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)

May 12, 2020 – 11:39 pm BST

PDB ID	:	5D96
Title	:	Oxidoreductase Fragment of Mouse QSOX1 in Complex with a FAb Fragment from an Antibody Targeting Mouse and Human QSOX1
Authors	:	Fass, D.; Grossman, I.
Deposited on	:	2015-08-18
Resolution	:	2.30 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

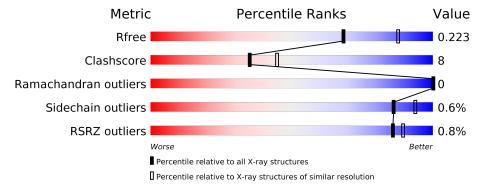|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.30 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	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 on 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	244	2% 79% 18% •
1	D	244	% 77% 18% ••
2	B	214	82% 18%
2	I	214	84% 15%
3	C	220	% 77% 19% ••
3	J	220	80% 17% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10605 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 Sulfhydryl oxidase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	234	Total 1807	C 1152	N 307	O 341	S 7	0	0	0
1	A	237	Total 1846	C 1177	N 314	O 348	S 7	0	1	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	32	GLY	-	expression tag	UNP Q8BND5
D	33	SER	-	expression tag	UNP Q8BND5
D	34	HIS	-	expression tag	UNP Q8BND5
D	35	MET	-	expression tag	UNP Q8BND5
A	32	GLY	-	expression tag	UNP Q8BND5
A	33	SER	-	expression tag	UNP Q8BND5
A	34	HIS	-	expression tag	UNP Q8BND5
A	35	MET	-	expression tag	UNP Q8BND5

- Molecule 2 is a protein called Light chain of Fab fragment from an antibody targeting mouse and human QSOX1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	214	Total 1656	C 1030	N 280	O 337	S 9	0	1	0
2	B	214	Total 1651	C 1027	N 279	O 336	S 9	0	0	0

- Molecule 3 is a protein called Heavy chain of Fab fragment from an antibody targeting mouse and human QSOX1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	215	Total 1612	C 1021	N 267	O 315	S 9	0	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	214	1608	1019	266	314	9	0	0	0

- Molecule 4 is water.

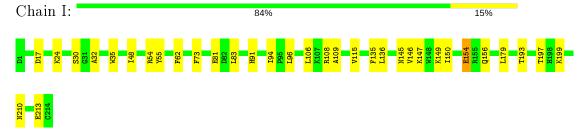| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	75	Total 75	O 75	0	0
4	I	77	Total 77	O 77	0	0
4	J	71	Total 71	O 71	0	0
4	A	81	Total 81	O 81	0	0
4	B	66	Total 66	O 66	0	0
4	C	55	Total 55	O 55	0	0

3 Residue-property plots ⓘ

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 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.
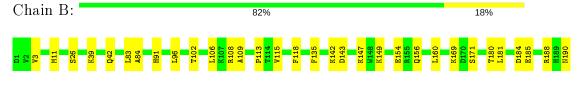
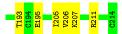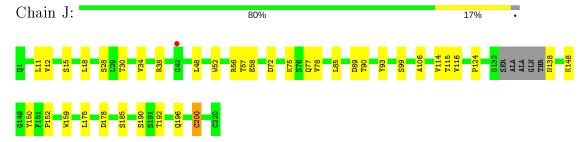- Molecule 1: Sulfhydryl oxidase 1

- Molecule 1: Sulfhydryl oxidase 1

- Molecule 2: Light chain of Fab fragment from an antibody targeting mouse and human QSOX1

- Molecule 2: Light chain of Fab fragment from an antibody targeting mouse and human QSOX1

T193 C194 E195 I205 V206 K207 R211 C214

● Molecule 3: Heavy chain of Fab fragment from an antibody targeting mouse and human QSOX1

Chain J:

80% 17%

Q1 L11 V12 S15 L18 S28 L29 T30 V34 R38 G42 L48 W52 R56 T57 E58 D72 K75 S76 Q77 V78 L85 D89 T90 Y93 S99 A106 V114 T115 V116 P124 G132 SER ALA ALA GLN THR N138 K148

G149 Y150 F151 P152 W159 L175 D178 S185 S190 S191 T192 Q196 C200 C220

● Molecule 3: Heavy chain of Fab fragment from an antibody targeting mouse and human QSOX1

Chain C:

%

77% 19%

Q1 L4 S7 G8 P9 G10 L11 V12 S15 L18 T21 V24 S40 K43 D54 G55 R56 Y59 K60 S61 K64 S65 R66 L67 S68 I69 L80 S84 L85 R92 S99 A106 Q110 A119 P124 P128 P131 GLY SER

ALA ALA GLN THR N138 Y150 F151 P152 W159 N160 L164 S177 D178 S191 S195 C200 H204 T209 K213 R218 D219 C220

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.80Å 204.77Å 44.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.42 – 2.30 27.42 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (27.42-2.30) 96.8 (27.42-2.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$<I/\sigma(I)>$ [1]	2.00 (at 2.31Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.167 , 0.228 0.168 , 0.223	Depositor DCC
R_{free} test set	1998 reflections (3.20%)	wwPDB-VP
Wilson B-factor (Å^2)	34.4	Xtriage
Anisotropy	0.193	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 24.1	EDS
L-test for twinning[2]	$<|L|> = 0.42, <L^2> = 0.25$	Xtriage
Estimated twinning fraction	0.420 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10605	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.13% 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 (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	0.27	0/1896	0.49	0/2589
1	D	0.28	0/1850	0.52	0/2522
2	B	0.26	0/1686	0.50	0/2287
2	I	0.28	0/1694	0.52	0/2298
3	C	0.27	0/1648	0.54	1/2250 (0.0%)
3	J	0.25	0/1652	0.53	0/2255
All	All	0.27	0/10426	0.52	1/14201 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	67	LEU	CA-CB-CG	5.59	128.17	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	1846	0	1790	31	1
1	D	1807	0	1758	42	0
2	B	1651	0	1585	28	0
2	I	1656	0	1591	23	1
3	C	1608	0	1592	24	0
3	J	1612	0	1595	27	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	81	0	0	4	0
4	B	66	0	0	1	0
4	C	55	0	0	2	1
4	D	75	0	0	4	0
4	I	77	0	0	1	0
4	J	71	0	0	1	1
All	All	10605	0	9911	164	2

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 8.

The worst 5 of 164 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:D:226:SER:OG	2:B:108:ARG:NH1	2.05	0.89
2:B:149:LYS:HB2	2:B:193:THR:HB	1.55	0.88
1:A:204:ARG:HG3	1:A:220:ARG:HD2	1.53	0.88
1:D:230:ASN:ND2	2:B:169:LYS:O	2.14	0.81
1:D:128:GLN:HA	1:D:128:GLN:HE21	1.48	0.77

All (2) 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)
4:J:301:HOH:O	4:C:333:HOH:O[2_554]	2.09	0.11
2:I:17:ASP:OD1	1:A:197:ARG:NH1[1_556]	2.13	0.07

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	236/244 (97%)	232 (98%)	4 (2%)	0	100	100
1	D	230/244 (94%)	226 (98%)	4 (2%)	0	100	100
2	B	212/214 (99%)	207 (98%)	5 (2%)	0	100	100
2	I	213/214 (100%)	207 (97%)	6 (3%)	0	100	100
3	C	210/220 (96%)	206 (98%)	4 (2%)	0	100	100
3	J	211/220 (96%)	209 (99%)	2 (1%)	0	100	100
All	All	1312/1356 (97%)	1287 (98%)	25 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	200/207 (97%)	199 (100%)	1 (0%)	88	95
1	D	195/207 (94%)	193 (99%)	2 (1%)	76	87
2	B	188/188 (100%)	188 (100%)	0	100	100
2	I	189/188 (100%)	188 (100%)	1 (0%)	88	95
3	C	187/191 (98%)	185 (99%)	2 (1%)	73	86
3	J	187/191 (98%)	186 (100%)	1 (0%)	88	95
All	All	1146/1172 (98%)	1139 (99%)	7 (1%)	86	94

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	200	CYS
3	C	138	ASN
1	A	163	THR
1	D	245	ARG
3	C	43	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	98	ASN
1	D	128	GLN
2	I	8	HIS
1	A	75	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 carbohydrates 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 (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(Å²)	Q<0.9
1	A	237/244 (97%)	-0.17	4 (1%)	70	76	20, 35, 51, 70	0
1	D	234/244 (95%)	-0.22	3 (1%)	77	81	25, 34, 51, 69	1 (0%)
2	B	214/214 (100%)	-0.30	0	100	100	22, 34, 49, 57	0
2	I	214/214 (100%)	-0.38	0	100	100	23, 32, 43, 53	0
3	C	214/220 (97%)	-0.27	2 (0%)	84	88	23, 35, 52, 65	0
3	J	215/220 (97%)	-0.31	1 (0%)	91	94	22, 33, 45, 54	0
All	All	1328/1356 (97%)	-0.27	10 (0%)	86	89	20, 34, 49, 70	1 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	271	GLY	4.9
3	C	219	ASP	3.6
1	A	273	THR	3.3
3	J	42	GLY	3.1
1	A	162	ASP	2.8

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 carbohydrates 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.