W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report (i)

Jun 17, 2024 – 01:20 AM EDT

PDB ID	:	3GA9
Title	:	Crystal structure of Bacillus anthracis transpeptidase enzyme CapD, crystal form II
Authors	:	Zhang, R.; Wu, R.; Richter, S.; Anderson, V.J.; Missiakas, D.; Joachimiak, A.
Deposited on	:	2009-02-16
Resolution	:	2.30 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1
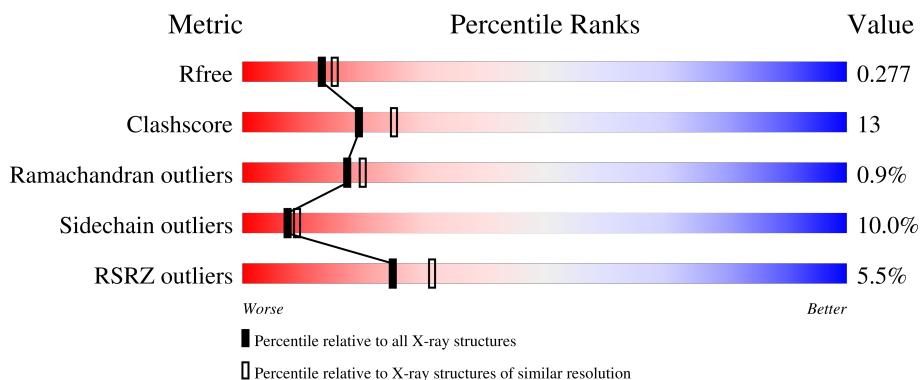
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION
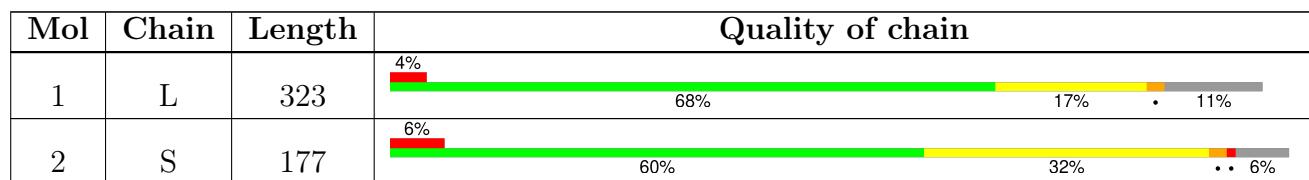
The reported resolution of this entry is 2.30 Å.

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	Percentile Ranks	Value
Rfree		0.277
Clashscore		13
Ramachandran outliers		0.9%
Sidechain outliers		10.0%
RSRZ outliers		5.5%

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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	L	323	4% / 68% 17% · 11%
2	S	177	6% / 60% 32% ·· 6%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 3605 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.
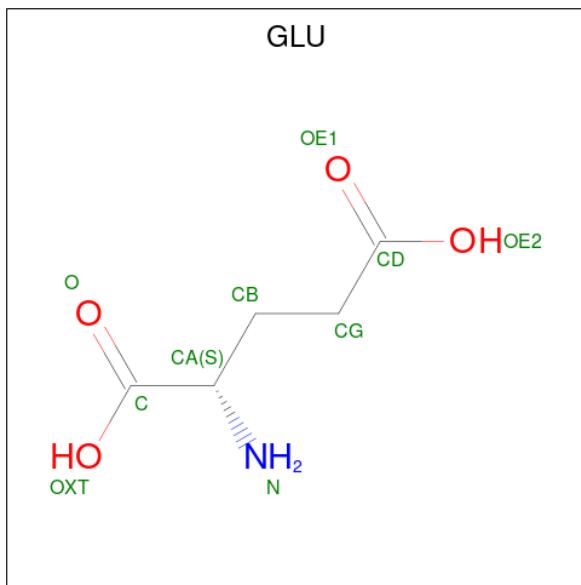
- Molecule 1 is a protein called Capsule biosynthesis protein capD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	286	Total	C	N	O	Se	0	0	0
			2208	1411	364	421	12			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	167	Total	C	N	O	Se	0	0	0
			1299	816	232	250	1			

- Molecule 3 is GLUTAMIC ACID (three-letter code: GLU) (formula: $C_5H_9NO_4$).

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		
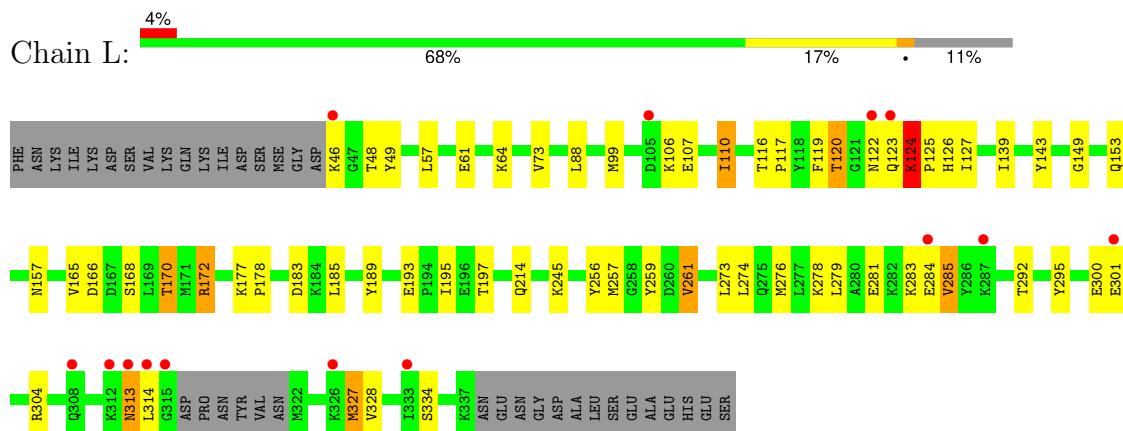
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	L	58	Total O 58 58		0	0
4	S	21	Total O 21 21		0	0
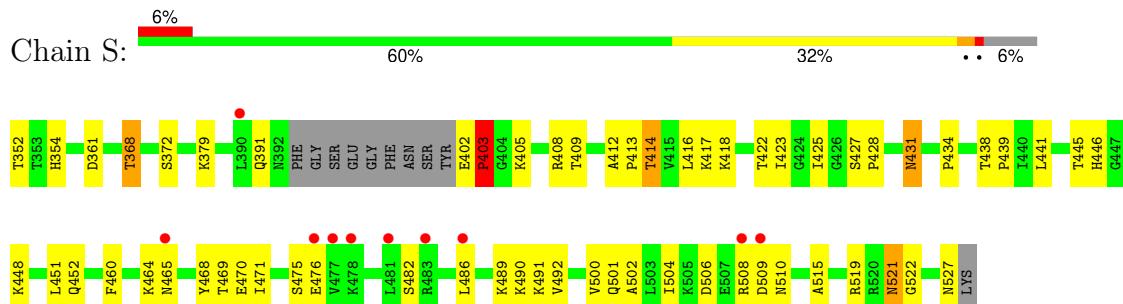
3 Residue-property plots ⓘ

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: Capsule biosynthesis protein capD

• Molecule 2: Capsule biosynthesis protein capD

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	47.70Å 72.00Å 128.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.74 – 2.30 44.74 – 2.28	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.74-2.30) 97.8 (44.74-2.28)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.28 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.5.0054	Depositor
R, R_{free}	0.197 , 0.252 0.213 , 0.277	Depositor DCC
R_{free} test set	1037 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	48.9	Xtriage
Anisotropy	0.076	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3605	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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 (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	L	0.77	0/2239	0.74	0/2995
2	S	0.78	0/1320	0.78	0/1775
All	All	0.77	0/3559	0.75	0/4770

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	L	2208	0	2243	56	0
2	S	1299	0	1299	50	0
3	S	19	0	14	0	0
4	L	58	0	0	2	0
4	S	21	0	0	3	0
All	All	3605	0	3556	91	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.

All (91) 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)
1:L:257:MSE:HE2	1:L:281:GLU:HB2	1.22	1.07
1:L:123:GLN:HG2	1:L:126:HIS:HE1	1.08	1.05
1:L:123:GLN:HG2	1:L:126:HIS:CE1	1.99	0.97
1:L:313:ASN:OD1	1:L:327:MSE:HE3	1.64	0.97
1:L:116:THR:HB	2:S:403:PRO:O	1.65	0.96
2:S:431:ASN:C	2:S:431:ASN:HD22	1.65	0.96
1:L:257:MSE:CE	1:L:281:GLU:HB2	2.05	0.87
1:L:257:MSE:HE2	1:L:281:GLU:CB	2.06	0.83
2:S:508:ARG:HA	2:S:509:ASP:HB3	1.59	0.83
1:L:124:LYS:HB2	1:L:125:PRO:HD2	1.62	0.81
1:L:48:THR:HG21	2:S:510:ASN:HB2	1.62	0.79
2:S:431:ASN:C	2:S:431:ASN:ND2	2.37	0.71
1:L:172:ARG:NH1	4:L:361:HOH:O	2.26	0.69
2:S:501:GLN:HE22	2:S:522:GLY:HA3	1.61	0.66
1:L:257:MSE:HE3	2:S:446:HIS:NE2	2.13	0.64
2:S:470:GLU:O	2:S:490:LYS:HE3	1.98	0.64
2:S:465:ASN:HA	2:S:486:LEU:HD23	1.81	0.62
1:L:301:GLU:HA	1:L:304:ARG:NH1	2.16	0.61
1:L:124:LYS:HB2	1:L:125:PRO:CD	2.26	0.61
1:L:123:GLN:CG	1:L:126:HIS:HE1	1.99	0.59
1:L:273:LEU:HD23	2:S:414:THR:HG21	1.85	0.58
1:L:106:LYS:HE2	1:L:143:TYR:HE1	1.69	0.57
1:L:279:LEU:O	1:L:283:LYS:HG2	2.04	0.57
2:S:402:GLU:HG3	2:S:403:PRO:HD2	1.87	0.57
2:S:448:LYS:HG2	2:S:448:LYS:O	2.05	0.57
2:S:448:LYS:O	2:S:448:LYS:CG	2.53	0.56
1:L:313:ASN:N	1:L:313:ASN:HD22	2.03	0.55
1:L:170:THR:HB	1:L:195:ILE:O	2.06	0.55
2:S:368:THR:HG22	2:S:413:PRO:HG3	1.89	0.55
1:L:314:LEU:HD22	2:S:408:ARG:HE	1.73	0.54
2:S:492:VAL:HG23	4:S:68:HOH:O	2.08	0.54
1:L:119:PHE:CZ	1:L:124:LYS:HB3	2.44	0.53
2:S:452:GLN:OE1	2:S:515:ALA:HA	2.09	0.53
2:S:354:HIS:HD2	2:S:368:THR:HB	1.72	0.53
1:L:153:GLN:HE21	1:L:157:ASN:HD21	1.56	0.52
1:L:48:THR:CG2	2:S:510:ASN:HD22	2.23	0.52
2:S:368:THR:CG2	2:S:413:PRO:HG3	2.41	0.51
1:L:120:THR:HB	4:L:355:HOH:O	2.11	0.51
2:S:425:ILE:HG13	2:S:502:ALA:HB2	1.91	0.50
2:S:470:GLU:OE2	2:S:519:ARG:NE	2.44	0.49
1:L:123:GLN:HB3	1:L:124:LYS:O	2.12	0.49
1:L:49:TYR:CZ	2:S:527:ASN:HB3	2.48	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:285:VAL:HG22	1:L:295:TYR:CD1	2.48	0.49
1:L:124:LYS:HE3	1:L:125:PRO:O	2.12	0.48
1:L:153:GLN:HE21	1:L:157:ASN:ND2	2.12	0.48
1:L:259:TYR:CZ	2:S:418:LYS:HE3	2.49	0.48
2:S:427:SER:HB2	2:S:500:VAL:HG22	1.96	0.48
2:S:508:ARG:O	2:S:508:ARG:HG2	2.14	0.48
1:L:256:TYR:O	1:L:259:TYR:HB2	2.14	0.47
1:L:177:LYS:HB3	1:L:178:PRO:HD3	1.96	0.47
2:S:412:ALA:O	2:S:414:THR:HG23	2.15	0.47
1:L:127:ILE:HD11	2:S:379:LYS:HD2	1.96	0.47
1:L:301:GLU:HG2	1:L:304:ARG:HH12	1.80	0.47
1:L:123:GLN:CG	1:L:126:HIS:CE1	2.85	0.46
1:L:99:MSE:SE	1:L:139:ILE:HD13	2.66	0.46
1:L:48:THR:HG23	4:S:72:HOH:O	2.15	0.46
2:S:354:HIS:HD2	2:S:368:THR:CG2	2.29	0.46
1:L:48:THR:HG22	2:S:510:ASN:HD22	1.80	0.46
1:L:276:MSE:SE	2:S:434:PRO:HB2	2.65	0.45
2:S:460:PHE:HA	2:S:468:TYR:O	2.16	0.45
1:L:256:TYR:CE2	2:S:445:THR:HG21	2.52	0.45
2:S:438:THR:HB	2:S:439:PRO:HD3	1.97	0.45
2:S:416:LEU:HB2	2:S:423:ILE:HB	1.98	0.45
2:S:464:LYS:O	2:S:465:ASN:HB2	2.17	0.45
1:L:110:ILE:HG13	1:L:139:ILE:HG12	1.99	0.45
2:S:451:LEU:HD22	2:S:504:ILE:HD11	1.98	0.45
1:L:285:VAL:HG22	1:L:285:VAL:O	2.16	0.44
2:S:354:HIS:CD2	2:S:368:THR:HB	2.52	0.44
1:L:73:VAL:HG21	1:L:143:TYR:O	2.17	0.44
2:S:354:HIS:HD2	2:S:368:THR:HG22	1.82	0.44
1:L:166:ASP:OD2	1:L:168:SER:OG	2.25	0.44
1:L:189:TYR:HA	1:L:193:GLU:O	2.18	0.44
1:L:122:ASN:N	1:L:122:ASN:OD1	2.50	0.43
1:L:285:VAL:HG22	1:L:295:TYR:CE1	2.52	0.43
2:S:468:TYR:HA	2:S:489:LYS:O	2.18	0.43
1:L:283:LYS:HG2	1:L:283:LYS:H	1.70	0.42
2:S:352:THR:HB	2:S:428:PRO:HB3	2.00	0.42
1:L:285:VAL:O	1:L:285:VAL:CG2	2.68	0.42
1:L:259:TYR:CD2	2:S:418:LYS:HG3	2.55	0.42
2:S:508:ARG:HA	2:S:509:ASP:CB	2.36	0.41
2:S:405:LYS:HG3	4:S:55:HOH:O	2.20	0.41
2:S:417:LYS:CG	2:S:422:THR:HG23	2.51	0.41
1:L:117:PRO:HG2	1:L:124:LYS:HE2	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:S:521:ASN:O	2:S:521:ASN:CG	2.59	0.41
1:L:261:VAL:HG12	2:S:416:LEU:HD12	2.03	0.41
2:S:508:ARG:O	2:S:508:ARG:CG	2.69	0.41
1:L:257:MSE:CE	2:S:446:HIS:CE1	3.04	0.40
1:L:149:GLY:HA2	1:L:214:GLN:HG3	2.02	0.40
2:S:475:SER:O	2:S:476:GLU:C	2.59	0.40
1:L:57:LEU:O	1:L:61:GLU:HG2	2.20	0.40
1:L:170:THR:HG21	1:L:197:THR:OG1	2.21	0.40

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	L	282/323 (87%)	273 (97%)	7 (2%)	2 (1%)	22	26
2	S	163/177 (92%)	149 (91%)	12 (7%)	2 (1%)	13	14
All	All	445/500 (89%)	422 (95%)	19 (4%)	4 (1%)	17	20

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	124	LYS
1	L	284	GLU
2	S	403	PRO
2	S	372	SER

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	L	237/257 (92%)	214 (90%)	23 (10%)	8	9
2	S	142/149 (95%)	128 (90%)	14 (10%)	8	9
All	All	379/406 (93%)	342 (90%)	37 (10%)	7	9

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	46	LYS
1	L	64	LYS
1	L	88	LEU
1	L	107	GLU
1	L	110	ILE
1	L	120	THR
1	L	124	LYS
1	L	165	VAL
1	L	170	THR
1	L	172	ARG
1	L	183	ASP
1	L	185	LEU
1	L	245	LYS
1	L	261	VAL
1	L	274	LEU
1	L	278	LYS
1	L	285	VAL
1	L	292	THR
1	L	300	GLU
1	L	313	ASN
1	L	327	MSE
1	L	328	VAL
1	L	334	SER
2	S	361	ASP
2	S	368	THR
2	S	391	GLN
2	S	403	PRO
2	S	409	THR
2	S	414	THR
2	S	431	ASN
2	S	441	LEU
2	S	469	THR
2	S	471	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	S	482	SER
2	S	491	LYS
2	S	506	ASP
2	S	521	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	L	68	ASN
1	L	126	HIS
1	L	157	ASN
1	L	313	ASN
1	L	325	ASN
2	S	354	HIS
2	S	392	ASN
2	S	431	ASN
2	S	435	GLN
2	S	456	ASN
2	S	465	ASN
2	S	501	GLN
2	S	510	ASN

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 ⓘ

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	1	3	7,8,9	0.99	0	4,9,11	0.65	0
3	GLU	S	2	3	8,9,9	1.28	0	8,11,11	1.49	1 (12%)

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	1	3	-	1/6/7/9	-
3	GLU	S	2	3	-	1/9/9/9	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	S	2	GLU	OXT-C-O	-3.51	116.12	124.08

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	S	1	GLU	CA-CB-CG-CD
3	S	2	GLU	CA-CB-CG-CD

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

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	L	274/323 (84%)	0.22	14 (5%) 28	35	34, 53, 83, 98	0
2	S	166/177 (93%)	0.39	10 (6%) 21	28	38, 58, 87, 94	0
All	All	440/500 (88%)	0.28	24 (5%) 25	31	34, 54, 85, 98	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	S	509	ASP	8.4
1	L	333	ILE	6.2
2	S	483	ARG	6.1
1	L	315	GLY	5.7
1	L	314	LEU	4.8
1	L	122	ASN	4.2
2	S	486	LEU	3.7
2	S	476	GLU	3.7
2	S	465	ASN	3.3
1	L	308	GLN	3.1
1	L	123	GLN	3.1
1	L	312	LYS	3.0
1	L	287	LYS	2.9
2	S	508	ARG	2.6
1	L	46	LYS	2.6
2	S	477	VAL	2.5
1	L	284	GLU	2.5
1	L	313	ASN	2.5
1	L	326	LYS	2.3
2	S	390	LEU	2.3
2	S	481	LEU	2.2
2	S	478	LYS	2.0
1	L	105	ASP	2.0
1	L	301	GLU	2.0

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 ⓘ

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($Å^2$)	Q<0.9
3	GLU	S	2	10/10	0.62	0.29	84,91,92,92	0
3	GLU	S	1	9/10	0.82	0.20	92,92,93,93	0

6.5 Other polymers ⓘ

There are no such residues in this entry.