Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 09:28 pm BST

PDB ID	:	2DI4
Title	:	Crystal structure of the FtsH protease domain
Authors	:	Suno, R.; Niwa, H.; Tsuchiya, D.; Zhang, X.; Yoshida, M.; Morikawa, K.
Deposited on	:	2006-03-28
Resolution	:	2.79 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

The following versions of software and data (see references ①) were used in the production of this report:

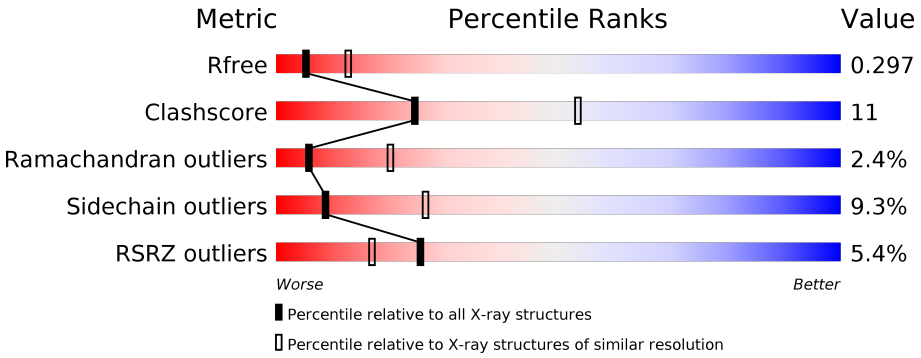
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.79 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	238	3% 53% 18% · 26%
1	B	238	5% 51% 19% · 26%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2645 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 Cell division protein ftsH homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	176	Total	C	N	O	S	0	0	0
			1318	835	219	258	6			
1	B	175	Total	C	N	O	S	0	0	0
			1310	835	218	251	6			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	397	PHE	-	CLONING ARTIFACT	UNP O67077
A	398	GLN	-	CLONING ARTIFACT	UNP O67077
A	399	GLY	-	CLONING ARTIFACT	UNP O67077
A	400	PRO	-	CLONING ARTIFACT	UNP O67077
A	401	LEU	-	CLONING ARTIFACT	UNP O67077
A	402	GLY	-	CLONING ARTIFACT	UNP O67077
A	403	SER	-	CLONING ARTIFACT	UNP O67077
A	404	HIS	-	CLONING ARTIFACT	UNP O67077
B	397	PHE	-	CLONING ARTIFACT	UNP O67077
B	398	GLN	-	CLONING ARTIFACT	UNP O67077
B	399	GLY	-	CLONING ARTIFACT	UNP O67077
B	400	PRO	-	CLONING ARTIFACT	UNP O67077
B	401	LEU	-	CLONING ARTIFACT	UNP O67077
B	402	GLY	-	CLONING ARTIFACT	UNP O67077
B	403	SER	-	CLONING ARTIFACT	UNP O67077
B	404	HIS	-	CLONING ARTIFACT	UNP O67077

- Molecule 2 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Hg	0	0
			1	1		
2	A	1	Total	Hg	0	0
			1	1		

- Molecule 3 is water.

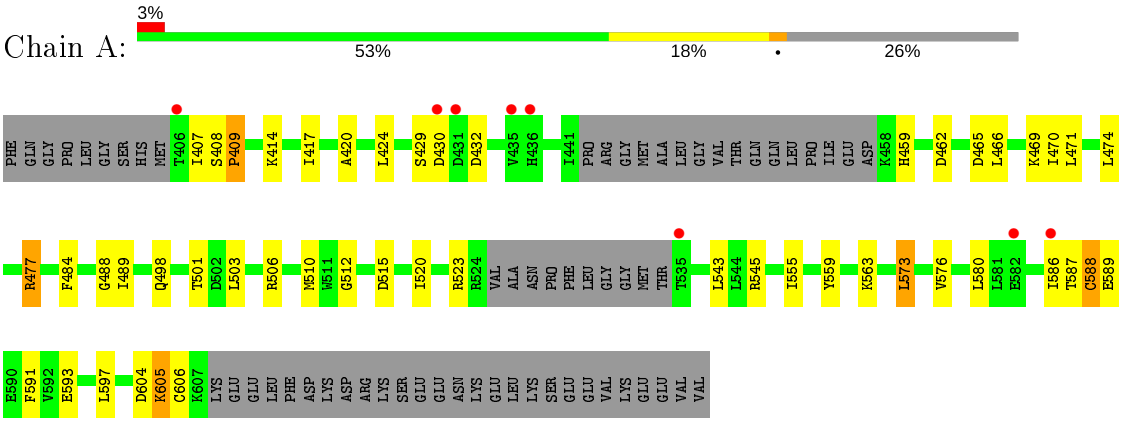
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	9	Total O 9 9	0	0
3	B	6	Total O 6 6	0	0

3 Residue-property plots (i)

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.

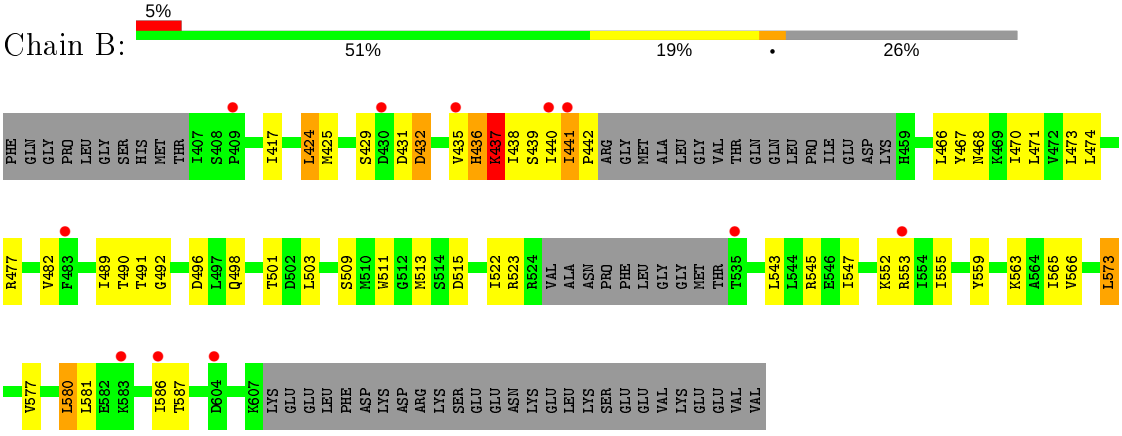
- Molecule 1: Cell division protein ftsH homolog

Chain A:



- Molecule 1: Cell division protein ftsH homolog

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	116.80Å 116.80Å 63.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.79 42.99 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.6 (20.00-2.79) 99.3 (42.99-2.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.80 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.251 , 0.299 0.251 , 0.297	Depositor DCC
R_{free} test set	882 reflections (7.11%)	wwPDB-VP
Wilson B-factor (Å^2)	44.8	Xtriage
Anisotropy	0.813	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 56.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.045 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2645	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 43.06 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.8630e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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

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	0.50	0/1335	0.62	0/1805
1	B	0.53	0/1328	0.72	3/1796 (0.2%)
All	All	0.51	0/2663	0.67	3/3601 (0.1%)

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	B	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	437	LYS	O-C-N	-12.42	102.83	122.70
1	B	437	LYS	CA-C-N	7.90	134.57	117.20
1	B	436	HIS	C-N-CA	5.99	136.67	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	437	LYS	Mainchain

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	1318	0	1250	25	0
1	B	1310	0	1257	34	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	9	0	0	0	0
3	B	6	0	0	0	0
All	All	2645	0	2507	56	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 11.

All (56) 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:522:ILE:HD11	1:B:547:ILE:CD1	2.07	0.84
1:B:490:THR:HG22	1:B:492:GLY:H	1.49	0.77
1:B:522:ILE:HD11	1:B:547:ILE:HD13	1.68	0.74
1:A:506:ARG:HD3	1:A:510:MET:SD	2.30	0.70
1:A:593:GLU:O	1:A:597:LEU:HG	1.97	0.65
1:B:439:SER:OG	1:B:442:PRO:HD2	1.97	0.65
1:B:559:TYR:CE2	1:B:563:LYS:HD2	2.35	0.61
1:A:420:ALA:HB1	1:A:573:LEU:HD21	1.83	0.61
1:A:477:ARG:CZ	1:A:489:ILE:HG23	2.31	0.60
1:B:522:ILE:HD11	1:B:547:ILE:HD11	1.83	0.59
1:A:501:THR:HA	1:A:555:ILE:HD11	1.84	0.58
1:A:420:ALA:HB1	1:A:573:LEU:CD2	2.33	0.58
1:B:438:ILE:HG23	1:B:586:ILE:HB	1.86	0.57
1:A:417:ILE:HD12	1:A:484:PHE:CZ	2.40	0.57
1:B:545:ARG:HG2	1:B:545:ARG:HH11	1.70	0.56
1:B:441:ILE:HB	1:B:442:PRO:CD	2.35	0.56
1:B:586:ILE:HG22	1:B:587:THR:O	2.05	0.56
1:A:559:TYR:CE2	1:A:563:LYS:HD2	2.41	0.55
1:B:501:THR:HA	1:B:555:ILE:HD11	1.89	0.55
1:B:441:ILE:HB	1:B:442:PRO:HD3	1.88	0.54
1:A:417:ILE:HD12	1:A:484:PHE:HZ	1.72	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:545:ARG:HG2	1:A:545:ARG:HH11	1.74	0.53
1:B:474:LEU:HD13	1:B:559:TYR:HB2	1.89	0.52
1:B:490:THR:HG22	1:B:491:THR:N	2.24	0.51
1:B:477:ARG:CZ	1:B:489:ILE:HG23	2.41	0.51
1:A:604:ASP:O	1:A:605:LYS:CB	2.59	0.50
1:B:573:LEU:O	1:B:577:VAL:HG23	2.11	0.50
1:B:466:LEU:O	1:B:470:ILE:HG12	2.11	0.50
1:A:414:LYS:HD3	1:A:484:PHE:CE1	2.47	0.50
1:B:435:VAL:C	1:B:436:HIS:CD2	2.85	0.50
1:B:417:ILE:HD13	1:B:577:VAL:HG21	1.94	0.50
1:B:424:LEU:HD13	1:B:565:ILE:HG22	1.95	0.49
1:A:466:LEU:O	1:A:470:ILE:HG12	2.13	0.48
1:B:432:ASP:HB2	1:B:468:ASN:HB3	1.95	0.47
1:B:425:MET:O	1:B:429:SER:HB2	2.14	0.47
1:A:588:CYS:SG	1:A:589:GLU:N	2.87	0.47
1:A:605:LYS:O	1:A:606:CYS:SG	2.73	0.46
1:A:429:SER:OG	1:A:432:ASP:HB3	2.16	0.46
1:B:440:ILE:HG22	1:B:440:ILE:O	2.15	0.46
1:A:520:ILE:HD11	1:B:552:LYS:HA	1.98	0.46
1:A:474:LEU:HD23	1:A:559:TYR:HA	1.99	0.45
1:B:473:LEU:O	1:B:496:ASP:HB3	2.16	0.45
1:B:440:ILE:HD11	1:B:581:LEU:HA	1.96	0.45
1:A:586:ILE:HG22	1:A:587:THR:O	2.15	0.45
1:A:576:VAL:HG11	1:A:591:PHE:CE1	2.52	0.44
1:B:509:SER:HB2	1:B:522:ILE:HD12	1.99	0.44
1:A:605:LYS:C	1:A:606:CYS:SG	2.96	0.43
1:A:462:ASP:HA	1:B:489:ILE:O	2.19	0.43
1:B:482:VAL:HG21	1:B:566:VAL:HB	2.01	0.43
1:B:431:ASP:O	1:B:432:ASP:C	2.56	0.43
1:B:440:ILE:HG13	1:B:580:LEU:HD22	2.01	0.42
1:A:512:GLY:H	1:B:477:ARG:HH12	1.67	0.42
1:B:490:THR:HG22	1:B:492:GLY:N	2.25	0.42
1:A:484:PHE:CB	1:A:488:GLY:HA3	2.50	0.42
1:A:408:SER:HA	1:A:409:PRO:HD2	1.86	0.41
1:B:511:TRP:O	1:B:513:MET:HG3	2.20	0.41

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	170/238 (71%)	154 (91%)	12 (7%)	4 (2%)	6 20
1	B	169/238 (71%)	150 (89%)	15 (9%)	4 (2%)	6 20
All	All	339/476 (71%)	304 (90%)	27 (8%)	8 (2%)	6 20

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	605	LYS
1	B	437	LYS
1	B	432	ASP
1	B	523	ARG
1	A	407	ILE
1	A	523	ARG
1	A	409	PRO
1	B	441	ILE

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	130/207 (63%)	116 (89%)	14 (11%)	6 19
1	B	129/207 (62%)	119 (92%)	10 (8%)	12 35
All	All	259/414 (63%)	235 (91%)	24 (9%)	9 26

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	424	LEU
1	A	430	ASP
1	A	459	HIS
1	A	465	ASP
1	A	469	LYS
1	A	471	LEU
1	A	477	ARG
1	A	498	GLN
1	A	503	LEU
1	A	515	ASP
1	A	543	LEU
1	A	573	LEU
1	A	580	LEU
1	A	588	CYS
1	B	424	LEU
1	B	467	TYR
1	B	471	LEU
1	B	498	GLN
1	B	503	LEU
1	B	515	ASP
1	B	543	LEU
1	B	553	ARG
1	B	573	LEU
1	B	580	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	436	HIS
1	B	436	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)

Of 2 ligands modelled in this entry, 2 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 (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

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	176/238 (73%)	0.39	8 (4%)	33	23	35, 47, 78, 80	0
1	B	175/238 (73%)	0.45	11 (6%)	20	12	33, 47, 73, 77	0
All	All	351/476 (73%)	0.42	19 (5%)	25	17	33, 47, 73, 80	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	431	ASP	4.7
1	A	435	VAL	4.2
1	A	406	THR	3.9
1	B	435	VAL	3.9
1	B	604	ASP	3.7
1	B	440	ILE	3.7
1	A	582	GLU	3.2
1	A	586	ILE	3.2
1	A	535	THR	3.0
1	B	586	ILE	2.9
1	A	430	ASP	2.8
1	B	535	THR	2.5
1	A	436	HIS	2.4
1	B	441	ILE	2.3
1	B	583	LYS	2.3
1	B	553	ARG	2.2
1	B	409	PRO	2.1
1	B	430	ASP	2.0
1	B	483	PHE	2.0

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 (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^{th} 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
2	HG	B	2001	1/1	0.97	0.08	92,92,92,92	0
2	HG	A	1001	1/1	0.98	0.25	25,25,25,25	1

6.5 Other polymers (i)

There are no such residues in this entry.