Full wwPDB X-ray Structure Validation Report ⓘ

Nov 16, 2024 – 08:21 AM EST

PDB ID	:	3G2M
Title	:	Crystal Structure of the Glycopeptide N-methyltransferase MtfA
Authors	:	Shi, R.; Matte, A.; Cygler, M.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)
Deposited on	:	2009-01-31
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 ⓘ 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 ⓘ) were used in the production of this report:

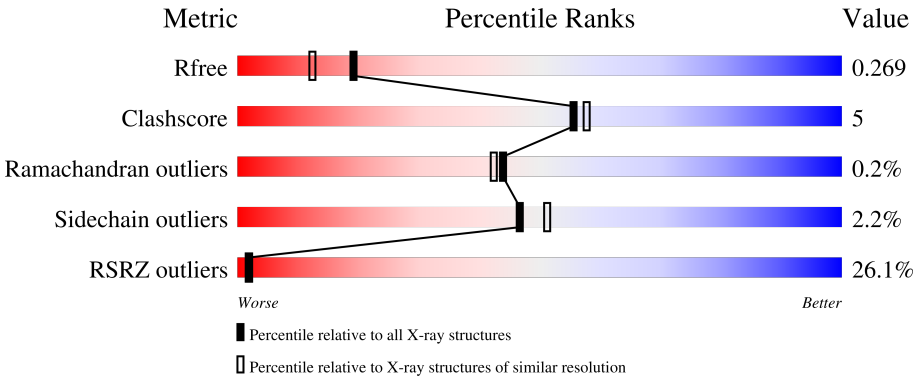
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (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	299	19% 70% 7% 22%
1	B	299	19% 66% 5% • 28%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3575 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 PCZA361.24.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	232	Total	C	N	O	S	Se	0	2	0
			1777	1119	315	335	3	5			
1	B	214	Total	C	N	O	S	Se	0	2	0
			1655	1034	299	316	1	5			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MSE	-	expression tag	UNP O52805
A	-17	GLY	-	expression tag	UNP O52805
A	-16	SER	-	expression tag	UNP O52805
A	-15	SER	-	expression tag	UNP O52805
A	-14	HIS	-	expression tag	UNP O52805
A	-13	HIS	-	expression tag	UNP O52805
A	-12	HIS	-	expression tag	UNP O52805
A	-11	HIS	-	expression tag	UNP O52805
A	-10	HIS	-	expression tag	UNP O52805
A	-9	HIS	-	expression tag	UNP O52805
A	-8	SER	-	expression tag	UNP O52805
A	-7	GLY	-	expression tag	UNP O52805
A	-6	GLY	-	expression tag	UNP O52805
A	-5	LEU	-	expression tag	UNP O52805
A	-4	VAL	-	expression tag	UNP O52805
A	-3	PRO	-	expression tag	UNP O52805
A	-2	ARG	-	expression tag	UNP O52805
A	-1	GLY	-	expression tag	UNP O52805
A	0	SER	-	expression tag	UNP O52805
B	-18	MSE	-	expression tag	UNP O52805
B	-17	GLY	-	expression tag	UNP O52805
B	-16	SER	-	expression tag	UNP O52805
B	-15	SER	-	expression tag	UNP O52805
B	-14	HIS	-	expression tag	UNP O52805
B	-13	HIS	-	expression tag	UNP O52805

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	HIS	-	expression tag	UNP O52805
B	-11	HIS	-	expression tag	UNP O52805
B	-10	HIS	-	expression tag	UNP O52805
B	-9	HIS	-	expression tag	UNP O52805
B	-8	SER	-	expression tag	UNP O52805
B	-7	GLY	-	expression tag	UNP O52805
B	-6	GLY	-	expression tag	UNP O52805
B	-5	LEU	-	expression tag	UNP O52805
B	-4	VAL	-	expression tag	UNP O52805
B	-3	PRO	-	expression tag	UNP O52805
B	-2	ARG	-	expression tag	UNP O52805
B	-1	GLY	-	expression tag	UNP O52805
B	0	SER	-	expression tag	UNP O52805

- Molecule 2 is water.

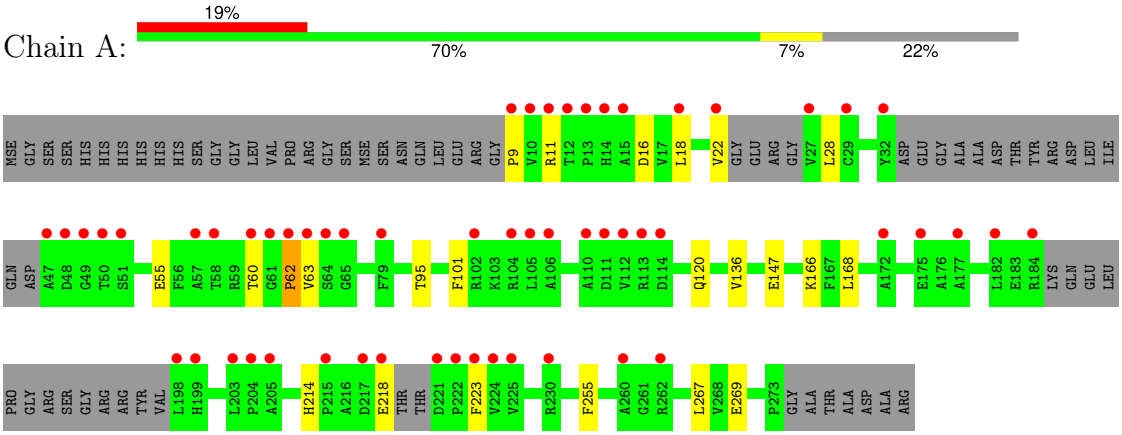
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	24	Total O 24 24	0	0
2	B	119	Total O 119 119	0	0

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.

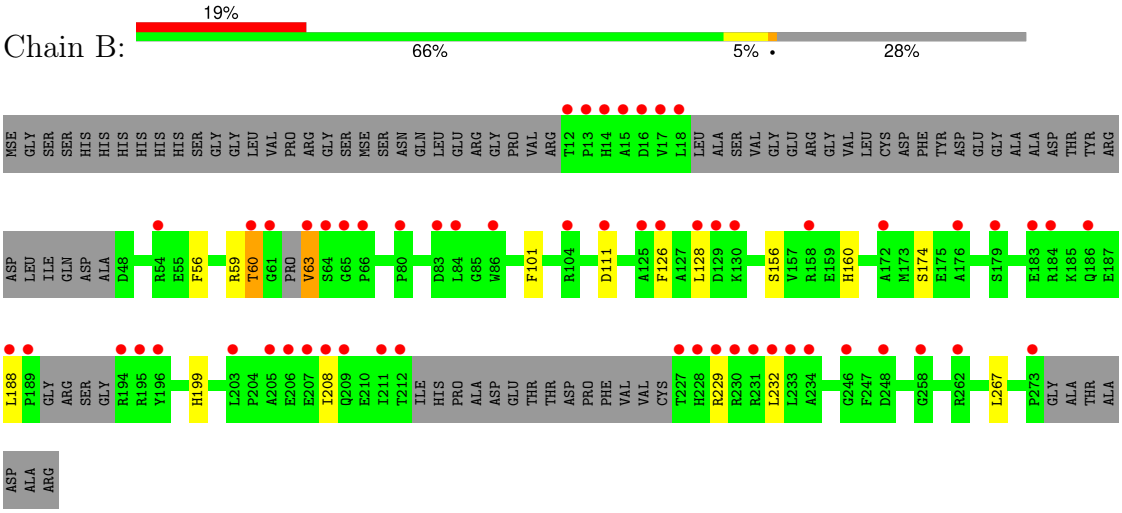
- Molecule 1: PCZA361.24

Chain A:



- Molecule 1: PCZA361.24

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	127.44Å 71.68Å 75.19Å 90.00° 103.02° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00 50.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.2 (50.00-2.00) 98.6 (50.00-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.52 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.238 , 0.271 0.237 , 0.269	Depositor DCC
R_{free} test set	2226 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	35.1	Xtriage
Anisotropy	0.412	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3575	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.18% 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

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	0.38	0/1808	0.56	0/2440
1	B	0.37	0/1680	0.54	0/2259
All	All	0.38	0/3488	0.55	0/4699

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1777	0	1771	19	0
1	B	1655	0	1648	13	0
2	A	24	0	0	0	0
2	B	119	0	0	2	0
All	All	3575	0	3419	31	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 5.

All (31) 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:62:PRO:CB	1:A:63:VAL:HA	1.88	1.02

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:62:PRO:HB2	1:A:63:VAL:HA	1.01	1.00
1:A:62:PRO:HB2	1:A:63:VAL:CA	1.93	0.99
1:B:128:LEU:H	1:B:160:HIS:HE1	1.21	0.86
1:B:128:LEU:H	1:B:160:HIS:CE1	2.08	0.71
1:B:229:ARG:HG2	1:B:229:ARG:HH11	1.58	0.69
1:A:11:ARG:HG2	1:A:16:ASP:OD1	1.95	0.65
1:B:128:LEU:N	1:B:160:HIS:HE1	1.97	0.62
1:A:60:THR:O	1:A:62:PRO:HD2	2.02	0.60
1:A:166:LYS:HG3	1:A:168:LEU:HD21	1.84	0.57
1:B:59:ARG:HG3	1:B:267:LEU:HD23	1.87	0.57
1:A:168:LEU:CD1	1:A:269:GLU:HG2	2.35	0.56
1:A:62:PRO:CB	1:A:63:VAL:CA	2.67	0.52
1:A:18:LEU:HD11	1:A:214:HIS:CD2	2.43	0.52
1:A:147:GLU:OE1	2:B:317:HOH:O	2.19	0.52
1:A:136:VAL:HG12	1:A:168:LEU:HB2	1.91	0.52
1:A:9:PRO:HG2	1:A:28:LEU:HG	1.92	0.50
1:A:95:THR:HG23	1:A:120[A]:GLN:HE22	1.76	0.50
1:A:218:GLU:HA	1:A:223:PHE:CE1	2.47	0.49
1:A:95:THR:CG2	1:A:120[A]:GLN:HE22	2.26	0.48
1:B:60:THR:CG2	1:B:63:VAL:HG13	2.44	0.48
1:A:55:GLU:HG3	1:A:255:PHE:CD1	2.49	0.47
1:B:229:ARG:HH11	1:B:229:ARG:CG	2.27	0.44
1:A:166:LYS:HG3	1:A:168:LEU:CD2	2.47	0.44
1:A:214:HIS:NE2	1:B:199:HIS:HE1	2.14	0.44
1:B:174:SER:HB3	1:B:232:LEU:HD21	2.00	0.43
1:A:55:GLU:HG3	1:A:255:PHE:CG	2.54	0.43
1:B:126:PHE:CE2	1:B:156:SER:HB3	2.53	0.43
1:B:60:THR:HG23	1:B:63:VAL:HG13	2.00	0.42
1:B:208:ILE:HD13	2:B:309:HOH:O	2.20	0.42
1:B:56:PHE:HA	1:B:267:LEU:HD21	2.03	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	224/299 (75%)	216 (96%)	7 (3%)	1 (0%)	30 27
1	B	206/299 (69%)	202 (98%)	4 (2%)	0	100 100
All	All	430/598 (72%)	418 (97%)	11 (3%)	1 (0%)	44 42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	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	189/233 (81%)	186 (98%)	3 (2%)	58 64
1	B	174/233 (75%)	169 (97%)	5 (3%)	37 39
All	All	363/466 (78%)	355 (98%)	8 (2%)	47 51

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	VAL
1	A	101	PHE
1	A	267	LEU
1	B	60	THR
1	B	63	VAL
1	B	101	PHE
1	B	111	ASP
1	B	188	LEU

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

Mol	Chain	Res	Type
1	B	160	HIS
1	B	199	HIS

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 oligosaccharides 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	227/299 (75%)	1.29	56 (24%) 2 2	21, 40, 56, 62	7 (3%)
1	B	209/299 (69%)	1.64	58 (27%) 2 1	21, 40, 55, 91	7 (3%)
All	All	436/598 (72%)	1.46	114 (26%) 2 2	21, 40, 56, 91	14 (3%)

All (114) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	199	HIS	9.5
1	B	17	VAL	8.4
1	B	18	LEU	8.0
1	A	198	LEU	7.2
1	B	12	THR	7.2
1	A	64	SER	6.8
1	B	15	ALA	6.5
1	A	104	ARG	6.1
1	B	230	ARG	5.7
1	B	130	LYS	5.5
1	A	47	ALA	5.3
1	B	188	LEU	5.3
1	B	211	ILE	5.2
1	B	104	ARG	5.1
1	A	62	PRO	5.0
1	B	184	ARG	5.0
1	B	234	ALA	4.9
1	B	54	ARG	4.8
1	B	248	ASP	4.3
1	B	189	PRO	4.3
1	B	212	THR	4.1
1	B	13	PRO	4.0
1	B	233	LEU	4.0
1	B	129	ASP	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	205	ALA	3.9
1	B	63	VAL	3.9
1	A	79	PHE	3.8
1	B	111	ASP	3.8
1	A	230	ARG	3.8
1	A	182	LEU	3.7
1	B	126	PHE	3.7
1	B	227	THR	3.5
1	A	203	LEU	3.5
1	B	14	HIS	3.5
1	B	66	PRO	3.4
1	B	196	TYR	3.4
1	A	224	VAL	3.4
1	B	61	GLY	3.4
1	B	232	LEU	3.4
1	A	10	VAL	3.3
1	A	22	VAL	3.3
1	A	13	PRO	3.3
1	B	206	GLU	3.3
1	B	16	ASP	3.2
1	A	184	ARG	3.2
1	A	50	THR	3.2
1	B	60	THR	3.1
1	A	32	TYR	3.1
1	B	183	GLU	3.1
1	A	63	VAL	3.1
1	A	172	ALA	3.1
1	A	60	THR	3.1
1	A	27	VAL	3.1
1	B	208	ILE	3.0
1	A	113	ARG	3.0
1	A	105	LEU	3.0
1	A	223	PHE	3.0
1	B	228	HIS	2.9
1	A	18	LEU	2.9
1	A	106	ALA	2.9
1	B	205	ALA	2.9
1	A	204	PRO	2.8
1	A	177	ALA	2.8
1	B	194	ARG	2.8
1	A	61	GLY	2.8
1	A	11	ARG	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	207	GLU	2.7
1	A	48	ASP	2.7
1	B	83	ASP	2.6
1	A	262	ARG	2.6
1	A	65	GLY	2.6
1	B	176	ALA	2.6
1	B	64	SER	2.6
1	A	112	VAL	2.5
1	A	225	VAL	2.5
1	A	218	GLU	2.5
1	B	172	ALA	2.5
1	A	111	ASP	2.5
1	A	15	ALA	2.5
1	A	215	PRO	2.4
1	A	49	GLY	2.4
1	B	84	LEU	2.4
1	B	258	GLY	2.4
1	B	186	GLN	2.4
1	A	9	PRO	2.4
1	B	158	ARG	2.4
1	A	114	ASP	2.3
1	A	51	SER	2.3
1	A	221	ASP	2.3
1	B	65	GLY	2.3
1	B	195	ARG	2.3
1	A	260	ALA	2.3
1	B	262	ARG	2.2
1	B	209	GLN	2.2
1	A	12	THR	2.2
1	A	110	ALA	2.2
1	A	222	PRO	2.2
1	B	80	PRO	2.2
1	B	231	ARG	2.2
1	B	203	LEU	2.2
1	B	125	ALA	2.2
1	B	179	SER	2.2
1	A	29	CYS	2.2
1	B	246	GLY	2.2
1	B	273	PRO	2.1
1	A	175	GLU	2.1
1	A	58	THR	2.1
1	A	217	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	128	LEU	2.1
1	A	14	HIS	2.1
1	A	57	ALA	2.0
1	B	86	TRP	2.0
1	A	102[A]	ARG	2.0
1	B	229	ARG	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.