



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 4, 2025 – 12:05 PM EST

PDB ID : 9EKE
Title : Structure of a C1r Zymogen Fragment Bound to SALO, Y51F Mutant
Authors : Duan, H.; Geisbrecht, B.V.
Deposited on : 2024-12-02
Resolution : 3.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

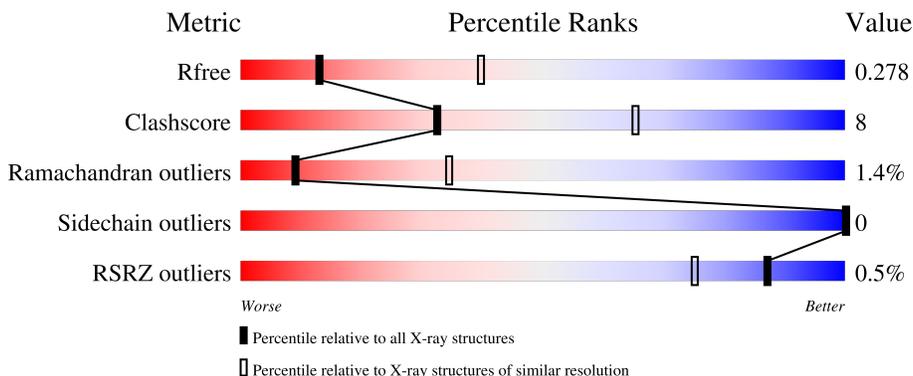
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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}	164625	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)

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 $\leq 5\%$. 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	104	
1	B	104	
2	C	409	
2	D	409	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 7515 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 Salivary anti-complement protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	89	720	450	110	153	7	0	0	0
1	B	89	720	450	110	153	7	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	51	PHE	TYR	engineered mutation	UNP Q5WPZ4
A	116	GLY	-	expression tag	UNP Q5WPZ4
A	117	SER	-	expression tag	UNP Q5WPZ4
A	118	GLY	-	expression tag	UNP Q5WPZ4
A	119	HIS	-	expression tag	UNP Q5WPZ4
A	120	HIS	-	expression tag	UNP Q5WPZ4
A	121	HIS	-	expression tag	UNP Q5WPZ4
A	122	HIS	-	expression tag	UNP Q5WPZ4
A	123	HIS	-	expression tag	UNP Q5WPZ4
A	124	HIS	-	expression tag	UNP Q5WPZ4
A	125	HIS	-	expression tag	UNP Q5WPZ4
A	126	HIS	-	expression tag	UNP Q5WPZ4
B	51	PHE	TYR	engineered mutation	UNP Q5WPZ4
B	116	GLY	-	expression tag	UNP Q5WPZ4
B	117	SER	-	expression tag	UNP Q5WPZ4
B	118	GLY	-	expression tag	UNP Q5WPZ4
B	119	HIS	-	expression tag	UNP Q5WPZ4
B	120	HIS	-	expression tag	UNP Q5WPZ4
B	121	HIS	-	expression tag	UNP Q5WPZ4
B	122	HIS	-	expression tag	UNP Q5WPZ4
B	123	HIS	-	expression tag	UNP Q5WPZ4
B	124	HIS	-	expression tag	UNP Q5WPZ4
B	125	HIS	-	expression tag	UNP Q5WPZ4
B	126	HIS	-	expression tag	UNP Q5WPZ4

- Molecule 2 is a protein called Complement C1r subcomponent.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	380	3040	1923	539	555	23	0	0	0
2	D	380	3035	1920	540	552	23	0	0	0

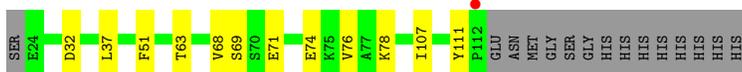
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	654	ALA	SER	engineered mutation	UNP P00736
C	706	GLY	-	expression tag	UNP P00736
C	707	SER	-	expression tag	UNP P00736
C	708	GLY	-	expression tag	UNP P00736
C	709	HIS	-	expression tag	UNP P00736
C	710	HIS	-	expression tag	UNP P00736
C	711	HIS	-	expression tag	UNP P00736
C	712	HIS	-	expression tag	UNP P00736
C	713	HIS	-	expression tag	UNP P00736
C	714	HIS	-	expression tag	UNP P00736
C	715	HIS	-	expression tag	UNP P00736
C	716	HIS	-	expression tag	UNP P00736
D	654	ALA	SER	engineered mutation	UNP P00736
D	706	GLY	-	expression tag	UNP P00736
D	707	SER	-	expression tag	UNP P00736
D	708	GLY	-	expression tag	UNP P00736
D	709	HIS	-	expression tag	UNP P00736
D	710	HIS	-	expression tag	UNP P00736
D	711	HIS	-	expression tag	UNP P00736
D	712	HIS	-	expression tag	UNP P00736
D	713	HIS	-	expression tag	UNP P00736
D	714	HIS	-	expression tag	UNP P00736
D	715	HIS	-	expression tag	UNP P00736
D	716	HIS	-	expression tag	UNP P00736

3 Residue-property plots [i](#)

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.

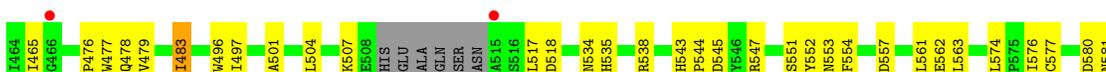
- Molecule 1: Salivary anti-complement protein



- Molecule 1: Salivary anti-complement protein

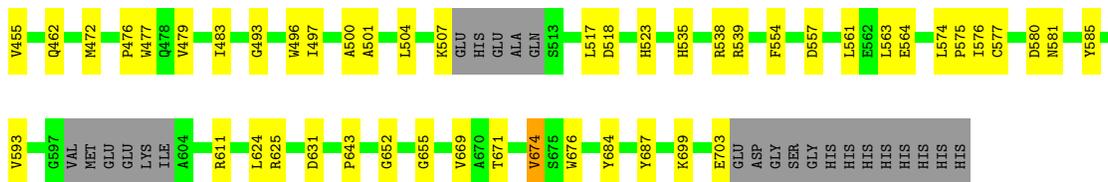


- Molecule 2: Complement C1r subcomponent



- Molecule 2: Complement C1r subcomponent





4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	117.81Å 117.81Å 195.59Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.36 – 3.10 49.36 – 3.10	Depositor EDS
% Data completeness (in resolution range)	86.4 (49.36-3.10) 86.6 (49.36-3.10)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.11 (at 3.07Å)	Xtrriage
Refinement program	PHENIX 1.21rc1_5127	Depositor
R, R_{free}	0.231 , 0.278 0.230 , 0.278	Depositor DCC
R_{free} test set	27162 reflections (7.90%)	wwPDB-VP
Wilson B-factor (Å ²)	75.7	Xtrriage
Anisotropy	0.051	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 47.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.034 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	7515	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.24	0/731	0.39	0/987
1	B	0.24	0/731	0.39	0/987
2	C	0.25	0/3117	0.49	0/4218
2	D	0.25	0/3113	0.48	0/4215
All	All	0.25	0/7692	0.47	0/10407

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 [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	720	0	670	8	0
1	B	720	0	670	7	0
2	C	3040	0	2924	57	0
2	D	3035	0	2922	47	0
All	All	7515	0	7186	114	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 8.

The worst 5 of 114 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)
2:C:507:LYS:NZ	2:C:547:ARG:O	2.15	0.78
2:C:335:ILE:HG22	2:C:355:THR:HG22	1.72	0.71
2:C:497:ILE:HB	2:C:561:LEU:HB2	1.78	0.66
2:C:553:ASN:HA	2:C:632:VAL:HG21	1.78	0.66
2:D:311:GLN:NE2	2:D:326:PRO:O	2.31	0.64

There are no symmetry-related clashes.

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	87/104 (84%)	85 (98%)	2 (2%)	0	100	100
1	B	87/104 (84%)	82 (94%)	4 (5%)	1 (1%)	12	39
2	C	370/409 (90%)	327 (88%)	36 (10%)	7 (2%)	6	27
2	D	372/409 (91%)	340 (91%)	27 (7%)	5 (1%)	10	36
All	All	916/1026 (89%)	834 (91%)	69 (8%)	13 (1%)	9	34

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	483	ILE
1	B	55	SER
2	D	441	GLY
2	C	360	ASP
2	C	422	SER

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	83/96 (86%)	83 (100%)	0	100	100
1	B	83/96 (86%)	83 (100%)	0	100	100
2	C	325/349 (93%)	325 (100%)	0	100	100
2	D	324/349 (93%)	324 (100%)	0	100	100
All	All	815/890 (92%)	815 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 oligosaccharides 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	89/104 (85%)	-0.30	1 (1%) 77 61	71, 92, 127, 151	0
1	B	89/104 (85%)	-0.22	1 (1%) 77 61	91, 124, 167, 237	0
2	C	380/409 (92%)	-0.17	3 (0%) 82 68	47, 88, 142, 194	0
2	D	380/409 (92%)	-0.34	0 100 100	43, 75, 119, 149	0
All	All	938/1026 (91%)	-0.25	5 (0%) 87 75	43, 86, 142, 237	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	466	GLY	2.9
1	B	54	ILE	2.7
1	A	112	PRO	2.1
2	C	515	ALA	2.1
2	C	651	GLN	2.0

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 monosaccharides 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.