wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 1, 2025 – 12:08 PM EST

PDB ID	:	9N20 / pdb_00009n20
Title	:	Structure of C3d Bound to a Fragment of FHR-2 and S. aureus Efb-C
Authors	:	Duan, H.; Geisbrecht, B.V.
Deposited on	:	2025-01-27
Resolution	:	3.30 Å(reported)

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

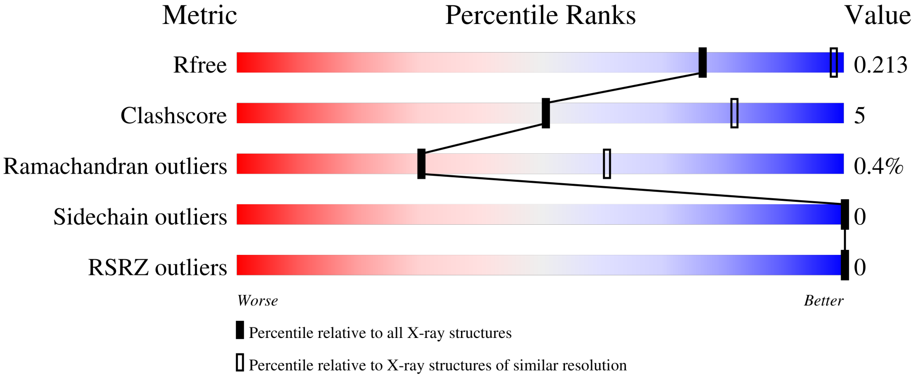
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

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 3.30 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}	164625	1085 (3.32-3.28)
Clashscore	180529	1128 (3.32-3.28)
Ramachandran outliers	177936	1125 (3.32-3.28)
Sidechain outliers	177891	1124 (3.32-3.28)
RSRZ outliers	164620	1085 (3.32-3.28)

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	297	
2	C	75	
3	B	131	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3808 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 Complement C3dg fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	294	2314	1486	388	431	9	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	991	GLY	-	expression tag	UNP P01024
A	992	SER	-	expression tag	UNP P01024
A	993	ARG	-	expression tag	UNP P01024
A	994	SER	-	expression tag	UNP P01024
A	995	THR	-	expression tag	UNP P01024
A	1010	ALA	CYS	engineered mutation	UNP P01024

- Molecule 2 is a protein called Fibrinogen-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	65	532	336	100	95	1	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	91	GLY	-	expression tag	UNP P68799
C	92	SER	-	expression tag	UNP P68799
C	93	THR	-	expression tag	UNP P68799

- Molecule 3 is a protein called Complement factor H-related protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	122	962	607	159	186	10	0	0	0

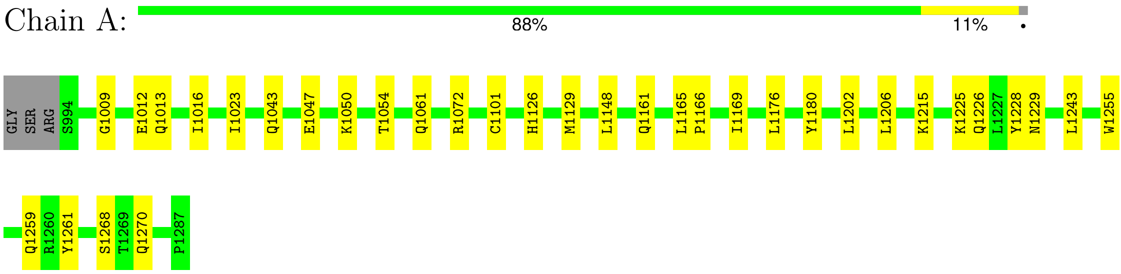
There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	140	GLY	-	expression tag	UNP P36980
B	141	SER	-	expression tag	UNP P36980
B	142	THR	-	expression tag	UNP P36980
B	143	GLY	-	expression tag	UNP P36980
B	144	SER	-	expression tag	UNP P36980

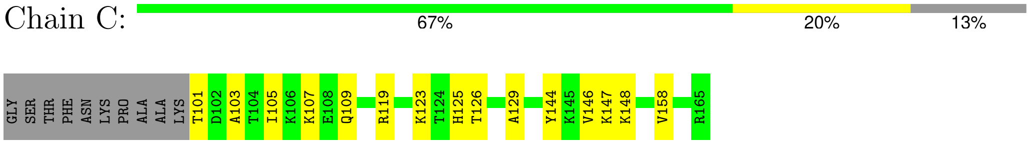
3 Residue-property plots ⓘ

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.

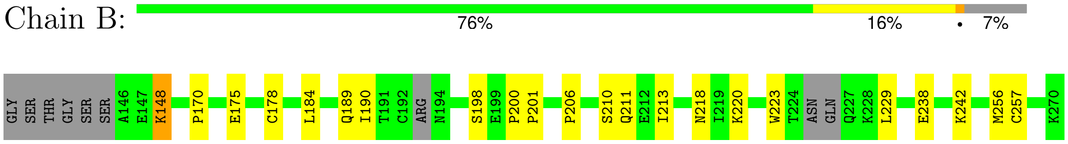
- Molecule 1: Complement C3dg fragment



- Molecule 2: Fibrinogen-binding protein



- Molecule 3: Complement factor H-related protein 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	145.58Å 145.58Å 76.86Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.74 – 3.30 48.74 – 3.30	Depositor EDS
% Data completeness (in resolution range)	94.4 (48.74-3.30) 94.4 (48.74-3.30)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.14 (at 3.33Å)	Xtriage
Refinement program	PHENIX (1.21rc1_5127: ???)	Depositor
R, R_{free}	0.178 , 0.214 0.178 , 0.213	Depositor DCC
R_{free} test set	1333 reflections (9.44%)	wwPDB-VP
Wilson B-factor (Å^2)	87.6	Xtriage
Anisotropy	0.080	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 60.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.065 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3808	wwPDB-VP
Average B, all atoms (Å^2)	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% 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	0.09	0/2363	0.26	0/3201
2	C	0.07	0/536	0.19	0/715
3	B	0.08	0/985	0.25	0/1332
All	All	0.08	0/3884	0.25	0/5248

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	A	2314	0	2318	19	0
2	C	532	0	572	8	0
3	B	962	0	921	12	0
All	All	3808	0	3811	39	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 5.

The worst 5 of 39 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:148:LYS:HG2	3:B:170:PRO:HD3	1.67	0.75

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1101:CYS:SG	1:A:1161:GLN:NE2	2.64	0.71
1:A:1148:LEU:HD11	1:A:1169:ILE:HG23	1.77	0.67
1:A:1009:GLY:HA3	1:A:1013:GLN:HB2	1.79	0.65
3:B:184:LEU:HD11	3:B:201:PRO:HB2	1.77	0.65

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	292/297 (98%)	286 (98%)	6 (2%)	0	100	100
2	C	63/75 (84%)	63 (100%)	0	0	100	100
3	B	116/131 (88%)	107 (92%)	7 (6%)	2 (2%)	7	31
All	All	471/503 (94%)	456 (97%)	13 (3%)	2 (0%)	30	61

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	148	LYS
3	B	206	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	A	243/245 (99%)	243 (100%)	0	100	100
2	C	59/66 (89%)	59 (100%)	0	100	100
3	B	111/118 (94%)	111 (100%)	0	100	100
All	All	413/429 (96%)	413 (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. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1277	GLN
3	B	177	GLN
3	B	246	HIS
3	B	218	ASN
1	A	1152	GLN

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 oligosaccharides 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 (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	294/297 (98%)	-0.64	0	100	100	50, 85, 133, 177	0
2	C	65/75 (86%)	-0.47	0	100	100	93, 132, 169, 181	0
3	B	122/131 (93%)	-0.32	0	100	100	81, 128, 189, 207	0
All	All	481/503 (95%)	-0.53	0	100	100	50, 101, 171, 207	0

There are no RSRZ outliers to report.

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 oligosaccharides 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.