wwPDB X-ray Structure Validation Summary Report (i)

Nov 13, 2024 – 08:08 AM EST

PDB ID	:	4KJP
Title	:	Structure of the CLC-ec1 deltaNC construct in the absence of halide
Authors	:	Lim, H.-H.; Miller, C.
Deposited on	:	2013-05-03
Resolution	:	3.20 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:

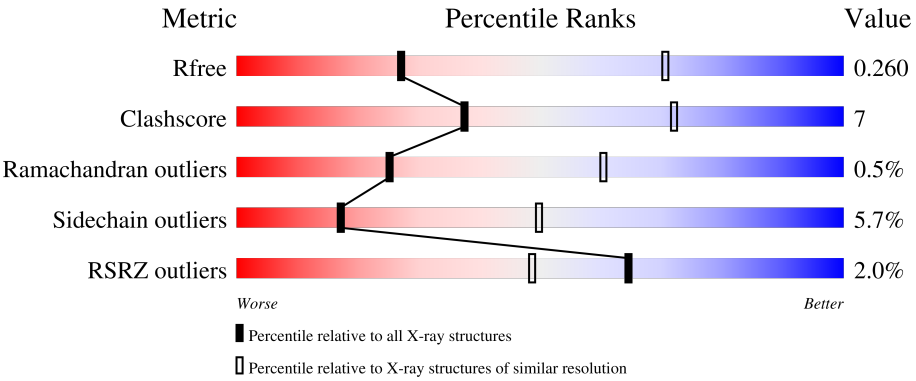
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.20 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}	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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	446	2% 77% 20%
1	B	446	3% 75% 22%
2	C	222	% 83% 16%
2	E	222	% 83% 15%
3	D	211	4% 77% 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	211	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13214 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 H(+)/Cl(-) exchange transporter ClcA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	443	Total	C	N	O	S	0	0	0
			3324	2185	558	561	20			
1	B	441	Total	C	N	O	S	0	0	0
			3304	2174	553	557	20			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	MET	-	initiating methionine	UNP P37019
A	461	LYS	-	expression tag	UNP P37019
B	16	MET	-	initiating methionine	UNP P37019
B	461	LYS	-	expression tag	UNP P37019

- Molecule 2 is a protein called heavy chain of Fab fragment.

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 light chain of Fab fragment.

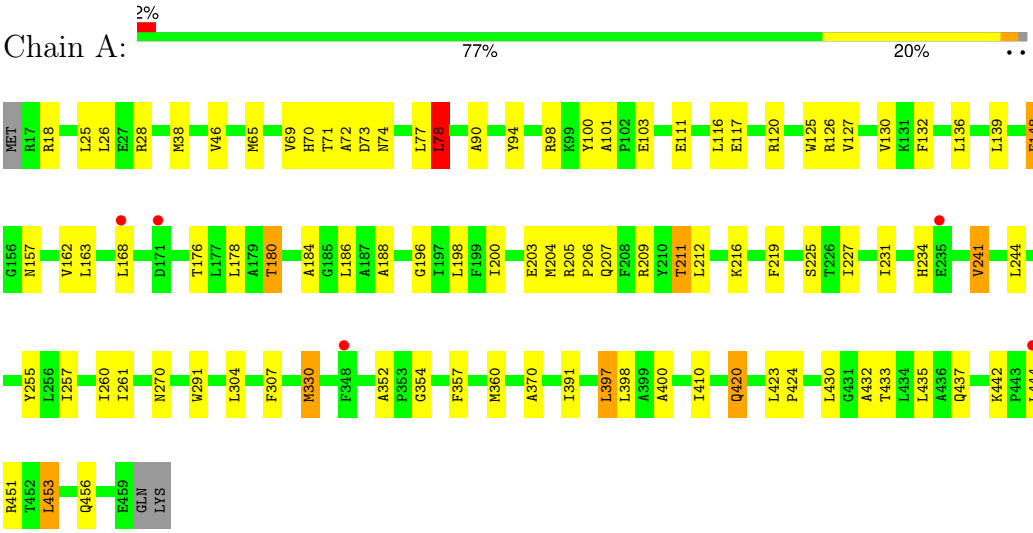
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			
3	F	211	Total	C	N	O	S	0	0	0
			1621	1008	271	334	8			

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.

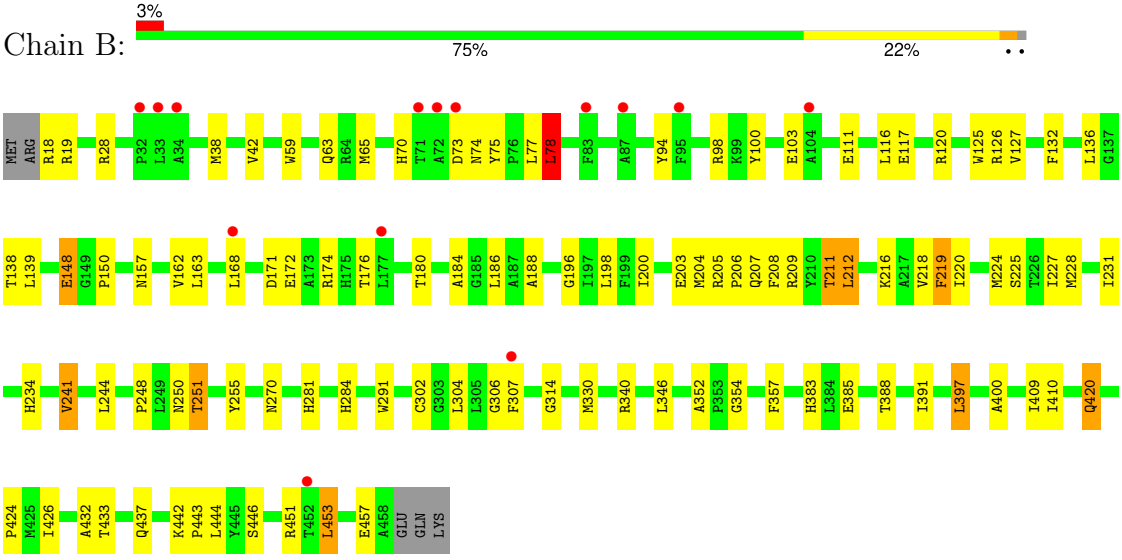
• Molecule 1: H(+)/Cl(-) exchange transporter ClcA

Chain A:



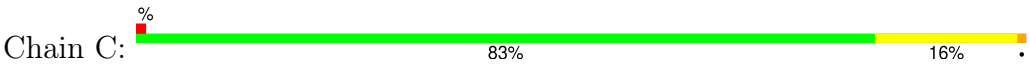
• Molecule 1: H(+)/Cl(-) exchange transporter ClcA

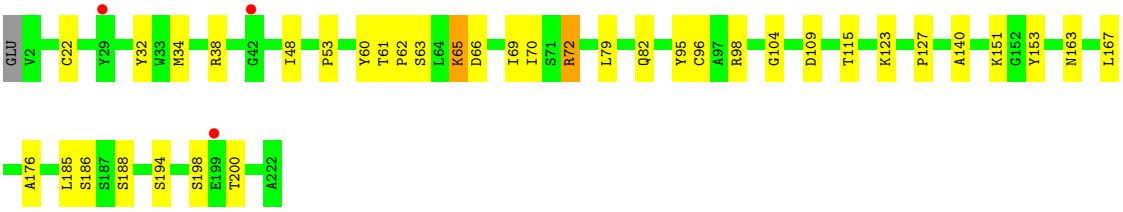
Chain B:



• Molecule 2: heavy chain of Fab fragment

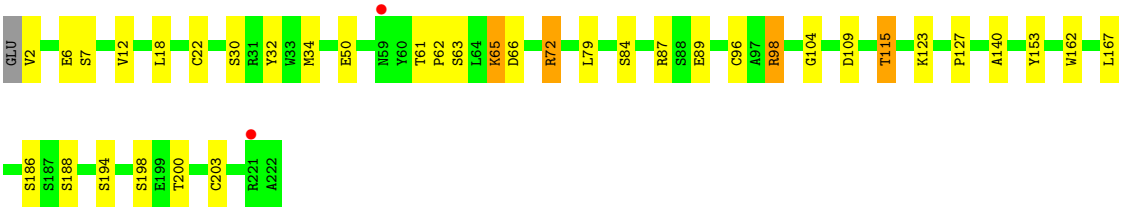
Chain C:





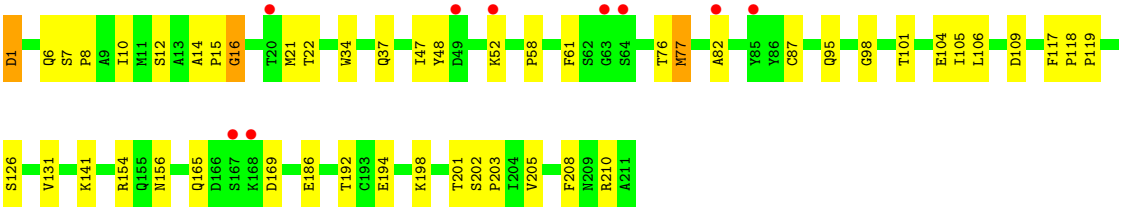
• Molecule 2: heavy chain of Fab fragment

Chain E:



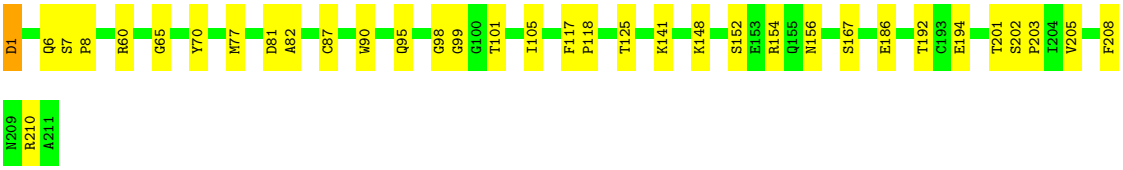
• Molecule 3: light chain of Fab fragment

Chain D:



• Molecule 3: light chain of Fab fragment

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	232.45Å 98.25Å 171.03Å 90.00° 132.08° 90.00°	Depositor
Resolution (Å)	24.89 – 3.20 24.89 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.1 (24.89-3.20) 99.2 (24.89-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 3.17Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.222 , 0.265 0.223 , 0.260	Depositor DCC
R_{free} test set	2378 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	89.4	Xtriage
Anisotropy	0.523	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 5.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.014 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	13214	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% 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	0.30	0/3396	0.43	1/4609 (0.0%)
1	B	0.30	0/3376	0.44	1/4583 (0.0%)
2	C	0.34	0/1721	0.47	0/2355
2	E	0.34	0/1721	0.46	0/2355
3	D	0.32	0/1660	0.48	0/2257
3	F	0.32	0/1660	0.47	0/2257
All	All	0.31	0/13534	0.45	2/18416 (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	78	LEU	CA-CB-CG	6.16	129.47	115.30
1	A	78	LEU	CA-CB-CG	5.93	128.94	115.30

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	3324	0	3476	56	0
1	B	3304	0	3457	63	0
2	C	1672	0	1654	16	0
2	E	1672	0	1654	16	0
3	D	1621	0	1546	28	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	1621	0	1546	20	0
All	All	13214	0	13333	183	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 7.

The worst 5 of 183 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)
3:F:6:GLN:HE22	3:F:87:CYS:H	1.31	0.78
3:F:6:GLN:HE21	3:F:98:GLY:HA3	1.50	0.76
1:A:216:LYS:NZ	1:B:437:GLN:OE1	2.20	0.75
2:E:98:ARG:NH1	2:E:109:ASP:OD2	2.20	0.74
3:F:1:ASP:OD2	3:F:1:ASP:N	2.24	0.70

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	441/446 (99%)	425 (96%)	16 (4%)	0	100 100
1	B	439/446 (98%)	421 (96%)	18 (4%)	0	100 100
2	C	219/222 (99%)	205 (94%)	11 (5%)	3 (1%)	9 40
2	E	219/222 (99%)	202 (92%)	14 (6%)	3 (1%)	9 40
3	D	209/211 (99%)	190 (91%)	16 (8%)	3 (1%)	9 40
3	F	209/211 (99%)	192 (92%)	17 (8%)	0	100 100
All	All	1736/1758 (99%)	1635 (94%)	92 (5%)	9 (0%)	25 60

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	62	PRO
2	E	62	PRO
2	C	65	LYS
3	D	16	GLY
3	D	126	SER

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	334/337 (99%)	309 (92%)	25 (8%)	11	40
1	B	332/337 (98%)	309 (93%)	23 (7%)	13	43
2	C	181/182 (100%)	172 (95%)	9 (5%)	20	54
2	E	181/182 (100%)	170 (94%)	11 (6%)	15	47
3	D	185/185 (100%)	180 (97%)	5 (3%)	40	69
3	F	185/185 (100%)	178 (96%)	7 (4%)	28	60
All	All	1398/1408 (99%)	1318 (94%)	80 (6%)	17	50

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	1	ASP
2	E	188	SER
3	D	141	LYS
2	E	84	SER
3	F	125	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	63	GLN
1	B	284	HIS
3	D	6	GLN
3	D	37	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	F	6	GLN

5.3.3 RNA

There are no RNA molecules 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 oligosaccharides 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^{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	443/446 (99%)	-0.17	7 (1%) 70 55	36, 56, 88, 119	0
1	B	441/446 (98%)	-0.04	14 (3%) 50 35	36, 61, 99, 133	0
2	C	221/222 (99%)	-0.28	3 (1%) 73 58	28, 54, 88, 121	0
2	E	221/222 (99%)	-0.27	2 (0%) 81 68	29, 54, 86, 126	0
3	D	211/211 (100%)	0.14	9 (4%) 40 27	38, 66, 94, 104	0
3	F	211/211 (100%)	-0.25	0 100 100	30, 50, 99, 114	0
All	All	1748/1758 (99%)	-0.13	35 (2%) 64 49	28, 57, 94, 133	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	104	ALA	5.3
1	B	73	ASP	5.0
3	D	167	SER	5.0
1	A	447	ALA	4.5
1	B	72	ALA	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 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.