



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 07:11 am GMT

PDB ID : 2H2S  
Title : Crystal Structure of E148A mutant of CLC-ec1 in SeCN-  
Authors : Nguitragool, W.; Miller, C.  
Deposited on : 2006-05-19  
Resolution : 3.10 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](mailto: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 ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

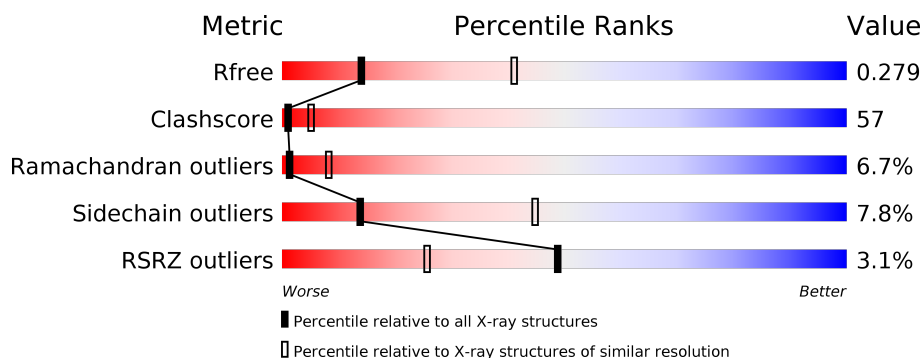
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	111664	1115 (3.12-3.08)
Clashscore	122126	1042 (3.10-3.10)
Ramachandran outliers	120053	1010 (3.10-3.10)
Sidechain outliers	120020	1010 (3.10-3.10)
RSRZ outliers	108989	1089 (3.12-3.08)

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. 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  $\leq 5\%$ . 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	465	<div> <div>2%</div> <div> <div></div> <div>32%</div> <div>57%</div> <div>7%</div> <div>5%</div> </div> </div>
1	B	465	<div> <div>4%</div> <div> <div></div> <div>33%</div> <div>55%</div> <div>8%</div> <div>5%</div> </div> </div>
2	C	221	<div> <div>5%</div> <div> <div></div> <div>38%</div> <div>51%</div> <div>10%</div> </div> </div>
2	E	221	<div> <div></div> <div> <div></div> <div>39%</div> <div>52%</div> <div>10%</div> </div> </div>
3	D	211	<div> <div>4%</div> <div> <div></div> <div>21%</div> <div>62%</div> <div>16%</div> <div>.</div> </div> </div>
3	F	211	<div> <div>3%</div> <div> <div></div> <div>29%</div> <div>60%</div> <div>9%</div> <div>.</div> </div> </div>

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
4	SEK	A	466	-	-	X	-
4	SEK	B	466	-	-	X	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13221 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 CLC Cl transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	0	0
			3329	2188	560	561	20			
1	B	441	Total	C	N	O	S	0	0	0
			3300	2172	553	555	20			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	148	ALA	GLU	ENGINEERED	UNP P37019
B	148	ALA	GLU	ENGINEERED	UNP P37019

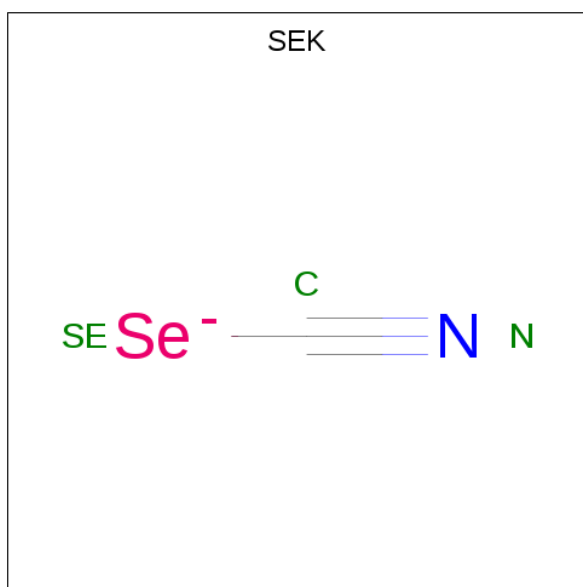
- Molecule 2 is a protein called FAB fragment, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	221	Total	C	N	O	S	0	0	0
			1672	1077	274	315	6			
2	E	221	Total	C	N	O	S	0	0	0
			1672	1077	274	315	6			

- Molecule 3 is a protein called FAB fragment, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			
3	F	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			

- Molecule 4 is SELENOCYANATE ION (three-letter code: SEK) (formula: CNSe).

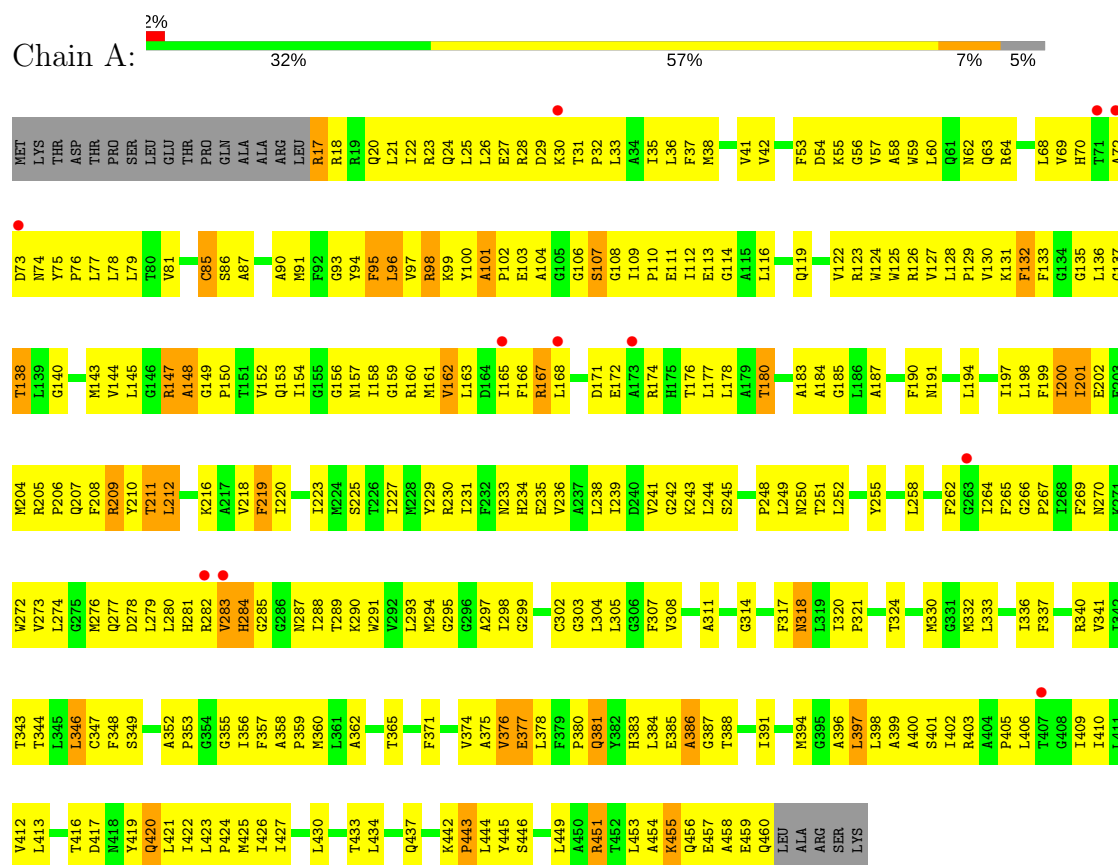


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Se 1	0	0
4	A	1	Total 1	Se 1	1	0
4	A	1	Total 1	Se 1	1	0
4	B	1	Total 1	Se 1	0	0
4	B	1	Total 1	Se 1	1	0
4	B	1	Total 1	Se 1	1	0

### 3 Residue-property plots

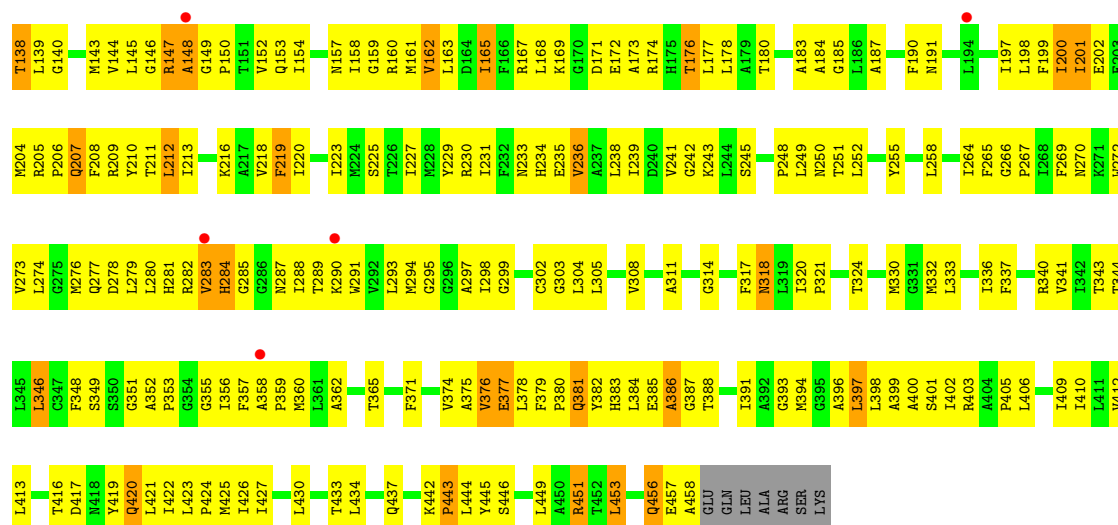
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: CLC Cl transporter

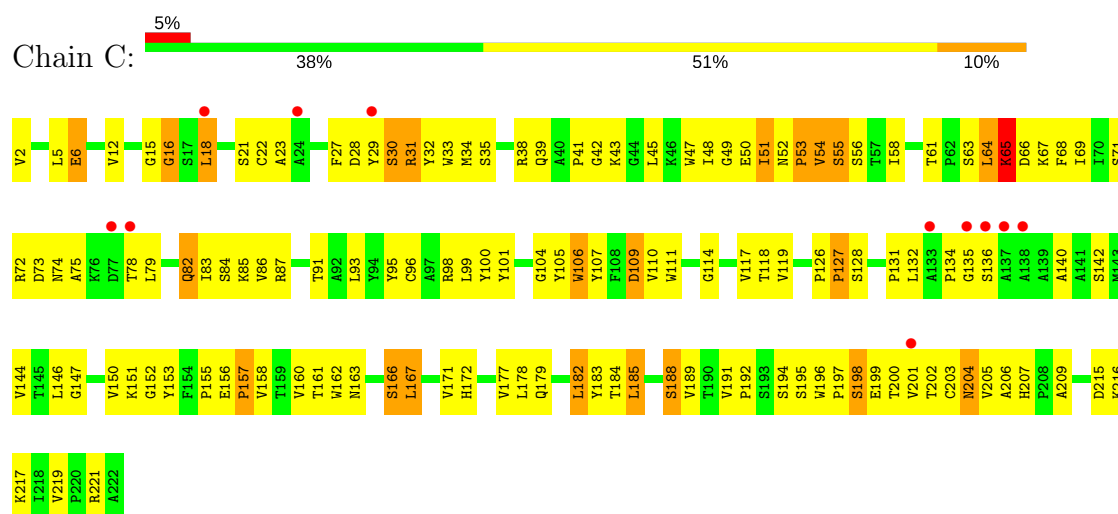


#### • Molecule 1: CLC Cl transporter

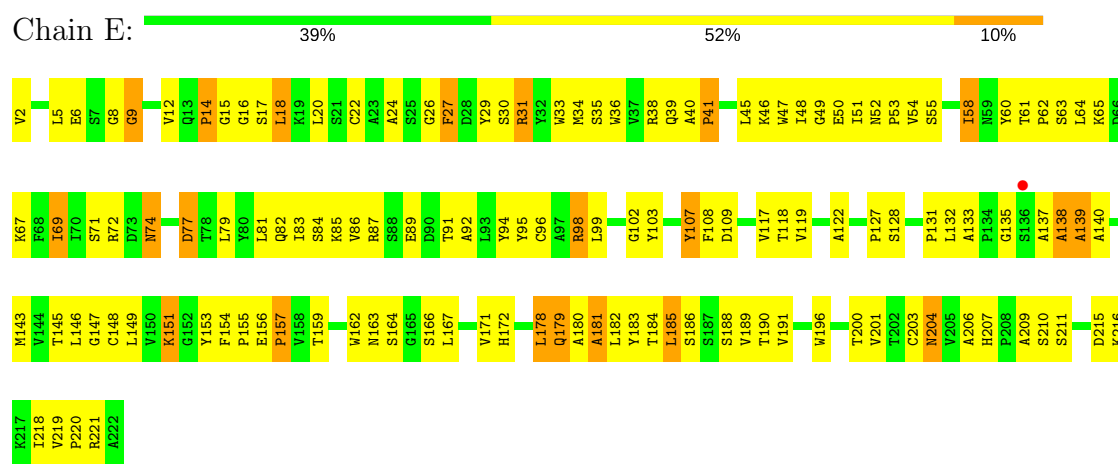




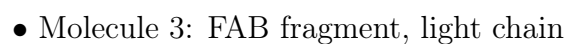
• Molecule 2: FAB fragment, heavy chain



• Molecule 2: FAB fragment, heavy chain



• Molecule 3: FAB fragment, light chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	216.93Å 118.69Å 148.91Å 90.00° 127.44° 90.00°	Depositor
Resolution (Å)	50.00 – 3.10 48.87 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.00-3.10) 99.3 (48.87-3.10)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.61 (at 3.12Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.281 , 0.282 0.277 , 0.279	Depositor DCC
$R_{free}$ test set	2714 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.0	Xtriage
Anisotropy	0.566	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 67.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	13221	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SEK

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	0/3401	0.67	0/4616
1	B	0.47	0/3372	0.67	0/4578
2	C	0.53	0/1721	0.79	0/2355
2	E	0.50	0/1721	0.74	0/2355
3	D	0.56	1/1660 (0.1%)	0.80	1/2257 (0.0%)
3	F	0.53	0/1660	0.76	1/2257 (0.0%)
All	All	0.51	1/13535 (0.0%)	0.72	2/18418 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	167	SER	N-CA	-5.62	1.35	1.46

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	167	SER	N-CA-CB	6.00	119.50	110.50
3	F	169	ASP	N-CA-C	5.92	126.98	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	3329	0	3483	400	0
1	B	3300	0	3456	386	0
2	C	1672	0	1653	190	0
2	E	1672	0	1654	163	0
3	D	1621	0	1546	263	0
3	F	1621	0	1546	203	0
4	A	3	0	0	5	0
4	B	3	0	0	5	0
All	All	13221	0	13338	1503	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 57.

The worst 5 of 1503 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:E:131:PRO:CG	2:E:216:LYS:HE3	1.29	1.56
2:E:131:PRO:CD	2:E:216:LYS:HE2	1.27	1.55
2:C:53:PRO:CA	2:C:72:ARG:NH2	1.68	1.53
2:E:131:PRO:CG	2:E:216:LYS:CE	1.82	1.50
3:D:88:GLN:HG3	3:D:97:PHE:CE1	1.60	1.33

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	442/465 (95%)	318 (72%)	101 (23%)	23 (5%)	2	14
1	B	439/465 (94%)	313 (71%)	101 (23%)	25 (6%)	2	12
2	C	219/221 (99%)	170 (78%)	29 (13%)	20 (9%)	1	4
2	E	219/221 (99%)	173 (79%)	35 (16%)	11 (5%)	2	15

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	209/211 (99%)	151 (72%)	33 (16%)	25 (12%)	0	2
3	F	209/211 (99%)	166 (79%)	30 (14%)	13 (6%)	1	10
All	All	1737/1794 (97%)	1291 (74%)	329 (19%)	117 (7%)	1	8

5 of 117 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	95	PHE
1	A	107	SER
1	A	132	PHE
1	A	148	ALA
1	A	167	ARG

### 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	334/352 (95%)	315 (94%)	19 (6%)	23	56
1	B	331/352 (94%)	312 (94%)	19 (6%)	23	56
2	C	181/181 (100%)	163 (90%)	18 (10%)	8	32
2	E	181/181 (100%)	165 (91%)	16 (9%)	11	39
3	D	185/185 (100%)	166 (90%)	19 (10%)	8	30
3	F	185/185 (100%)	167 (90%)	18 (10%)	9	33
All	All	1397/1436 (97%)	1288 (92%)	109 (8%)	14	45

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	172	HIS
3	D	70	TYR
3	F	119	PRO
2	C	182	LEU
2	C	204	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	207	GLN
1	B	327	ASN
3	F	36	GLN
1	B	270	ASN
1	B	277	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 are modelled with single atom - 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 ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	444/465 (95%)	0.18	11 (2%) 57 34	78, 78, 78, 78	0
1	B	441/465 (94%)	0.30	17 (3%) 39 19	78, 78, 78, 78	0
2	C	221/221 (100%)	0.22	11 (4%) 29 13	78, 78, 78, 78	0
2	E	221/221 (100%)	0.11	1 (0%) 90 81	78, 78, 78, 78	0
3	D	211/211 (100%)	0.22	9 (4%) 35 17	78, 78, 78, 78	0
3	F	211/211 (100%)	0.18	6 (2%) 53 29	78, 78, 78, 78	0
All	All	1749/1794 (97%)	0.21	55 (3%) 49 26	78, 78, 78, 78	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	137	ALA	4.5
2	C	136	SER	4.2
1	B	75	TYR	4.0
2	C	77	ASP	3.9
2	C	135	GLY	3.8

### 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 carbohydrates 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	SEK	A	467	1/3	-	-	78,78,78,78	1
4	SEK	B	467	1/3	-	-	78,78,78,78	1
4	SEK	B	468	1/3	-	-	78,78,78,78	1
4	SEK	A	466	1/3	0.60	0.30	78,78,78,78	0
4	SEK	B	466	1/3	0.79	0.43	78,78,78,78	0
4	SEK	A	468	1/3	-	-	78,78,78,78	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.