wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 07:50 AM EDT

PDB ID	:	4MQX
Title	:	CLC-ec1 Fab Complex Cysless A399C-A432C mutant
Authors	:	Basilio, D.; Noack, K.; Picollo, A.; Accardi, A.
Deposited on	:	2013-09-16
Resolution	:	3.52 Å(reported)

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

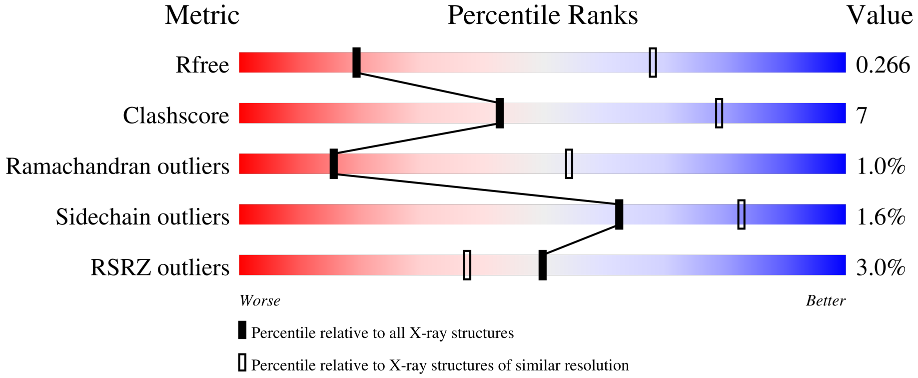
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

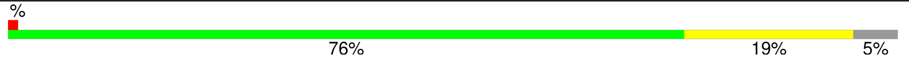
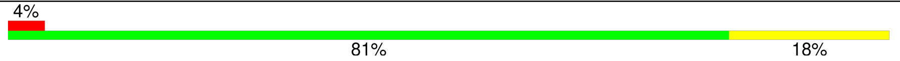
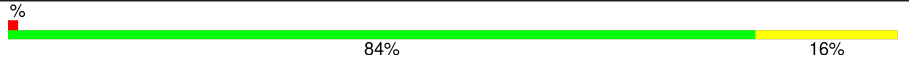
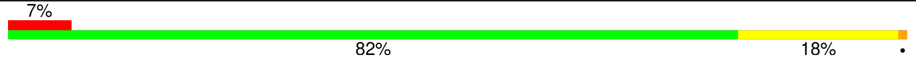
The reported resolution of this entry is 3.52 Å.

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	1161 (3.60-3.44)
Clashscore	141614	1244 (3.60-3.44)
Ramachandran outliers	138981	1206 (3.60-3.44)
Sidechain outliers	138945	1207 (3.60-3.44)
RSRZ outliers	127900	1080 (3.60-3.44)

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	465	
1	B	465	
2	C	222	
2	E	222	
3	D	211	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	211	

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	CL	A	502	-	-	-	X

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 13227 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 560	O 564	S 19	0	0	0
1	B	441	Total 3304	C 2174	N 553	O 558	S 19	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	ALA	CYS	ENGINEERED MUTATION	UNP P37019
A	302	ALA	CYS	ENGINEERED MUTATION	UNP P37019
A	347	SER	CYS	ENGINEERED MUTATION	UNP P37019
A	399	CYS	ALA	ENGINEERED MUTATION	UNP P37019
A	432	CYS	ALA	ENGINEERED MUTATION	UNP P37019
B	85	ALA	CYS	ENGINEERED MUTATION	UNP P37019
B	302	ALA	CYS	ENGINEERED MUTATION	UNP P37019
B	347	SER	CYS	ENGINEERED MUTATION	UNP P37019
B	399	CYS	ALA	ENGINEERED MUTATION	UNP P37019
B	432	CYS	ALA	ENGINEERED MUTATION	UNP P37019

- Molecule 2 is a protein called ecCLC.

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

- Molecule 3 is a protein called ERIC.

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

Continued on next page...

Continued from previous page...

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

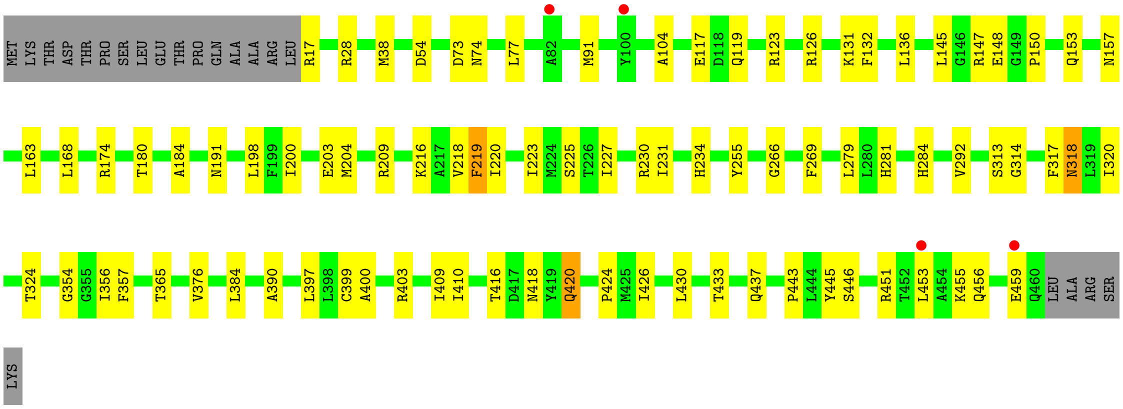
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Cl 2 2	0	0
4	B	2	Total Cl 2 2	0	0

3 Residue-property plots

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.

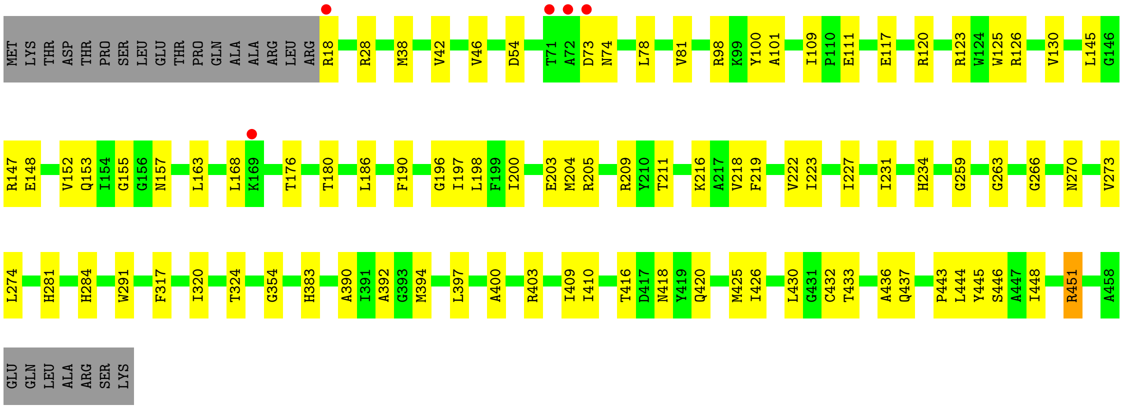
- Molecule 1: H(+)/Cl(-) exchange transporter ClcA

Chain A:



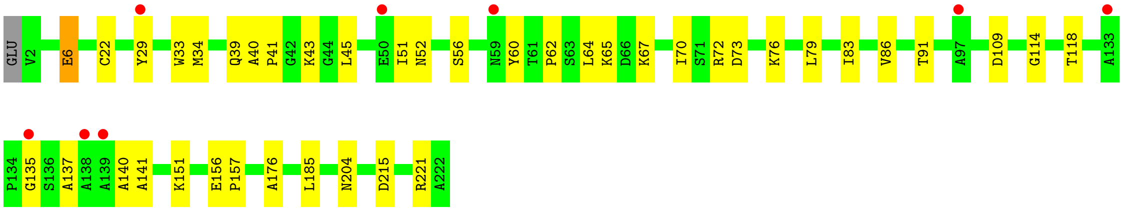
- Molecule 1: H(+)/Cl(-) exchange transporter ClcA

Chain B:



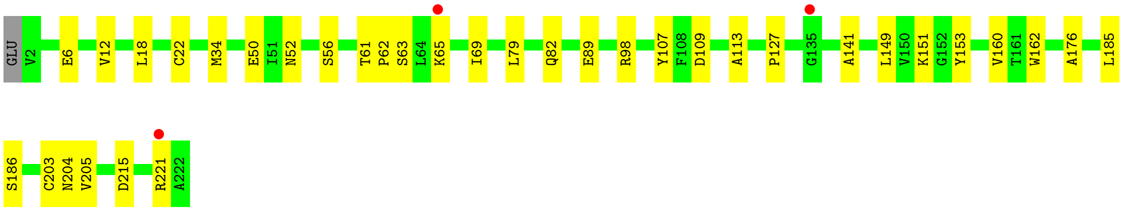
- Molecule 2: ecCLC

Chain C:



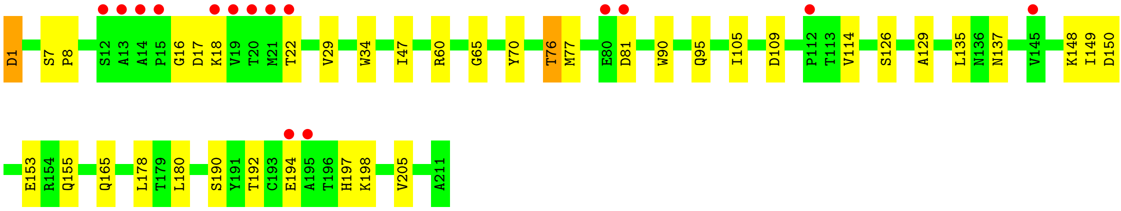
- Molecule 2: ecCLC

Chain E:



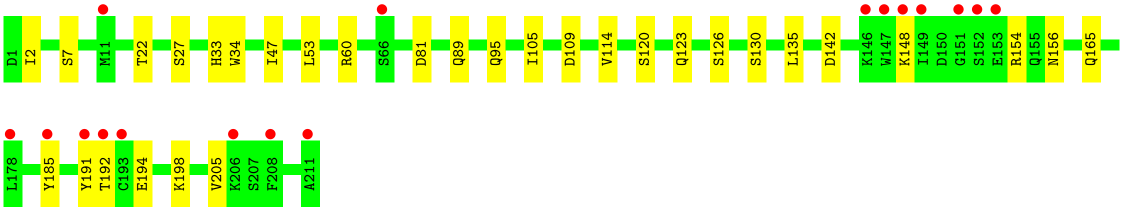
- Molecule 3: ERIC

Chain D:



- Molecule 3: ERIC

Chain F:



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	233.51Å 94.52Å 170.60Å 90.00° 131.75° 90.00°	Depositor
Resolution (Å)	43.56 – 3.52 44.30 – 3.52	Depositor EDS
% Data completeness (in resolution range)	98.7 (43.56-3.52) 92.9 (44.30-3.52)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 3.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.223 , 0.258 0.227 , 0.266	Depositor DCC
R_{free} test set	2005 reflections (5.84%)	wwPDB-VP
Wilson B-factor (Å^2)	112.3	Xtriage
Anisotropy	0.612	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 24.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	13227	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% 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 (i)

5.1 Standard geometry (i)

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

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	0.22	0/3405	0.38	0/4621
1	B	0.22	0/3376	0.38	0/4583
2	C	0.21	0/1721	0.39	0/2355
2	E	0.22	0/1721	0.40	0/2355
3	D	0.22	0/1660	0.39	0/2257
3	F	0.22	0/1660	0.39	0/2257
All	All	0.22	0/13543	0.39	0/18428

There are no bond length outliers.

There are no bond angle outliers.

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	3482	61	0
1	B	3304	0	3455	59	0
2	C	1672	0	1654	20	0
2	E	1672	0	1654	21	0
3	D	1621	0	1546	21	0
3	F	1621	0	1546	17	0
4	A	2	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	2	0	0	2	0
All	All	13227	0	13337	180	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 180 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:B:445:TYR:OH	4:B:501:CL:CL	2.41	0.74
1:A:38:MET:HG3	1:A:168:LEU:HD11	1.72	0.72
2:C:64:LEU:HB2	2:C:67:LYS:HB2	1.74	0.70
1:B:117:GLU:OE1	1:B:209:ARG:NH1	2.27	0.67
1:B:109:ILE:HG12	1:B:152:VAL:HG11	1.77	0.67

There are no symmetry-related clashes.

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	442/465 (95%)	427 (97%)	14 (3%)	1 (0%)	47	80
1	B	439/465 (94%)	413 (94%)	25 (6%)	1 (0%)	47	80
2	C	219/222 (99%)	197 (90%)	15 (7%)	7 (3%)	4	31
2	E	219/222 (99%)	194 (89%)	22 (10%)	3 (1%)	11	47
3	D	209/211 (99%)	190 (91%)	16 (8%)	3 (1%)	11	47
3	F	209/211 (99%)	190 (91%)	17 (8%)	2 (1%)	15	54
All	All	1737/1796 (97%)	1611 (93%)	109 (6%)	17 (1%)	15	54

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	65	LYS
2	E	141	ALA
3	F	142	ASP
2	C	65	LYS
2	E	62	PRO

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	335/353 (95%)	326 (97%)	9 (3%)	44	73
1	B	332/353 (94%)	323 (97%)	9 (3%)	44	73
2	C	181/182 (100%)	179 (99%)	2 (1%)	73	87
2	E	181/182 (100%)	180 (99%)	1 (1%)	86	94
3	D	185/185 (100%)	183 (99%)	2 (1%)	73	87
3	F	185/185 (100%)	185 (100%)	0	100	100
All	All	1399/1440 (97%)	1376 (98%)	23 (2%)	62	83

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	420	GLN
1	B	451	ARG
1	B	444	LEU
2	C	6	GLU
1	A	399	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	284	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - 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 ⓘ

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, 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%)	-0.13	4 (0%)	84	73	69, 91, 120, 141	0
1	B	441/465 (94%)	-0.13	5 (1%)	80	69	61, 86, 124, 154	0
2	C	221/222 (99%)	-0.03	8 (3%)	42	32	62, 87, 122, 151	0
2	E	221/222 (99%)	-0.19	3 (1%)	75	62	62, 87, 119, 148	0
3	D	211/211 (100%)	0.30	15 (7%)	16	13	72, 96, 114, 125	0
3	F	211/211 (100%)	0.36	17 (8%)	12	10	64, 84, 119, 129	0
All	All	1749/1796 (97%)	-0.01	52 (2%)	50	37	61, 89, 120, 154	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	13	ALA	5.1
2	C	139	ALA	4.9
3	F	192	THR	4.5
3	F	149	ILE	4.0
3	D	15	PRO	3.8

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 ⓘ

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, 95th 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($Å^2$)	Q<0.9
4	CL	A	502	1/1	0.37	0.60	109,109,109,109	0
4	CL	A	501	1/1	0.85	0.43	92,92,92,92	0
4	CL	B	501	1/1	0.91	0.10	83,83,83,83	0
4	CL	B	502	1/1	0.91	0.29	78,78,78,78	0

6.5 Other polymers ⓘ

There are no such residues in this entry.