wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 08:16 AM EDT

PDB ID	:	1KPK
Title	:	Crystal Structure of the ClC Chloride Channel from E. coli
Authors	:	Dutzler, R.; Campbell, E.B.; Cadene, M.; Chait, B.T.; MacKinnon, R.
Deposited on	:	2001-12-31
Resolution	:	3.50 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 (i)) were used in the production of this report:

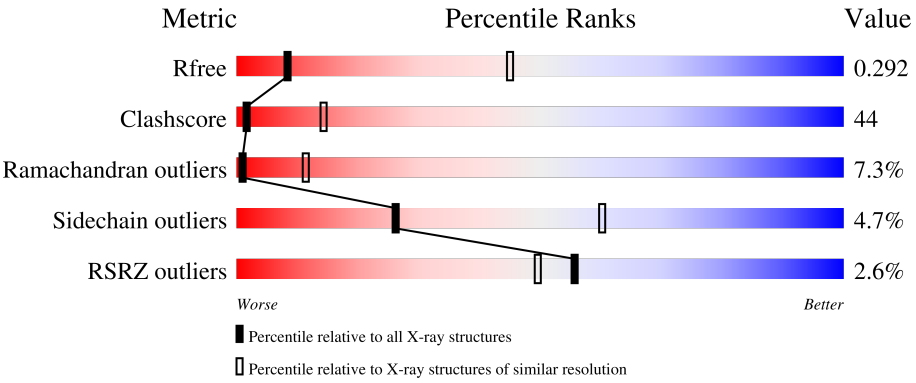
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.50 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.



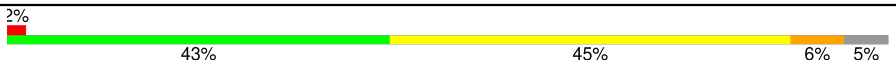
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	473	2% 42% 46% 6% 5%
1	B	473	2% 41% 47% 7% 5%
1	C	473	2% 42% 46% 6% 5%
1	D	473	4% 42% 46% 6% 5%
1	E	473	3% 42% 47% 6% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	473	

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 20274 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 putative channel transporter.

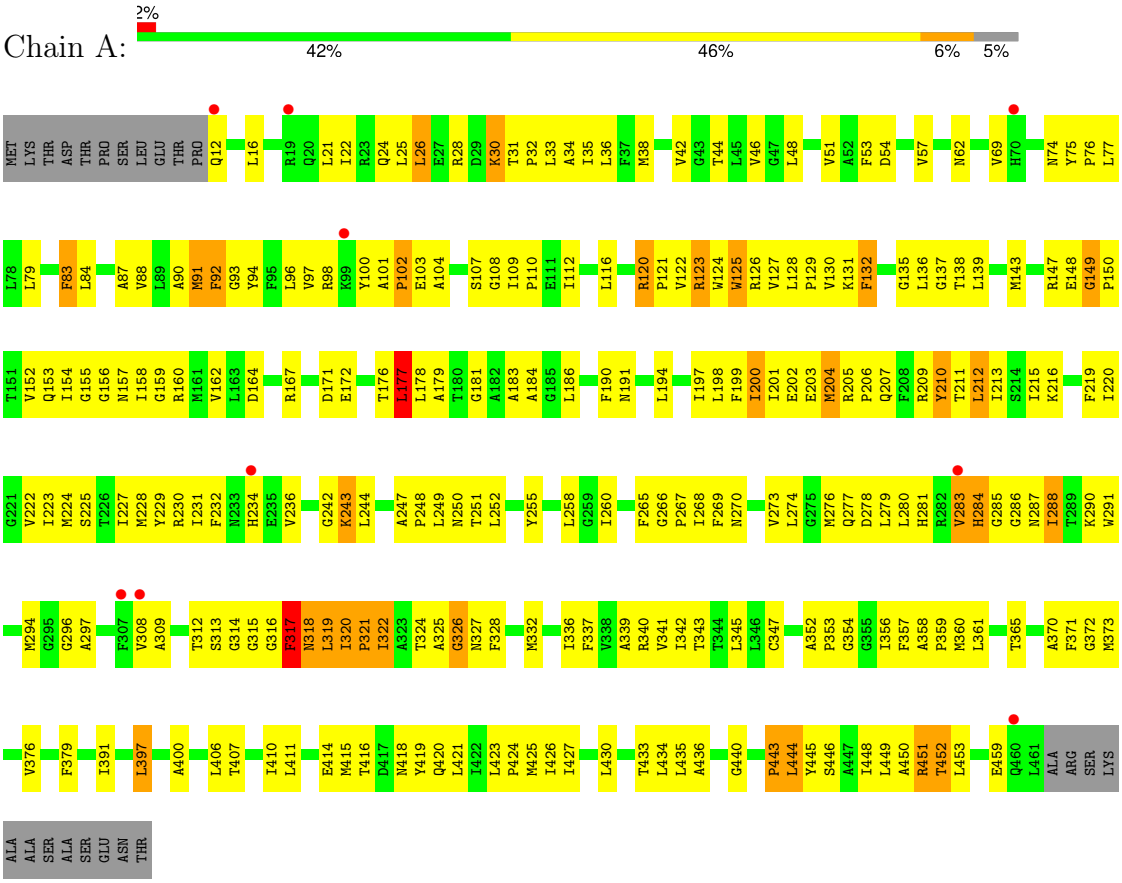
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	B	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	C	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	D	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	E	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	F	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			

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.

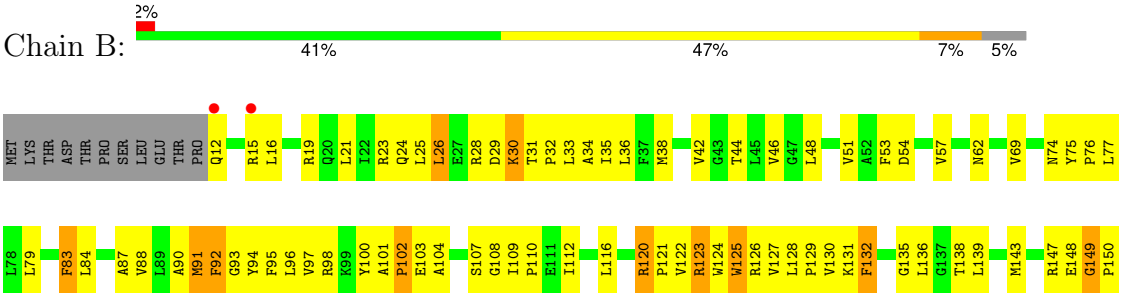
• Molecule 1: putative channel transporter

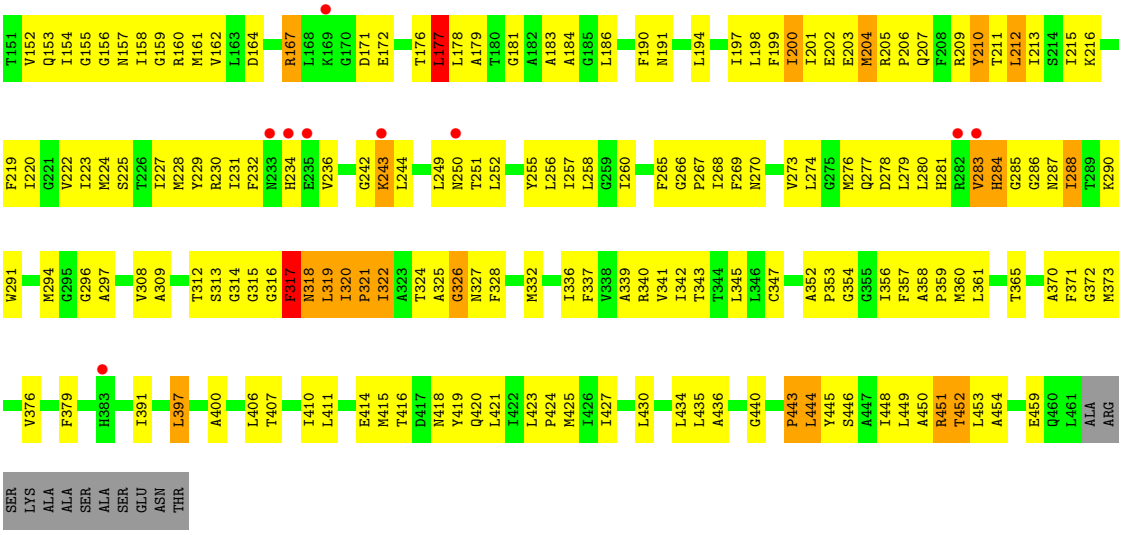
Chain A:



• Molecule 1: putative channel transporter

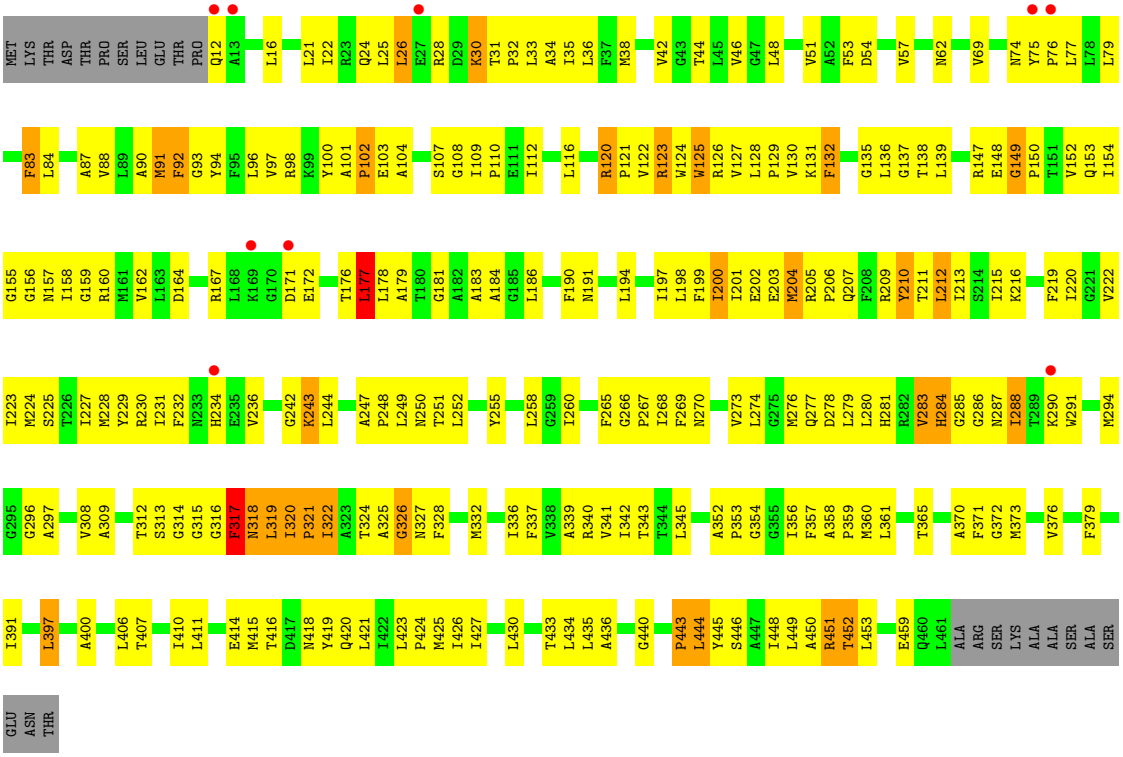
Chain B:





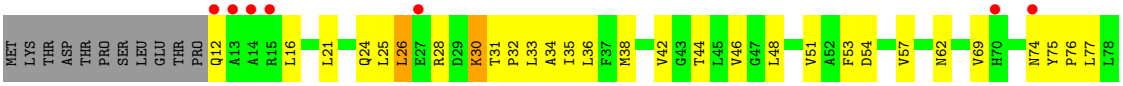
• Molecule 1: putative channel transporter

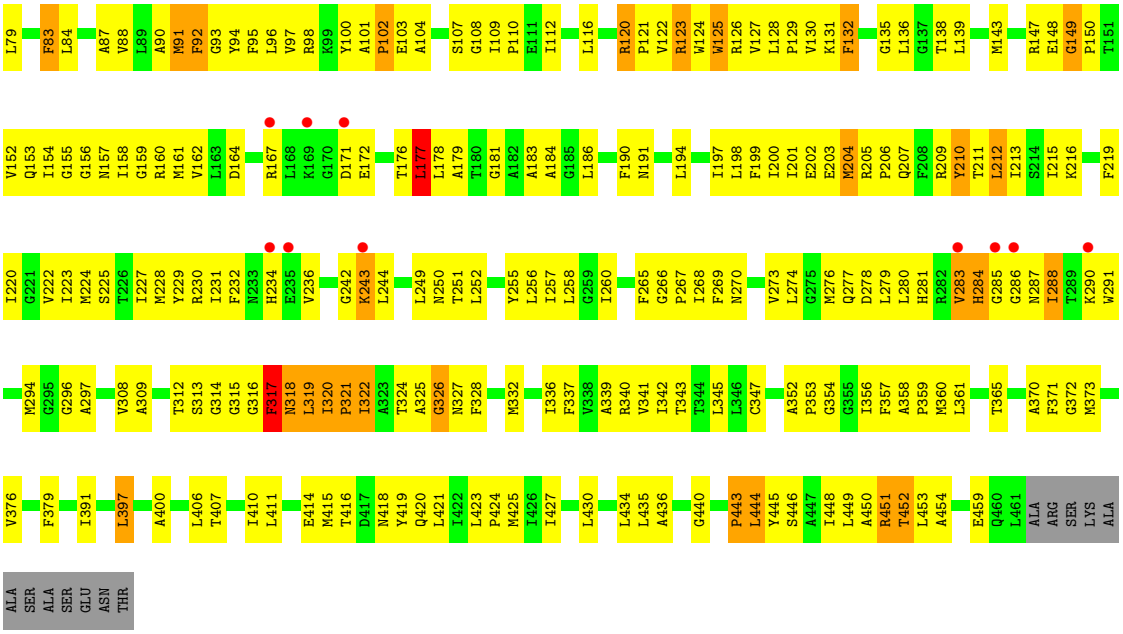
Chain C:



• Molecule 1: putative channel transporter

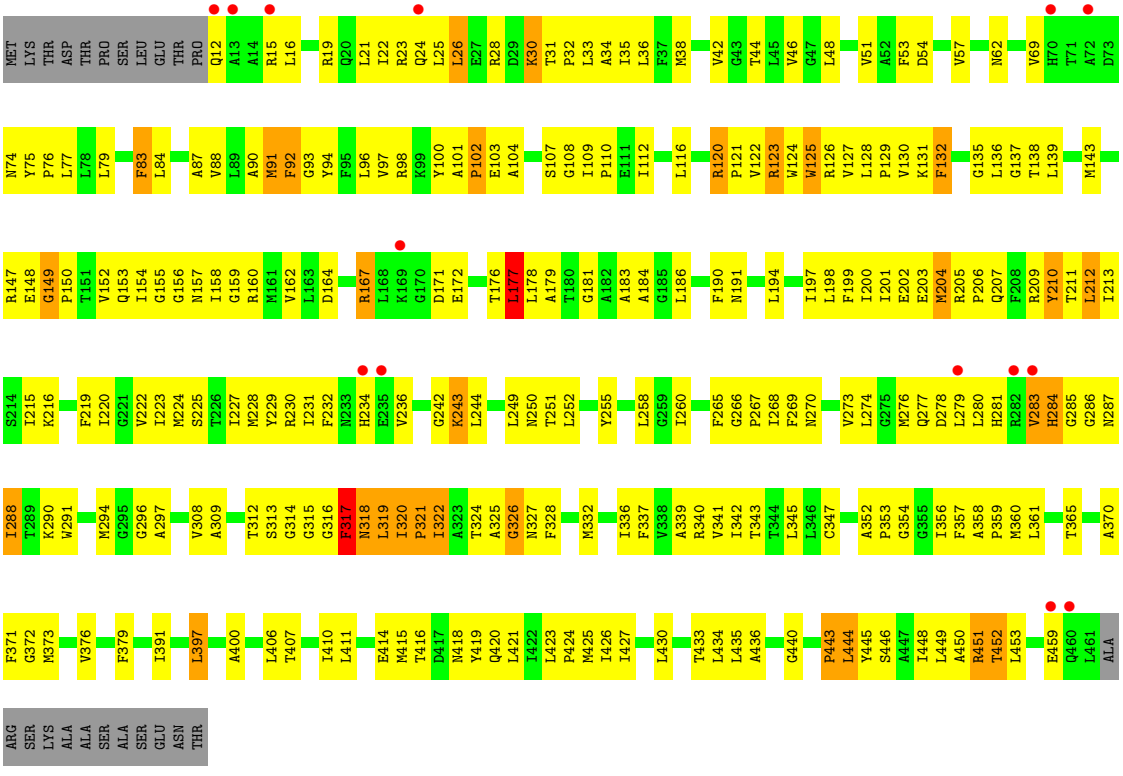
Chain D:





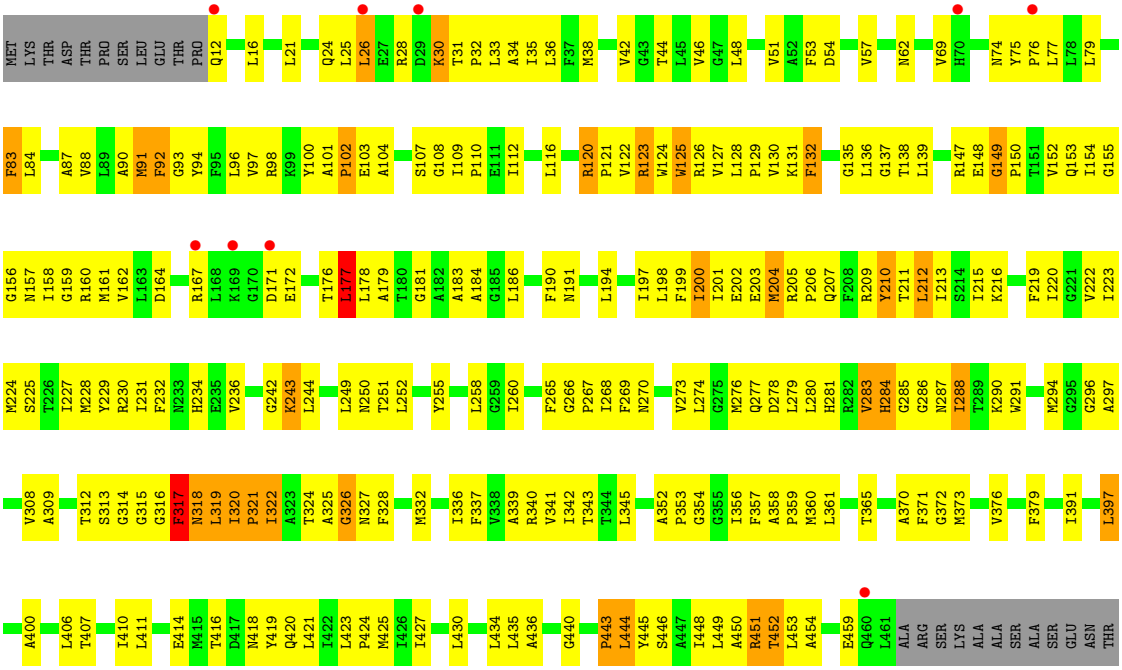
• Molecule 1: putative channel transporter

Chain E:



• Molecule 1: putative channel transporter

Chain F:


MET LYS THR ASP THR PRO SER LEU GLU THR PRO Q12 L16 L21 Q24 L25 L26 E27 R28 D29 K30 T31 P32 L33 A34 I35 L36 F37 M38 V42 G43 T44 L45 V46 G47 L48 V51 A52 F53 D54 V57 N62 V69 H70 N74 Y75 P76 L77 L78 L79
F83 L84 A87 V88 L89 A90 M91 F92 G93 Y94 F95 L96 V97 R98 K99 Y100 A101 P102 E103 A104 S107 G108 I109 P110 E111 I112 L116 R120 P121 V122 R123 W124 W125 R126 V127 L128 P129 V130 K131 F132 G135 L136 G137 T138 L139 R147 E148 G149 P150 T151 V152 Q153 I154 G155
G156 N157 I158 G159 R160 M161 V162 L163 D164 R167 L168 K169 G170 D171 E172 T176 L177 L178 A179 T180 G181 A182 A183 A184 G185 L186 F190 N191 L194 I197 L198 F199 I200 I201 E202 E203 M204 R205 P206 Q207 F208 R209 Y210 T211 L212 I213 S214 I215 K216 F219 I220 G221 V222 I223
M224 S225 T226 I227 M228 Y229 R230 I231 F232 N233 H234 E235 V236 G242 K243 L244 L249 N250 T251 L252 Y255 L258 G259 I260 F265 G266 P267 I268 F269 N270 V273 L274 G275 M276 Q277 D278 L279 L280 H281 R282 V283 H284 G285 G286 N287 T288 T289 K290 W291 M294 G295 G296 A297
V308 A309 T312 S313 G314 G315 G316 F317 N318 L319 I320 P321 I322 A323 T324 A325 G326 N327 F328 M332 I336 F337 V338 A339 R340 V341 I342 T343 T344 L345 A352 P353 G354 G355 I356 F357 A358 P359 M360 L361 T365 A370 F371 G372 M373 V376 F379 I391 L397
A400 L406 T407 I410 L411 E414 M415 T416 D417 N418 Y419 Q420 L421 I422 L423 P424 M425 I426 I427 L430 L434 L435 A436 G440 P443 L444 Y445 S446 A447 I448 L449 A450 R451 T452 L453 A454 E459 Q460 L461 ALA ARG SER LYS ALA ALA SER ALA SER GLU ASN THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.66Å 152.53Å 263.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.50 19.97 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.4 (20.00-3.50) 99.4 (19.97-3.50)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.58 (at 3.52Å)	Xtriage
Refinement program	CNS 0.3	Depositor
R, R_{free}	0.290 , 0.301 0.285 , 0.292	Depositor DCC
R_{free} test set	5321 reflections (9.90%)	wwPDB-VP
Wilson B-factor (Å^2)	125.9	Xtriage
Anisotropy	0.318	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 56.1	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	0.91	EDS
Total number of atoms	20274	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.85% 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

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	0.49	0/3451	0.76	2/4683 (0.0%)
1	B	0.49	0/3451	0.76	2/4683 (0.0%)
1	C	0.49	0/3451	0.76	2/4683 (0.0%)
1	D	0.49	0/3451	0.76	2/4683 (0.0%)
1	E	0.49	0/3451	0.76	2/4683 (0.0%)
1	F	0.49	0/3451	0.76	2/4683 (0.0%)
All	All	0.49	0/20706	0.76	12/28098 (0.0%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	320	ILE	N-CA-C	5.97	127.11	111.00
1	F	320	ILE	N-CA-C	5.97	127.11	111.00
1	A	320	ILE	N-CA-C	5.96	127.10	111.00
1	C	320	ILE	N-CA-C	5.95	127.06	111.00
1	B	320	ILE	N-CA-C	5.95	127.06	111.00

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	3379	0	3537	338	0
1	B	3379	0	3537	336	52

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3379	0	3537	336	1
1	D	3379	0	3537	337	4
1	E	3379	0	3537	332	49
1	F	3379	0	3537	335	0
All	All	20274	0	21222	1844	53

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 44.

The worst 5 of 1844 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:D:200:ILE:HD13	1:D:204:MET:HG3	1.38	1.06
1:B:200:ILE:HD13	1:B:204:MET:HG3	1.38	1.05
1:F:200:ILE:HD13	1:F:204:MET:HG3	1.38	1.05
1:E:200:ILE:HD13	1:E:204:MET:HG3	1.38	1.05
1:F:322:ILE:N	1:F:322:ILE:HD12	1.74	1.03

The worst 5 of 53 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)
1:B:15:ARG:CA	1:E:12:GLN:NE2[4_455]	0.70	1.50
1:B:16:LEU:CG	1:E:19:ARG:CZ[4_455]	0.73	1.47
1:B:19:ARG:NH1	1:E:16:LEU:CB[4_455]	0.88	1.32
1:B:12:GLN:CD	1:E:15:ARG:CB[4_455]	0.91	1.29
1:B:19:ARG:NH1	1:E:16:LEU:CG[4_455]	0.96	1.24

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	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	1	11
1	B	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	1	11
1	C	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	1	11
1	D	448/473 (95%)	340 (76%)	76 (17%)	32 (7%)	1	12
1	E	448/473 (95%)	340 (76%)	76 (17%)	32 (7%)	1	12
1	F	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	1	11
All	All	2688/2838 (95%)	2040 (76%)	452 (17%)	196 (7%)	1	11

5 of 196 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	319	LEU
1	B	319	LEU
1	C	319	LEU
1	D	319	LEU
1	E	319	LEU

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	339/358 (95%)	323 (95%)	16 (5%)	26	60
1	B	339/358 (95%)	323 (95%)	16 (5%)	26	60
1	C	339/358 (95%)	323 (95%)	16 (5%)	26	60
1	D	339/358 (95%)	323 (95%)	16 (5%)	26	60
1	E	339/358 (95%)	323 (95%)	16 (5%)	26	60
1	F	339/358 (95%)	323 (95%)	16 (5%)	26	60
All	All	2034/2148 (95%)	1938 (95%)	96 (5%)	26	60

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	243	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	177	LEU
1	D	317	PHE
1	E	62	ASN
1	E	317	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 72 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	418	ASN
1	F	437	GLN
1	E	437	GLN
1	F	270	ASN
1	C	62	ASN

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)

There are no ligands in this entry.

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	450/473 (95%)	-0.37	9 (2%)	65	60	51, 100, 147, 151	0
1	B	450/473 (95%)	-0.37	11 (2%)	59	53	51, 100, 147, 151	0
1	C	450/473 (95%)	-0.34	9 (2%)	65	60	51, 100, 147, 151	0
1	D	450/473 (95%)	-0.33	17 (3%)	40	36	51, 100, 147, 151	0
1	E	450/473 (95%)	-0.31	14 (3%)	49	43	51, 100, 147, 151	0
1	F	450/473 (95%)	-0.31	9 (2%)	65	60	51, 100, 147, 151	0
All	All	2700/2838 (95%)	-0.34	69 (2%)	56	49	51, 100, 147, 151	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	12	GLN	8.3
1	E	12	GLN	7.6
1	B	12	GLN	5.7
1	F	12	GLN	5.4
1	D	169	LYS	5.0

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

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.