W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 19, 2023 – 03:34 AM EDT

PDB ID	:	5DQO
Title	:	Crystal structure of Y347F mutant of human primase p58 iron-sulfur cluster domain
Authors	:	Salay, L.E.; Thompson, M.K.; Chazin, W.J.
Deposited on	:	2015-09-15
Resolution	:	2.30 Å(reported)

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

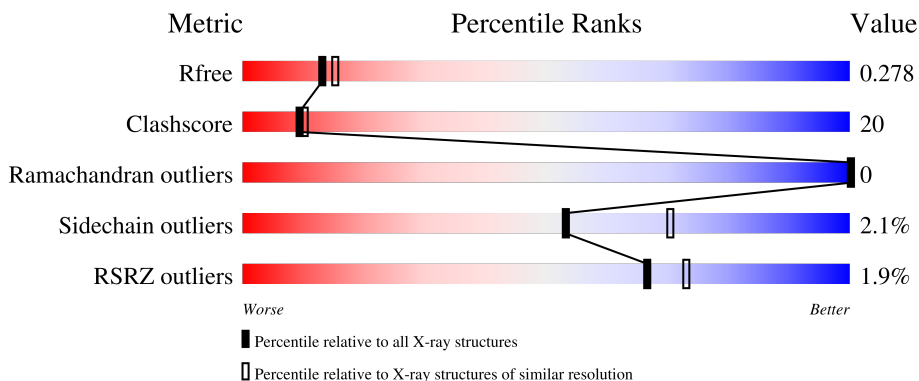|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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:
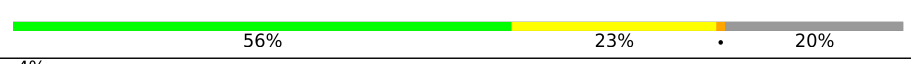*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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
Rfree		0.278
Clashscore		20
Ramachandran outliers		0
Sidechain outliers		2.1%
RSRZ outliers		1.9%

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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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% 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	197	% 50% 29% • 20%
1	B	197	2% 53% 24% • 22%
1	C	197	56% 23% • 20%
1	D	197	4% 51% 28% • 19%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4911 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 DNA primase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	158	Total	C	N	O	S	0	1	0
			1229	786	219	215	9			
1	B	154	Total	C	N	O	S	0	0	0
			1193	752	215	217	9			
1	C	158	Total	C	N	O	S	0	1	0
			1236	790	218	219	9			
1	D	159	Total	C	N	O	S	0	0	0
			1207	762	216	220	9			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	268	GLY	-	expression tag	UNP P49643
A	269	SER	-	expression tag	UNP P49643
A	270	PRO	-	expression tag	UNP P49643
A	271	ASN	-	expression tag	UNP P49643
A	347	PHE	TYR	engineered mutation	UNP P49643
B	268	GLY	-	expression tag	UNP P49643
B	269	SER	-	expression tag	UNP P49643
B	270	PRO	-	expression tag	UNP P49643
B	271	ASN	-	expression tag	UNP P49643
B	347	PHE	TYR	engineered mutation	UNP P49643
C	268	GLY	-	expression tag	UNP P49643
C	269	SER	-	expression tag	UNP P49643
C	270	PRO	-	expression tag	UNP P49643
C	271	ASN	-	expression tag	UNP P49643
C	347	PHE	TYR	engineered mutation	UNP P49643
D	268	GLY	-	expression tag	UNP P49643
D	269	SER	-	expression tag	UNP P49643
D	270	PRO	-	expression tag	UNP P49643
D	271	ASN	-	expression tag	UNP P49643
D	347	PHE	TYR	engineered mutation	UNP P49643

- Molecule 2 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 8	Fe 4	S 4	0	0
2	B	1	Total 8	Fe 4	S 4	0	0
2	C	1	Total 8	Fe 4	S 4	0	0
2	D	1	Total 8	Fe 4	S 4	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total 4	O 4	0	0
3	B	3	Total 3	O 3	0	0
3	C	5	Total 5	O 5	0	0
3	D	2	Total 2	O 2	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: DNA primase large subunit

- Molecule 1: DNA primase large subunit

- Molecule 1: DNA primase large subunit

- Molecule 1: DNA primase large subunit

Chain D:

4%

51%   28%   •   19%

GLY SER PRO ASN S272 L273 D274 Q275 I276 D277 L278 L279 S280 S283 F284 P285 P286 C287 Q290 L291 L295 H300 L301 M307 Q308 Y309 L313 L318 T319 L320 L324 Q325 F326 W327 LYS GLN GLU PHE ILE LYS GLY LYS MET ASP PRO ASP LYS PHE ASP LYS GLY TYR

S346 F347 N348 I349 R350 H351 S352 F353 GLY LYS GLU GLY LYS ARG THR ASP TYR T363 P364 F365 S366 C367 I370 N374 H382 G383 C384 P385 F386 R387 D390 P391 Y401 K402 I403 S404 P405 G406 G407 I408 G417 T418 H419 Y420 C424 Q425 K426 Y427 F428 E429 M430 I431

V434 C437 G438 F439 S440 L441 P444 F447 F448 C449 R453 G457 GLY LYS ASP ILE LYS LYS GLU

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	60.40Å 60.40Å 246.73Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.14 – 2.30	Depositor
	44.14 – 2.30	EDS
% Data completeness (in resolution range)	95.3 (44.14-2.30)	Depositor
	95.3 (44.14-2.30)	EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.19 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.251 , 0.275	Depositor
	0.253 , 0.278	DCC
R_{free} test set	2045 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	58.1	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 47.1	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l 0.300 for h,-h-k,-l 0.037 for -k,-h,-l	Xtriage
Reported twinning fraction	0.365 for H, K, L 0.232 for -K, -H, -L 0.221 for -h,-k,l 0.182 for K, H, -L	Depositor
Outliers	0 of 42683 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4911	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.22% 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.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SF4

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	1.41	0/1267	0.97	0/1711
1	B	1.27	0/1223	1.00	3/1653 (0.2%)
1	C	1.38	0/1273	1.09	4/1717 (0.2%)
1	D	1.21	0/1238	1.00	2/1675 (0.1%)
All	All	1.32	0/5001	1.02	9/6756 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	324	LEU	N-CA-C	11.80	142.85	111.00
1	C	323	ALA	CB-CA-C	-10.58	94.23	110.10
1	C	324	LEU	CB-CA-C	-9.40	92.34	110.20
1	D	424	CYS	CA-CB-SG	8.46	129.22	114.00
1	B	424	CYS	CA-CB-SG	6.99	126.58	114.00

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	1229	0	1164	54	0
1	B	1193	0	1101	44	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1236	0	1173	42	0
1	D	1207	0	1088	56	0
2	A	8	0	0	1	0
2	B	8	0	0	1	0
2	C	8	0	0	0	0
2	D	8	0	0	0	0
3	A	4	0	0	1	0
3	B	3	0	0	0	0
3	C	5	0	0	0	0
3	D	2	0	0	0	0
All	All	4911	0	4526	186	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 20.

The worst 5 of 186 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:449:CYS:SG	1:B:449:CYS:SG	1.35	1.33
1:B:301:LEU:O	1:B:306:ARG:NH1	1.76	1.17
1:D:295:LEU:HD13	1:D:301:LEU:HD12	1.35	1.03
1:B:367:CYS:O	1:B:371:ILE:HD12	1.59	0.99
1:B:286:PRO:HD2	2:B:501:SF4:S3	2.07	0.94

There are no symmetry-related clashes.

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	153/197 (78%)	147 (96%)	6 (4%)	0	100	100
1	B	148/197 (75%)	140 (95%)	8 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	153/197 (78%)	148 (97%)	5 (3%)	0	100	100
1	D	153/197 (78%)	145 (95%)	8 (5%)	0	100	100
All	All	607/788 (77%)	580 (96%)	27 (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	131/176 (74%)	128 (98%)	3 (2%)	50	67
1	B	127/176 (72%)	125 (98%)	2 (2%)	62	78
1	C	133/176 (76%)	130 (98%)	3 (2%)	50	67
1	D	124/176 (70%)	121 (98%)	3 (2%)	49	66
All	All	515/704 (73%)	504 (98%)	11 (2%)	53	70

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	449	CYS
1	D	384	CYS
1	D	437	CYS
1	D	418	THR
1	B	456	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	322	GLN
1	D	300	HIS

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 ⓘ

4 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	SF4	D	501	1	0,12,12	-	-	-		
2	SF4	A	501	1	0,12,12	-	-	-		
2	SF4	B	501	1	0,12,12	-	-	-		
2	SF4	C	501	1	0,12,12	-	-	-		

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
2	SF4	D	501	1	-	-	0/6/5/5
2	SF4	A	501	1	-	-	0/6/5/5
2	SF4	B	501	1	-	-	0/6/5/5
2	SF4	C	501	1	-	-	0/6/5/5

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	SF4	1	0
2	B	501	SF4	1	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 ⓘ

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	158/197 (80%)	0.36	2 (1%)	77	81	35, 50, 65, 83	5 (3%)
1	B	154/197 (78%)	0.33	3 (1%)	66	73	36, 55, 73, 81	4 (2%)
1	C	158/197 (80%)	0.24	0	100	100	32, 53, 69, 83	3 (1%)
1	D	159/197 (80%)	0.37	7 (4%)	34	41	33, 56, 71, 92	3 (1%)
All	All	629/788 (79%)	0.32	12 (1%)	66	73	32, 54, 72, 92	15 (2%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	347	PHE	3.1
1	B	383	GLY	2.9
1	D	424	CYS	2.9
1	B	365	PHE	2.9
1	A	270	PRO	2.8

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 ⓘ

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
2	SF4	B	501	8/8	0.91	0.16	51,59,63,63	0
2	SF4	A	501	8/8	0.96	0.18	51,57,67,77	0
2	SF4	C	501	8/8	0.96	0.18	44,57,58,60	0
2	SF4	D	501	8/8	0.96	0.17	53,58,64,67	0

6.5 Other polymers ⓘ

There are no such residues in this entry.