wwPDB X-ray Structure Validation Summary Report (i)

Feb 20, 2024 – 04:39 AM EST

PDB ID	:	4L3A
Title	:	Crystal structure of Internalin K (InlK) from Listeria monocytogenes
Authors	:	Neves, D.
Deposited on	:	2013-06-05
Resolution	:	2.59 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 (i) 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 (1)) were used in the production of this report:

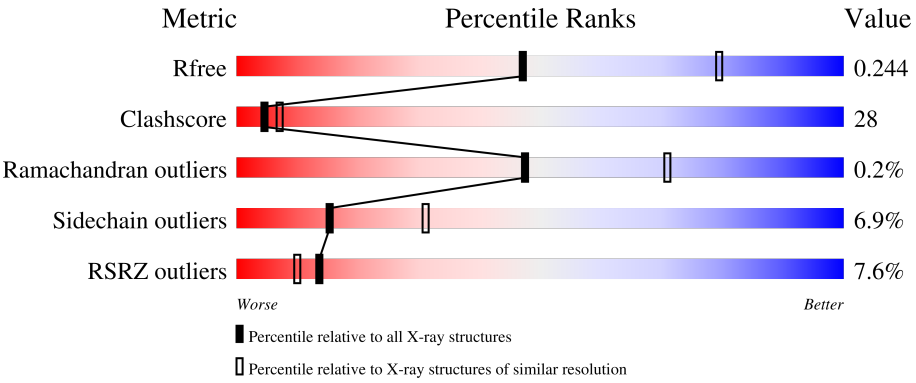
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.59 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	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	590	5% 54% 24% . 21%
1	B	590	6% 48% 16% . 33%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6832 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 Internalin K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	469	Total	C	N	O	S	0	0	0
			3631	2275	586	761	9			
1	B	395	Total	C	N	O	S	0	0	0
			3070	1919	500	643	8			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q8Y7I7
A	2	GLY	-	expression tag	UNP Q8Y7I7
A	3	SER	-	expression tag	UNP Q8Y7I7
A	4	SER	-	expression tag	UNP Q8Y7I7
A	5	HIS	-	expression tag	UNP Q8Y7I7
A	6	HIS	-	expression tag	UNP Q8Y7I7
A	7	HIS	-	expression tag	UNP Q8Y7I7
A	8	HIS	-	expression tag	UNP Q8Y7I7
A	9	HIS	-	expression tag	UNP Q8Y7I7
A	10	HIS	-	expression tag	UNP Q8Y7I7
A	11	SER	-	expression tag	UNP Q8Y7I7
A	12	SER	-	expression tag	UNP Q8Y7I7
A	13	GLY	-	expression tag	UNP Q8Y7I7
A	14	LEU	-	expression tag	UNP Q8Y7I7
A	15	VAL	-	expression tag	UNP Q8Y7I7
A	16	PRO	-	expression tag	UNP Q8Y7I7
A	17	ARG	-	expression tag	UNP Q8Y7I7
A	18	GLY	-	expression tag	UNP Q8Y7I7
A	19	SER	-	expression tag	UNP Q8Y7I7
A	20	HIS	-	expression tag	UNP Q8Y7I7
A	21	MET	-	expression tag	UNP Q8Y7I7
A	22	ALA	-	expression tag	UNP Q8Y7I7
A	23	SER	-	expression tag	UNP Q8Y7I7
A	24	MET	-	expression tag	UNP Q8Y7I7
A	25	THR	-	expression tag	UNP Q8Y7I7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	GLY	-	expression tag	UNP Q8Y7I7
A	27	GLY	-	expression tag	UNP Q8Y7I7
A	28	GLN	-	expression tag	UNP Q8Y7I7
A	29	GLN	-	expression tag	UNP Q8Y7I7
A	30	MET	-	expression tag	UNP Q8Y7I7
A	31	GLY	-	expression tag	UNP Q8Y7I7
A	32	ARG	-	expression tag	UNP Q8Y7I7
A	33	ASP	-	expression tag	UNP Q8Y7I7
A	576	VAL	-	expression tag	UNP Q8Y7I7
A	577	ASP	-	expression tag	UNP Q8Y7I7
A	578	LYS	-	expression tag	UNP Q8Y7I7
A	579	LEU	-	expression tag	UNP Q8Y7I7
A	580	ALA	-	expression tag	UNP Q8Y7I7
A	581	ALA	-	expression tag	UNP Q8Y7I7
A	582	ALA	-	expression tag	UNP Q8Y7I7
A	583	LEU	-	expression tag	UNP Q8Y7I7
A	584	GLU	-	expression tag	UNP Q8Y7I7
A	585	HIS	-	expression tag	UNP Q8Y7I7
A	586	HIS	-	expression tag	UNP Q8Y7I7
A	587	HIS	-	expression tag	UNP Q8Y7I7
A	588	HIS	-	expression tag	UNP Q8Y7I7
A	589	HIS	-	expression tag	UNP Q8Y7I7
A	590	HIS	-	expression tag	UNP Q8Y7I7
B	1	MET	-	expression tag	UNP Q8Y7I7
B	2	GLY	-	expression tag	UNP Q8Y7I7
B	3	SER	-	expression tag	UNP Q8Y7I7
B	4	SER	-	expression tag	UNP Q8Y7I7
B	5	HIS	-	expression tag	UNP Q8Y7I7
B	6	HIS	-	expression tag	UNP Q8Y7I7
B	7	HIS	-	expression tag	UNP Q8Y7I7
B	8	HIS	-	expression tag	UNP Q8Y7I7
B	9	HIS	-	expression tag	UNP Q8Y7I7
B	10	HIS	-	expression tag	UNP Q8Y7I7
B	11	SER	-	expression tag	UNP Q8Y7I7
B	12	SER	-	expression tag	UNP Q8Y7I7
B	13	GLY	-	expression tag	UNP Q8Y7I7
B	14	LEU	-	expression tag	UNP Q8Y7I7
B	15	VAL	-	expression tag	UNP Q8Y7I7
B	16	PRO	-	expression tag	UNP Q8Y7I7
B	17	ARG	-	expression tag	UNP Q8Y7I7
B	18	GLY	-	expression tag	UNP Q8Y7I7
B	19	SER	-	expression tag	UNP Q8Y7I7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	HIS	-	expression tag	UNP Q8Y7I7
B	21	MET	-	expression tag	UNP Q8Y7I7
B	22	ALA	-	expression tag	UNP Q8Y7I7
B	23	SER	-	expression tag	UNP Q8Y7I7
B	24	MET	-	expression tag	UNP Q8Y7I7
B	25	THR	-	expression tag	UNP Q8Y7I7
B	26	GLY	-	expression tag	UNP Q8Y7I7
B	27	GLY	-	expression tag	UNP Q8Y7I7
B	28	GLN	-	expression tag	UNP Q8Y7I7
B	29	GLN	-	expression tag	UNP Q8Y7I7
B	30	MET	-	expression tag	UNP Q8Y7I7
B	31	GLY	-	expression tag	UNP Q8Y7I7
B	32	ARG	-	expression tag	UNP Q8Y7I7
B	33	ASP	-	expression tag	UNP Q8Y7I7
B	576	VAL	-	expression tag	UNP Q8Y7I7
B	577	ASP	-	expression tag	UNP Q8Y7I7
B	578	LYS	-	expression tag	UNP Q8Y7I7
B	579	LEU	-	expression tag	UNP Q8Y7I7
B	580	ALA	-	expression tag	UNP Q8Y7I7
B	581	ALA	-	expression tag	UNP Q8Y7I7
B	582	ALA	-	expression tag	UNP Q8Y7I7
B	583	LEU	-	expression tag	UNP Q8Y7I7
B	584	GLU	-	expression tag	UNP Q8Y7I7
B	585	HIS	-	expression tag	UNP Q8Y7I7
B	586	HIS	-	expression tag	UNP Q8Y7I7
B	587	HIS	-	expression tag	UNP Q8Y7I7
B	588	HIS	-	expression tag	UNP Q8Y7I7
B	589	HIS	-	expression tag	UNP Q8Y7I7
B	590	HIS	-	expression tag	UNP Q8Y7I7

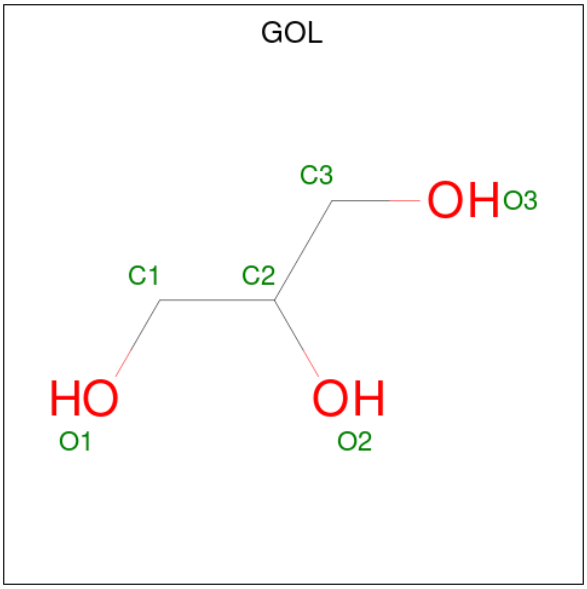
- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 6 3 3	0	0

- Molecule 5 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total K 1 1	0	0

- Molecule 6 is water.

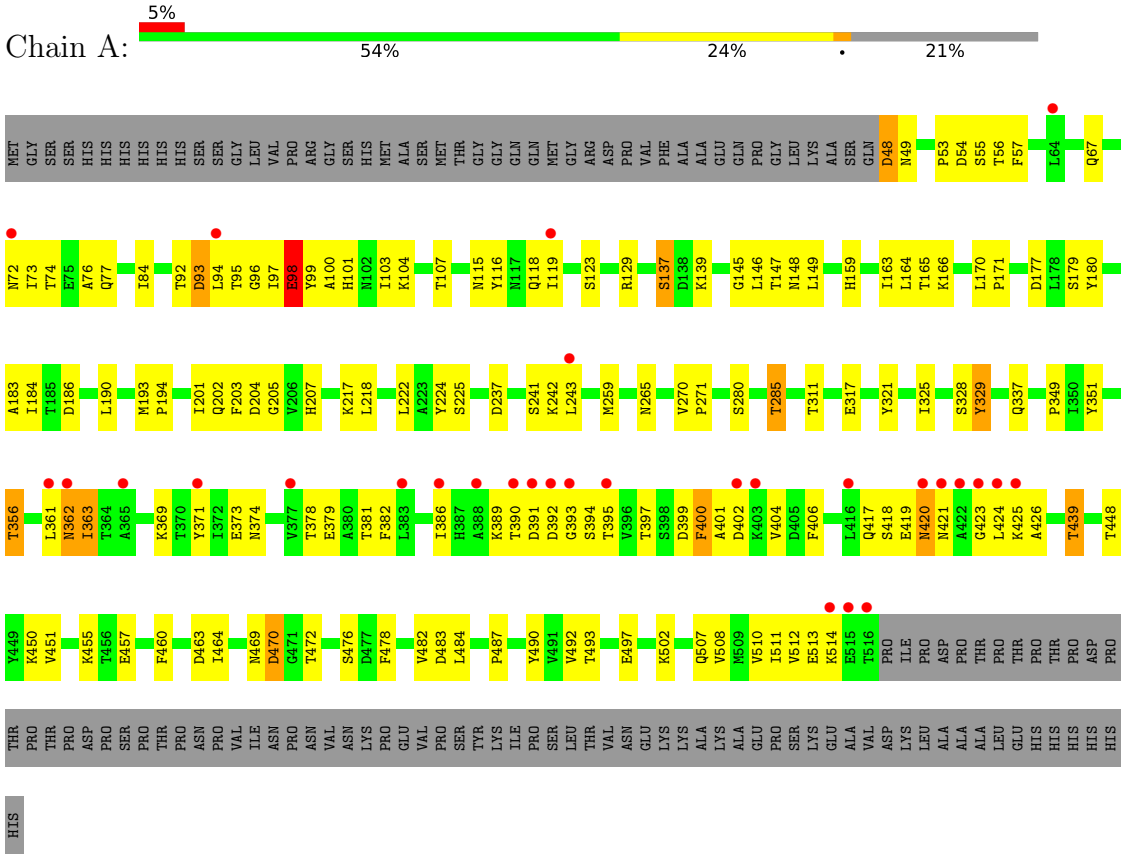
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	46	Total O 46 46	0	0
6	B	74	Total O 74 74	0	0

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.

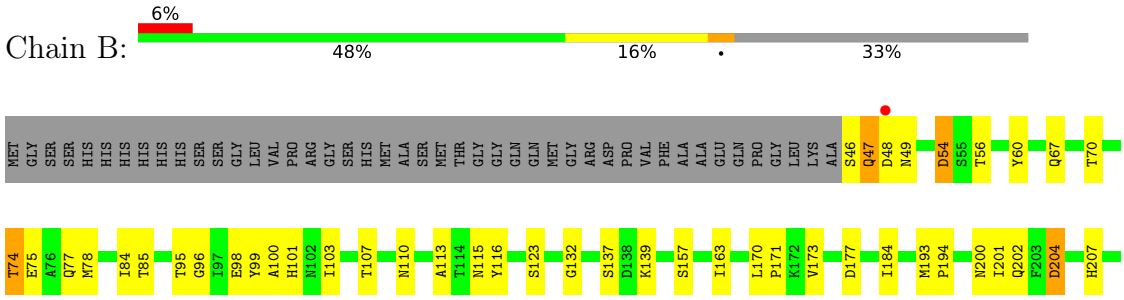
• Molecule 1: Internalin K

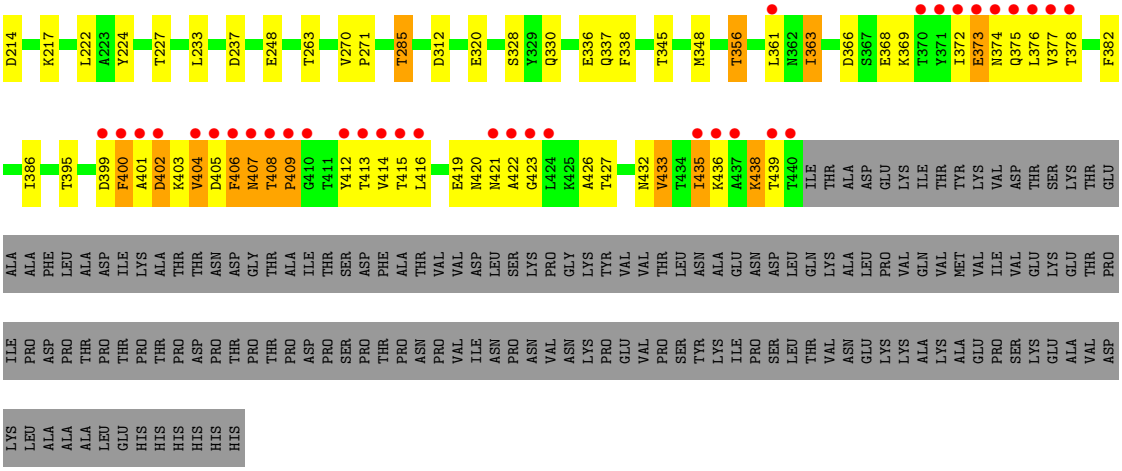
Chain A:



• Molecule 1: Internalin K

Chain B:




D214 K217 L222 A223 Y224 T227 L233 D237 E248 T263 V270 P271 T285 D312 E320 S328 Y329 Q330 E336 Q337 F338 T345 M348 T356 L361 N362 I363 D366 S367 E368 K369 T370 V371 I372 E373 N374 Q375 L376 V377 T378 F382
I386 T395 D399 F400 A401 D402 K403 V404 D405 F406 N407 T408 P409 G410 T411 Y412 T413 V414 T415 L416 E419 N420 N421 A422 G423 L424 K425 A426 T427 N432 V433 T434 I435 K436 A437 K438 T439 T440 ILE THR ALA ASP GLU LYS ILE THR TYR LYS VAL ASP THR SER LYS THR GLU
ALA ALA PHE LEU ALA ASP ILE LYS ALA THR THR ASN ASP GLY THR ALA ILE THR SER ASP PHE ALA THR VAL VAL ASP LEU SER LYS PRO GLY LYS TYR VAL THR VAL LEU ASN ALA GLU ASN ASP LEU GLN LYS ALA LEU PRO VAL GLN VAL MET VAL ILE VAL GLU LYS GLU THR PRO
ILE PRO ASP PRO THR PRO THR PRO THR PRO ASP PRO THR PRO THR PRO ASP PRO PRO SER PRO THR PRO ASN PRO VAL ILE ASN PRO ASN VAL ASN LYS PRO GLU VAL PRO SER TYR LYS ILE PRO SER LEU THR VAL ASN GLU LYS LYS ALA LYS ALA GLU PRO SER LYS GLU ALA VAL ASP
LYS LEU ALA ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α, β, γ	115.34Å 115.34Å 190.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.35 – 2.59 49.35 – 2.59	Depositor EDS
% Data completeness (in resolution range)	85.4 (49.35-2.59) 85.4 (49.35-2.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.64 (at 2.58Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.228 , 0.243 0.230 , 0.244	Depositor DCC
R_{free} test set	3810 reflections (9.83%)	wwPDB-VP
Wilson B-factor (Å^2)	49.4	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 34.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6832	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.85% 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: K, GOL, MG, NA

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	0.58	0/3688	0.75	0/5026
1	B	0.67	0/3121	0.81	2/4252 (0.0%)
All	All	0.62	0/6809	0.78	2/9278 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	222	LEU	CB-CG-CD1	-5.45	101.73	111.00
1	B	204	ASP	C-N-CA	-5.28	111.22	122.30

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	3631	0	3568	243	0
1	B	3070	0	2986	131	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
4	B	6	0	8	2	1
5	B	1	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	46	0	0	40	0
6	B	74	0	0	10	0
All	All	6832	0	6562	374	1

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 28.

The worst 5 of 374 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:B:372:ILE:HD11	1:B:438:LYS:CD	1.37	1.52
1:B:372:ILE:CD1	1:B:438:LYS:HD2	1.48	1.42
1:A:418:SER:O	1:A:424:LEU:HG	1.30	1.28
1:A:390:THR:CG2	1:A:392:ASP:OD1	1.81	1.28
1:A:418:SER:O	1:A:424:LEU:CG	1.81	1.27

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:602:GOL:O2	4:B:602:GOL:O2[5_554]	1.99	0.21

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	467/590 (79%)	434 (93%)	32 (7%)	1 (0%)	47 71
1	B	393/590 (67%)	361 (92%)	31 (8%)	1 (0%)	41 64
All	All	860/1180 (73%)	795 (92%)	63 (7%)	2 (0%)	47 71

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	98	GLU
1	B	409	PRO

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	418/522 (80%)	399 (96%)	19 (4%)	27 52
1	B	354/522 (68%)	320 (90%)	34 (10%)	8 16
All	All	772/1044 (74%)	719 (93%)	53 (7%)	15 31

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	233	LEU
1	B	356	THR
1	B	427	THR
1	B	248	GLU
1	B	336	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	337	GLN
1	B	67	GLN
1	B	374	ASN
1	B	330	GLN
1	A	101	HIS

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)

Of 6 ligands modelled in this entry, 5 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
4	GOL	B	602	-	5,5,5	0.50	0	5,5,5	0.53	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	GOL	B	602	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	602	GOL	O1-C1-C2-C3
4	B	602	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	602	GOL	2	1

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

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, 95^{th} 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	469/590 (79%)	0.44	30 (6%)	19	14	45, 74, 108, 143	0
1	B	395/590 (66%)	0.45	36 (9%)	9	6	38, 53, 122, 147	0
All	All	864/1180 (73%)	0.45	66 (7%)	13	10	38, 65, 115, 147	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	400	PHE	13.2
1	B	372	ILE	10.0
1	A	516	THR	8.1
1	B	404	VAL	7.9
1	B	401	ALA	6.5

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

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, 95^{th} 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	MG	B	601	1/1	0.81	0.47	58,58,58,58	1
5	K	B	603	1/1	0.91	0.11	81,81,81,81	0
2	MG	A	601	1/1	0.93	0.42	54,54,54,54	0
4	GOL	B	602	6/6	0.94	0.22	63,68,71,75	0
2	MG	A	602	1/1	0.94	0.29	66,66,66,66	0
3	NA	A	603	1/1	0.99	0.35	64,64,64,64	0

6.5 Other polymers (i)

There are no such residues in this entry.