WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 13, 2020 – 07:39 am BST

PDB ID	:	3OMO
Title	:	Fragment-Based Design of novel Estrogen Receptor Ligands
Authors	:	Moecklinghoff, S.; van Otterlo, W.A.; Rose, R.; Fuchs, S.; Dominguez Seoane, M.; Waldmann, H.; Ottmann, C.; Brunsveld, L.
Deposited on	:	2010-08-27
Resolution	:	2.21 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 following versions of software and data (see references (i)) were used in the production of this report:

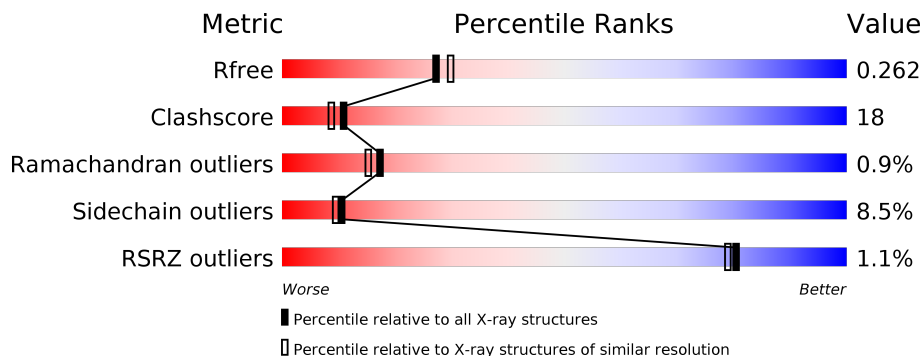|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.21 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	Percentile Ranks	Value
Rfree		0.262
Clashscore		18
Ramachandran outliers		0.9%
Sidechain outliers		8.5%
RSRZ outliers		1.1%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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 on 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	240	67% 23% 5% · 5%
1	B	240	69% 20% 5% 7%
2	C	19	21% 26% 5% 47%
2	D	19	26% 21% 5% 47%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4130 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 Estrogen receptor beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total 1821	C 1171	N 304	O 326	S 20	0	2	0
1	B	224	Total 1766	C 1141	N 293	O 312	S 20	0	0	0

- Molecule 2 is a protein called Nuclear receptor coactivator 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	10	Total 88	C 56	N 21	O 11	0	0	0
2	D	10	Total 88	C 56	N 21	O 11	0	0	0

- Molecule 3 is 2-(trifluoroacetyl)-1,2,3,4-tetrahydroisoquinolin-6-ol (three-letter code: WV7) (formula: $C_{11}H_{10}F_3NO_2$).
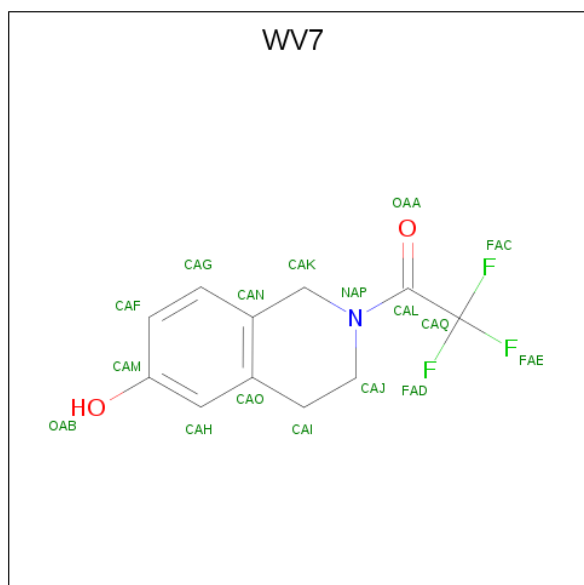
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 17	C 11	F 3	N 1	O 2	0	0
3	B	1	Total 17	C 11	F 3	N 1	O 2	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	166	Total 166	O 166	0	0
4	B	156	Total 156	O 156	0	0
4	C	5	Total 5	O 5	0	0
4	D	6	Total 6	O 6	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 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.
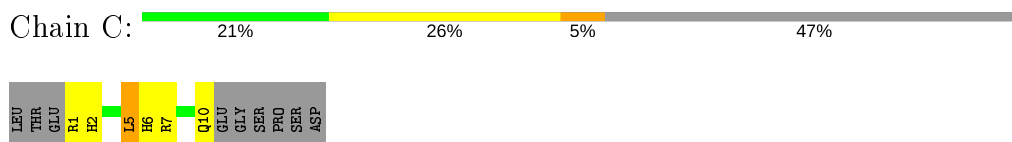
- Molecule 1: Estrogen receptor beta
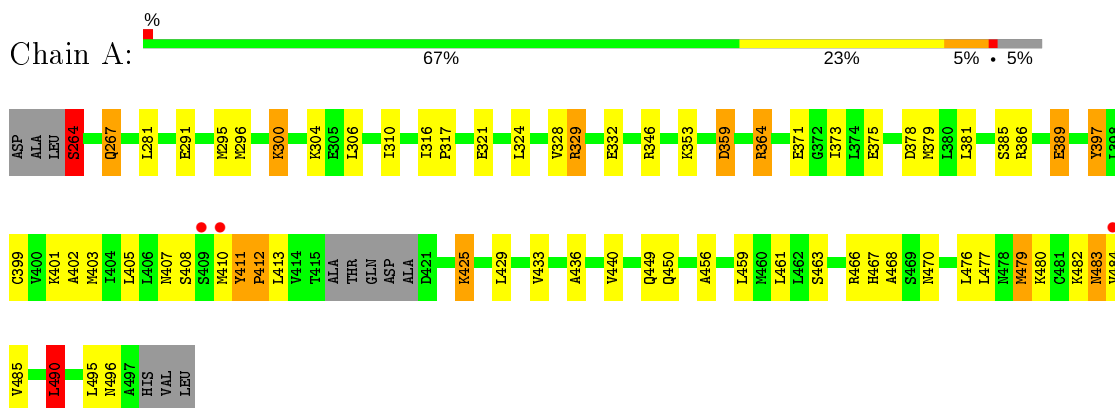
- Molecule 1: Estrogen receptor beta
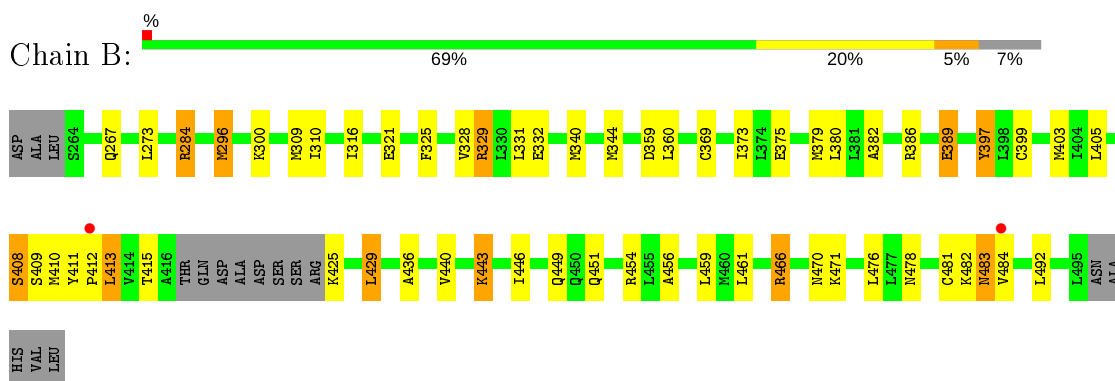
- Molecule 2: Nuclear receptor coactivator 1

- Molecule 2: Nuclear receptor coactivator 1

LEU THR GLU R1 H2 L5 H6 Q10 GLU GLY SER PRO SER ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	70.57Å 70.57Å 109.99Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.54 – 2.21 19.54 – 2.21	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.54-2.21) 96.6 (19.54-2.21)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.21Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.208 , 0.264 0.207 , 0.262	Depositor DCC
R_{free} test set	1475 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	17.9	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 15.6	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.027 for -h,-k,l 0.477 for h,-h-k,-l 0.026 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4130	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% 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: WV7

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	1.05	2/1854 (0.1%)	0.97	5/2506 (0.2%)
1	B	1.08	1/1799 (0.1%)	1.04	6/2432 (0.2%)
2	C	0.99	0/89	1.15	0/118
2	D	0.84	0/89	1.21	0/118
All	All	1.06	3/3831 (0.1%)	1.01	11/5174 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	397	TYR	CD2-CE2	5.38	1.47	1.39
1	A	389	GLU	CG-CD	5.28	1.59	1.51
1	A	402	ALA	CA-CB	5.06	1.63	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	284	ARG	NE-CZ-NH1	-12.95	113.83	120.30
1	B	454	ARG	NE-CZ-NH2	-12.12	114.24	120.30
1	A	364	ARG	NE-CZ-NH1	8.50	124.55	120.30
1	B	454	ARG	NE-CZ-NH1	7.74	124.17	120.30
1	A	364	ARG	NE-CZ-NH2	-7.59	116.50	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	264	SER	Peptide

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	1821	0	1889	73	1
1	B	1766	0	1843	71	0
2	C	88	0	96	7	1
2	D	88	0	96	7	0
3	A	17	0	9	0	0
3	B	17	0	10	1	0
4	A	166	0	0	8	0
4	B	156	0	0	19	0
4	C	5	0	0	0	0
4	D	6	0	0	1	0
All	All	4130	0	3943	141	1

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 18.

The worst 5 of 141 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:A:479:MET:SD	1:A:485:VAL:CG1	2.09	1.39
1:B:296:MET:HE2	4:B:175:HOH:O	1.20	1.34
1:A:479:MET:SD	1:A:485:VAL:HG11	1.66	1.34
1:A:449:GLN:HG3	4:A:99:HOH:O	1.23	1.29
1:B:408:SER:HB2	4:B:527:HOH:O	1.26	1.26

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:359:ASP:OD2	2:C:6:HIS:CE1[2_665]	2.16	0.04

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	227/240 (95%)	218 (96%)	6 (3%)	3 (1%)	12	9
1	B	220/240 (92%)	209 (95%)	10 (4%)	1 (0%)	29	30
2	C	8/19 (42%)	6 (75%)	2 (25%)	0	100	100
2	D	8/19 (42%)	6 (75%)	2 (25%)	0	100	100
All	All	463/518 (89%)	439 (95%)	20 (4%)	4 (1%)	17	15

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	411	TYR
1	B	483	ASN
1	A	483	ASN
1	A	412	PRO

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	208/215 (97%)	191 (92%)	17 (8%)	11	10
1	B	201/215 (94%)	184 (92%)	17 (8%)	10	9

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	9/18 (50%)	8 (89%)	1 (11%)	6	5
2	D	9/18 (50%)	8 (89%)	1 (11%)	6	5
All	All	427/466 (92%)	391 (92%)	36 (8%)	10	10

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	495	LEU
1	B	369	CYS
1	B	483	ASN
1	B	329	ARG
1	B	373	ILE

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

Mol	Chain	Res	Type
1	A	267	GLN
1	A	449	GLN
1	A	450	GLN
1	A	470	ASN
1	B	407	ASN

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)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	WV7	A	1	-	18,18,18	1.14	2 (11%)	27,27,27	1.81	7 (25%)
3	WV7	B	1	-	18,18,18	1.27	2 (11%)	27,27,27	2.06	13 (48%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	WV7	A	1	-	-	0/10/19/19	0/2/2/2
3	WV7	B	1	-	-	0/10/19/19	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1	WV7	CAJ-NAP	3.06	1.52	1.47
3	A	1	WV7	CAJ-CAI	2.56	1.56	1.51
3	A	1	WV7	CAJ-NAP	2.34	1.51	1.47
3	B	1	WV7	CAQ-CAL	-2.03	1.48	1.54

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1	WV7	FAC-CAQ-CAL	-4.66	101.15	111.87
3	B	1	WV7	FAC-CAQ-CAL	-4.46	101.61	111.87
3	B	1	WV7	CAM-CAH-CAO	-3.74	116.67	120.83
3	B	1	WV7	CAQ-CAL-NAP	3.14	123.06	118.01
3	B	1	WV7	CAG-CAN-CAO	3.09	122.78	118.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1	WV7	1	0

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	229/240 (95%)	-0.64	3 (1%)	77	75	8, 16, 38, 56	0
1	B	224/240 (93%)	-0.66	2 (0%)	84	83	7, 16, 35, 56	0
2	C	10/19 (52%)	0.09	0	100	100	24, 33, 38, 40	0
2	D	10/19 (52%)	-0.05	0	100	100	26, 36, 42, 47	0
All	All	473/518 (91%)	-0.62	5 (1%)	80	79	7, 16, 40, 56	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	484	VAL	4.0
1	B	412	PRO	3.4
1	A	484	VAL	3.4
1	A	409	SER	2.4
1	A	410	MET	2.2

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)

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
3	WV7	A	1	17/17	0.94	0.10	8,14,26,30	0
3	WV7	B	1	17/17	0.94	0.10	8,14,23,28	0

6.5 Other polymers ⓘ

There are no such residues in this entry.