF4	C	750	1	-	-	0/6/5/5
3	SF4	P	900	2	-	-	0/6/5/5
3	SF4	O	900	2	-	-	0/6/5/5
3	SF4	A	700	1	-	-	0/6/5/5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	N	950	ACT	O-C	2.11	1.29	1.20

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	O	950	ACT	1	0
4	D	800	XCC	2	0
7	M	950	ACT	2	0
4	B	800	XCC	1	0
4	A	800	XCC	3	0
4	C	800	XCC	2	0
7	P	950	ACT	1	0
7	N	950	ACT	2	0

## 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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.