WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 20, 2024 – 04:40 PM EDT

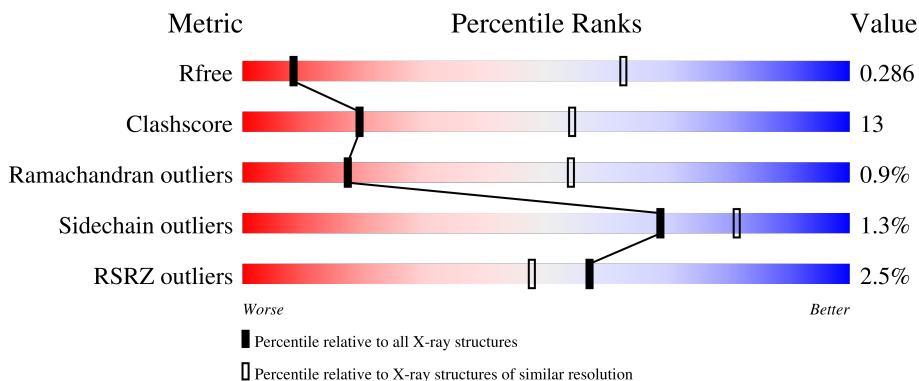|		
PDB ID	:	8W18
Title	:	The crystal structure of the Michaelis-Menten complex of a C1s/C1-INH at 3.94 Angstroms
Authors	:	Garrigues, R.J.; Garcia, B.L.
Deposited on	:	2024-02-15
Resolution	:	3.94 Å(reported)

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 3.94 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		0.286
Clashscore		13
Ramachandran outliers		0.9%
Sidechain outliers		1.3%
RSRZ outliers		2.5%

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	1036 (4.20-3.68)
Clashscore	141614	1009 (4.18-3.70)
Ramachandran outliers	138981	1057 (4.20-3.68)
Sidechain outliers	138945	1049 (4.20-3.68)
RSRZ outliers	127900	1007 (4.24-3.64)

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	I	395	3% · 81% · 15% ·
2	B	251	% · 70% · 27% ··
3	A	151	3% · 66% · 29% ··

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6029 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 Plasma protease C1 inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	378	Total	C	N	O	S	0	0	0
			2982	1915	494	557	16			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	106	GLY	-	expression tag	UNP P05155
I	107	SER	-	expression tag	UNP P05155
I	108	THR	-	expression tag	UNP P05155
I	109	GLY	-	expression tag	UNP P05155
I	110	SER	-	expression tag	UNP P05155

- Molecule 2 is a protein called Complement C1s subcomponent light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	248	Total	C	N	O	S	0	0	0
			1913	1216	328	356	13			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	632	ALA	SER	engineered mutation	UNP P09871

- Molecule 3 is a protein called Complement C1s subcomponent heavy chain.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	147	Total	C	N	O	S	0	0	0
			1134	712	184	227	11			

There are 5 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	287	GLY	-	expression tag	UNP P09871
A	288	SER	-	expression tag	UNP P09871
A	289	THR	-	expression tag	UNP P09871
A	290	GLY	-	expression tag	UNP P09871
A	291	SER	-	expression tag	UNP P09871

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.
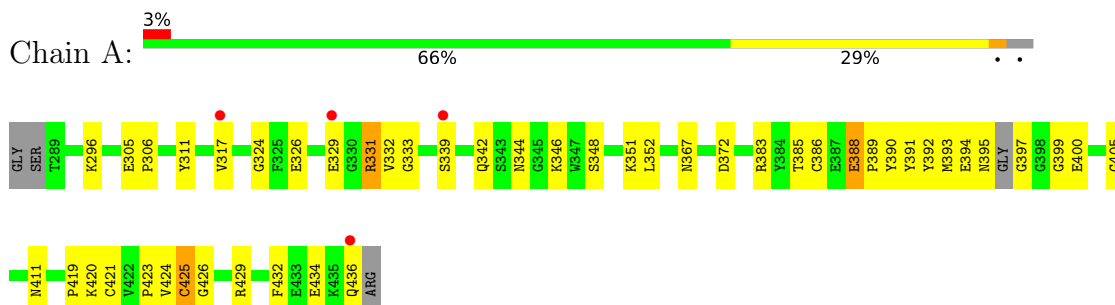
• Molecule 1: Plasma protease C1 inhibitor
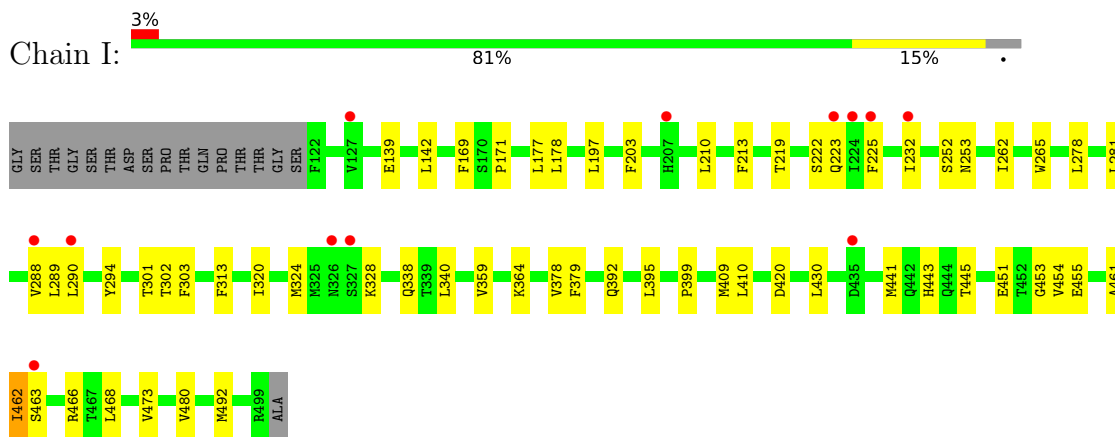
Chain I:

• Molecule 2: Complement C1s subcomponent light chain
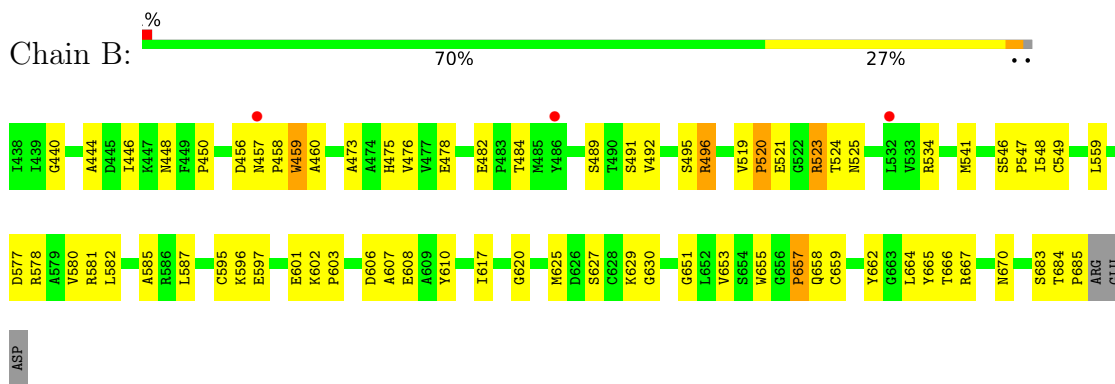
Chain B:

• Molecule 3: Complement C1s subcomponent heavy chain

Chain A:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	168.63Å 168.63Å 88.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.42 – 3.94 47.42 – 3.94	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.42-3.94) 99.9 (47.42-3.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 4.00Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.262 , 0.281 0.269 , 0.286	Depositor DCC
R_{free} test set	1173 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	145.3	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 84.9	EDS
L-test for twinning[2]	$< |L| > = 0.40, < L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6029	wwPDB-VP
Average B, all atoms (Å^2)	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.95% 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	I	0.28	0/3046	0.46	0/4133
2	B	0.31	0/1962	0.51	0/2665
3	A	0.38	0/1164	0.54	0/1581
All	All	0.31	0/6172	0.49	0/8379

There are no bond length outliers.

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	I	2982	0	3025	57	0
2	B	1913	0	1885	82	0
3	A	1134	0	1056	42	0
All	All	6029	0	5966	155	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 13.

The worst 5 of 155 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:A:393:MET:HG2	3:A:394:GLU:H	1.07	1.15

Continued on next page...

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

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:456:ASP:HB2	2:B:484:THR:HB	1.53	0.90
3:A:393:MET:CG	3:A:394:GLU:H	1.86	0.88
1:I:468:LEU:HD23	2:B:629:LYS:HB3	1.56	0.88
1:I:468:LEU:HG	2:B:630:GLY:N	1.89	0.88

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	I	376/395 (95%)	345 (92%)	29 (8%)	2 (0%)	29	66
2	B	246/251 (98%)	219 (89%)	23 (9%)	4 (2%)	9	44
3	A	143/151 (95%)	125 (87%)	17 (12%)	1 (1%)	22	60
All	All	765/797 (96%)	689 (90%)	69 (9%)	7 (1%)	17	54

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	462	ILE
2	B	459	TRP
2	B	596	LYS
2	B	657	PRO
2	B	520	PRO

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	I	343/356 (96%)	340 (99%)	3 (1%)	78	87
2	B	204/207 (99%)	202 (99%)	2 (1%)	76	85
3	A	128/130 (98%)	124 (97%)	4 (3%)	40	63
All	All	675/693 (97%)	666 (99%)	9 (1%)	69	81

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	425	CYS
3	A	429	ARG
2	B	496	ARG
2	B	523	ARG
3	A	331	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	I	378/395 (95%)	0.13	12 (3%)	47	37	116, 160, 197, 227	0
2	B	248/251 (98%)	0.16	3 (1%)	79	70	125, 150, 180, 219	0
3	A	147/151 (97%)	0.25	4 (2%)	54	44	142, 168, 206, 222	0
All	All	773/797 (96%)	0.16	19 (2%)	57	47	116, 158, 197, 227	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	225	PHE	6.6
1	I	288	VAL	4.3
3	A	436	GLN	4.1
3	A	317	VAL	4.0
1	I	290	LEU	3.6

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.