Full wwPDB X-ray Structure Validation Report (i)

Jan 7, 2024 – 09:50 pm GMT

PDB ID	:	6EUU
Title	:	The Transcriptional Regulator PrfA from Listeria Monocytogenes in complex with a ring-fused 2-pyridone (KSK29)
Authors	:	Begum, A.; Hall, M.; Grundstrom, C.; Kulen, M.; Lindgren, M.; Johansson, J.; Almqvist, F.; Sauer, U.H.; Sauer-Eriksson, A.E.
Deposited on	:	2017-10-31
Resolution	:	2.60 Å(reported)

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

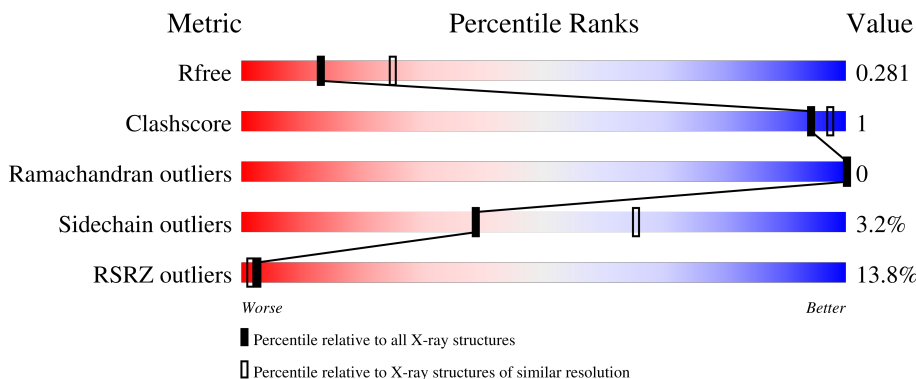|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.60 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		0.281
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		3.2%
RSRZ outliers		13.8%

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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	239	19% / 90% 5% 5%
1	B	239	7% / 90% 5% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NA	A	301	-	-	-	X

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 3796 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 Listeriolysin positive regulatory factor A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	227	Total	C	N	O	S	0	2	0
			1879	1228	295	348	8			
1	B	228	Total	C	N	O	S	0	0	0
			1865	1219	291	348	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q4TVQ0
A	0	ALA	-	expression tag	UNP Q4TVQ0
B	-1	GLY	-	expression tag	UNP Q4TVQ0
B	0	ALA	-	expression tag	UNP Q4TVQ0

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		
2	B	1	Total	Na	0	0
			1	1		

- Molecule 3 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).

IPA

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 4	C 3	O 1	0	0
3	A	1	Total 4	C 3	O 1	0	0
3	B	1	Total 4	C 3	O 1	0	0
3	B	1	Total 4	C 3	O 1	0	0

- Molecule 4 is (3 {R})-8-cyclopropyl-5-oxidanylidene-7-(quinolin-3-ylmethyl)-2,3-dihydro-[1, 3]thiazolo[3,2-a]pyridine-3-carboxylic acid (three-letter code: K29) (formula: $C_{21}H_{18}N_2O_3S$).

K29
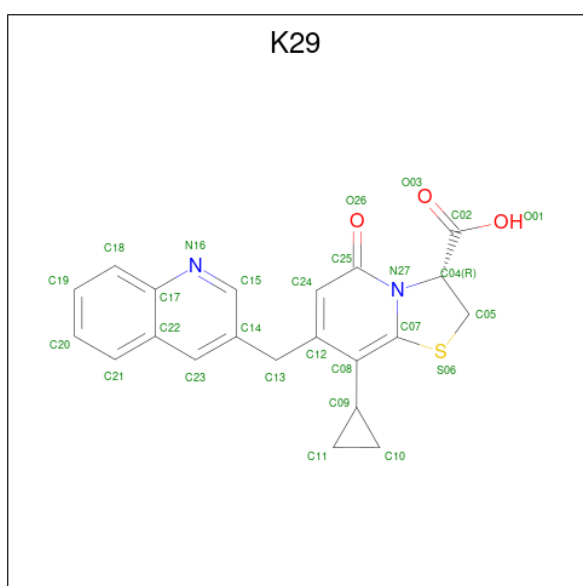
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	S	0	0
			27	21	2	3	1		

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	5	Total	O	0	0
			5	5		
5	B	2	Total	O	0	0
			2	2		

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: Listeriolysin positive regulatory factor A
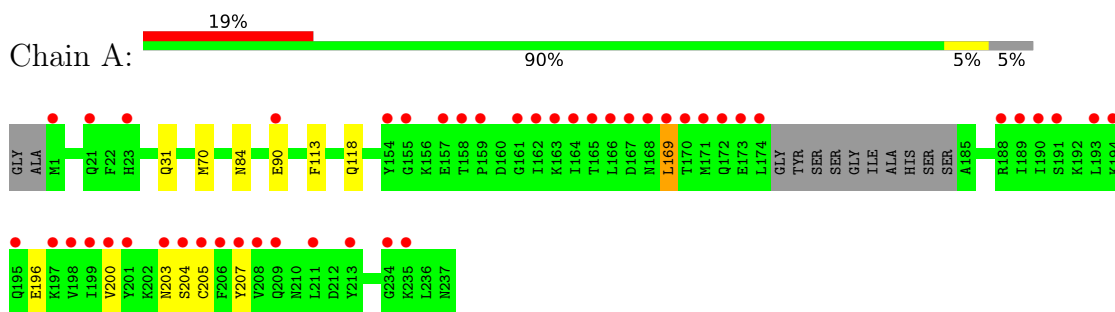
• Molecule 1: Listeriolysin positive regulatory factor A
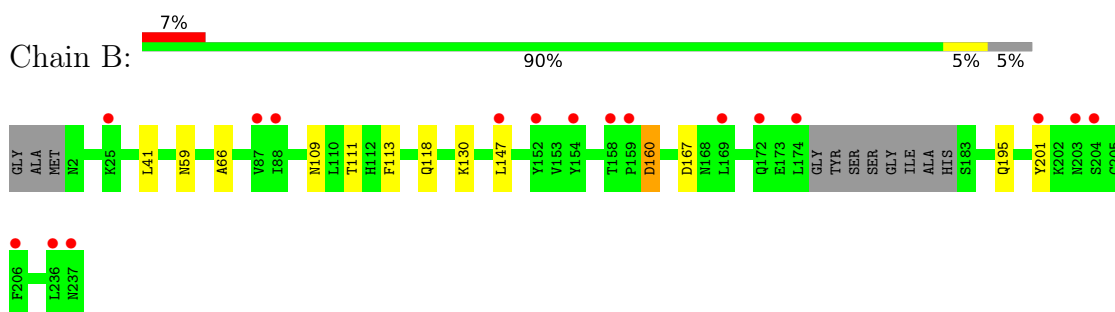
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.54Å 81.18Å 59.20Å 90.00° 111.88° 90.00°	Depositor
Resolution (Å)	45.50 – 2.60 45.50 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.50-2.60) 100.0 (45.50-2.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.37 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.213 , 0.271 0.216 , 0.281	Depositor DCC
R_{free} test set	723 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	65.5	Xtriage
Anisotropy	0.544	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 75.1	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3796	wwPDB-VP
Average B, all atoms (Å^2)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.12% 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: K29, NA, IPA

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	0.27	0/1921	0.41	0/2592
1	B	0.26	0/1906	0.40	0/2572
All	All	0.26	0/3827	0.41	0/5164

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	1879	0	1883	7	0
1	B	1865	0	1871	3	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	8	0	16	1	0
3	B	8	0	16	0	0
4	B	27	0	0	0	0
5	A	5	0	0	0	0
5	B	2	0	0	0	0
All	All	3796	0	3786	10	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 1.

All (10) 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:A:200:VAL:O	1:A:207:TYR:N	2.42	0.51
1:A:70:MET:O	3:A:302:IPA:H13	2.12	0.49
1:B:160:ASP:OD1	1:B:160:ASP:N	2.47	0.47
1:A:90:GLU:OE2	1:A:90:GLU:HA	2.15	0.46
1:A:31:GLN:N	1:A:84:ASN:OD1	2.51	0.43
1:B:109:ASN:OD1	1:B:111:THR:N	2.52	0.42
1:B:41:LEU:O	1:B:66:ALA:HA	2.19	0.41
1:A:203:ASN:O	1:A:204:SER:HB3	2.21	0.41
1:A:169:LEU:N	1:A:169:LEU:CD2	2.84	0.41
1:A:196:GLU:O	1:A:196:GLU:HG3	2.21	0.41

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	225/239 (94%)	215 (96%)	10 (4%)	0	100	100
1	B	224/239 (94%)	213 (95%)	11 (5%)	0	100	100
All	All	449/478 (94%)	428 (95%)	21 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	A	207/212 (98%)	203 (98%)	4 (2%)	57	79
1	B	206/212 (97%)	197 (96%)	9 (4%)	28	53
All	All	413/424 (97%)	400 (97%)	13 (3%)	39	66

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	113	PHE
1	A	118	GLN
1	A	169	LEU
1	A	205	CYS
1	B	59	ASN
1	B	113	PHE
1	B	118	GLN
1	B	130	LYS
1	B	147	LEU
1	B	160	ASP
1	B	167	ASP
1	B	195	GLN
1	B	201	TYR

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 (i)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 for Mogul analysis.

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	IPA	B	303	-	3,3,3	0.54	0	3,3,3	0.18	0
4	K29	B	301	-	29,31,31	1.06	0	36,46,46	1.22	1 (2%)
3	IPA	A	302	-	3,3,3	0.57	0	3,3,3	0.16	0
3	IPA	B	304	-	3,3,3	0.56	0	3,3,3	0.19	0
3	IPA	A	303	-	3,3,3	0.54	0	3,3,3	0.21	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
4	K29	B	301	-	-	2/12/23/23	0/5/5/5

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	301	K29	C13-C12-C24	-2.77	119.70	123.43

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	301	K29	O01-C02-C04-C05
4	B	301	K29	O03-C02-C04-C05

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	302	IPA	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand K29 B 301

Bond lengths

Bond angles

Torsions

Rings
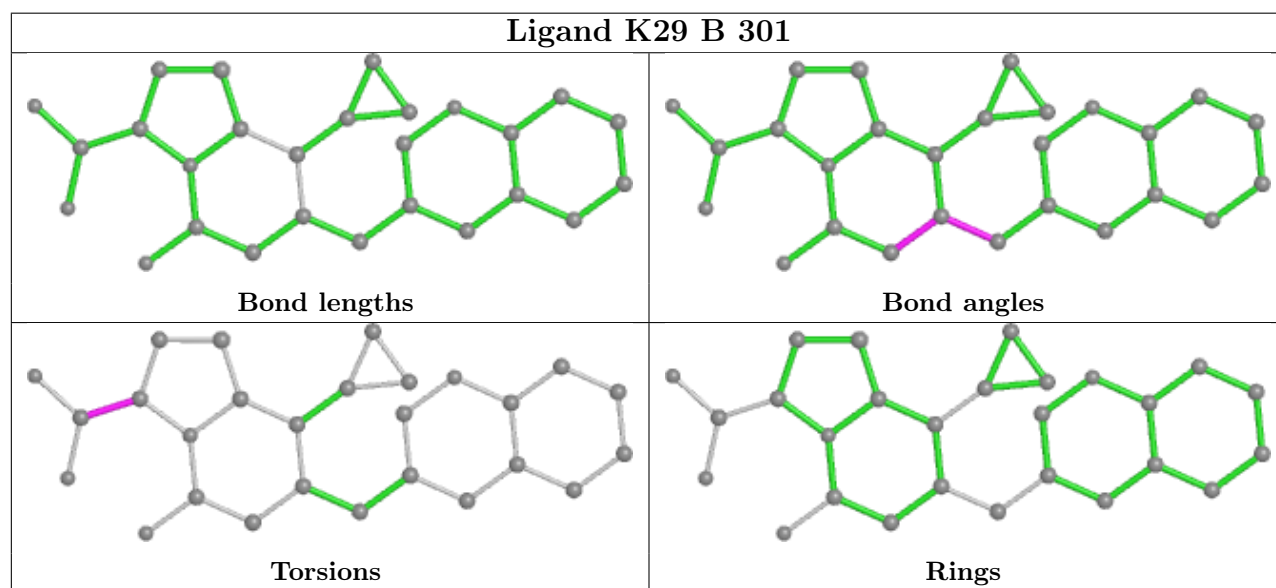
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	A	227/239 (94%)	1.09	46 (20%)	1	0	47, 79, 158, 183	0
1	B	228/239 (95%)	0.53	17 (7%)	14	10	47, 81, 139, 159	0
All	All	455/478 (95%)	0.81	63 (13%)	2	1	47, 81, 150, 183	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	163	LYS	11.2
1	A	168	ASN	8.3
1	A	199	ILE	8.3
1	A	165	THR	8.1
1	A	200	VAL	6.9
1	A	169	LEU	5.7
1	B	201	TYR	5.7
1	A	173	GLU	5.0
1	B	204	SER	4.9
1	A	207	TYR	4.8
1	A	201	TYR	4.7
1	A	190	ILE	4.7
1	A	155	GLY	4.4
1	A	161	GLY	4.4
1	A	170	THR	4.3
1	A	154	TYR	4.3
1	A	157	GLU	4.2
1	A	166	LEU	4.2
1	A	171	MET	4.1
1	A	208	VAL	4.0
1	A	194	LYS	4.0
1	A	195	GLN	3.8
1	A	206	PHE	3.8
1	A	164	ILE	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	197	LYS	3.7
1	A	158	THR	3.7
1	A	203	ASN	3.7
1	A	209	GLN	3.7
1	B	169	LEU	3.7
1	A	205	CYS	3.4
1	B	172	GLN	3.4
1	A	189	ILE	3.3
1	B	237	ASN	3.2
1	B	206	PHE	3.2
1	B	87	VAL	3.1
1	A	159	PRO	3.1
1	A	204	SER	3.1
1	A	172	GLN	3.1
1	B	159	PRO	3.0
1	A	167	ASP	2.9
1	A	174	LEU	2.8
1	A	193	LEU	2.8
1	B	158	THR	2.8
1	B	174	LEU	2.7
1	A	1	MET	2.6
1	A	191	SER	2.6
1	A	234	GLY	2.6
1	A	23[A]	HIS	2.5
1	A	198	VAL	2.5
1	A	211	LEU	2.5
1	A	213	TYR	2.5
1	A	188	ARG	2.4
1	B	147	LEU	2.4
1	B	88	ILE	2.4
1	B	236	LEU	2.4
1	B	25	LYS	2.3
1	A	235	LYS	2.2
1	A	90	GLU	2.2
1	B	152	TYR	2.1
1	B	154	TYR	2.1
1	A	21	GLN	2.0
1	A	162	ILE	2.0
1	B	203	ASN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.
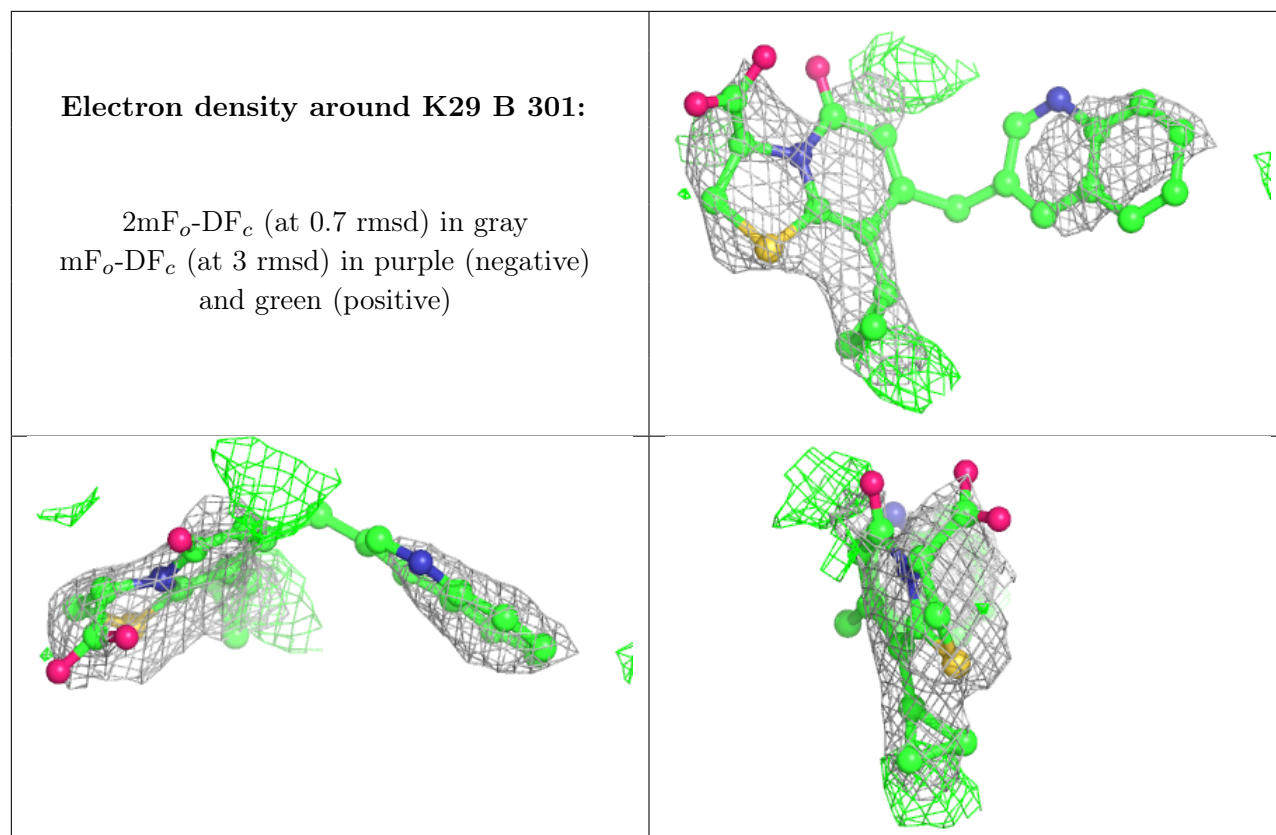
6.3 Carbohydrates (i)

There are no 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
2	NA	A	301	1/1	0.30	0.73	92,92,92,92	0
2	NA	B	302	1/1	0.30	0.19	87,87,87,87	0
4	K29	B	301	27/27	0.72	0.38	83,116,131,134	27
3	IPA	B	304	4/4	0.80	0.36	81,86,89,100	0
3	IPA	B	303	4/4	0.91	0.33	75,77,83,113	0
3	IPA	A	302	4/4	0.94	0.39	53,57,62,62	0
3	IPA	A	303	4/4	0.96	0.34	73,76,79,81	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around K29 B 301:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.