W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 10:04 pm BST

PDB ID	:	6HF1
Title	:	Mutant oxidoreductase fragment of mouse QSOX1 in complex with an antibody Fab
Authors	:	Grossman-Haham, I.; Fass, D.
Deposited on	:	2018-08-21
Resolution	:	1.94 Å(reported)

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

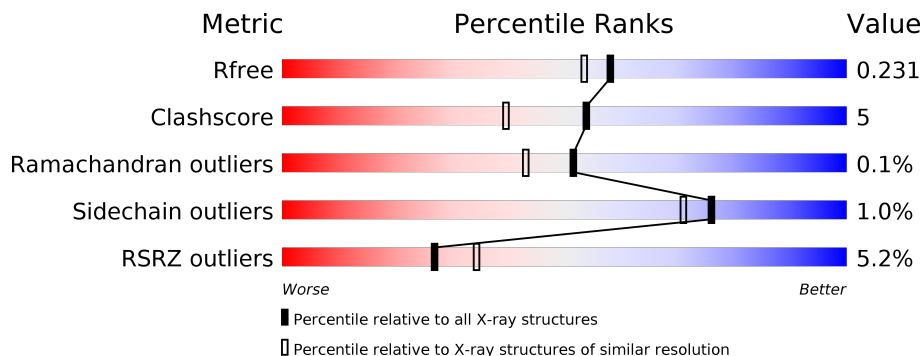|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.94 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		0.231
Clashscore		5
Ramachandran outliers		0.1%
Sidechain outliers		1.0%
RSRZ outliers		5.2%

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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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	238	6% / 85% / 15%
1	D	238	8% / 85% / 15%
2	C	215	4% / 88% / 9% •
3	B	212	9% / 87% / 12%
3	E	212	% / 92% / 8%
4	F	216	% / 87% / 10% •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 10954 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	A	238	Total 1854	C 1179	N 318	O 350	S 7	0	2	0
1	D	238	Total 1847	C 1173	N 318	O 349	S 7	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	75	ALA	HIS	engineered mutation	UNP Q8BND5
A	122	THR	PRO	engineered mutation	UNP Q8BND5
D	75	ALA	HIS	engineered mutation	UNP Q8BND5
D	122	THR	PRO	engineered mutation	UNP Q8BND5

- Molecule 2 is a protein called Fab 316 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	209	Total 1603	C 1024	N 261	O 313	S 5	0	2	0

- Molecule 3 is a protein called Fab 316 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	211	Total 1616	C 1015	N 267	O 327	S 7	0	1	0
3	E	212	Total 1627	C 1022	N 269	O 329	S 7	0	2	0

- Molecule 4 is a protein called Fab 316 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	210	Total 1603	C 1024	N 261	O 313	S 5	0	1	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	126	Total 126	O 126	0	0
5	C	133	Total 133	O 133	0	0
5	B	163	Total 163	O 163	0	0
5	D	101	Total 101	O 101	0	0
5	F	155	Total 155	O 155	0	0
5	E	126	Total 126	O 126	0	0

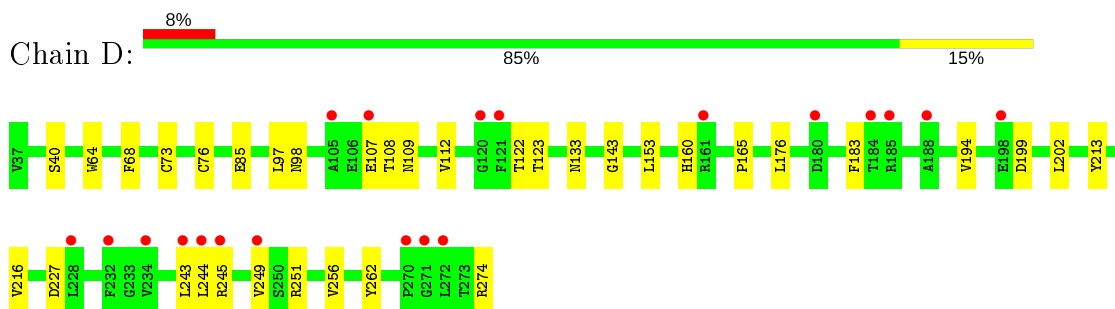# 3 Residue-property plots ⓘ

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.
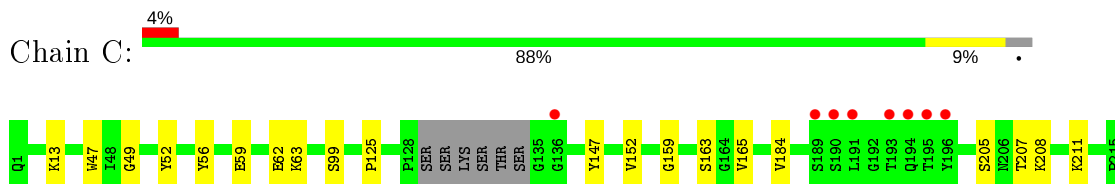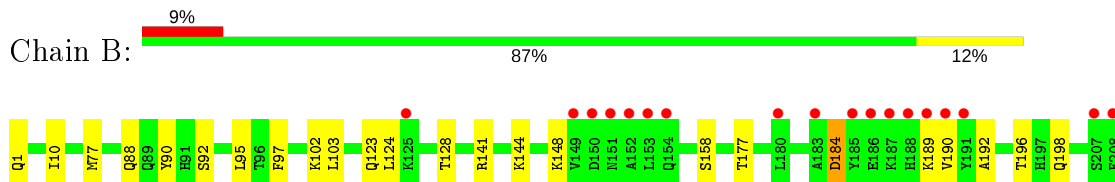
- Molecule 1: Sulfhydryl oxidase 1
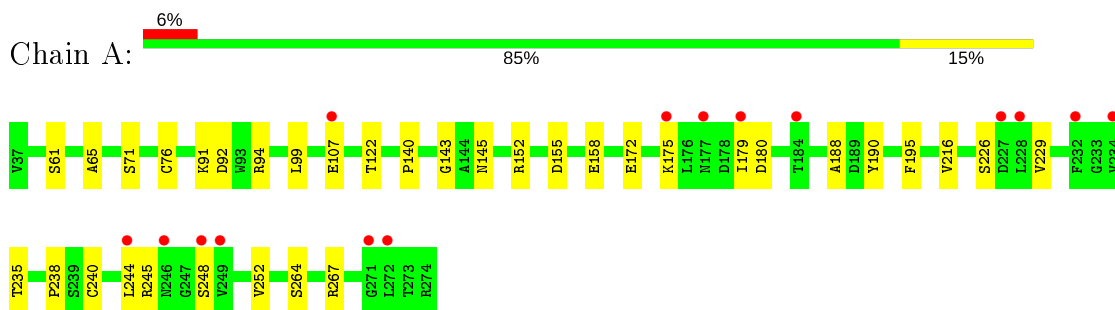
- Molecule 1: Sulfhydryl oxidase 1

- Molecule 2: Fab 316 heavy chain

- Molecule 3: Fab 316 light chain

N209 R210 G211 GLU

● Molecule 3: Fab 316 light chain

%

Chain E:

92% 8%

Q1 P8 A9 I10 M11 K18 Q88 Q89 Y90 H91 L95 K102 L103 E104 L105 K144 K148 N151 A152 L153 N157 T196 E212
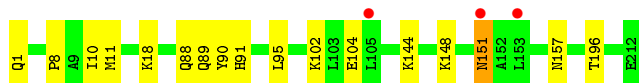
● Molecule 4: Fab 316 heavy chain

%

Chain F:

87% 10% •

Q1 V12 K23 R40 Q43 W47 I48 G49 Y52 N57 G66 L86 S99 S122 S129 SER LYS SER THR SER GLY G136 K145 S163 L172 Y178 T193 Q194 T195 N201 S205 K208 R212 K216
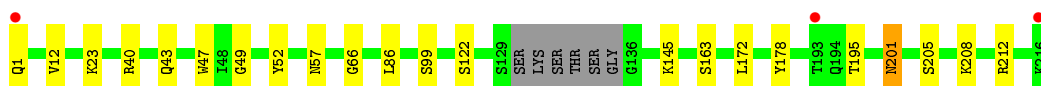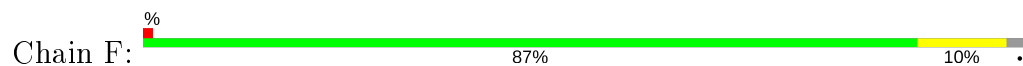
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.54Å 112.72Å 193.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.89 – 1.94 48.89 – 1.94	Depositor EDS
% Data completeness (in resolution range)	97.6 (48.89-1.94) 97.6 (48.89-1.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.15 (at 1.94Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.194 , 0.231 0.194 , 0.231	Depositor DCC
R_{free} test set	5215 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	36.0	Xtriage
Anisotropy	0.282	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 45.9	EDS
L-test for twinning[2]	$< |L| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10954	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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 ⓘ

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	0.40	0/1904	0.58	0/2598
1	D	0.38	0/1891	0.54	0/2580
2	C	0.41	0/1654	0.58	0/2258
3	B	0.38	0/1659	0.56	0/2255
3	E	0.42	0/1673	0.59	0/2274
4	F	0.40	0/1651	0.58	0/2255
All	All	0.40	0/10432	0.57	0/14220

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	C	0	1
4	F	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	99	SER	Peptide
4	F	99	SER	Peptide

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	1854	0	1815	29	1
1	D	1847	0	1801	24	0
2	C	1603	0	1559	13	2
3	B	1616	0	1564	23	0
3	E	1627	0	1574	11	0
4	F	1603	0	1552	15	0
5	A	126	0	0	10	1
5	B	163	0	0	4	0
5	C	133	0	0	3	0
5	D	101	0	0	2	0
5	E	126	0	0	4	0
5	F	155	0	0	5	0
All	All	10954	0	9865	106	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 5.

The worst 5 of 106 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)
3:B:10:ILE:HD13	3:B:102:LYS:HB3	1.34	1.06
3:B:144:LYS:NZ	5:B:301:HOH:O	1.92	1.02
1:A:61:SER:O	5:A:302:HOH:O	1.83	0.95
3:B:1:GLN:N	3:B:1:GLN:OE1	2.09	0.86
3:B:10:ILE:HD12	3:B:102:LYS:O	1.80	0.82

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)
2:C:56:TYR:OH	5:A:301:HOH:O[4_555]	1.93	0.27
1:A:107:GLU:OE1	2:C:56:TYR:OH[4_455]	2.14	0.06

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	238/238 (100%)	231 (97%)	7 (3%)	0	100	100
1	D	236/238 (99%)	230 (98%)	6 (2%)	0	100	100
2	C	207/215 (96%)	202 (98%)	5 (2%)	0	100	100
3	B	210/212 (99%)	204 (97%)	6 (3%)	0	100	100
3	E	212/212 (100%)	204 (96%)	7 (3%)	1 (0%)	29	17
4	F	207/216 (96%)	203 (98%)	4 (2%)	0	100	100
All	All	1310/1331 (98%)	1274 (97%)	35 (3%)	1 (0%)	51	43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	151	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	201/201 (100%)	200 (100%)	1 (0%)	88	88
1	D	199/201 (99%)	196 (98%)	3 (2%)	65	56
2	C	179/183 (98%)	177 (99%)	2 (1%)	73	67
3	B	184/184 (100%)	183 (100%)	1 (0%)	88	88
3	E	185/184 (100%)	185 (100%)	0	100	100
4	F	178/184 (97%)	174 (98%)	4 (2%)	52	39

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1126/1137 (99%)	1115 (99%)	11 (1%)	76 71

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

Mol	Chain	Res	Type
1	D	40	SER
1	D	133	ASN
4	F	57	ASN
3	B	184	ASP
4	F	1	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	5	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)

There are no ligands in this entry.

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 (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(Å^2)	Q<0.9
1	A	238/238 (100%)	0.70	15 (6%)	20	26	26, 42, 67, 74	0
1	D	238/238 (100%)	0.55	20 (8%)	11	16	31, 44, 71, 79	0
2	C	209/215 (97%)	0.44	8 (3%)	40	48	23, 38, 67, 83	0
3	B	211/212 (99%)	0.49	20 (9%)	8	12	24, 37, 65, 72	0
3	E	212/212 (100%)	0.14	3 (1%)	75	80	30, 41, 54, 70	0
4	F	210/216 (97%)	0.23	3 (1%)	75	80	27, 36, 52, 65	0
All	All	1318/1331 (99%)	0.43	69 (5%)	27	34	23, 39, 66, 83	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	193	THR	5.4
1	D	184	THR	4.5
3	B	183	ALA	4.5
3	B	190	VAL	4.4
3	B	149	VAL	4.2

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.