Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 05:28 PM GMT

PDB ID	:	4IHE
Title	:	Crystal Structure of Uncleaved ThnT T282A
Authors	:	Buller, A.R.; Schildbach, J.F.; Townsend, C.A.
Deposited on	:	2012-12-18
Resolution	:	1.50 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
http://wwpdb.org/validation/2016/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:

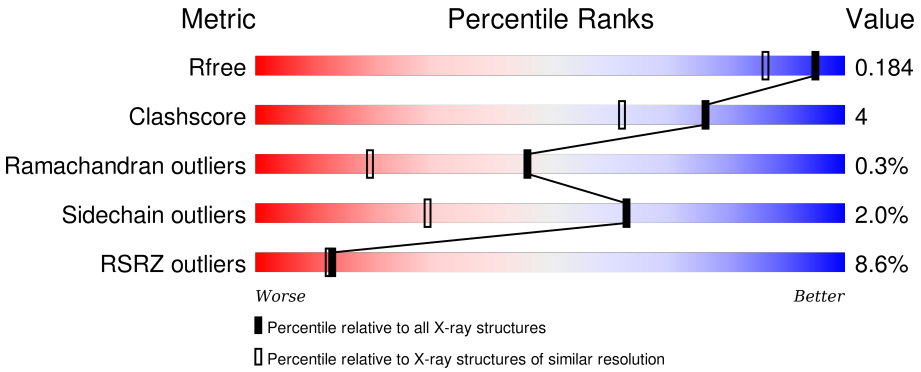
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.50 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}	91344	2072 (1.50-1.50)
Clashscore	102246	2274 (1.50-1.50)
Ramachandran outliers	100387	2218 (1.50-1.50)
Sidechain outliers	100360	2216 (1.50-1.50)
RSRZ outliers	91569	2075 (1.50-1.50)

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. 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	419	8% 79% 5% 15%
1	B	419	6% 79% 6% • 14%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5802 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 ThnT protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	358	Total	C	N	O	0	18	0
			2574	1593	485	496			
1	B	360	Total	C	N	O	0	17	0
			2584	1601	487	496			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q83XN4
A	-18	GLY	-	EXPRESSION TAG	UNP Q83XN4
A	-17	SER	-	EXPRESSION TAG	UNP Q83XN4
A	-16	SER	-	EXPRESSION TAG	UNP Q83XN4
A	-15	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	-14	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	-13	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	-12	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	-11	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	-10	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	-9	SER	-	EXPRESSION TAG	UNP Q83XN4
A	-8	SER	-	EXPRESSION TAG	UNP Q83XN4
A	-7	GLY	-	EXPRESSION TAG	UNP Q83XN4
A	-6	LEU	-	EXPRESSION TAG	UNP Q83XN4
A	-5	VAL	-	EXPRESSION TAG	UNP Q83XN4
A	-4	PRO	-	EXPRESSION TAG	UNP Q83XN4
A	-3	ARG	-	EXPRESSION TAG	UNP Q83XN4
A	-2	GLY	-	EXPRESSION TAG	UNP Q83XN4
A	-1	SER	-	EXPRESSION TAG	UNP Q83XN4
A	0	HIS	-	EXPRESSION TAG	UNP Q83XN4
A	282	ALA	THR	ENGINEERED MUTATION	UNP Q83XN4
B	-19	MET	-	EXPRESSION TAG	UNP Q83XN4
B	-18	GLY	-	EXPRESSION TAG	UNP Q83XN4
B	-17	SER	-	EXPRESSION TAG	UNP Q83XN4
B	-16	SER	-	EXPRESSION TAG	UNP Q83XN4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	-14	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	-13	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	-12	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	-11	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	-10	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	-9	SER	-	EXPRESSION TAG	UNP Q83XN4
B	-8	SER	-	EXPRESSION TAG	UNP Q83XN4
B	-7	GLY	-	EXPRESSION TAG	UNP Q83XN4
B	-6	LEU	-	EXPRESSION TAG	UNP Q83XN4
B	-5	VAL	-	EXPRESSION TAG	UNP Q83XN4
B	-4	PRO	-	EXPRESSION TAG	UNP Q83XN4
B	-3	ARG	-	EXPRESSION TAG	UNP Q83XN4
B	-2	GLY	-	EXPRESSION TAG	UNP Q83XN4
B	-1	SER	-	EXPRESSION TAG	UNP Q83XN4
B	0	HIS	-	EXPRESSION TAG	UNP Q83XN4
B	282	ALA	THR	ENGINEERED MUTATION	UNP Q83XN4

- Molecule 2 is water.

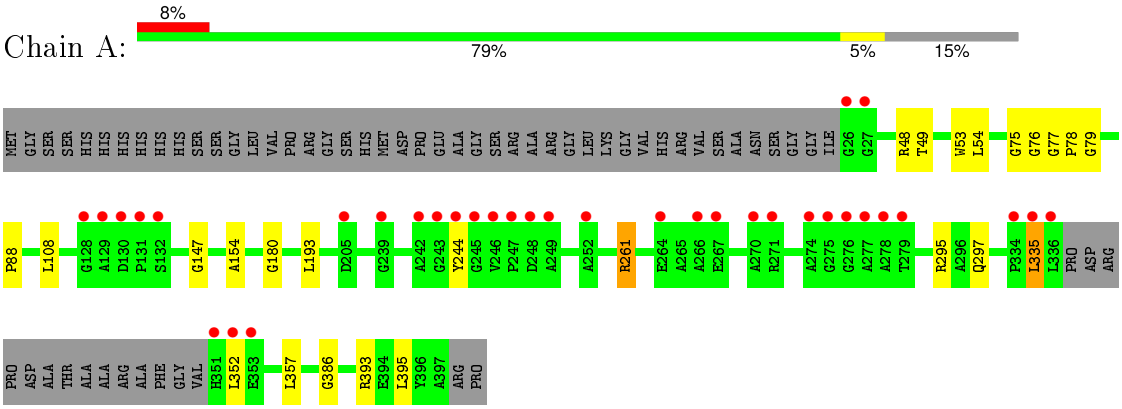
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	287	Total O 289 289	0	7
2	B	351	Total O 355 355	0	11

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 errors 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.

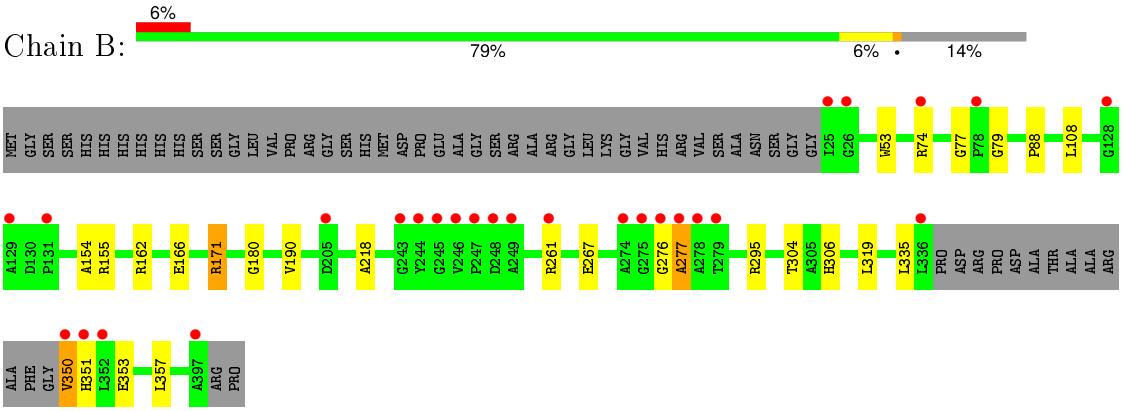
- Molecule 1: ThnT protein

Chain A:



- Molecule 1: ThnT protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	140.75Å 67.96Å 73.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.59 – 1.50 35.56 – 1.50	Depositor EDS
% Data completeness (in resolution range)	95.6 (33.59-1.50) 95.7 (35.56-1.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.165 , 0.180 0.171 , 0.184	Depositor DCC
R_{free} test set	5672 reflections (5.51%)	DCC
Wilson B-factor (Å^2)	18.1	Xtriage
Anisotropy	0.215	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 45.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 108456 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5802	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.28	0/2641	0.52	0/3612
1	B	0.30	0/2648	0.55	0/3622
All	All	0.29	0/5289	0.54	0/7234

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	2
1	B	0	2
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	77[B]	GLY	Mainchain,Peptide
1	B	77[B]	GLY	Mainchain,Peptide

5.2 Too-close contacts

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	2574	0	2578	21	0
1	B	2584	0	2594	19	0
2	A	289	0	0	7	0
2	B	355	0	0	4	0
All	All	5802	0	5172	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 4.

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:A:75[A]:GLY:HA2	2:A:450:HOH:O	1.56	1.02
1:B:304[B]:THR:HG23	2:B:526:HOH:O	1.69	0.90
1:B:171:ARG:HE	1:B:171:ARG:HA	1.42	0.83
1:B:304[B]:THR:CG2	2:B:526:HOH:O	2.31	0.76
1:A:49[B]:THR:HG21	2:A:500:HOH:O	1.91	0.69
1:A:76[A]:GLY:N	2:A:450:HOH:O	2.34	0.60
1:B:350:VAL:CG1	1:B:353:GLU:H	2.14	0.60
1:A:88:PRO:HA	1:A:295[B]:ARG:CD	2.32	0.58
1:B:180:GLY:HA2	2:B:622:HOH:O	2.02	0.58
1:B:79[B]:GLY:HA2	1:B:108:LEU:HD11	1.85	0.58
1:B:88:PRO:HA	1:B:295[B]:ARG:CD	2.34	0.57
1:B:350:VAL:HG13	1:B:353:GLU:H	1.71	0.54
1:A:357[B]:LEU:HD13	2:A:604:HOH:O	2.09	0.53
1:A:335:LEU:HD21	1:A:357[B]:LEU:CD2	2.40	0.52
1:A:244:TYR:HB3	1:A:393:ARG:HD3	1.91	0.52
1:B:88:PRO:HG3	1:B:295[B]:ARG:HG2	1.91	0.52
1:B:88:PRO:HA	1:B:295[B]:ARG:HD2	1.92	0.52
1:A:79[B]:GLY:HA2	1:A:108:LEU:HD11	1.90	0.52
1:A:75[A]:GLY:CA	2:A:450:HOH:O	2.34	0.51
1:B:276:GLY:O	1:B:277:ALA:HB3	2.11	0.50
1:A:180:GLY:HA2	2:A:465:HOH:O	2.11	0.50
1:A:53:TRP:CD2	1:A:154:ALA:HB1	2.47	0.50
1:B:335:LEU:HD21	1:B:357[B]:LEU:CD2	2.42	0.49
1:B:53:TRP:CD2	1:B:154:ALA:HB1	2.49	0.47
1:B:74[B]:ARG:HG3	1:B:306:HIS:CE1	2.50	0.47
1:B:335:LEU:HD21	1:B:357[B]:LEU:HG	1.97	0.46
1:A:193:LEU:HD21	1:A:395:LEU:HD12	1.99	0.45
1:B:267:GLU:HG3	2:B:646:HOH:O	2.17	0.45
1:A:335:LEU:HD21	1:A:357[B]:LEU:HD21	1.99	0.44
1:A:88:PRO:HG3	1:A:295[B]:ARG:HG2	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:261:ARG:CG	1:A:261:ARG:HH21	2.32	0.43
1:A:357[A]:LEU:HD23	1:B:319:LEU:HD11	2.00	0.42
1:B:162[B]:ARG:NE	1:B:166:GLU:OE2	2.52	0.42
1:A:357[B]:LEU:CD1	2:A:604:HOH:O	2.67	0.42
1:A:147:GLY:O	1:A:386:GLY:HA3	2.19	0.42
1:B:190:VAL:HG23	1:B:218:ALA:HB3	2.02	0.41
1:A:297:GLN:NE2	1:A:357[B]:LEU:HD11	2.35	0.41
1:A:49[B]:THR:HG22	1:A:54:LEU:HD12	2.02	0.41
1:A:88:PRO:HA	1:A:295[B]:ARG:HD2	2.01	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	372/419 (89%)	365 (98%)	5 (1%)	2 (0%)	34	10
1	B	373/419 (89%)	364 (98%)	8 (2%)	1 (0%)	46	19
All	All	745/838 (89%)	729 (98%)	13 (2%)	3 (0%)	46	14

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	277	ALA
1	A	78[A]	PRO
1	A	78[B]	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	241/272 (89%)	237 (98%)	4 (2%)	68	37
1	B	242/272 (89%)	237 (98%)	5 (2%)	61	27
All	All	483/544 (89%)	474 (98%)	9 (2%)	63	31

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	ARG
1	A	261	ARG
1	A	335	LEU
1	A	352	LEU
1	B	155	ARG
1	B	171	ARG
1	B	261	ARG
1	B	350	VAL
1	B	351	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 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	358/419 (85%)	0.56	35 (9%)	10	9	12, 19, 43, 70	0
1	B	360/419 (85%)	0.44	27 (7%)	17	17	10, 16, 36, 58	0
All	All	718/838 (85%)	0.50	62 (8%)	13	12	10, 17, 42, 70	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	274	ALA	10.5
1	B	129	ALA	9.2
1	B	128	GLY	9.1
1	B	274	ALA	8.0
1	B	277	ALA	8.0
1	A	336	LEU	7.2
1	A	352	LEU	7.1
1	A	248	ASP	6.9
1	A	247	PRO	6.6
1	A	275	GLY	6.5
1	A	277	ALA	6.4
1	B	247	PRO	6.3
1	A	276	GLY	6.0
1	B	248	ASP	5.9
1	A	129	ALA	5.8
1	A	249	ALA	5.8
1	A	128	GLY	5.8
1	B	25	ILE	5.5
1	B	243	GLY	5.2
1	A	351	HIS	5.1
1	B	350	VAL	5.1
1	A	242	ALA	5.1
1	B	278	ALA	4.9
1	A	243	GLY	4.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	246	VAL	4.8
1	B	276	GLY	4.5
1	A	278	ALA	4.4
1	B	131	PRO	4.3
1	B	352	LEU	4.3
1	B	279	THR	4.1
1	A	131	PRO	4.1
1	B	351	HIS	3.8
1	A	246	VAL	3.6
1	A	244	TYR	3.6
1	A	245	GLY	3.5
1	B	397	ALA	3.4
1	A	267	GLU	3.4
1	B	275	GLY	3.4
1	A	334	PRO	3.4
1	A	335	LEU	3.3
1	A	271	ARG	3.3
1	B	245	GLY	3.2
1	B	261	ARG	3.2
1	A	279	THR	2.9
1	B	244	TYR	2.9
1	A	270	ALA	2.7
1	A	26	GLY	2.7
1	B	205	ASP	2.6
1	A	264	GLU	2.5
1	A	266	ALA	2.5
1	B	249	ALA	2.5
1	B	336	LEU	2.4
1	B	26	GLY	2.4
1	B	78[A]	PRO	2.3
1	A	205	ASP	2.3
1	A	132	SER	2.3
1	A	252	ALA	2.2
1	B	74[A]	ARG	2.1
1	A	130	ASP	2.1
1	A	353	GLU	2.1
1	A	27	GLY	2.0
1	A	239	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 carbohydrates 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.