



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 04:40 PM GMT

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 validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

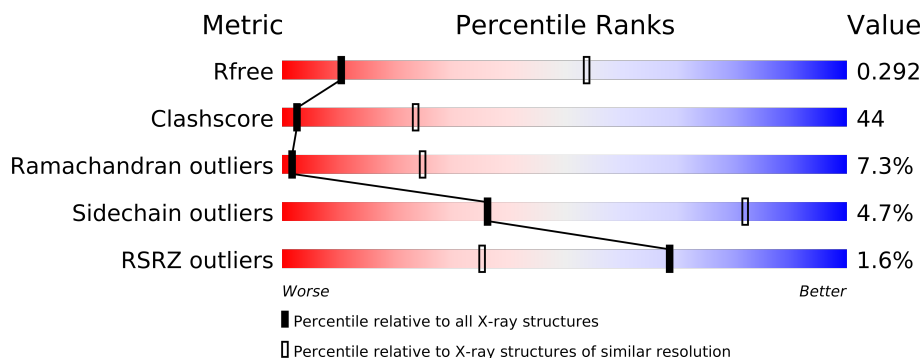
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

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))
$R_{free}$	66092	1243 (3.70-3.30)
Clashscore	79885	1039 (3.66-3.34)
Ramachandran outliers	78287	1000 (3.66-3.34)
Sidechain outliers	78261	1000 (3.66-3.34)
RSRZ outliers	66119	1243 (3.70-3.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	473	
1	B	473	
1	C	473	
1	D	473	
1	E	473	
1	F	473	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 20274 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

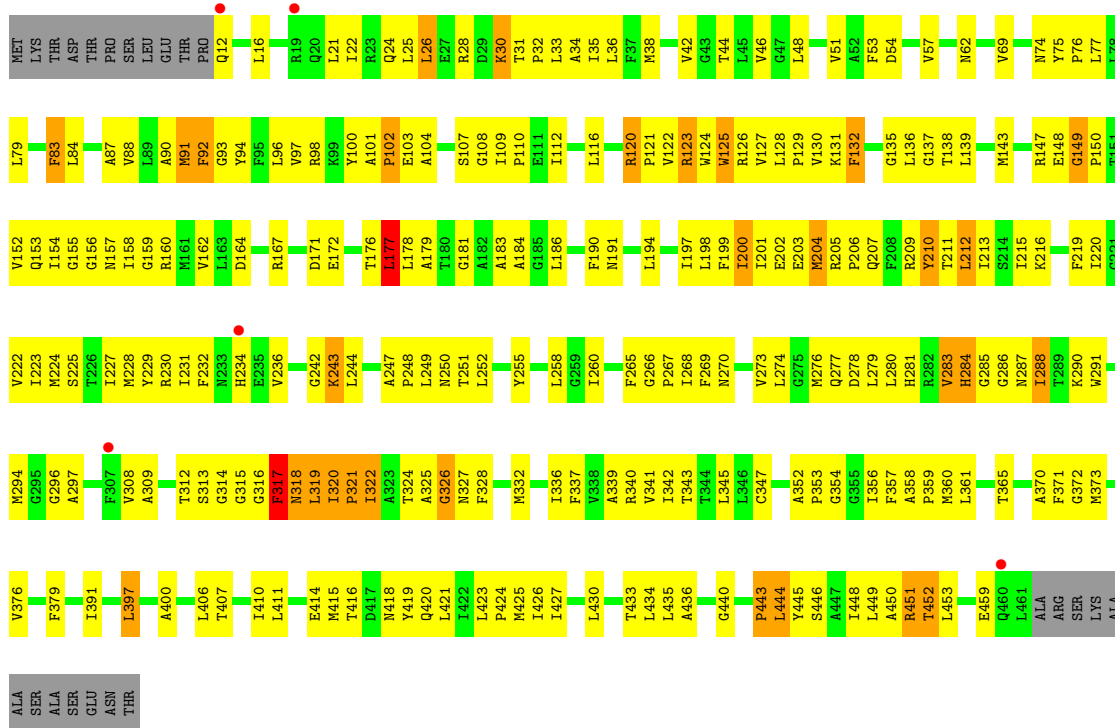
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	B	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	C	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	D	450	Total	C	N	O	S	0	0	0
			3379	2219	570	570	20			
1	E	450	Total	C	N	O	S	0	0	0
	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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

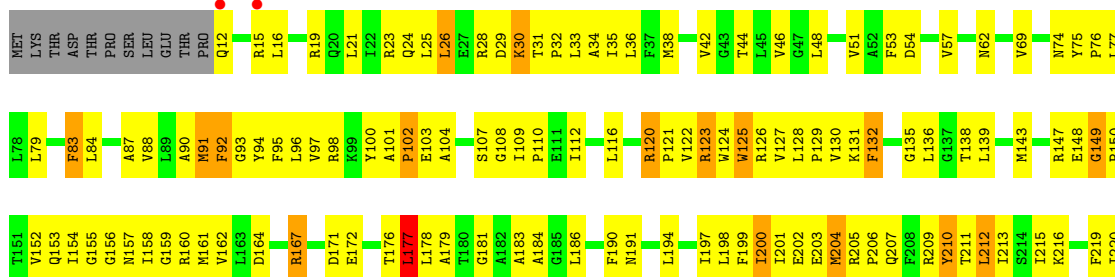
- Molecule 1: putative channel transporter

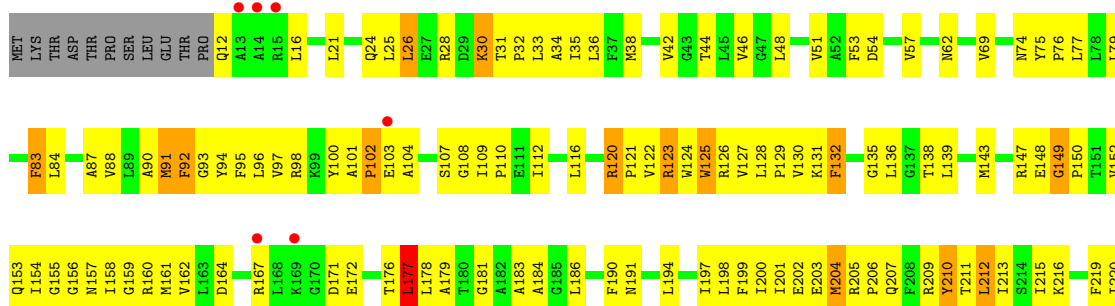
Chain A: 

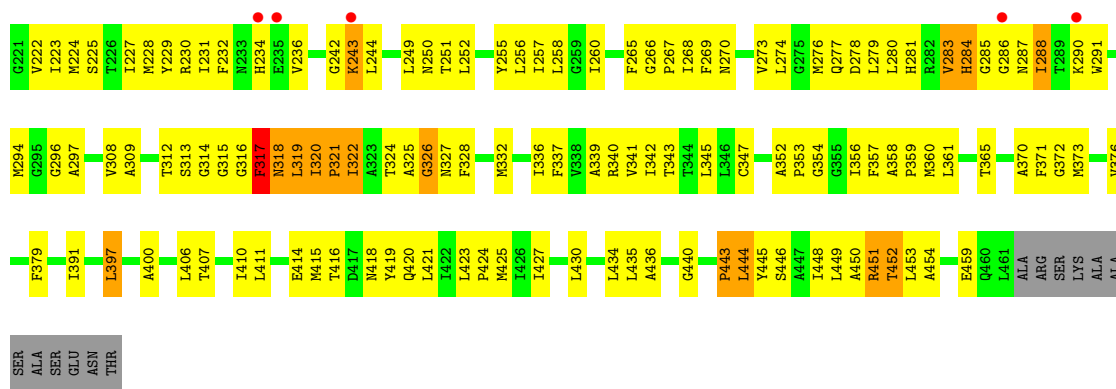


- Molecule 1: putative channel transporter

Chain B: 

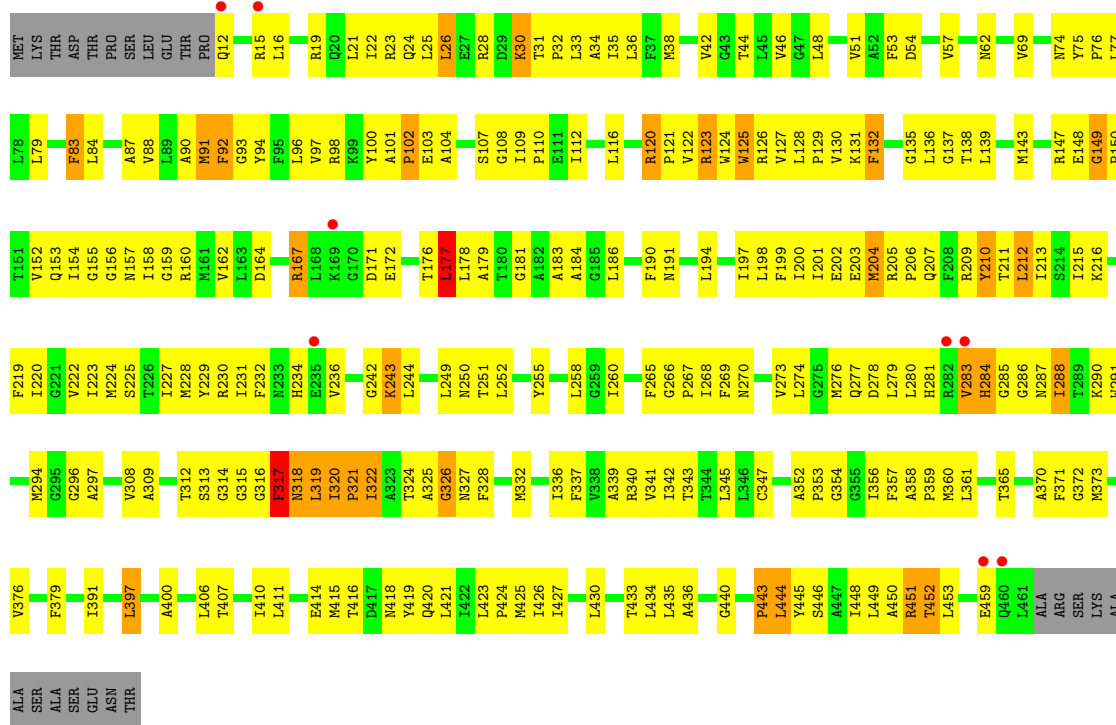






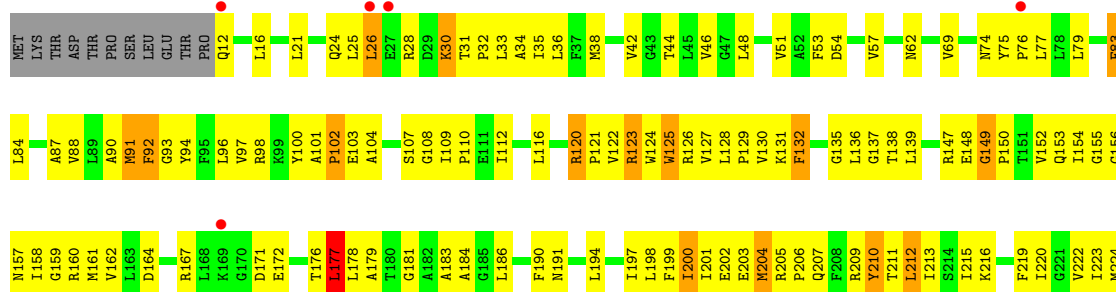
- Molecule 1: putative channel transporter

Chain E:



- Molecule 1: putative channel transporter

Chain F:



A400	L406	T407	I410	L411	E414	M415	T416	D417	M418	Y419	Q420	L421	L422	L423	P424	M425	I426	L427	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	
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			
S225	I226	I227	M228	Y229	R230	I231	F232	M233	H234	E235	V236	G242	K243	L244	L249	M250	T251	L252	Y255	L258	G259	I260	F265	G266	P267	I268	F269	M270	V273	L274	G275	M276	Q277	D278	L279	M280	H281	R282	V283	H284	G285	G286	I287	I288	T289	K290	W291	M294	G295	G296	A297

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	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%)	DCC
Wilson B-factor (Å <sup>2</sup> )	125.9	Xtriage
Anisotropy	0.318	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 49.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 53763 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	20274	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.



## 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(°)	Ideal(°)
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 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3379	0	3537	338	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3379	0	3537	336	52
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

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 44.

The worst 5 of 1844 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:200:ILE:HD13	1:D:204:MET:HG3	1.38	1.06
1:F:200:ILE:HD13	1:F:204:MET:HG3	1.38	1.05
1:B:200:ILE:HD13	1:B: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:HD12	1:F:322:ILE:N	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	Distance(Å)	Clash(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

### 5.3.1 Protein backbone ⓘ

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%)	2	24
1	B	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	2	24

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	2	24
1	D	448/473 (95%)	340 (76%)	76 (17%)	32 (7%)	2	25
1	E	448/473 (95%)	340 (76%)	76 (17%)	32 (7%)	2	25
1	F	448/473 (95%)	340 (76%)	75 (17%)	33 (7%)	2	24
All	All	2688/2838 (95%)	2040 (76%)	452 (17%)	196 (7%)	2	25

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 ⓘ

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%)	36	82
1	B	339/358 (95%)	323 (95%)	16 (5%)	36	82
1	C	339/358 (95%)	323 (95%)	16 (5%)	36	82
1	D	339/358 (95%)	323 (95%)	16 (5%)	36	82
1	E	339/358 (95%)	323 (95%)	16 (5%)	36	82
1	F	339/358 (95%)	323 (95%)	16 (5%)	36	82
All	All	2034/2148 (95%)	1938 (95%)	96 (5%)	36	82

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

Mol	Chain	Res	Type
1	C	317	PHE
1	D	132	PHE
1	F	212	LEU
1	C	322	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	30	LYS

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

Mol	Chain	Res	Type
1	C	327	ASN
1	D	207	GLN
1	F	277	GLN
1	C	418	ASN
1	D	62	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	450/473 (95%)	-0.09	5 (1%)	77 44	51, 100, 147, 151	0
1	B	450/473 (95%)	-0.05	6 (1%)	74 40	51, 100, 147, 151	0
1	C	450/473 (95%)	-0.06	7 (1%)	68 35	51, 100, 147, 151	0
1	D	450/473 (95%)	-0.03	11 (2%)	56 26	51, 100, 147, 151	0
1	E	450/473 (95%)	-0.03	8 (1%)	65 33	51, 100, 147, 151	0
1	F	450/473 (95%)	-0.04	5 (1%)	77 44	51, 100, 147, 151	0
All	All	2700/2838 (95%)	-0.05	42 (1%)	68 35	51, 100, 147, 151	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	12	GLN	6.3
1	A	12	GLN	5.7
1	B	12	GLN	5.5
1	D	15	ARG	4.2
1	D	169	LYS	4.1

### 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 carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.