wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 02:57 PM EDT

PDB ID	:	5BUT
Title	:	Crystal structure of inactive conformation of KtrAB K+ transporter
Authors	:	Vieira-Pires, R.S.; Morais-Cabral, J.H.
Deposited on	:	2015-06-04
Resolution	:	5.97 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:

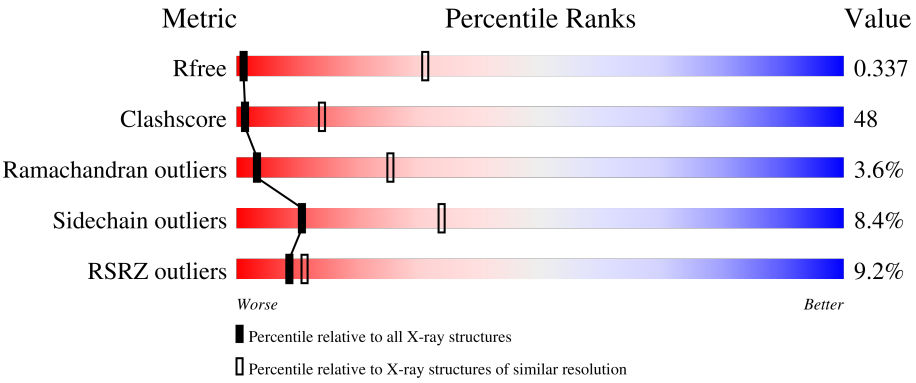
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 5.97 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.



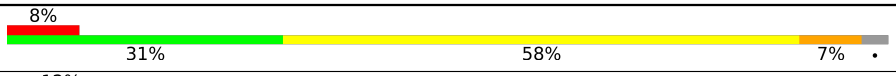
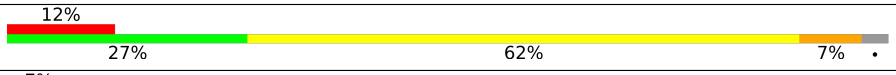
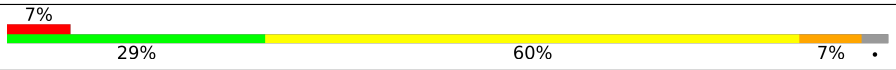
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1000 (8.00-3.88)
Clashscore	141614	1049 (8.00-3.90)
Ramachandran outliers	138981	1016 (8.00-3.86)
Sidechain outliers	138945	1017 (8.00-3.82)
RSRZ outliers	127900	1014 (8.00-3.78)

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	288	12% 38% 47% 8% 7%
1	C	288	5% 37% 48% 8% 7%
1	E	288	9% 40% 45% 8% • 7%
1	G	288	14% 38% 47% 8% 7%
2	I	445	5% 30% 60% 7% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	445	
2	K	445	
2	L	445	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 21476 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 Ktr system potassium uptake protein A,Ktr system potassium uptake protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	268	Total	C	N	O	S	0	0	0
			2096	1336	362	394	4			
1	C	268	Total	C	N	O	S	0	0	0
			2096	1336	362	394	4			
1	E	268	Total	C	N	O	S	0	0	0
			2096	1336	362	394	4			
1	G	268	Total	C	N	O	S	0	0	0
			2096	1336	362	394	4			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	145	LEU	-	linker	UNP O32080
A	146	GLU	-	linker	UNP O32080
A	147	GLY	-	linker	UNP O32080
A	148	SER	-	linker	UNP O32080
A	283	LEU	-	expression tag	UNP O32080
A	284	GLU	-	expression tag	UNP O32080
A	285	LEU	-	expression tag	UNP O32080
A	286	VAL	-	expression tag	UNP O32080
A	287	PRO	-	expression tag	UNP O32080
A	288	ARG	-	expression tag	UNP O32080
A	22	VAL	CYS	engineered mutation	UNP O32080
C	145	LEU	-	linker	UNP O32080
C	146	GLU	-	linker	UNP O32080
C	147	GLY	-	linker	UNP O32080
C	148	SER	-	linker	UNP O32080
C	283	LEU	-	expression tag	UNP O32080
C	284	GLU	-	expression tag	UNP O32080
C	285	LEU	-	expression tag	UNP O32080
C	286	VAL	-	expression tag	UNP O32080
C	287	PRO	-	expression tag	UNP O32080

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	288	ARG	-	expression tag	UNP O32080
C	22	VAL	CYS	engineered mutation	UNP O32080
E	145	LEU	-	linker	UNP O32080
E	146	GLU	-	linker	UNP O32080
E	147	GLY	-	linker	UNP O32080
E	148	SER	-	linker	UNP O32080
E	283	LEU	-	expression tag	UNP O32080
E	284	GLU	-	expression tag	UNP O32080
E	285	LEU	-	expression tag	UNP O32080
E	286	VAL	-	expression tag	UNP O32080
E	287	PRO	-	expression tag	UNP O32080
E	288	ARG	-	expression tag	UNP O32080
E	22	VAL	CYS	engineered mutation	UNP O32080
G	145	LEU	-	linker	UNP O32080
G	146	GLU	-	linker	UNP O32080
G	147	GLY	-	linker	UNP O32080
G	148	SER	-	linker	UNP O32080
G	283	LEU	-	expression tag	UNP O32080
G	284	GLU	-	expression tag	UNP O32080
G	285	LEU	-	expression tag	UNP O32080
G	286	VAL	-	expression tag	UNP O32080
G	287	PRO	-	expression tag	UNP O32080
G	288	ARG	-	expression tag	UNP O32080
G	22	VAL	CYS	engineered mutation	UNP O32080

- Molecule 2 is a protein called Ktr system potassium uptake protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	431	Total	C	N	O	S	0	0	0
			3272	2180	507	570	15			
2	J	431	Total	C	N	O	S	0	0	0
			3272	2180	507	570	15			
2	K	431	Total	C	N	O	S	0	0	0
			3272	2180	507	570	15			
2	L	431	Total	C	N	O	S	0	0	0
			3272	2180	507	570	15			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	103	ALA	GLY	conflict	UNP O32081
I	104	ALA	LYS	conflict	UNP O32081

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	105	ALA	LYS	conflict	UNP O32081
I	218	GLN	ASN	conflict	UNP O32081
I	229	ALA	LYS	conflict	UNP O32081
I	261	ALA	HIS	conflict	UNP O32081
I	262	ALA	ILE	conflict	UNP O32081
I	429	ALA	LYS	conflict	UNP O32081
J	103	ALA	GLY	conflict	UNP O32081
J	104	ALA	LYS	conflict	UNP O32081
J	105	ALA	LYS	conflict	UNP O32081
J	218	GLN	ASN	conflict	UNP O32081
J	229	ALA	LYS	conflict	UNP O32081
J	261	ALA	HIS	conflict	UNP O32081
J	262	ALA	ILE	conflict	UNP O32081
J	429	ALA	LYS	conflict	UNP O32081
K	103	ALA	GLY	conflict	UNP O32081
K	104	ALA	LYS	conflict	UNP O32081
K	105	ALA	LYS	conflict	UNP O32081
K	218	GLN	ASN	conflict	UNP O32081
K	229	ALA	LYS	conflict	UNP O32081
K	261	ALA	HIS	conflict	UNP O32081
K	262	ALA	ILE	conflict	UNP O32081
K	429	ALA	LYS	conflict	UNP O32081
L	103	ALA	GLY	conflict	UNP O32081
L	104	ALA	LYS	conflict	UNP O32081
L	105	ALA	LYS	conflict	UNP O32081
L	218	GLN	ASN	conflict	UNP O32081
L	229	ALA	LYS	conflict	UNP O32081
L	261	ALA	HIS	conflict	UNP O32081
L	262	ALA	ILE	conflict	UNP O32081
L	429	ALA	LYS	conflict	UNP O32081

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

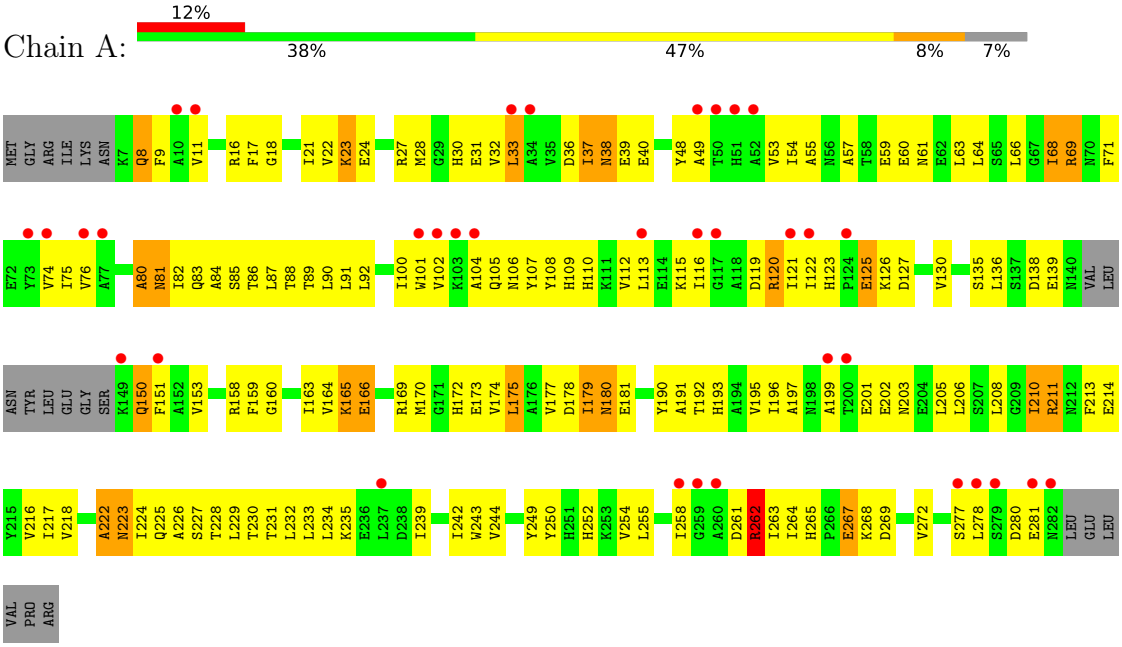
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	I	1	Total K 1 1	0	0
3	J	1	Total K 1 1	0	0
3	K	1	Total K 1 1	0	0
3	L	1	Total K 1 1	0	0

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.

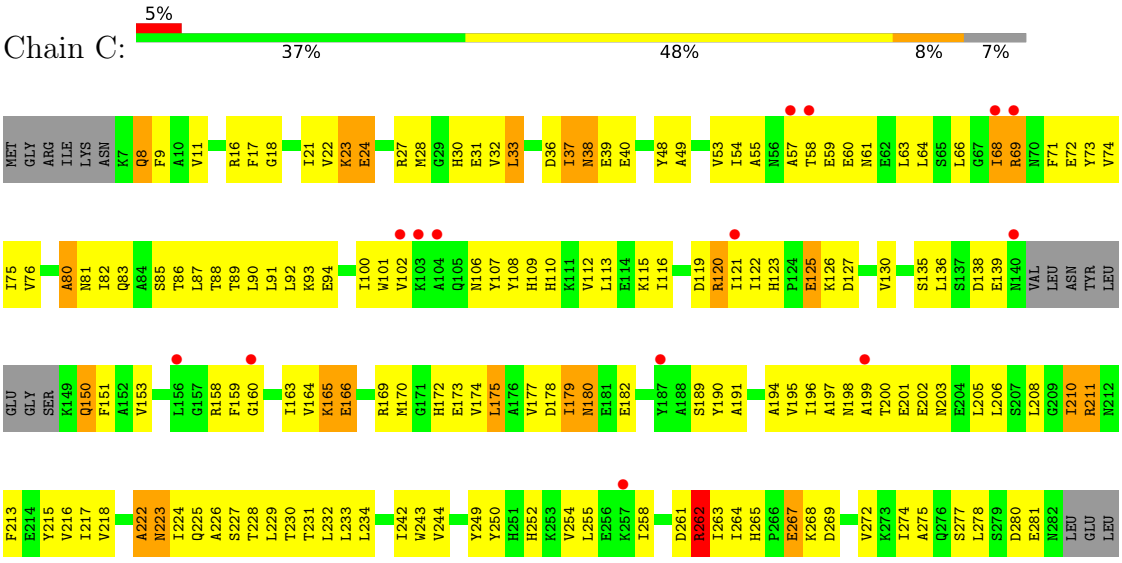
• Molecule 1: Ktr system potassium uptake protein A,Ktr system potassium uptake protein A

Chain A:



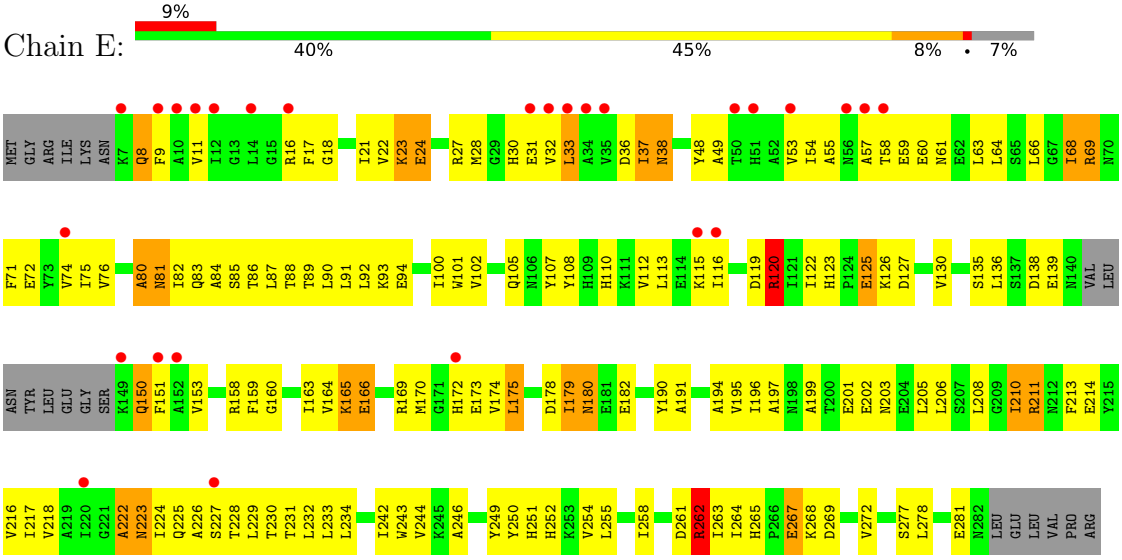
• Molecule 1: Ktr system potassium uptake protein A,Ktr system potassium uptake protein A

Chain C:



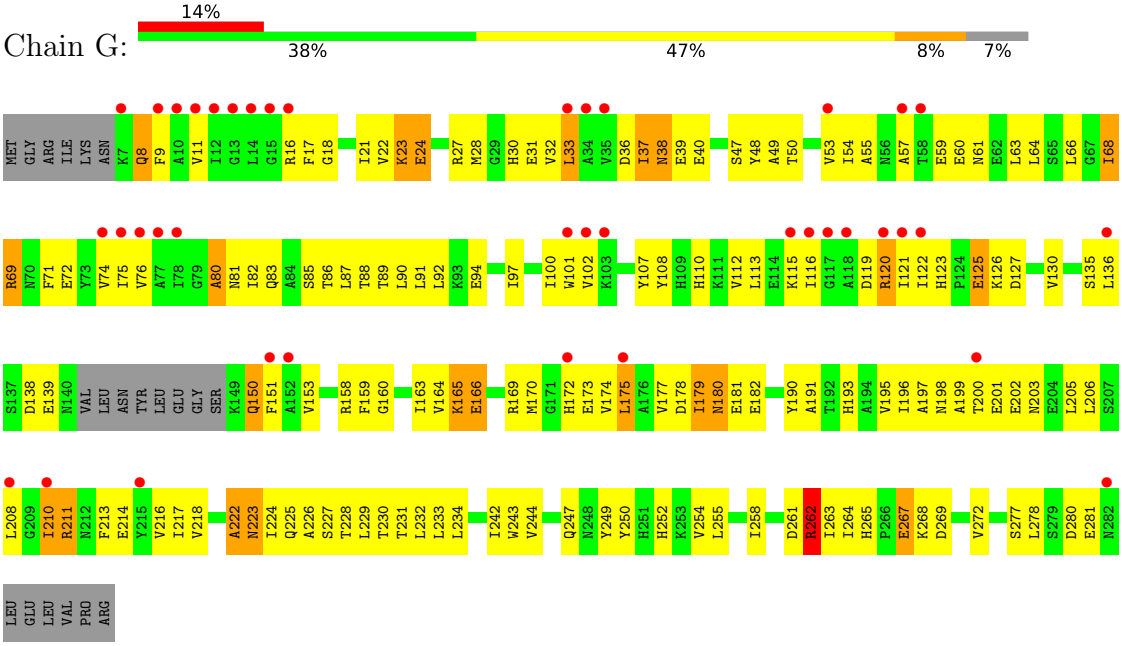
• Molecule 1: Ktr system potassium uptake protein A,Ktr system potassium uptake protein A

Chain E:



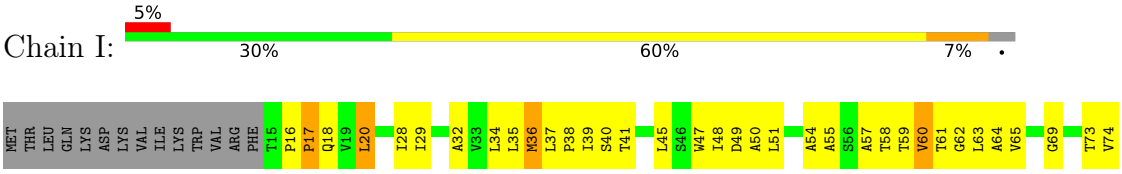
• Molecule 1: Ktr system potassium uptake protein A,Ktr system potassium uptake protein A

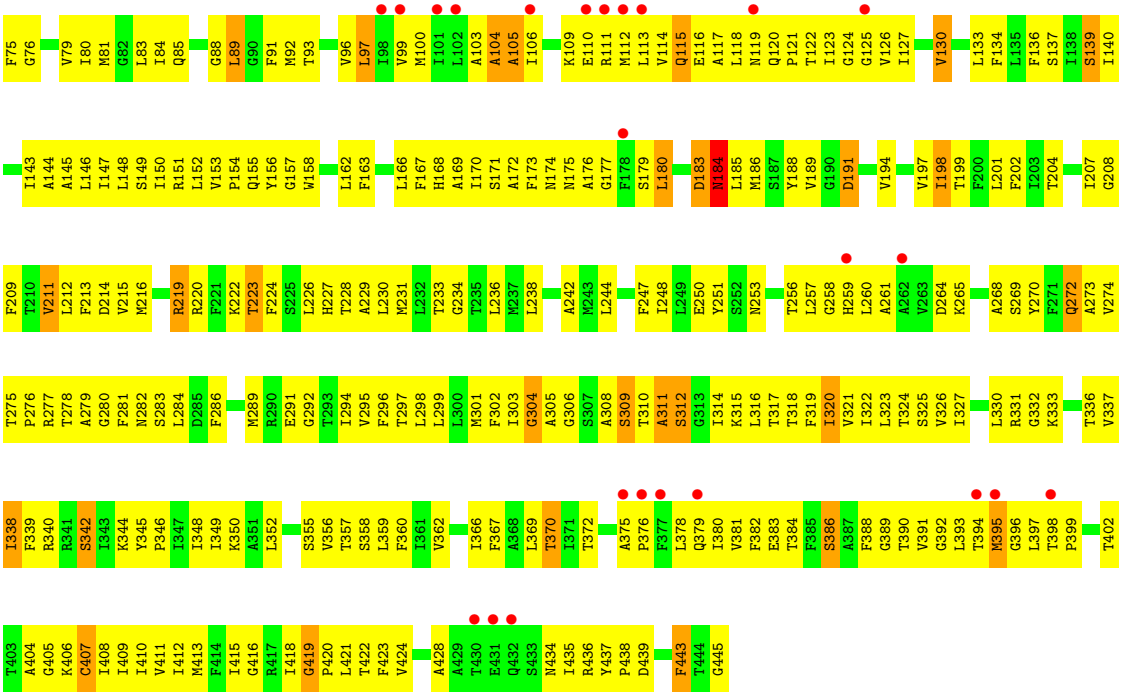
Chain G:



• Molecule 2: Ktr system potassium uptake protein B

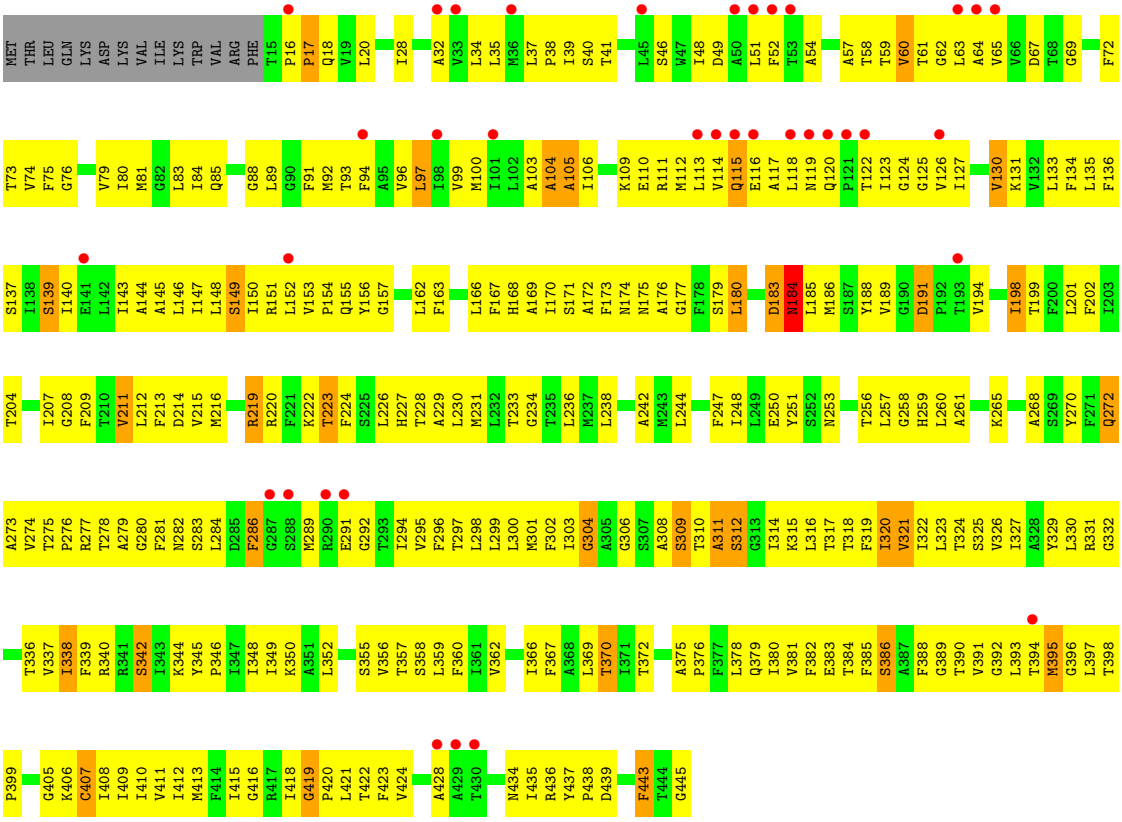
Chain I:





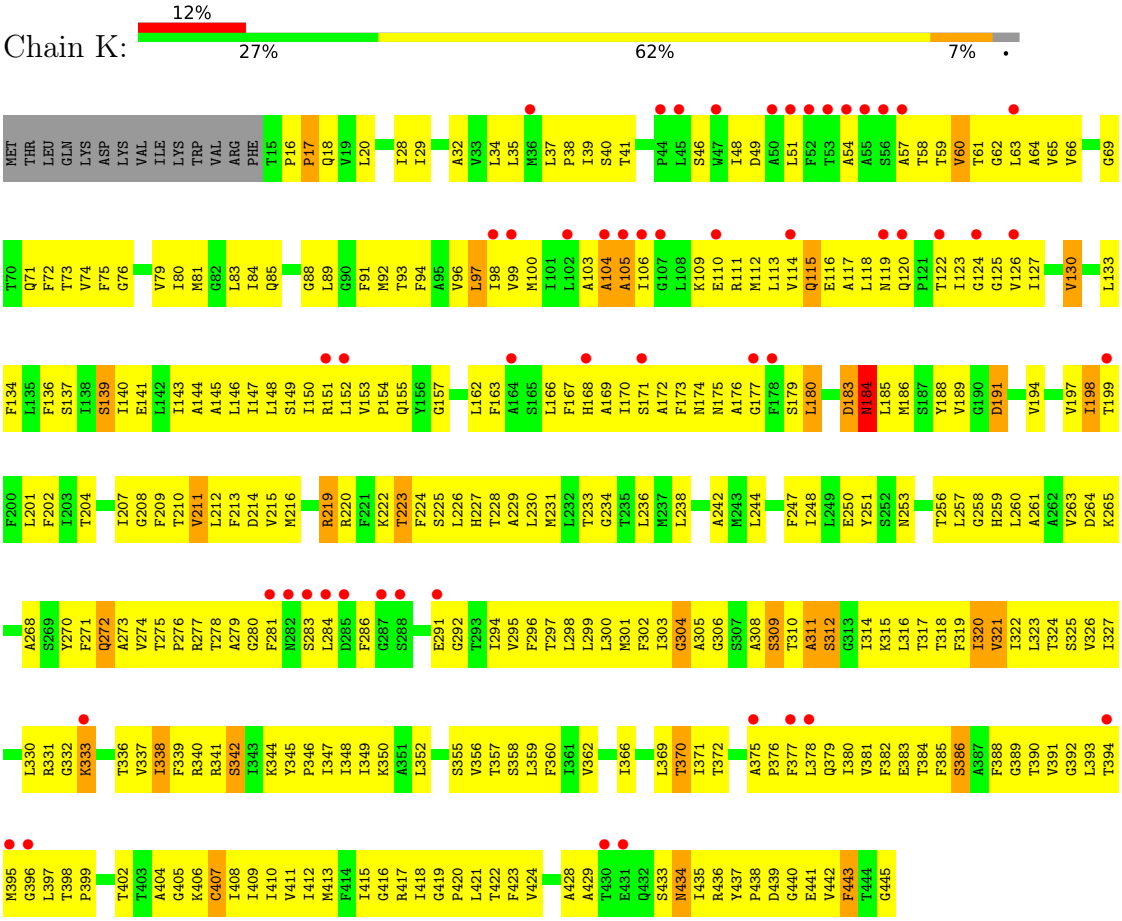
• Molecule 2: Ktr system potassium uptake protein B

Chain J:



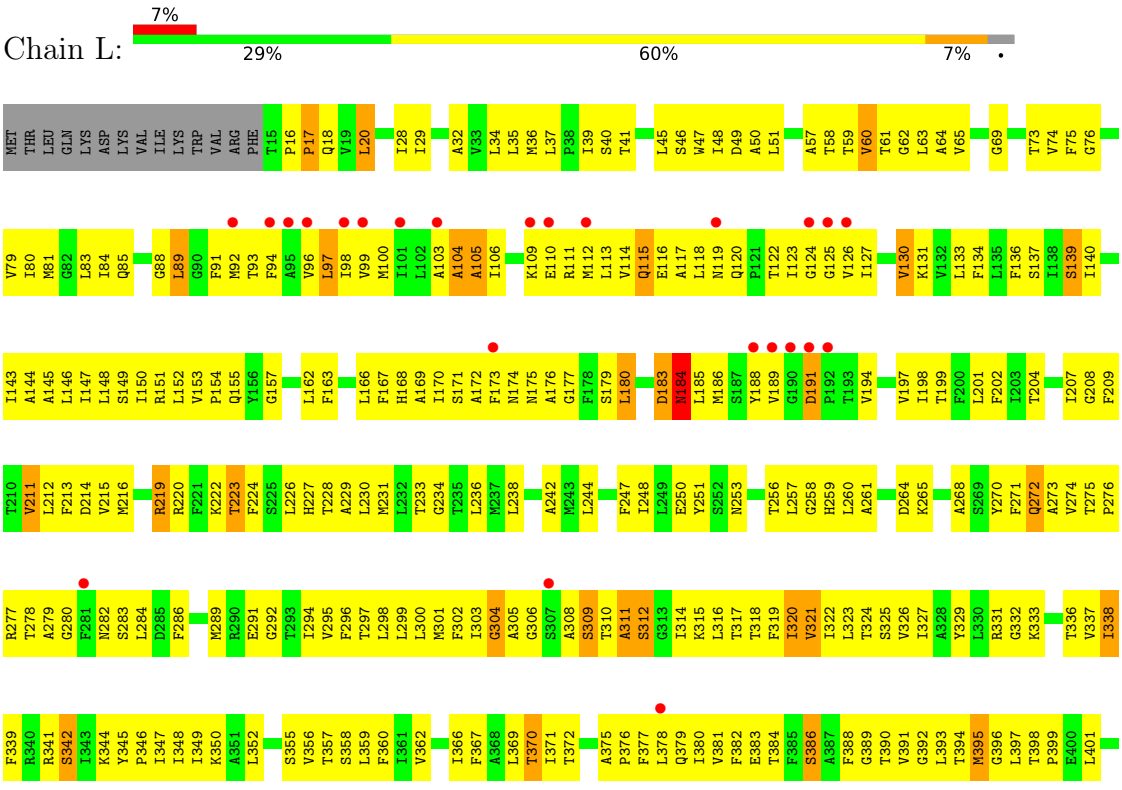
• Molecule 2: Ktr system potassium uptake protein B

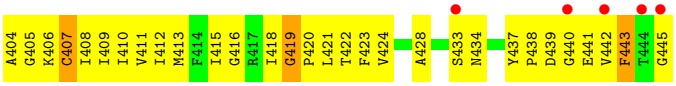
Chain K:



• Molecule 2: Ktr system potassium uptake protein B

Chain L:




A404
G405
K406
C407
I408
I409
I410
V411
I412
M413
F414
I415
G416
R417
I418
G419
P420
L421
T422
F423
V424
A428
S433
N434
Y437
P438
D439
G440
E441
V442
F443
T444
G445

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	307.06Å 79.41Å 205.65Å 90.00° 98.10° 90.00°	Depositor
Resolution (Å)	200.00 – 5.97 58.35 – 5.97	Depositor EDS
% Data completeness (in resolution range)	96.7 (200.00-5.97) 98.1 (58.35-5.97)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 6.17Å)	Xtriage
Refinement program	CNS DENR	Depositor
R, R_{free}	0.325 , 0.339 0.338 , 0.337	Depositor DCC
R_{free} test set	2336 reflections (9.90%)	wwPDB-VP
Wilson B-factor (Å^2)	316.7	Xtriage
Anisotropy	0.488	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 272.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	21476	wwPDB-VP
Average B, all atoms (Å^2)	342.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.94% 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

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	0.23	0/2130	0.52	4/2882 (0.1%)
1	C	0.25	0/2130	0.75	5/2882 (0.2%)
1	E	0.23	0/2130	0.52	4/2882 (0.1%)
1	G	0.25	0/2130	0.75	5/2882 (0.2%)
2	I	0.28	0/3343	0.50	0/4550
2	J	0.28	0/3343	0.50	0/4550
2	K	0.28	0/3343	0.50	0/4550
2	L	0.28	0/3343	0.50	0/4550
All	All	0.27	0/21892	0.56	18/29728 (0.1%)

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	120	ARG	NE-CZ-NH1	-21.01	109.80	120.30
1	G	120	ARG	NE-CZ-NH1	-20.98	109.81	120.30
1	C	120	ARG	NE-CZ-NH2	19.76	130.18	120.30
1	G	120	ARG	NE-CZ-NH2	19.75	130.18	120.30
1	C	120	ARG	CD-NE-CZ	9.78	137.29	123.60

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	2096	0	2120	181	0
1	C	2096	0	2120	177	0
1	E	2096	0	2120	145	0
1	G	2096	0	2120	156	0
2	I	3272	0	3442	384	0
2	J	3272	0	3442	379	0
2	K	3272	0	3442	405	0
2	L	3272	0	3442	386	0
3	I	1	0	0	0	0
3	J	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
All	All	21476	0	22248	2100	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 48.

The worst 5 of 2100 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:J:28:ILE:HG12	2:J:58:THR:HG21	1.37	1.05
2:K:28:ILE:HG12	2:K:58:THR:HG21	1.38	1.04
1:C:194:ALA:O	2:K:438:PRO:HD3	1.63	0.99
2:J:227:HIS:HA	2:J:230:LEU:HD12	1.46	0.96
2:L:133:LEU:HD12	2:L:134:PHE:H	1.29	0.96

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	264/288 (92%)	228 (86%)	32 (12%)	4 (2%)	10 45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	264/288 (92%)	232 (88%)	28 (11%)	4 (2%)	10 45
1	E	264/288 (92%)	230 (87%)	30 (11%)	4 (2%)	10 45
1	G	264/288 (92%)	231 (88%)	29 (11%)	4 (2%)	10 45
2	I	429/445 (96%)	323 (75%)	84 (20%)	22 (5%)	2 19
2	J	429/445 (96%)	330 (77%)	78 (18%)	21 (5%)	2 20
2	K	429/445 (96%)	329 (77%)	80 (19%)	20 (5%)	2 21
2	L	429/445 (96%)	325 (76%)	83 (19%)	21 (5%)	2 20
All	All	2772/2932 (94%)	2228 (80%)	444 (16%)	100 (4%)	3 25

5 of 100 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	81	ASN
1	A	223	ASN
1	C	81	ASN
1	C	223	ASN
1	E	81	ASN

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	222/240 (92%)	200 (90%)	22 (10%)	8 26
1	C	222/240 (92%)	198 (89%)	24 (11%)	6 24
1	E	222/240 (92%)	199 (90%)	23 (10%)	7 25
1	G	222/240 (92%)	197 (89%)	25 (11%)	6 22
2	I	354/368 (96%)	328 (93%)	26 (7%)	14 39
2	J	354/368 (96%)	329 (93%)	25 (7%)	14 39
2	K	354/368 (96%)	330 (93%)	24 (7%)	16 41
2	L	354/368 (96%)	329 (93%)	25 (7%)	14 39
All	All	2304/2432 (95%)	2110 (92%)	194 (8%)	11 34

5 of 194 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	386	SER
2	J	439	ASP
2	I	443	PHE
2	J	184	ASN
2	K	130	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	195	ASN
2	K	119	ASN
2	I	239	ASN
2	J	175	ASN
2	K	379	GLN

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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	268/288 (93%)	0.58	35 (13%) 3 7	343, 343, 343, 343	0
1	C	268/288 (93%)	0.25	14 (5%) 27 27	343, 343, 343, 343	0
1	E	268/288 (93%)	0.65	27 (10%) 7 9	343, 343, 343, 343	0
1	G	268/288 (93%)	0.65	40 (14%) 2 5	343, 343, 343, 343	0
2	I	431/445 (96%)	0.39	24 (5%) 24 24	343, 343, 343, 343	0
2	J	431/445 (96%)	0.34	36 (8%) 11 13	343, 343, 343, 343	0
2	K	431/445 (96%)	0.54	52 (12%) 4 8	343, 343, 343, 343	0
2	L	431/445 (96%)	0.28	29 (6%) 17 17	343, 343, 343, 343	0
All	All	2796/2932 (95%)	0.44	257 (9%) 9 11	343, 343, 343, 343	0

The worst 5 of 257 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	105	ALA	7.1
2	J	119	ASN	7.0
2	I	431	GLU	6.3
2	K	104	ALA	5.8
1	A	281	GLU	5.5

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, 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
3	K	I	501	1/1	0.88	0.46	342,342,342,342	0
3	K	K	501	1/1	0.89	0.24	342,342,342,342	0
3	K	L	501	1/1	0.92	0.25	342,342,342,342	0
3	K	J	501	1/1	0.94	0.34	342,342,342,342	0

6.5 Other polymers (i)

There are no such residues in this entry.