WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Aug 18, 2022 – 02:10 pm BST

PDB ID	:	7Q69
Title	:	Crystal structure of Chaetomium thermophilum C30S Ahp1 in the pre-reaction state
Authors	:	Ravichandran, K.E.; Wilk, P.; Grudnik, P.; Glatt, S.
Deposited on	:	2021-11-05
Resolution	:	1.85 Å(reported)

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

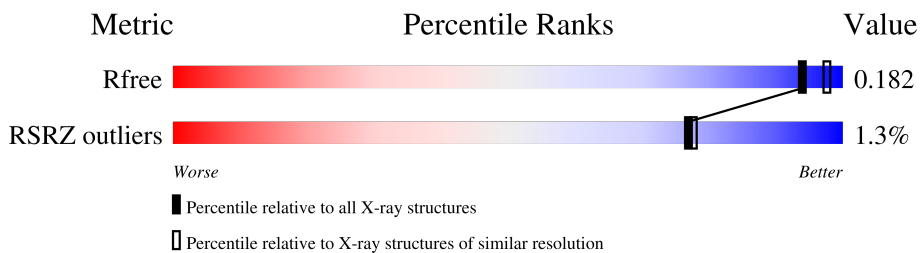|		
MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 1.85 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		0.182
RSRZ outliers		1.3%

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	2469 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 11035 atoms, of which 5032 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 Thioredoxin domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	174	Total	C	H	N	O	S	10	5	0
			2581	816	1291	218	253	3			
1	B	169	Total	C	H	N	O	S	0	3	0
			2528	802	1259	214	250	3			
1	C	167	Total	C	H	N	O	S	13	1	0
			2475	786	1234	210	243	2			
1	D	166	Total	C	H	N	O	S	20	3	0
			2471	785	1233	209	242	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP G0S1P8
A	0	ALA	-	expression tag	UNP G0S1P8
A	30	SER	CYS	engineered mutation	UNP G0S1P8
B	-1	GLY	-	expression tag	UNP G0S1P8
B	0	ALA	-	expression tag	UNP G0S1P8
B	30	SER	CYS	engineered mutation	UNP G0S1P8
C	-1	GLY	-	expression tag	UNP G0S1P8
C	0	ALA	-	expression tag	UNP G0S1P8
C	30	SER	CYS	engineered mutation	UNP G0S1P8
D	-1	GLY	-	expression tag	UNP G0S1P8
D	0	ALA	-	expression tag	UNP G0S1P8
D	30	SER	CYS	engineered mutation	UNP G0S1P8

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	C	1	Total 5	O 4	S 1	0	0
2	C	1	Total 5	O 4	S 1	0	0
2	C	1	Total 5	O 4	S 1	0	0
2	D	1	Total 5	O 4	S 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

Full wwPDB X-ray Structure Validation Report

GOL

C3

OH O3

C1   C2

HO

O1

OH

O2

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total 11	C 3	H 5	O 3	0	0
3	C	1	Total 11	C 3	H 5	O 3	0	0
3	D	1	Total 11	C 3	H 5	O 3	0	0

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	250	Total 250	O 250	0	0
4	B	241	Total 241	O 241	0	0
4	C	253	Total 253	O 253	0	0
4	D	153	Total 153	O 153	0	0

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

W O R L D W I D E

wwPDB

PROTEIN DATA BANK

3 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.14Å 101.73Å 69.42Å 90.00° 107.33° 90.00°	Depositor
Resolution (Å)	46.07 – 1.85 46.07 – 1.85	Depositor EDS
% Data completeness (in resolution range)	96.1 (46.07-1.85) 96.2 (46.07-1.85)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.01 (at 1.84Å)	Xtriage
Refinement program	PHENIX 1.18.2-3874	Depositor
R, R_{free}	0.157 , 0.183 0.156 , 0.182	Depositor DCC
R_{free} test set	3284 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å²)	18.9	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	(Not available) , (Not available)	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	11035	wwPDB-VP
Average B, all atoms (Å²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% 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.

4 Model quality ⓘ

4.1 Standard geometry ⓘ

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

4.2 Too-close contacts ⓘ

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

4.3 Torsion angles ⓘ

4.3.1 Protein backbone ⓘ

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

4.3.2 Protein sidechains ⓘ

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

4.3.3 RNA ⓘ

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

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

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

4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

4.6 Ligand geometry ⓘ

13 ligands are modelled in this entry.

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
2	SO4	B	204	-	4,4,4	0.16	0	6,6,6	0.05	0
3	GOL	B	201	-	5,5,5	0.99	1 (20%)	5,5,5	0.88	0
3	GOL	C	201	-	5,5,5	1.06	1 (20%)	5,5,5	1.04	0
2	SO4	A	201	-	4,4,4	0.17	0	6,6,6	0.11	0
2	SO4	B	202	-	4,4,4	0.15	0	6,6,6	0.09	0
2	SO4	C	203	-	4,4,4	0.13	0	6,6,6	0.16	0
2	SO4	C	204	-	4,4,4	0.14	0	6,6,6	0.09	0
3	GOL	D	201	-	5,5,5	0.63	0	5,5,5	1.03	0
2	SO4	A	203	-	4,4,4	0.17	0	6,6,6	0.10	0
2	SO4	A	202	-	4,4,4	0.15	0	6,6,6	0.08	0
2	SO4	C	202	-	4,4,4	0.13	0	6,6,6	0.07	0
2	SO4	B	203	-	4,4,4	0.16	0	6,6,6	0.11	0
2	SO4	D	202	-	4,4,4	0.14	0	6,6,6	0.06	0

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	GOL	C	201	-	-	2/4/4/4	-
3	GOL	D	201	-	-	4/4/4/4	-
3	GOL	B	201	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	201	GOL	O2-C2	-2.13	1.37	1.43
3	C	201	GOL	O2-C2	-2.00	1.37	1.43

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	201	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	C	201	GOL	O1-C1-C2-C3
3	D	201	GOL	O1-C1-C2-C3
3	D	201	GOL	C1-C2-C3-O3
3	D	201	GOL	O2-C2-C3-O3
3	B	201	GOL	O1-C1-C2-O2
3	D	201	GOL	O1-C1-C2-O2
3	C	201	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers ⓘ

There are no such residues in this entry.

4.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

5 Fit of model and data ⓘ

5.1 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(Å²)	Q<0.9
1	A	174/174 (100%)	-0.41	0	100	100	12, 19, 38, 68	1 (0%)
1	B	169/174 (97%)	-0.43	2 (1%)	79	79	11, 17, 31, 53	0
1	C	167/174 (95%)	-0.44	1 (0%)	89	89	13, 20, 35, 55	1 (0%)
1	D	166/174 (95%)	0.01	6 (3%)	42	40	16, 34, 54, 72	3 (1%)
All	All	676/696 (97%)	-0.32	9 (1%)	77	78	11, 21, 46, 72	5 (0%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1	MET	3.1
1	D	72	LEU	2.9
1	D	76	ARG	2.6
1	D	70	ALA	2.5
1	D	77[A]	GLN	2.4
1	D	164	THR	2.4
1	D	74	GLN	2.3
1	C	3	PRO	2.2
1	B	2	ALA	2.2

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

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

5.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GOL	B	201	6/6	0.88	0.19	27,34,46,46	0
3	GOL	D	201	6/6	0.88	0.25	50,57,80,96	0
2	SO4	B	204	5/5	0.90	0.18	57,66,82,84	0
2	SO4	C	204	5/5	0.92	0.19	47,50,61,63	5
2	SO4	B	203	5/5	0.92	0.11	38,41,52,53	5
2	SO4	A	202	5/5	0.92	0.15	34,43,55,59	5
3	GOL	C	201	6/6	0.93	0.14	31,42,51,51	0
2	SO4	A	203	5/5	0.93	0.20	40,44,63,66	0
2	SO4	D	202	5/5	0.95	0.09	49,50,58,59	5
2	SO4	C	202	5/5	0.95	0.17	35,36,41,42	5
2	SO4	C	203	5/5	0.96	0.09	31,34,40,40	5
2	SO4	A	201	5/5	0.97	0.10	33,36,41,42	5
2	SO4	B	202	5/5	0.98	0.09	28,31,37,40	5

5.5 Other polymers ⓘ

There are no such residues in this entry.