wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 25, 2023 – 11:09 AM EDT

PDB ID : 5VU3

Title : Crystal structure of the competence-damaged protein (CinA) superfamily protein ECL_02051 from Enterobacter cloacae

Authors : Stogios, P.J.; Skarina, T.; McChesney, C.; Sandoval, J.; Di Leo, R.; Savchenko, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)

Deposited on : 2017-05-18
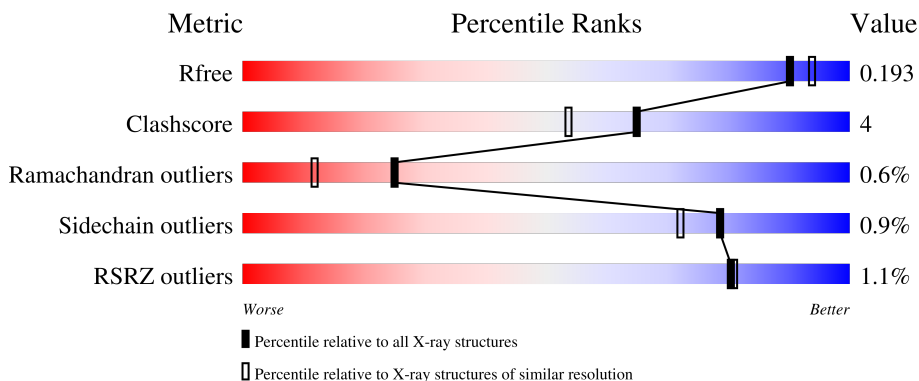
Resolution : 1.87 Å(reported)

---

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
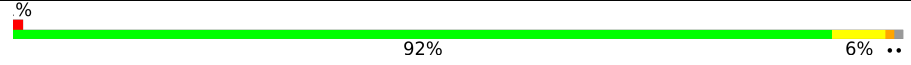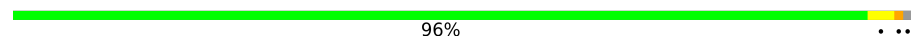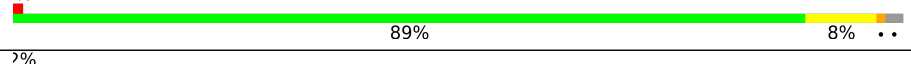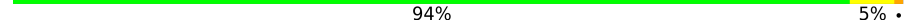*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.87 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		0.193
Clashscore		4
Ramachandran outliers		0.6%
Sidechain outliers		0.9%
RSRZ outliers		1.1%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
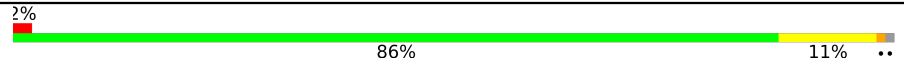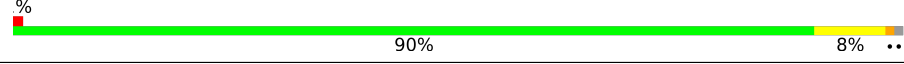
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	170	92% 6% ••
1	B	170	96% • ••
1	C	170	89% 8% ••
1	D	170	94% 5% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	170	2% 86% 11% ··
1	F	170	2% 86% 12% ··
1	G	170	% 90% 8% ··
1	H	170	89% 7% · ·

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 11834 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 Competence damage-inducible protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	169	Total	C	N	O	S	0	5	0
			1257	781	206	262	8			
1	B	169	Total	C	N	O	S	0	7	0
			1271	790	212	261	8			
1	C	166	Total	C	N	O	S	0	8	0
			1256	783	209	258	6			
1	D	170	Total	C	N	O	S	0	4	0
			1263	784	210	261	8			
1	E	168	Total	C	N	O	S	0	5	0
			1258	781	208	262	7			
1	F	169	Total	C	N	O	S	0	2	0
			1247	775	206	258	8			
1	G	168	Total	C	N	O	S	0	2	0
			1239	769	205	258	7			
1	H	165	Total	C	N	O	S	0	4	0
			1227	764	202	255	6			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A0A0H3CM13
B	0	GLY	-	expression tag	UNP A0A0H3CM13
C	0	GLY	-	expression tag	UNP A0A0H3CM13
D	0	GLY	-	expression tag	UNP A0A0H3CM13
E	0	GLY	-	expression tag	UNP A0A0H3CM13
F	0	GLY	-	expression tag	UNP A0A0H3CM13
G	0	GLY	-	expression tag	UNP A0A0H3CM13
H	0	GLY	-	expression tag	UNP A0A0H3CM13

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

ACT

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 4	C 2	O 2	0	0

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).

PG4
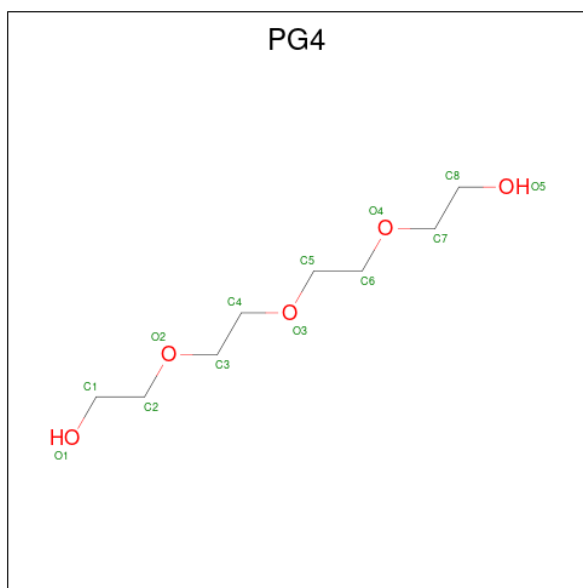
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	G	1	Total 13	C 8	O 5	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	281	Total 284	O 284	0	3
4	B	266	Total 269	O 269	0	3
4	C	260	Total 262	O 262	0	2
4	D	240	Total 241	O 241	0	1
4	E	181	Total 185	O 185	0	4
4	F	172	Total 174	O 174	0	2
4	G	204	Total 205	O 205	0	1
4	H	175	Total 179	O 179	0	4

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: Competence damage-inducible protein A
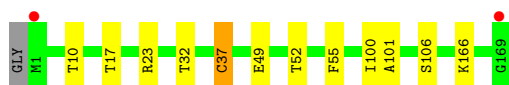
Chain A:

• Molecule 1: Competence damage-inducible protein A
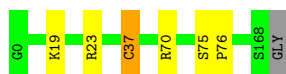
Chain B:

• Molecule 1: Competence damage-inducible protein A

Chain C:

• Molecule 1: Competence damage-inducible protein A

Chain D:

• Molecule 1: Competence damage-inducible protein A

Chain E:

• Molecule 1: Competence damage-inducible protein A

Chain F:

2%

86%   12%

G0 H7 A21 I25 D26 R30 E35 S36 C37 L42 N51 T52 F55 V61 V62 F63 E87 Q88 T89 M93 I100 V105 N132 G135 T139 A161 T165 S168 GLY

• Molecule 1: Competence damage-inducible protein A

Chain G:

%

90%   8%

GLY MET M2 S36 C37 V62 F63 S64 D65 K68 L79 E87 V90 M93 S110 E119 C128 S140 R141 G169

• Molecule 1: Competence damage-inducible protein A

Chain H:

89%   7%

GLY MET MET SER N4 I11 E15 E35 S36 C37 L42 T52 F55 I107 G118 E119 T122 C148 L163 S168 GLY

WORLDWIDE
wwPDB
PROTEIN DATA BANK
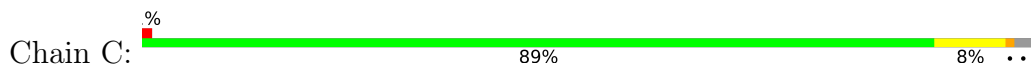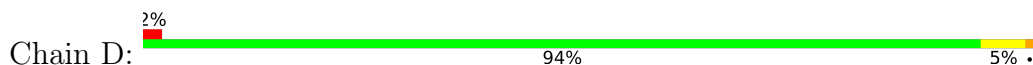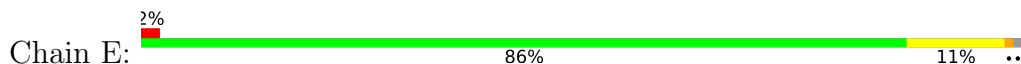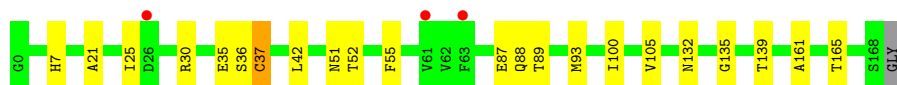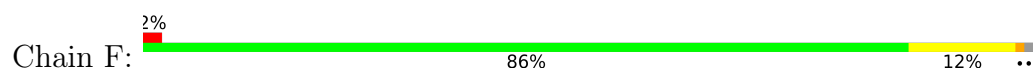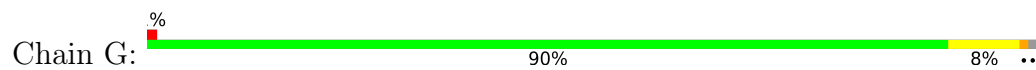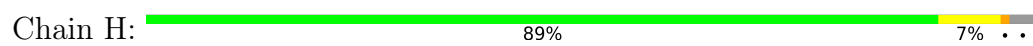
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	44.53Å 52.83Å 132.16Å 85.27° 87.19° 68.98°	Depositor
Resolution (Å)	29.87 – 1.87 29.87 – 1.87	Depositor EDS
% Data completeness (in resolution range)	94.0 (29.87-1.87) 90.1 (29.87-1.87)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 1.87Å)	Xtriage
Refinement program	PHENIX (dev_2733)	Depositor
R, R_{free}	0.151 , 0.193 0.151 , 0.193	Depositor DCC
R_{free} test set	2000 reflections (2.29%)	wwPDB-VP
Wilson B-factor (Å^2)	24.4	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 56.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	11834	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.98% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PG4, ACT

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	0.34	0/1286	0.52	0/1742
1	B	0.31	0/1306	0.50	0/1768
1	C	0.31	0/1294	0.52	0/1755
1	D	0.30	0/1289	0.48	0/1745
1	E	0.30	0/1287	0.49	0/1743
1	F	0.27	0/1267	0.47	0/1718
1	G	0.30	0/1259	0.50	0/1706
1	H	0.27	0/1253	0.48	0/1701
All	All	0.30	0/10241	0.50	0/13878

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	1257	0	1255	10	0
1	B	1271	0	1279	4	0
1	C	1256	0	1266	15	0
1	D	1263	0	1261	7	0
1	E	1258	0	1249	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1247	0	1240	14	0
1	G	1239	0	1226	13	0
1	H	1227	0	1221	10	0
2	A	4	0	3	0	0
3	G	13	0	18	2	0
4	A	284	0	0	4	0
4	B	269	0	0	0	0
4	C	262	0	0	6	0
4	D	241	0	0	2	0
4	E	185	0	0	3	0
4	F	174	0	0	4	0
4	G	205	0	0	2	0
4	H	179	0	0	3	0
All	All	11834	0	10018	83	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 4.

The worst 5 of 83 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:G:90:VAL:HA	1:G:93:MET:HE2	1.58	0.84
1:A:100:ILE:O	1:C:70[B]:ARG:NH1	2.20	0.70
1:H:148[B]:CYS:SG	4:H:343:HOH:O	2.50	0.68
1:C:81[B]:ARG:NH2	4:C:203:HOH:O	2.28	0.64
1:E:65:ASP:OD1	4:E:201[A]:HOH:O	2.15	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	172/170 (101%)	164 (95%)	7 (4%)	1 (1%)	25	12
1	B	174/170 (102%)	166 (95%)	7 (4%)	1 (1%)	25	12
1	C	172/170 (101%)	164 (95%)	7 (4%)	1 (1%)	25	12
1	D	172/170 (101%)	166 (96%)	5 (3%)	1 (1%)	25	12
1	E	171/170 (101%)	164 (96%)	6 (4%)	1 (1%)	25	12
1	F	169/170 (99%)	160 (95%)	8 (5%)	1 (1%)	25	12
1	G	168/170 (99%)	161 (96%)	6 (4%)	1 (1%)	25	12
1	H	167/170 (98%)	160 (96%)	6 (4%)	1 (1%)	25	12
All	All	1365/1360 (100%)	1305 (96%)	52 (4%)	8 (1%)	25	12

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	37	CYS
1	B	37	CYS
1	C	37	CYS
1	D	37	CYS
1	E	37	CYS

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	138/133 (104%)	138 (100%)	0	100	100
1	B	140/133 (105%)	139 (99%)	1 (1%)	84	79
1	C	139/133 (104%)	136 (98%)	3 (2%)	52	36
1	D	137/133 (103%)	137 (100%)	0	100	100
1	E	137/133 (103%)	133 (97%)	4 (3%)	42	26
1	F	135/133 (102%)	134 (99%)	1 (1%)	84	79
1	G	134/133 (101%)	133 (99%)	1 (1%)	84	79
1	H	134/133 (101%)	133 (99%)	1 (1%)	84	79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1094/1064 (103%)	1083 (99%)	11 (1%)	78 69

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	116	GLU
1	F	30	ARG
1	H	119	GLU
1	G	140	SER
1	E	2	MET

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

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	PG4	G	201	-	12,12,12	0.59	0	11,11,11	1.66	3 (27%)
2	ACT	A	201	-	3,3,3	1.27	0	3,3,3	1.48	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	PG4	G	201	-	-	4/10/10/10	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	201	PG4	O3-C4-C3	3.08	124.30	110.39
3	G	201	PG4	O2-C3-C4	-2.98	96.95	110.39
3	G	201	PG4	C5-O3-C4	-2.46	102.64	113.29

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	201	PG4	C6-C5-O3-C4
3	G	201	PG4	C5-C6-O4-C7
3	G	201	PG4	O2-C3-C4-O3
3	G	201	PG4	O3-C5-C6-O4

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	201	PG4	2	0

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(Å²)	Q<0.9
1	A	169/170 (99%)	-0.56	2 (1%)	79	79	16, 25, 48, 96	0
1	B	169/170 (99%)	-0.57	0	100	100	15, 25, 43, 69	0
1	C	166/170 (97%)	-0.42	2 (1%)	79	79	16, 25, 57, 98	0
1	D	170/170 (100%)	-0.39	3 (1%)	68	68	19, 29, 54, 98	0
1	E	168/170 (98%)	-0.34	4 (2%)	59	57	20, 29, 64, 141	0
1	F	169/170 (99%)	-0.07	3 (1%)	68	68	24, 41, 66, 80	0
1	G	168/170 (98%)	-0.31	1 (0%)	89	89	24, 37, 62, 103	0
1	H	165/170 (97%)	-0.20	0	100	100	23, 35, 65, 97	0
All	All	1344/1360 (98%)	-0.36	15 (1%)	80	81	15, 31, 62, 141	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	3	SER	4.8
1	E	169	GLY	4.8
1	G	169	GLY	4.4
1	E	2	MET	4.2
1	C	119	GLU	3.5

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 (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(Å^2)	Q<0.9
3	PG4	G	201	13/13	0.80	0.16	48,53,66,68	0
2	ACT	A	201	4/4	0.91	0.10	46,46,47,50	0

6.5 Other polymers (i)

There are no such residues in this entry.