



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 03:00 PM UTC

PDB ID : 3G9K / pdb_00003g9k
Title : Crystal structure of Bacillus anthracis transpeptidase enzyme CapD
Authors : Zhang, R.; Wu, R.; Richter, S.; Anderson, V.J.; Missiakas, D.; Joachimiak, A.
Deposited on : 2009-02-13
Resolution : 1.79 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
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)	:	2.49

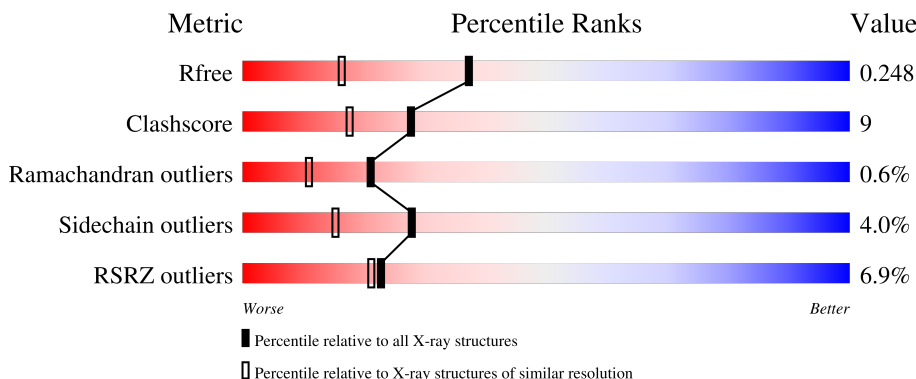
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.79 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}	180053	7662 (1.80-1.80)
Clashscore	190562	8479 (1.80-1.80)
Ramachandran outliers	187476	8391 (1.80-1.80)
Sidechain outliers	187428	8390 (1.80-1.80)
RSRZ outliers	180081	7663 (1.80-1.80)

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	D	323	
1	L	323	
2	F	177	
2	S	177	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7489 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.

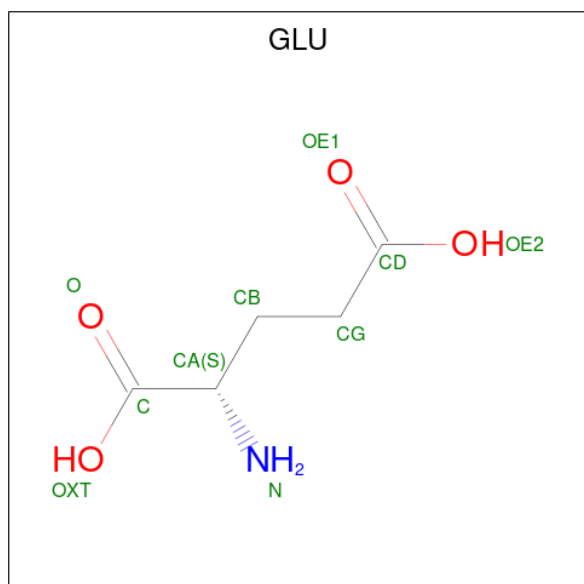
- Molecule 1 is a protein called Capsule biosynthesis protein capD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	280	Total	C	N	O	Se	0	0	0
			2160	1381	355	412	12			
1	D	280	Total	C	N	O	Se	0	0	0
			2163	1385	355	412	11			

- Molecule 2 is a protein called Capsule biosynthesis protein capD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	166	Total	C	N	O	Se	0	1	0
			1301	817	235	248	1			
2	F	166	Total	C	N	O	Se	0	0	0
			1291	813	231	246	1			

- Molecule 3 is GLUTAMIC ACID (CCD ID: GLU) (formula: C₅H₉NO₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	S	1	Total	C	N	O	0	0
			9	5	1	3		
3	S	1	Total	C	N	O	0	0
			10	5	1	4		

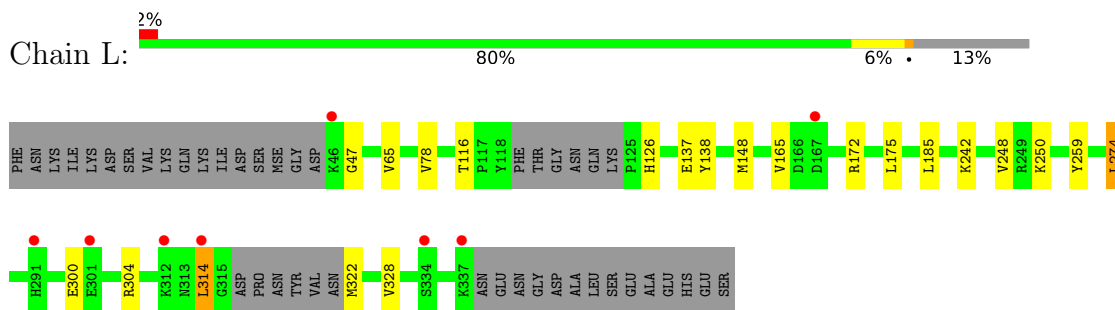
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	190	Total	O	0	0
			190	190		
4	S	72	Total	O	0	0
			72	72		
4	D	186	Total	O	0	0
			186	186		
4	F	107	Total	O	0	0
			107	107		

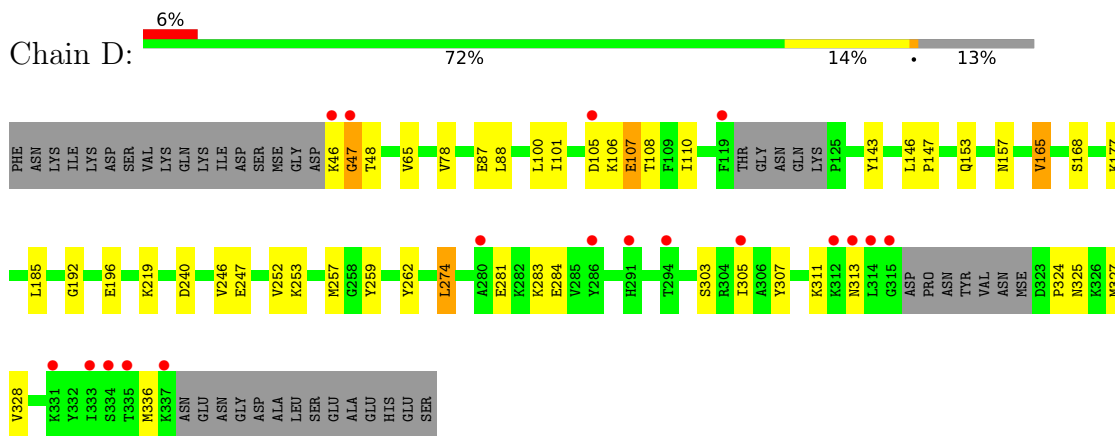
3 Residue-property plots [i](#)

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.

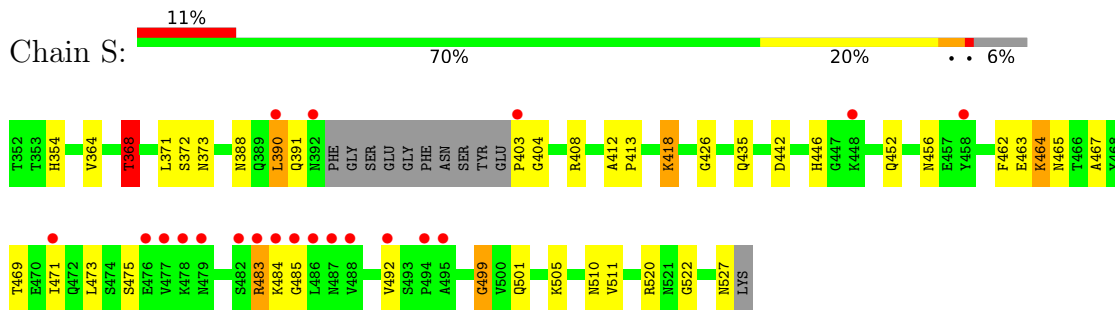
- Molecule 1: Capsule biosynthesis protein capD



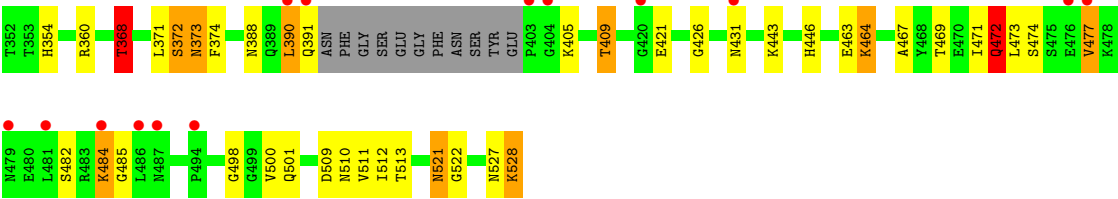
- Molecule 1: Capsule biosynthesis protein capD



- Molecule 2: Capsule biosynthesis protein capD



- Molecule 2: Capsule biosynthesis protein capD



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	52.67Å 120.62Å 77.38Å 90.00° 90.87° 90.00°	Depositor
Resolution (Å)	40.69 – 1.79 40.69 – 1.79	Depositor EDS
% Data completeness (in resolution range)	97.6 (40.69-1.79) 97.6 (40.69-1.79)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.197 , 0.242 0.202 , 0.248	Depositor DCC
R_{free} test set	4413 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	25.1	Xtriage
Anisotropy	0.641	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 49.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.079 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7489	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% 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	D	0.97	2/2193 (0.1%)	0.95	1/2932 (0.0%)
1	L	0.89	1/2189 (0.0%)	0.95	0/2926
2	F	1.08	5/1312 (0.4%)	1.01	2/1762 (0.1%)
2	S	1.03	2/1322 (0.2%)	0.97	4/1776 (0.2%)
All	All	0.98	10/7016 (0.1%)	0.96	7/9396 (0.1%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	240	ASP	C-N	8.33	1.44	1.33
2	F	467	ALA	CA-CB	-7.14	1.43	1.53
2	F	368	THR	CA-CB	5.97	1.62	1.53
2	F	472	GLN	C-N	5.96	1.41	1.33
2	S	368	THR	CA-CB	5.80	1.62	1.53

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	S	373	ASN	N-CA-C	-8.09	96.41	108.79
2	F	484	LYS	N-CA-C	6.87	121.36	112.92
2	S	464	LYS	CB-CA-C	-6.23	109.37	116.54
2	F	368	THR	CB-CA-C	5.84	119.85	110.16
2	S	368	THR	CB-CA-C	5.52	119.32	110.16

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	D	2163	0	2197	39	0
1	L	2160	0	2197	15	0
2	F	1291	0	1301	51	0
2	S	1301	0	1306	36	0
3	S	19	0	14	1	0
4	D	186	0	0	2	0
4	F	107	0	0	12	0
4	L	190	0	0	1	0
4	S	72	0	0	5	0
All	All	7489	0	7015	126	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 126 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:F:469:THR:HG22	2:F:471:ILE:O	1.41	1.20
2:F:511:VAL:HG13	2:F:528:LYS:HD2	1.30	1.12
1:D:257:MSE:HE3	2:F:446:HIS:NE2	1.64	1.11
2:F:521:ASN:O	2:F:521:ASN:ND2	1.86	1.05
2:F:469:THR:CG2	2:F:471:ILE:O	2.05	1.03

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	D	274/323 (85%)	267 (97%)	6 (2%)	1 (0%)	30 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	274/323 (85%)	272 (99%)	2 (1%)	0	100	100
2	F	162/177 (92%)	153 (94%)	7 (4%)	2 (1%)	10	3
2	S	163/177 (92%)	155 (95%)	6 (4%)	2 (1%)	10	3
All	All	873/1000 (87%)	847 (97%)	21 (2%)	5 (1%)	21	11

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	372	SER
2	F	405	LYS
2	S	404	GLY
1	D	47	GLY
2	S	499	GLY

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	D	232/257 (90%)	224 (97%)	8 (3%)	32	20
1	L	232/257 (90%)	227 (98%)	5 (2%)	45	34
2	F	141/149 (95%)	132 (94%)	9 (6%)	16	6
2	S	142/149 (95%)	134 (94%)	8 (6%)	19	8
All	All	747/812 (92%)	717 (96%)	30 (4%)	28	15

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	107	GLU
2	F	477	VAL
1	D	185	LEU
2	F	528	LYS
2	F	409	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	354	HIS
2	F	446	HIS
2	F	510	ASN
2	F	456	ASN
2	F	388	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	GLU	S	2	3	8,9,9	0.94	0	8,11,11	1.26	1 (12%)
3	GLU	S	1	3	7,8,9	0.92	0	4,9,11	1.10	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GLU	S	2	3	-	3/9/9/9	-
3	GLU	S	1	3	-	2/6/7/9	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	S	2	GLU	OXT-C-O	-3.03	117.20	124.08

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	S	2	GLU	C-CA-CB-CG
3	S	2	GLU	OXT-C-CA-N
3	S	2	GLU	O-C-CA-N
3	S	1	GLU	OE2-CD-CG-CB
3	S	1	GLU	OE1-CD-CG-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	S	2	GLU	1	0

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	D	269/323 (83%)	0.43	18 (6%) 24 22	14, 25, 44, 58	0
1	L	268/323 (82%)	0.40	8 (2%) 52 52	14, 25, 43, 64	0
2	F	165/177 (93%)	0.54	14 (8%) 16 14	14, 24, 50, 56	0
2	S	165/177 (93%)	0.59	20 (12%) 8 7	9, 25, 55, 62	1 (0%)
All	All	867/1000 (86%)	0.47	60 (6%) 23 21	9, 25, 49, 64	1 (0%)

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	314	LEU	4.8
2	F	404	GLY	4.5
2	S	494	PRO	4.3
2	S	486	LEU	4.0
2	S	477	VAL	4.0

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	GLU	S	1	9/10	0.62	0.16	50,51,52,56	0
3	GLU	S	2	10/10	0.73	0.15	38,48,49,49	0

6.5 Other polymers [i](#)

There are no such residues in this entry.