Full wwPDB X-ray Structure Validation Report (i)

Aug 27, 2023 – 02:54 AM EDT

PDB ID	:	3FG7
Title	:	The crystal structure of villin domain 6
Authors	:	Wang, H.; Burtnick, L.D.; Robinson, R.C.
Deposited on	:	2008-12-05
Resolution	:	2.00 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

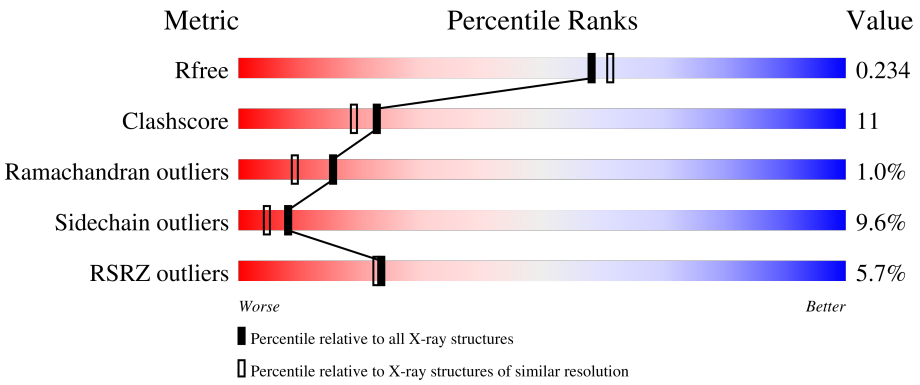
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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	398	2% 19% 6% • 73%
1	B	398	% 22% • • 73%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1910 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 Villin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	106	875	569	141	164	1	0	0	0
1	B	106	875	569	141	164	1	0	0	0

There are 74 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	337	MET	-	expression tag	UNP P09327
A	338	ALA	-	expression tag	UNP P09327
A	339	GLU	-	expression tag	UNP P09327
A	340	GLU	-	expression tag	UNP P09327
A	341	HIS	-	expression tag	UNP P09327
A	342	HIS	-	expression tag	UNP P09327
A	343	HIS	-	expression tag	UNP P09327
A	344	HIS	-	expression tag	UNP P09327
A	345	HIS	-	expression tag	UNP P09327
A	346	HIS	-	expression tag	UNP P09327
A	347	HIS	-	expression tag	UNP P09327
A	348	HIS	-	expression tag	UNP P09327
A	349	LEU	-	expression tag	UNP P09327
A	350	GLU	-	expression tag	UNP P09327
A	351	VAL	-	expression tag	UNP P09327
A	352	LEU	-	expression tag	UNP P09327
A	353	PHE	-	expression tag	UNP P09327
A	354	GLN	-	expression tag	UNP P09327
A	355	GLY	-	expression tag	UNP P09327
A	356	PRO	-	expression tag	UNP P09327
A	357	GLY	-	expression tag	UNP P09327
A	358	ARG	-	expression tag	UNP P09327
A	359	PRO	-	expression tag	UNP P09327
A	721	GLY	-	expression tag	UNP P09327
A	722	ILE	-	expression tag	UNP P09327

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	723	HIS	-	expression tag	UNP P09327
A	724	VAL	-	expression tag	UNP P09327
A	725	VAL	-	expression tag	UNP P09327
A	726	PRO	-	expression tag	UNP P09327
A	727	ASN	-	expression tag	UNP P09327
A	728	LEU	-	expression tag	UNP P09327
A	729	SER	-	expression tag	UNP P09327
A	730	PRO	-	expression tag	UNP P09327
A	731	LEU	-	expression tag	UNP P09327
A	732	SER	-	expression tag	UNP P09327
A	733	ASN	-	expression tag	UNP P09327
A	734	ASN	-	expression tag	UNP P09327
B	337	MET	-	expression tag	UNP P09327
B	338	ALA	-	expression tag	UNP P09327
B	339	GLU	-	expression tag	UNP P09327
B	340	GLU	-	expression tag	UNP P09327
B	341	HIS	-	expression tag	UNP P09327
B	342	HIS	-	expression tag	UNP P09327
B	343	HIS	-	expression tag	UNP P09327
B	344	HIS	-	expression tag	UNP P09327
B	345	HIS	-	expression tag	UNP P09327
B	346	HIS	-	expression tag	UNP P09327
B	347	HIS	-	expression tag	UNP P09327
B	348	HIS	-	expression tag	UNP P09327
B	349	LEU	-	expression tag	UNP P09327
B	350	GLU	-	expression tag	UNP P09327
B	351	VAL	-	expression tag	UNP P09327
B	352	LEU	-	expression tag	UNP P09327
B	353	PHE	-	expression tag	UNP P09327
B	354	GLN	-	expression tag	UNP P09327
B	355	GLY	-	expression tag	UNP P09327
B	356	PRO	-	expression tag	UNP P09327
B	357	GLY	-	expression tag	UNP P09327
B	358	ARG	-	expression tag	UNP P09327
B	359	PRO	-	expression tag	UNP P09327
B	721	GLY	-	expression tag	UNP P09327
B	722	ILE	-	expression tag	UNP P09327
B	723	HIS	-	expression tag	UNP P09327
B	724	VAL	-	expression tag	UNP P09327
B	725	VAL	-	expression tag	UNP P09327
B	726	PRO	-	expression tag	UNP P09327
B	727	ASN	-	expression tag	UNP P09327

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	728	LEU	-	expression tag	UNP P09327
B	729	SER	-	expression tag	UNP P09327
B	730	PRO	-	expression tag	UNP P09327
B	731	LEU	-	expression tag	UNP P09327
B	732	SER	-	expression tag	UNP P09327
B	733	ASN	-	expression tag	UNP P09327
B	734	ASN	-	expression tag	UNP P09327

- Molecule 2 is water.

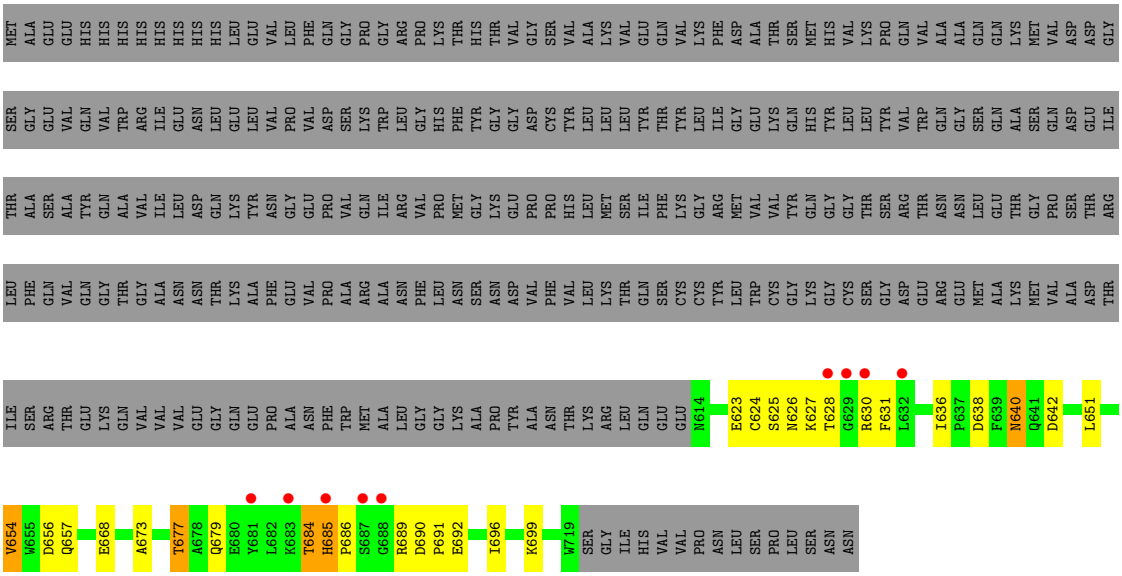
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	78	78	78	0	0
			Total	O		
2	B	82	82	82	0	0

3 Residue-property plots (i)

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.

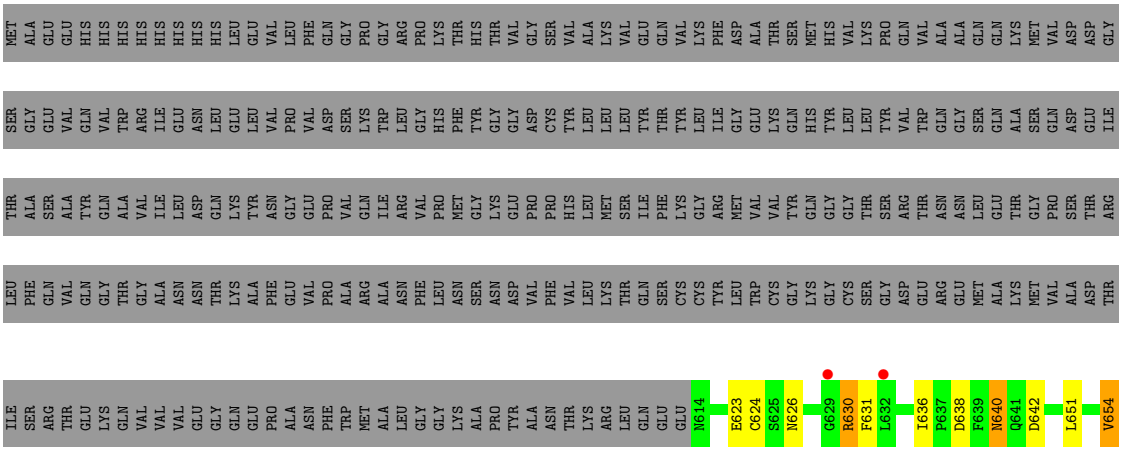
- Molecule 1: Villin-1

Chain A:



- Molecule 1: Villin-1

Chain B:



W655
D656
Q657
A673
T677
Y681
T684
H685
P686
R689
E692
I696
W719
SER
GLY
ILE
HIS
VAL
VAL
PRO
ASN
LEU
SER
PRO
LEU
SER
ASN
ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	47.82Å 47.82Å 98.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	23.91 – 2.00 23.91 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (23.91-2.00) 99.7 (23.91-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.92 (at 2.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.197 , 0.246 0.191 , 0.234	Depositor DCC
R_{free} test set	811 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	34.7	Xtriage
Anisotropy	0.104	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 51.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.480 for -h,-k,l 0.076 for h,-h-k,-l 0.068 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1910	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.87% 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)

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	0.64	0/905	0.77	1/1236 (0.1%)
1	B	0.63	0/905	0.76	1/1236 (0.1%)
All	All	0.63	0/1810	0.77	2/2472 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	654	VAL	CG1-CB-CG2	5.04	118.97	110.90
1	B	654	VAL	CG1-CB-CG2	5.03	118.94	110.90

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	875	0	823	26	0
1	B	875	0	823	13	0
2	A	78	0	0	5	0
2	B	82	0	0	0	0
All	All	1910	0	1646	39	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 (39) 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:631:PHE:HE2	1:B:677:THR:HG21	1.19	1.05
1:A:631:PHE:HE2	1:A:677:THR:HG21	1.24	1.02
1:B:631:PHE:CE2	1:B:677:THR:HG21	2.10	0.84
1:A:679:GLN:NE2	2:A:162:HOH:O	2.15	0.79
1:B:631:PHE:HE2	1:B:677:THR:CG2	1.96	0.78
1:A:631:PHE:CE2	1:A:677:THR:HG21	2.16	0.74
1:B:624:CYS:HB3	1:B:631:PHE:CZ	2.28	0.69
1:A:684:THR:HG23	1:A:685:HIS:CD2	2.29	0.68
1:A:624:CYS:HB3	1:A:631:PHE:CZ	2.29	0.68
1:B:673:ALA:O	1:B:677:THR:HG23	1.95	0.67
1:A:673:ALA:O	1:A:677:THR:HG23	1.96	0.65
1:A:624:CYS:HB3	1:A:631:PHE:HZ	1.59	0.65
1:B:624:CYS:HB3	1:B:631:PHE:HZ	1.61	0.65
1:A:631:PHE:HE2	1:A:677:THR:CG2	2.04	0.64
1:A:685:HIS:CD2	1:A:685:HIS:N	2.73	0.57
1:A:699:LYS:NZ	2:A:101:HOH:O	2.38	0.56
1:B:623:GLU:HB2	1:B:636:ILE:HD11	1.88	0.56
1:B:631:PHE:CE2	1:B:677:THR:CG2	2.80	0.56
1:A:685:HIS:CD2	1:A:685:HIS:H	2.24	0.55
1:A:685:HIS:H	1:A:685:HIS:HD2	1.54	0.55
1:A:623:GLU:HB2	1:A:636:ILE:HD11	1.90	0.53
1:A:631:PHE:CE2	1:A:677:THR:CG2	2.87	0.52
1:B:686:PRO:HD2	1:B:689:ARG:HD2	1.95	0.49
1:A:657:GLN:NE2	1:A:696:ILE:HD11	2.30	0.47
1:A:684:THR:HG23	1:A:685:HIS:HD2	1.77	0.47
1:A:699:LYS:NZ	2:A:129:HOH:O	2.47	0.46
1:A:685:HIS:N	2:A:86:HOH:O	2.49	0.45
1:B:657:GLN:NE2	1:B:696:ILE:HD11	2.32	0.45
1:B:692:GLU:O	1:B:692:GLU:HG3	2.18	0.45
1:A:685:HIS:C	2:A:86:HOH:O	2.55	0.44
1:A:657:GLN:NE2	1:A:696:ILE:CD1	2.81	0.43
1:A:625:SER:OG	1:A:627:LYS:HG2	2.19	0.43
1:A:686:PRO:HD2	1:A:689:ARG:HD2	2.01	0.42
1:A:692:GLU:O	1:A:692:GLU:HG3	2.20	0.42
1:A:627:LYS:HG3	1:A:628:THR:N	2.34	0.41
1:B:630:ARG:H	1:B:630:ARG:HG3	1.73	0.41
1:B:640:ASN:HD22	1:B:642:ASP:H	1.67	0.41
1:A:690:ASP:HA	1:A:691:PRO:HD3	1.95	0.41
1:A:640:ASN:HD22	1:A:642:ASP:H	1.68	0.40

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	104/398 (26%)	97 (93%)	6 (6%)	1 (1%)	15 9
1	B	104/398 (26%)	98 (94%)	5 (5%)	1 (1%)	15 9
All	All	208/796 (26%)	195 (94%)	11 (5%)	2 (1%)	15 9

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	638	ASP
1	A	638	ASP

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	94/342 (28%)	84 (89%)	10 (11%)	6 3
1	B	94/342 (28%)	86 (92%)	8 (8%)	10 6
All	All	188/684 (28%)	170 (90%)	18 (10%)	8 5

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	626	ASN
1	A	630	ARG
1	A	640	ASN
1	A	651	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	654	VAL
1	A	656	ASP
1	A	668	GLU
1	A	677	THR
1	A	684	THR
1	A	685	HIS
1	B	626	ASN
1	B	630	ARG
1	B	640	ASN
1	B	651	LEU
1	B	654	VAL
1	B	656	ASP
1	B	677	THR
1	B	684	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	626	ASN
1	A	640	ASN
1	A	657	GLN
1	A	685	HIS
1	B	626	ASN
1	B	640	ASN
1	B	657	GLN

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 monosaccharides 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	106/398 (26%)	0.46	9 (8%)	10	10	21, 34, 74, 76	0
1	B	106/398 (26%)	0.24	3 (2%)	53	51	21, 34, 74, 76	0
All	All	212/796 (26%)	0.35	12 (5%)	23	23	21, 34, 74, 76	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	687	SER	4.1
1	B	629	GLY	3.3
1	A	632	LEU	3.3
1	A	630	ARG	3.1
1	B	632	LEU	3.0
1	B	681	TYR	2.6
1	A	681	TYR	2.6
1	A	629	GLY	2.6
1	A	628	THR	2.5
1	A	683	LYS	2.4
1	A	685	HIS	2.2
1	A	688	GLY	2.0

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 monosaccharides 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.