W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2023 – 12:24 PM EDT

PDB ID	:	7S23
Title	:	Crystal structure of alpha-COP-WD40 domain, Y139A mutant
Authors	:	Dey, D.; Singh, S.; Khan, S.; Martin, M.; Schnicker, N.; Gakhar, L.; Pierce, B.; Hasan, S.S.
Deposited on	:	2021-09-03
Resolution	:	1.49 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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 1.49 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.
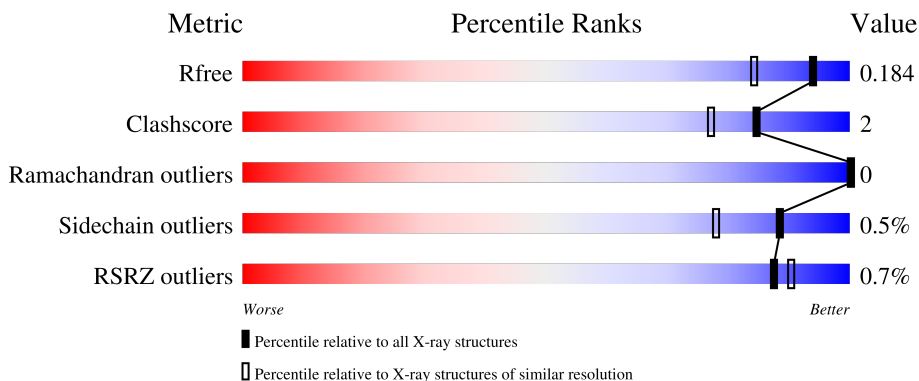
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	A	338	86% 6% 9%
1	B	338	87% 6% 7%
1	C	338	84% 6% 10%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 8425 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 Coatomer subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	308	Total 2503	C 1600	N 450	O 435	S 18	0	3	0
1	B	313	Total 2523	C 1610	N 453	O 442	S 18	0	0	0
1	C	305	Total 2464	C 1577	N 441	O 429	S 17	0	1	0

There are 51 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ACE	-	acetylation	UNP Q96WV5
A	139	ALA	TYR	engineered mutation	UNP Q96WV5
A	181	LYS	LEU	conflict	UNP Q96WV5
A	185	LYS	LEU	conflict	UNP Q96WV5
A	192	LYS	ILE	conflict	UNP Q96WV5
A	196	LYS	LEU	conflict	UNP Q96WV5
A	197	LYS	PHE	conflict	UNP Q96WV5
A	328	SER	-	expression tag	UNP Q96WV5
A	329	ALA	-	expression tag	UNP Q96WV5
A	330	TRP	-	expression tag	UNP Q96WV5
A	331	SER	-	expression tag	UNP Q96WV5
A	332	HIS	-	expression tag	UNP Q96WV5
A	333	PRO	-	expression tag	UNP Q96WV5
A	334	GLN	-	expression tag	UNP Q96WV5
A	335	PHE	-	expression tag	UNP Q96WV5
A	336	GLU	-	expression tag	UNP Q96WV5
A	337	LYS	-	expression tag	UNP Q96WV5
B	0	ACE	-	acetylation	UNP Q96WV5
B	139	ALA	TYR	engineered mutation	UNP Q96WV5
B	181	LYS	LEU	conflict	UNP Q96WV5
B	185	LYS	LEU	conflict	UNP Q96WV5
B	192	LYS	ILE	conflict	UNP Q96WV5
B	196	LYS	LEU	conflict	UNP Q96WV5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	197	LYS	PHE	conflict	UNP Q96WV5
B	328	SER	-	expression tag	UNP Q96WV5
B	329	ALA	-	expression tag	UNP Q96WV5
B	330	TRP	-	expression tag	UNP Q96WV5
B	331	SER	-	expression tag	UNP Q96WV5
B	332	HIS	-	expression tag	UNP Q96WV5
B	333	PRO	-	expression tag	UNP Q96WV5
B	334	GLN	-	expression tag	UNP Q96WV5
B	335	PHE	-	expression tag	UNP Q96WV5
B	336	GLU	-	expression tag	UNP Q96WV5
B	337	LYS	-	expression tag	UNP Q96WV5
C	0	ACE	-	acetylation	UNP Q96WV5
C	139	ALA	TYR	engineered mutation	UNP Q96WV5
C	181	LYS	LEU	conflict	UNP Q96WV5
C	185	LYS	LEU	conflict	UNP Q96WV5
C	192	LYS	ILE	conflict	UNP Q96WV5
C	196	LYS	LEU	conflict	UNP Q96WV5
C	197	LYS	PHE	conflict	UNP Q96WV5
C	328	SER	-	expression tag	UNP Q96WV5
C	329	ALA	-	expression tag	UNP Q96WV5
C	330	TRP	-	expression tag	UNP Q96WV5
C	331	SER	-	expression tag	UNP Q96WV5
C	332	HIS	-	expression tag	UNP Q96WV5
C	333	PRO	-	expression tag	UNP Q96WV5
C	334	GLN	-	expression tag	UNP Q96WV5
C	335	PHE	-	expression tag	UNP Q96WV5
C	336	GLU	-	expression tag	UNP Q96WV5
C	337	LYS	-	expression tag	UNP Q96WV5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	330	Total 330	O 330	0	0
2	B	302	Total 302	O 302	0	0
2	C	303	Total 303	O 303	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: Coatomer subunit alpha

• Molecule 1: Coatomer subunit alpha

• Molecule 1: Coatomer subunit alpha

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	37.31Å 171.55Å 71.28Å 90.00° 99.71° 90.00°	Depositor
Resolution (Å)	65.02 – 1.49 85.78 – 1.49	Depositor EDS
% Data completeness (in resolution range)	97.8 (65.02-1.49) 97.8 (85.78-1.49)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 1.49Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.140 , 0.184 0.140 , 0.184	Depositor DCC
R_{free} test set	7008 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	13.5	Xtriage
Anisotropy	0.368	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.9	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	8425	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 24.08 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.1084e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ACE

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	0.32	1/2582 (0.0%)	0.54	0/3506
1	B	0.33	2/2596 (0.1%)	0.55	0/3528
1	C	0.35	1/2539 (0.0%)	0.55	0/3451
All	All	0.33	4/7717 (0.1%)	0.54	0/10485

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	1	MET	C-N	8.60	1.53	1.34
1	B	1	MET	C-N	7.43	1.51	1.34
1	A	1	MET	C-N	-5.48	1.21	1.34
1	B	67	LEU	C-O	-5.32	1.13	1.23

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	A	2503	0	2421	12	0
1	B	2523	0	2440	11	0
1	C	2464	0	2378	12	0
2	A	330	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	302	0	0	2	1
2	C	303	0	0	0	1
All	All	8425	0	7239	34	1

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.

The worst 5 of 34 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:A:34:ARG:HD3	2:A:611:HOH:O	1.73	0.88
1:C:236:VAL:HB	1:C:252:CYS:HB2	1.81	0.63
1:C:6:LYS:HZ1	1:C:325:LYS:HE2	1.73	0.54
1:C:228:LEU:HG	1:C:264:PHE:CZ	2.43	0.54
1:A:76:LYS:NZ	1:A:78:ASN:HD21	2.05	0.53

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:405:HOH:O	2:C:637:HOH:O[2_746]	2.18	0.02

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	307/338 (91%)	294 (96%)	13 (4%)	0	100	100
1	B	309/338 (91%)	297 (96%)	12 (4%)	0	100	100
1	C	302/338 (89%)	289 (96%)	13 (4%)	0	100	100
All	All	918/1014 (90%)	880 (96%)	38 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	270/297 (91%)	269 (100%)	1 (0%)	91	82
1	B	273/297 (92%)	272 (100%)	1 (0%)	91	82
1	C	265/297 (89%)	263 (99%)	2 (1%)	81	66
All	All	808/891 (91%)	804 (100%)	4 (0%)	88	78

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	288	ARG
1	B	288	ARG
1	C	12	SER
1	C	288	ARG

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	A	307/338 (90%)	-0.63	2 (0%)	87	90	11, 17, 30, 50	0
1	B	312/338 (92%)	-0.58	3 (0%)	82	85	11, 18, 34, 65	0
1	C	304/338 (89%)	-0.61	1 (0%)	94	95	11, 18, 33, 64	0
All	All	923/1014 (91%)	-0.61	6 (0%)	87	90	11, 18, 33, 65	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	177	VAL	4.3
1	B	174	ALA	2.9
1	A	332	HIS	2.7
1	B	171	MET	2.6
1	A	174	ALA	2.5

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 (i)

There are no such residues in this entry.