WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 02:03 AM GMT

PDB ID : 2FEC
Title : Structure of the E203Q mutant of the Cl-/H+ exchanger CLC-ec1 from E.Coli
Authors : Accardi, A.; Walden, M.P.; Nguitragool, W.; Jayaram, H.; Williams, C.; Miller, C.
Deposited on : 2005-12-15
Resolution : 3.97 Å(reported)

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

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 3.97 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.
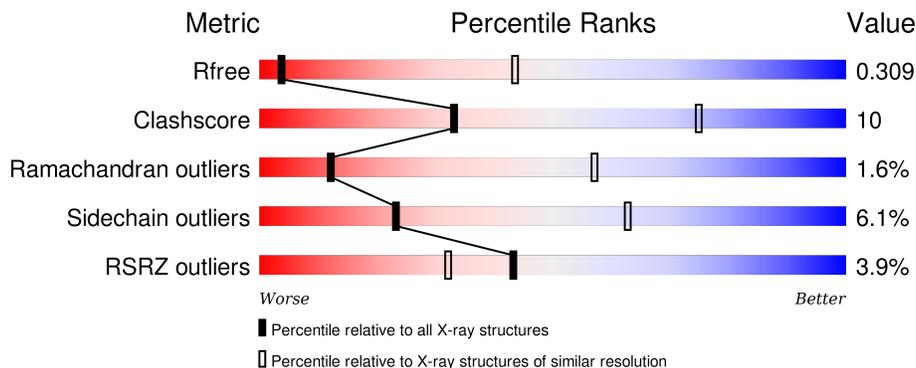
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1011 (4.38-3.54)
Clashscore	102246	1025 (4.32-3.60)
Ramachandran outliers	100387	1008 (4.34-3.58)
Sidechain outliers	100360	1027 (4.36-3.56)
RSRZ outliers	91569	1015 (4.38-3.54)

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 >=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 <=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	
1	B	465	
2	I	222	
2	J	222	
3	L	211	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	O	211	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13223 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	444	Total 3333	C 2190	N 561	O 562	S 20	0	0	0
1	B	441	Total 3304	C 2174	N 554	O 556	S 20	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	GLN	GLU	ENGINEERED	UNP P37019
B	203	GLN	GLU	ENGINEERED	UNP P37019

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	O	211	Total 1621	C 1008	N 271	O 334	S 8	0	0	0
3	L	211	Total 1621	C 1008	N 271	O 334	S 8	0	0	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 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: H(+)/Cl(-) exchange transporter clcA
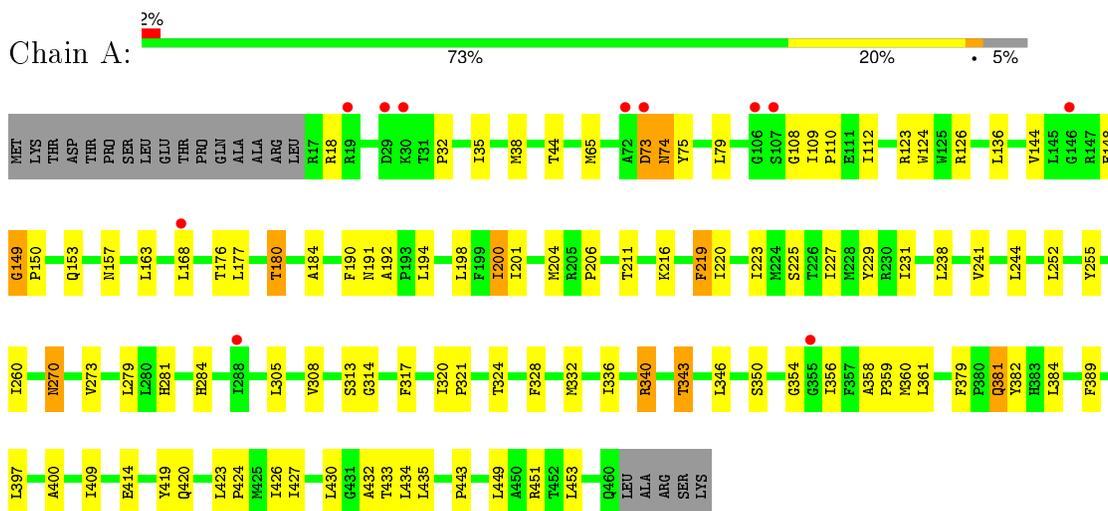
- Molecule 1: H(+)/Cl(-) exchange transporter clcA

- Molecule 2: Fab fragment, heavy chain

Chain J:

5%

71% 27%

GLU V2 R3 E6 V12 S17 L18 K19 C22 G26 S30 R31 Y32 W33 M34 S35 P41 G42 K43 W47 I48 G49 E50 I51 N52 P53 V54 N59 Y60 T61 P62 S63 L64 K65 I69 I70 D77 T78 L79 Q82 V86 T91 A92 Y95

R98 T116 V117 T118 V119 S120 S121 T125 P126 P127 S128 G135 S136 A137 A138 A139 A140 A141 G147 K151 G152 Y153 E156 P157 W162 N163 L167 F174 Q179 Y183 T184 L185 S186 W196 P197 S198 E199 T200 V201 T202 C203 H207 P208 A209 T212

V219 P220 R221 A222

• Molecule 2: Fab fragment, heavy chain

Chain I:

%

72% 26%

GLU V2 L5 E6 V12 Q13 P14 L20 S21 C22 A23 A24 F27 S30 R31 Y32 W33 M34 S35 R38 I48 N52 P53 V54 S55 S56 T61 P62 S63 L64 K65 D66 K67 F68 I69 L79 Y80 L81 Q82 I83 S84 K85 E89 D90 T91 L99 Y107

F108 D109 V110 W111 T115 T116 V117 T118 V119 Y130 P131 L132 G135 S136 A137 A140 A141 S142 K151 P155 E156 P157 S164 Q179 S193 W196 P197 S198 E199 T200 H207 P208 A209 K217 A222

• Molecule 3: Fab fragment, light chain

Chain O:

5%

74% 25%

D1 T5 Q6 S7 P8 A9 I10 T20 S28 V29 S30 Y31 I32 H33 W34 Y35 Y48 D49 T50 S51 T54 S55 G56 V57 P58 V59 R60 F61 S62 G63 S64 T68 S69 Y70 T73 I74 N75 M77 Y85 Y86 C87 W90 Q95 T96 L103 A110 A111

P112 F117 P118 P119 L124 T125 S126 K148 I149 D150 G151 S152 E153 W162 D169 M174 T179 N189 S190 Y191 T192 C193 E194 A195 T196 H197 K198 T201 I204 V205 R210 A211

• Molecule 3: Fab fragment, light chain

Chain L:

5%

65% 32%

D1 Q6 S7 P8 T22 V29 S30 Y31 I32 H33 W34 Y35 Q36 Q37 K38 T41 S42 P43 K44 R45 W46 I47 Y48 D49 T50 S51 K52 S55 G56 V57 P58 V59 S66 G67 T68 S69 Y70 A79 E80 D81 T84 Y85 Y86 C87 Q88 Q89 W90 H93 P94 Q95 T96

F97 G98 G99 K102 L106 A110 T113 V114 S115 I116 F117 P118 P119 S120 Q123 L124 T125 S126 G127 G128 A129 S130 N136 N137 F138 K146 W147 K148 E153 R154 N156 G157 V158 K168 D169 S170 S173 T177 L178 T179 L180 D183 E186 N189

T192 C193 E194 V205 F208 N209 R210 A211

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	231.31Å 98.50Å 170.41Å 90.00° 131.51° 90.00°	Depositor
Resolution (Å)	128.04 – 3.97 33.31 – 3.97	Depositor EDS
% Data completeness (in resolution range)	97.8 (128.04-3.97) 97.9 (33.31-3.97)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$<I/\sigma(I)>$ [1]	3.88 (at 3.99Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, R_{free}	0.269 , 0.326 0.253 , 0.309	Depositor DCC
R_{free} test set	1258 reflections (5.39%)	DCC
Wilson B-factor (Å^2)	141.0	Xtriage
Anisotropy	0.433	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 77.5	EDS
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
L-test for twinning [2]	$<|L|> = 0.50, <L^2> = 0.33$	Xtriage
Outliers	0 of 24578 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	13223	wwPDB-VP
Average B, all atoms (Å^2)	155.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.43	0/3405	0.53	0/4621
1	B	0.49	1/3376 (0.0%)	0.55	0/4583
2	I	0.40	0/1721	0.55	0/2355
2	J	0.72	4/1721 (0.2%)	0.59	2/2355 (0.1%)
3	L	0.38	0/1660	0.54	0/2257
3	O	0.42	0/1660	0.52	0/2257
All	All	0.48	5/13543 (0.0%)	0.55	2/18428 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	19	LYS	CE-NZ	13.79	1.83	1.49
2	J	199	GLU	CD-OE2	13.32	1.40	1.25
2	J	199	GLU	CD-OE1	12.20	1.39	1.25
1	B	73	ASP	C-O	10.60	1.43	1.23
2	J	19	LYS	CD-CE	8.95	1.73	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	19	LYS	CD-CE-NZ	-7.37	94.75	111.70
2	J	185	LEU	CA-CB-CG	5.02	126.86	115.30

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	3333	0	3486	63	0
1	B	3304	0	3459	86	0
2	I	1672	0	1654	34	0
2	J	1672	0	1654	35	0
3	L	1621	0	1546	47	0
3	O	1621	0	1546	30	0
All	All	13223	0	13345	273	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 10.

The worst 5 of 273 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:19:LYS:CE	2:J:19:LYS:NZ	1.83	1.41
3:L:7:SER:HB3	3:L:8:PRO:HD3	1.41	1.00
3:L:146:LYS:HB3	3:L:194:GLU:HB2	1.44	1.00
1:B:381:GLN:HE21	1:B:381:GLN:H	1.15	0.92
1:A:220:ILE:HG12	1:B:430:LEU:HD21	1.59	0.84

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%)	411 (93%)	25 (6%)	6 (1%)	14	59
1	B	439/465 (94%)	406 (92%)	29 (7%)	4 (1%)	21	66
2	I	219/222 (99%)	192 (88%)	20 (9%)	7 (3%)	5	43
2	J	219/222 (99%)	193 (88%)	24 (11%)	2 (1%)	21	66
3	L	209/211 (99%)	181 (87%)	22 (10%)	6 (3%)	6	46

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	O	209/211 (99%)	181 (87%)	25 (12%)	3 (1%)	14	59
All	All	1737/1796 (97%)	1564 (90%)	145 (8%)	28 (2%)	12	57

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	73	ASP
1	B	73	ASP
3	O	55	SER
3	L	7	SER
1	A	74	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	335/353 (95%)	317 (95%)	18 (5%)	27	67
1	B	332/353 (94%)	309 (93%)	23 (7%)	19	59
2	I	181/182 (100%)	173 (96%)	8 (4%)	35	72
2	J	181/182 (100%)	165 (91%)	16 (9%)	12	48
3	L	185/185 (100%)	173 (94%)	12 (6%)	21	61
3	O	185/185 (100%)	176 (95%)	9 (5%)	31	69
All	All	1399/1440 (97%)	1313 (94%)	86 (6%)	23	63

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	435	LEU
2	J	98	ARG
3	L	106	LEU
1	B	451	ARG
2	J	30	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	318	ASN
1	B	381	GLN
3	L	136	ASN
1	B	327	ASN
1	B	383	HIS

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)

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 (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, 95th 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	444/465 (95%)	-0.08	11 (2%)	61	48	117, 150, 183, 202	0
1	B	441/465 (94%)	0.03	21 (4%)	34	25	121, 158, 194, 210	0
2	I	221/222 (99%)	-0.00	3 (1%)	78	68	131, 162, 185, 194	0
2	J	221/222 (99%)	0.13	12 (5%)	29	21	131, 146, 169, 176	0
3	L	211/211 (100%)	0.17	10 (4%)	35	26	115, 139, 188, 197	0
3	O	211/211 (100%)	0.17	11 (5%)	31	23	152, 167, 181, 183	0
All	All	1749/1796 (97%)	0.04	68 (3%)	43	32	115, 156, 185, 210	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	147	GLY	8.4
1	B	355	GLY	5.6
1	B	73	ASP	4.5
3	O	197	HIS	4.3
1	B	107	SER	4.2

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.