



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 09:50 AM EDT

PDB ID : 8D7K
Title : Bifunctional Inhibition of Neutrophil Elastase and Cathepsin G by Eap2 from *S. aureus*
Authors : Gido, C.D.; Herdendorf, T.J.; Geisbrecht, B.V.
Deposited on : 2022-06-07
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) : 1.13
EDS : 2.36
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) : 2.36

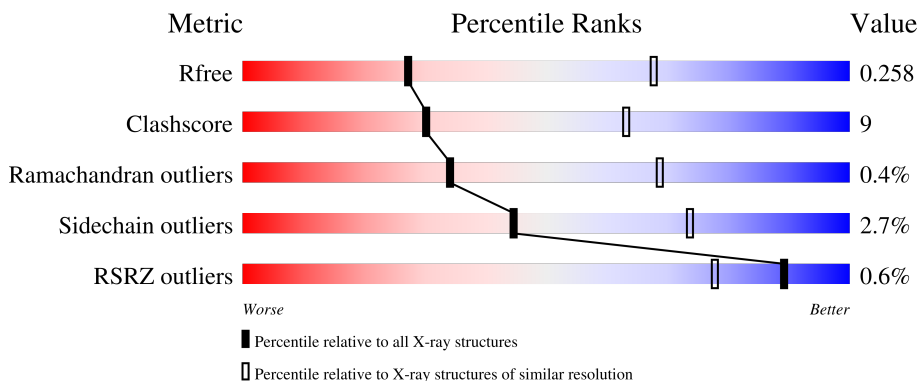
1 Overall quality at a glance

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}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (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	C	223	 77% 22% .
1	F	223	 72% 27% .
1	I	223	 77% 23%
1	L	223	 78% 22%
2	B	100	 75% 20% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	100	 71% 24% . .
2	H	100	 75% 21% . .
2	K	100	 % 72% 24% . .
3	A	218	 % 77% 22% .
3	D	218	 76% 23%
3	G	218	 2% 78% 21% .
3	J	218	 % 78% 22%

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 16684 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 Cathepsin G, C-terminal truncated form.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	223	1780	1093	369	308	10	0	0	0
1	F	223	1780	1093	369	308	10	0	0	0
1	I	223	1780	1093	369	308	10	0	0	0
1	L	223	1780	1093	369	308	10	0	0	0

- Molecule 2 is a protein called Extracellular Adherence Protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	97	756	474	131	151	0	0	0
2	E	97	756	474	131	151	0	0	0
2	H	97	756	474	131	151	0	0	0
2	K	97	756	474	131	151	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	155	GLY	-	expression tag	UNP Q99QS1
B	156	SER	-	expression tag	UNP Q99QS1
B	157	THR	-	expression tag	UNP Q99QS1
E	155	GLY	-	expression tag	UNP Q99QS1
E	156	SER	-	expression tag	UNP Q99QS1
E	157	THR	-	expression tag	UNP Q99QS1
H	155	GLY	-	expression tag	UNP Q99QS1
H	156	SER	-	expression tag	UNP Q99QS1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	157	THR	-	expression tag	UNP Q99QS1
K	155	GLY	-	expression tag	UNP Q99QS1
K	156	SER	-	expression tag	UNP Q99QS1
K	157	THR	-	expression tag	UNP Q99QS1

- Molecule 3 is a protein called Neutrophil elastase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	218	1635	1026	316	282	11	0	0	0
3	D	218	1635	1026	316	282	11	0	0	0
3	G	218	1635	1026	316	282	11	0	0	0
3	J	218	1635	1026	316	282	11	0	0	0



• Molecule 2: Extracellular Adherence Protein



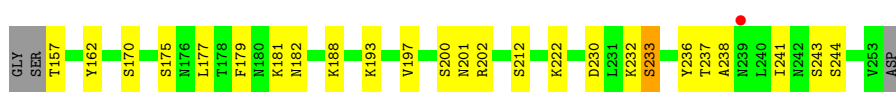
• Molecule 2: Extracellular Adherence Protein



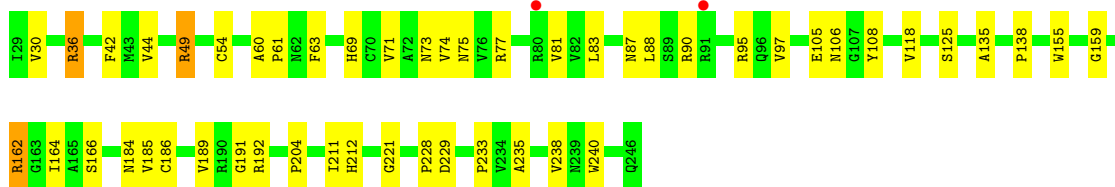
• Molecule 2: Extracellular Adherence Protein



• Molecule 2: Extracellular Adherence Protein

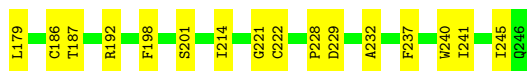


• Molecule 3: Neutrophil elastase

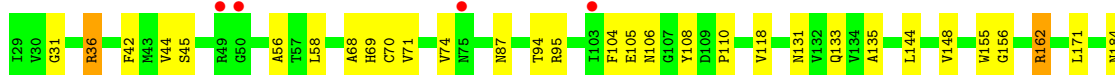
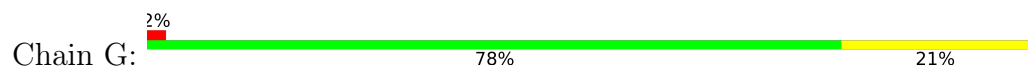


• Molecule 3: Neutrophil elastase

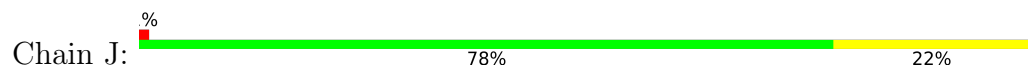




- Molecule 3: Neutrophil elastase



- Molecule 3: Neutrophil elastase



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	97.59Å 107.44Å 105.42Å 90.00° 91.52° 90.00°	Depositor
Resolution (Å)	43.83 – 3.10 47.06 – 3.10	Depositor EDS
% Data completeness (in resolution range)	95.9 (43.83-3.10) 95.6 (47.06-3.10)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.15 (at 3.12Å)	Xtriage
Refinement program	PHENIX v1.9.2	Depositor
R, R_{free}	0.184 , 0.258 0.186 , 0.258	Depositor DCC
R_{free} test set	1986 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	51.0	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 42.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.020 for -h,-l,-k 0.001 for -h,l,k 0.032 for h,-k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	16684	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% 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	C	0.45	0/1814	0.76	0/2447
1	F	0.46	0/1814	0.74	0/2447
1	I	0.47	0/1814	0.74	0/2447
1	L	0.41	0/1814	0.73	0/2447
2	B	0.45	0/762	0.70	0/1029
2	E	0.60	0/762	0.72	0/1029
2	H	0.54	0/762	0.73	0/1029
2	K	0.49	0/762	0.69	0/1029
3	A	0.43	0/1665	0.70	0/2263
3	D	0.49	0/1665	0.71	0/2263
3	G	0.40	0/1665	0.65	0/2263
3	J	0.38	0/1665	0.65	0/2263
All	All	0.45	0/16964	0.71	0/22956

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	C	1780	0	1793	39	0
1	F	1780	0	1793	45	0
1	I	1780	0	1793	35	0
1	L	1780	0	1793	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	756	0	788	16	0
2	E	756	0	788	30	0
2	H	756	0	788	17	0
2	K	756	0	788	18	0
3	A	1635	0	1652	30	0
3	D	1635	0	1652	43	0
3	G	1635	0	1652	31	0
3	J	1635	0	1652	33	0
All	All	16684	0	16932	311	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 9.

The worst 5 of 311 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:E:239:ASN:HD22	3:D:69:HIS:CD2	1.64	1.14
2:E:239:ASN:ND2	3:D:69:HIS:CD2	2.40	0.89
2:E:239:ASN:ND2	3:D:69:HIS:HD2	1.72	0.87
3:G:144:LEU:HD22	3:G:148:VAL:HG11	1.56	0.85
2:E:239:ASN:HD22	3:D:69:HIS:HD2	0.87	0.84

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	C	221/223 (99%)	213 (96%)	7 (3%)	1 (0%)	29 64
1	F	221/223 (99%)	205 (93%)	14 (6%)	2 (1%)	17 52
1	I	221/223 (99%)	214 (97%)	7 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	221/223 (99%)	212 (96%)	9 (4%)	0	100	100
2	B	95/100 (95%)	92 (97%)	3 (3%)	0	100	100
2	E	95/100 (95%)	93 (98%)	2 (2%)	0	100	100
2	H	95/100 (95%)	91 (96%)	4 (4%)	0	100	100
2	K	95/100 (95%)	91 (96%)	4 (4%)	0	100	100
3	A	216/218 (99%)	197 (91%)	17 (8%)	2 (1%)	17	52
3	D	216/218 (99%)	197 (91%)	17 (8%)	2 (1%)	17	52
3	G	216/218 (99%)	199 (92%)	16 (7%)	1 (0%)	29	64
3	J	216/218 (99%)	201 (93%)	14 (6%)	1 (0%)	29	64
All	All	2128/2164 (98%)	2005 (94%)	114 (5%)	9 (0%)	34	69

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	161	ASN
1	C	130	GLY
1	F	130	GLY
3	A	162	ARG
1	F	131	LEU

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	C	190/190 (100%)	185 (97%)	5 (3%)	46	74
1	F	190/190 (100%)	183 (96%)	7 (4%)	34	66
1	I	190/190 (100%)	186 (98%)	4 (2%)	53	79
1	L	190/190 (100%)	186 (98%)	4 (2%)	53	79
2	B	90/92 (98%)	84 (93%)	6 (7%)	16	46
2	E	90/92 (98%)	86 (96%)	4 (4%)	28	61
2	H	90/92 (98%)	85 (94%)	5 (6%)	21	52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	K	90/92 (98%)	85 (94%)	5 (6%)	21	52
3	A	172/172 (100%)	169 (98%)	3 (2%)	60	83
3	D	172/172 (100%)	169 (98%)	3 (2%)	60	83
3	G	172/172 (100%)	171 (99%)	1 (1%)	86	94
3	J	172/172 (100%)	171 (99%)	1 (1%)	86	94
All	All	1808/1816 (100%)	1760 (97%)	48 (3%)	44	74

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	64	ASN
2	H	237	THR
1	I	69	LEU
2	H	206	ASP
3	G	36	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	133	GLN
2	K	182	ASN
2	K	183	GLN
1	F	166	GLN
3	A	75	ASN

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 monosaccharides 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	C	223/223 (100%)	-0.29	2 (0%) 84 69	19, 43, 75, 93	0
1	F	223/223 (100%)	-0.31	0 100 100	17, 39, 64, 80	0
1	I	223/223 (100%)	-0.38	0 100 100	18, 35, 61, 68	0
1	L	223/223 (100%)	-0.25	0 100 100	24, 49, 71, 82	0
2	B	97/100 (97%)	-0.33	0 100 100	23, 34, 54, 57	0
2	E	97/100 (97%)	-0.47	0 100 100	15, 29, 45, 52	0
2	H	97/100 (97%)	-0.42	0 100 100	14, 29, 44, 58	0
2	K	97/100 (97%)	-0.40	1 (1%) 82 67	27, 43, 68, 83	0
3	A	218/218 (100%)	-0.34	2 (0%) 84 69	22, 46, 65, 91	0
3	D	218/218 (100%)	-0.36	0 100 100	20, 35, 58, 82	0
3	G	218/218 (100%)	0.06	5 (2%) 60 39	33, 57, 85, 97	0
3	J	218/218 (100%)	0.05	2 (0%) 84 69	27, 72, 97, 108	0
All	All	2152/2164 (99%)	-0.26	12 (0%) 89 78	14, 43, 79, 108	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	75	ASN	4.1
1	C	127	ALA	3.2
3	G	50	GLY	3.0
3	G	49	ARG	2.6
3	G	246	GLN	2.6

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.