wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 04:33 am BST

PDB ID	:	1Z0T
Title	:	Crystal Structure of A. fulgidus Lon proteolytic domain
Authors	:	Dauter, Z.; Botos, I.; LaRonde-LeBlanc, N.; Wlodawer, A.
Deposited on	:	2005-03-02
Resolution	:	3.00 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 (1)) were used in the production of this report:

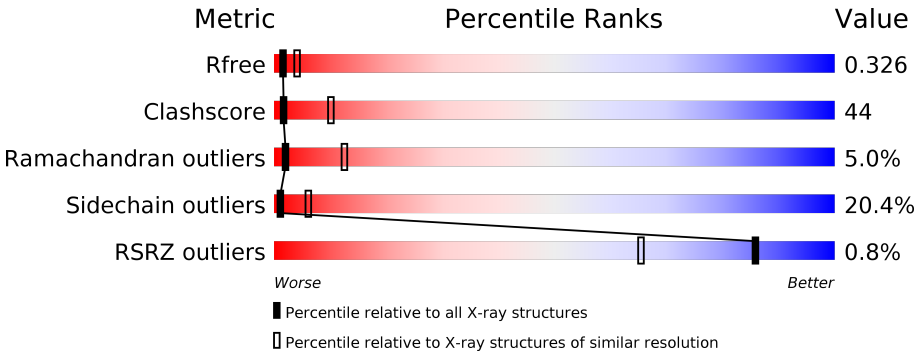
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	205	43% 34% 14% • 5%
1	B	205	2% 29% 40% 20% 7% •
1	C	205	42% 33% 15% 5% •
1	D	205	41% 35% 14% 6% •
1	E	205	% 25% 40% 21% 9% 5%
1	F	205	34% 40% 15% 6% •

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9870 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 Putative protease La homolog type.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	195	Total	C	N	O	S	0	0	0
			1457	921	247	284	5			
1	B	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			
1	C	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			
1	D	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			
1	E	195	Total	C	N	O	S	0	0	0
			1457	921	247	284	5			
1	F	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			

- Molecule 2 is water.

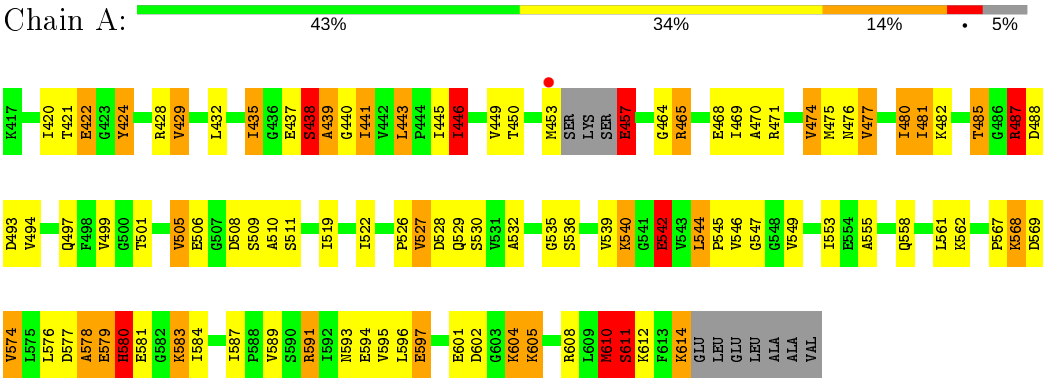
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	159	Total	O	0	0
			159	159		
2	B	164	Total	O	0	0
			164	164		
2	C	208	Total	O	0	0
			208	208		
2	D	190	Total	O	0	0
			190	190		
2	E	145	Total	O	0	0
			145	145		
2	F	226	Total	O	0	0
			226	226		

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.

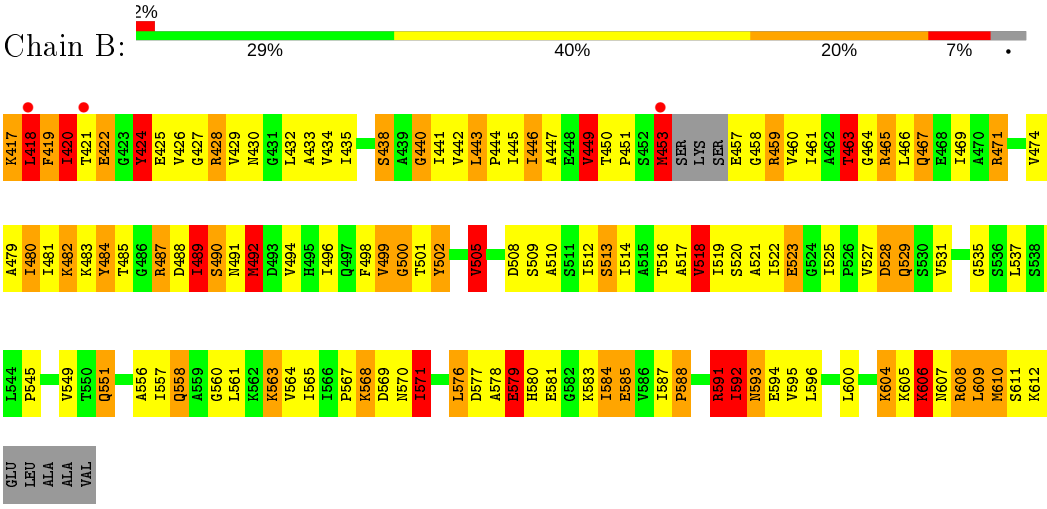
- Molecule 1: Putative protease La homolog type

Chain A:



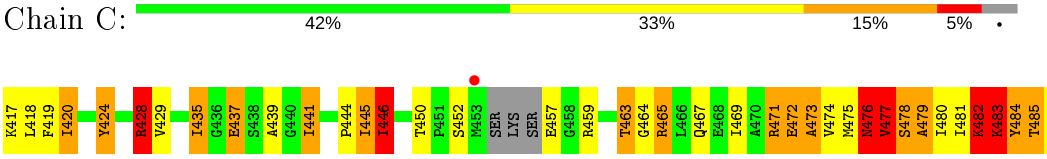
- Molecule 1: Putative protease La homolog type

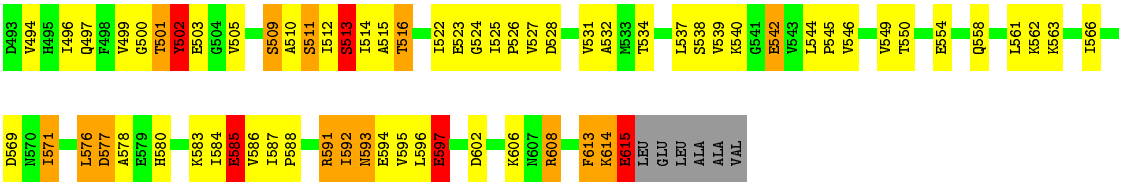
Chain B:



- Molecule 1: Putative protease La homolog type

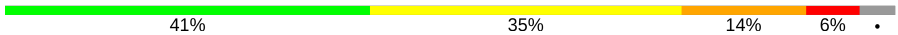
Chain C:

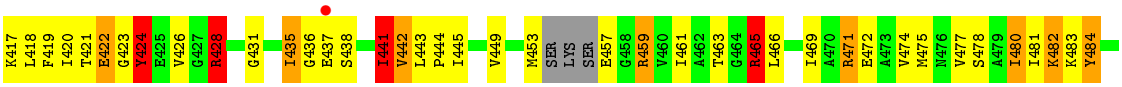


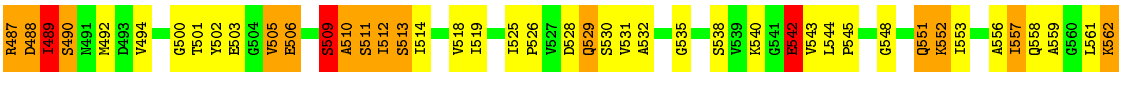


• Molecule 1: Putative protease La homolog type

Chain D:

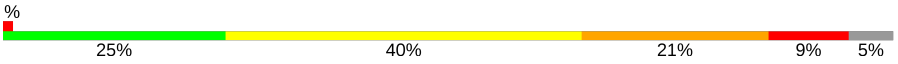


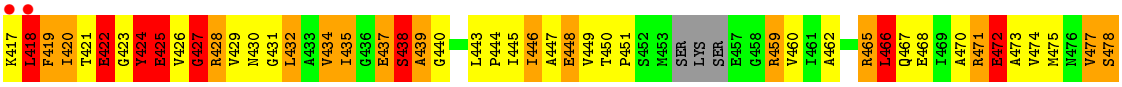


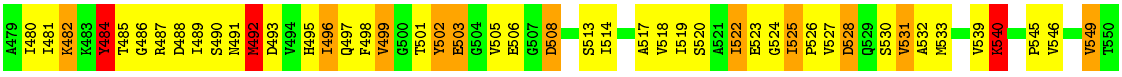


• Molecule 1: Putative protease La homolog type

Chain E:

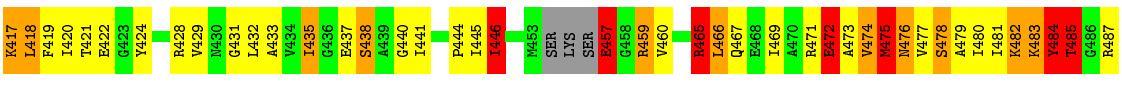


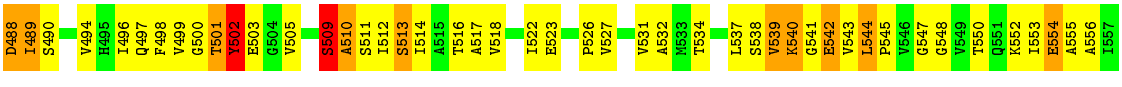


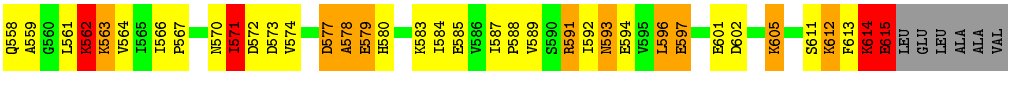


• Molecule 1: Putative protease La homolog type

Chain F:







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.25Å 90.55Å 147.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.00 14.98 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.00-3.00) 100.0 (14.98-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.94 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.200 , 0.330 0.200 , 0.326	Depositor DCC
R_{free} test set	1210 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	41.7	Xtriage
Anisotropy	0.218	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 48.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.55$, $< L^2 > = 0.39$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9870	wwPDB-VP
Average B, all atoms (Å^2)	14.0	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 54.90 % 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 3.4165e-05. 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

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	1.93	32/1470 (2.2%)	1.58	26/1983 (1.3%)
1	B	1.90	27/1479 (1.8%)	1.54	14/1995 (0.7%)
1	C	1.98	34/1479 (2.3%)	1.49	18/1995 (0.9%)
1	D	1.94	36/1479 (2.4%)	1.58	17/1995 (0.9%)
1	E	1.91	35/1470 (2.4%)	1.64	28/1983 (1.4%)
1	F	1.88	26/1479 (1.8%)	1.53	16/1995 (0.8%)
All	All	1.92	190/8856 (2.1%)	1.56	119/11946 (1.0%)

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	2
1	C	0	1
1	D	0	2
1	E	0	2
1	F	0	1
All	All	0	8

The worst 5 of 190 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	612	LYS	CE-NZ	13.53	1.82	1.49
1	C	597	GLU	CG-CD	11.91	1.69	1.51
1	F	484	TYR	CB-CG	10.78	1.67	1.51
1	A	597	GLU	CG-CD	10.61	1.67	1.51
1	C	472	GLU	CB-CG	10.32	1.71	1.52

The worst 5 of 119 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	471	ARG	NE-CZ-NH2	-16.20	112.20	120.30
1	D	587	ILE	C-N-CD	11.68	152.92	128.40
1	F	563	LYS	CD-CE-NZ	-11.20	85.94	111.70
1	A	471	ARG	NE-CZ-NH1	9.81	125.20	120.30
1	D	465	ARG	NE-CZ-NH2	-9.72	115.44	120.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	500	GLY	Peptide
1	B	606	LYS	Peptide
1	C	452	SER	Peptide
1	D	438	SER	Peptide
1	D	614	LYS	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	1457	0	1523	100	0
1	B	1466	0	1529	178	0
1	C	1466	0	1529	135	0
1	D	1466	0	1529	96	0
1	E	1457	0	1523	172	0
1	F	1466	0	1529	140	0
2	A	159	0	0	13	0
2	B	164	0	0	33	0
2	C	208	0	0	36	0
2	D	190	0	0	17	0
2	E	145	0	0	31	0
2	F	226	0	0	35	0
All	All	9870	0	9162	795	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 44.

The worst 5 of 795 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:C:435:ILE:CG1	1:C:435:ILE:CD1	1.76	1.61
1:C:420:ILE:CB	1:C:420:ILE:CG2	1.78	1.60
1:F:522:ILE:CG1	1:F:522:ILE:CD1	1.79	1.58
1:B:463:THR:CG2	1:B:463:THR:CB	1.74	1.54
1:D:612:LYS:CE	1:D:612:LYS:NZ	1.70	1.54

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	191/205 (93%)	175 (92%)	13 (7%)	3 (2%)	9 40
1	B	192/205 (94%)	146 (76%)	29 (15%)	17 (9%)	1 3
1	C	192/205 (94%)	163 (85%)	22 (12%)	7 (4%)	3 19
1	D	192/205 (94%)	163 (85%)	21 (11%)	8 (4%)	3 16
1	E	191/205 (93%)	149 (78%)	26 (14%)	16 (8%)	1 4
1	F	192/205 (94%)	165 (86%)	21 (11%)	6 (3%)	4 23
All	All	1150/1230 (94%)	961 (84%)	132 (12%)	57 (5%)	2 12

5 of 57 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	439	ALA
1	B	438	SER
1	B	440	GLY
1	B	528	ASP
1	B	529	GLN

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	160/168 (95%)	130 (81%)	30 (19%)	1 8
1	B	161/168 (96%)	121 (75%)	40 (25%)	0 3
1	C	161/168 (96%)	129 (80%)	32 (20%)	1 7
1	D	161/168 (96%)	135 (84%)	26 (16%)	2 12
1	E	160/168 (95%)	122 (76%)	38 (24%)	0 3
1	F	161/168 (96%)	130 (81%)	31 (19%)	1 8
All	All	964/1008 (96%)	767 (80%)	197 (20%)	1 6

5 of 197 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	539	VAL
1	D	480	ILE
1	F	501	THR
1	C	542	GLU
1	C	614	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	551	GLN
1	E	607	ASN
1	E	497	GLN
1	C	476	ASN
1	D	558	GLN

5.3.3 RNA (i)

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 carbohydrates 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	195/205 (95%)	-0.73	1 (0%) 91 75	2, 6, 29, 44	0
1	B	196/205 (95%)	-0.36	4 (2%) 65 36	2, 15, 49, 67	0
1	C	196/205 (95%)	-0.72	1 (0%) 91 75	2, 5, 28, 48	0
1	D	196/205 (95%)	-0.73	1 (0%) 91 75	2, 5, 25, 42	0
1	E	195/205 (95%)	-0.40	2 (1%) 82 59	2, 13, 44, 61	0
1	F	196/205 (95%)	-0.67	0 100 100	2, 7, 34, 68	0
All	All	1174/1230 (95%)	-0.60	9 (0%) 86 65	2, 9, 38, 68	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	453	MET	3.8
1	B	615	GLU	3.1
1	E	417	LYS	3.0
1	C	453	MET	2.3
1	B	421	THR	2.2

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 carbohydrates 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.