Full wwPDB X-ray Structure Validation Report (i)

May 19, 2016 – 05:36 PM EDT

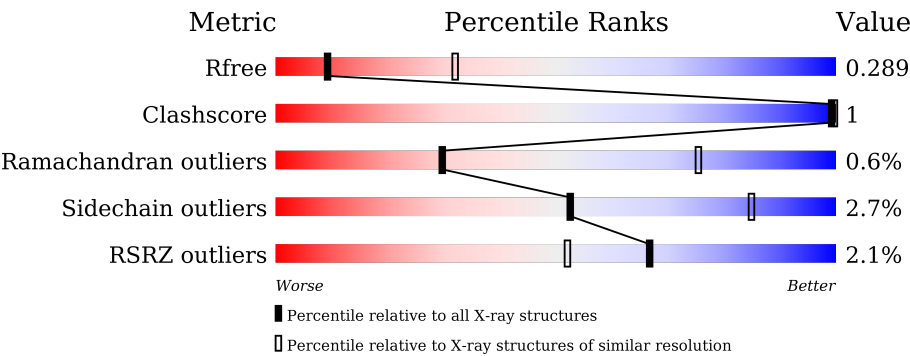|		
PDB ID	:	5FXY
Title	:	Structure of the human RBBP4:MTA1(464-546) complex
Authors	:	Millard, C.J.; Varma, N.; Fairall, L.; Schwabe, J.W.R.
Deposited on	:	2016-03-03
Resolution	:	3.20 Å(reported)

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 3.20 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	Percentile Ranks	Value
Rfree		0.289
Clashscore		1
Ramachandran outliers		0.6%
Sidechain outliers		2.7%
RSRZ outliers		2.1%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	425	88% • 10%
1	C	425	88% • 11%
1	E	425	2% 88% • 11%
1	G	425	4% 83% 7% 10%
2	B	85	% 74% 7% 19%
2	D	85	% 71% 11% 19%

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	85	2% / 75% / 5% / 20%
2	H	85	2% / 76% / • / 20%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14366 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 HISTONE-BINDING PROTEIN RBBP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total 3026	C 1912	N 514	O 590	S 10	0	0	0
1	C	380	Total 3019	C 1907	N 513	O 589	S 10	0	0	0
1	E	380	Total 3019	C 1907	N 513	O 589	S 10	0	0	0
1	G	381	Total 3026	C 1912	N 514	O 590	S 10	0	0	0

- Molecule 2 is a protein called METASTASIS-ASSOCIATED PROTEIN MTA1.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	69	Total 570	C 364	N 116	O 88	S 2	0	0	0
2	D	69	Total 574	C 366	N 116	O 90	S 2	0	0	0
2	F	68	Total 563	C 360	N 112	O 89	S 2	0	0	0
2	H	68	Total 569	C 363	N 115	O 89	S 2	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	462	GLY	-	EXPRESSION TAG	UNP Q13330
B	463	ALA	-	EXPRESSION TAG	UNP Q13330
D	462	GLY	-	EXPRESSION TAG	UNP Q13330
D	463	ALA	-	EXPRESSION TAG	UNP Q13330
F	462	GLY	-	EXPRESSION TAG	UNP Q13330
F	463	ALA	-	EXPRESSION TAG	UNP Q13330
H	462	GLY	-	EXPRESSION TAG	UNP Q13330
H	463	ALA	-	EXPRESSION TAG	UNP Q13330

3 Residue-property plots ⓘ

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: HISTONE-BINDING PROTEIN RBBP4
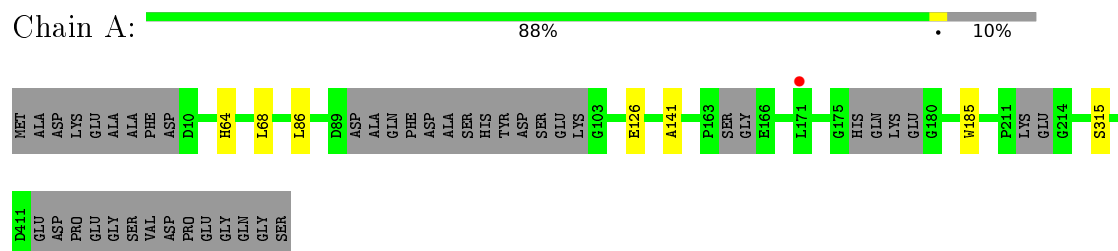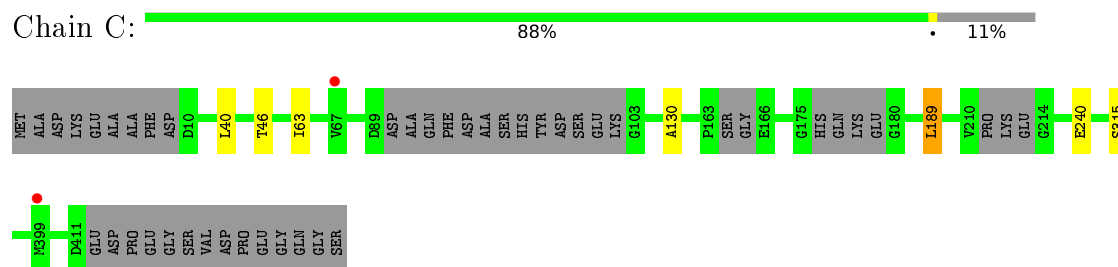
Chain A:

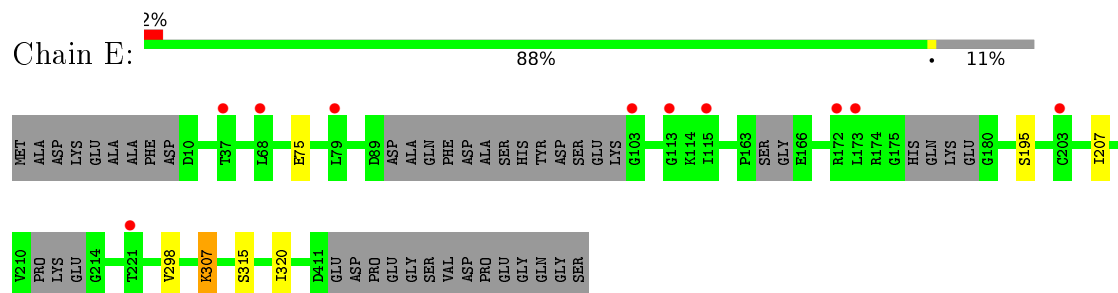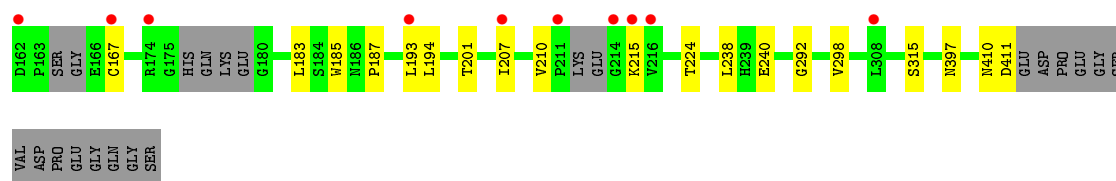- Molecule 1: HISTONE-BINDING PROTEIN RBBP4

Chain C:

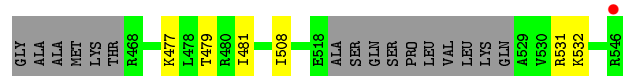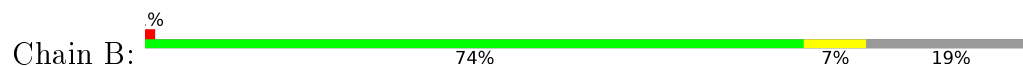- Molecule 1: HISTONE-BINDING PROTEIN RBBP4

Chain E:
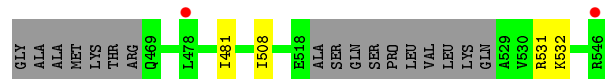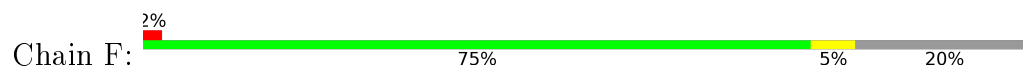

- Molecule 1: HISTONE-BINDING PROTEIN RBBP4

Chain G:

• Molecule 2: METASTASIS-ASSOCIATED PROTEIN MTA1

Chain B:

• Molecule 2: METASTASIS-ASSOCIATED PROTEIN MTA1

Chain D:

• Molecule 2: METASTASIS-ASSOCIATED PROTEIN MTA1

Chain F:

• Molecule 2: METASTASIS-ASSOCIATED PROTEIN MTA1

Chain H:
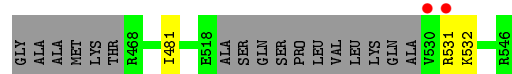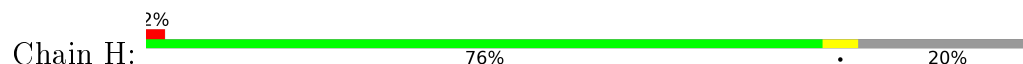
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.29Å 150.07Å 95.59Å 90.00° 94.54° 90.00°	Depositor
Resolution (Å)	95.29 – 3.20 95.29 – 3.20	Depositor EDS
% Data completeness (in resolution range)	97.5 (95.29-3.20) 97.5 (95.29-3.20)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.87 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.249 , 0.291 0.248 , 0.289	Depositor DCC
R_{free} test set	1752 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	80.8	Xtriage
Anisotropy	0.355	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 42.0	EDS
L-test for twinning[2]	$<|L|> = 0.47$, $<L^2> = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	14366	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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 (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	0.34	0/3107	0.53	0/4233
1	C	0.34	0/3099	0.52	0/4221
1	E	0.35	0/3099	0.54	0/4221
1	G	0.35	0/3107	0.54	0/4233
2	B	0.39	0/584	0.61	0/791
2	D	0.43	0/588	0.67	1/796 (0.1%)
2	F	0.39	0/577	0.62	0/782
2	H	0.39	0/583	0.66	0/789
All	All	0.35	0/14744	0.55	1/20066 (0.0%)

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	G	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	544	HIS	C-N-CD	5.06	139.03	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	139	ILE	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	3026	0	2876	2	0
1	C	3019	0	2869	2	0
1	E	3019	0	2869	1	0
1	G	3026	0	2876	11	0
2	B	570	0	597	0	0
2	D	574	0	601	0	0
2	F	563	0	588	0	0
2	H	569	0	596	0	0
All	All	14366	0	13872	16	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 1.

All (16) 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:G:139:ILE:HG13	1:G:140:ILE:N	2.26	0.50
1:G:139:ILE:HG12	1:G:152:PHE:C	2.32	0.50
1:G:292:GLY:HA2	1:G:298:VAL:HG12	1.93	0.50
1:G:201:THR:CG2	1:G:224:THR:HG22	2.43	0.49
1:G:292:GLY:CA	1:G:298:VAL:HG12	2.44	0.48
1:E:298:VAL:HG23	1:E:320:ILE:HD13	1.97	0.47
1:C:189:LEU:HD21	1:C:240:GLU:HB2	1.99	0.45
1:G:139:ILE:HD11	1:G:152:PHE:HB2	1.99	0.44
1:G:201:THR:HG22	1:G:224:THR:HG22	1.99	0.44
1:G:149:VAL:HG11	1:G:193:LEU:HD21	1.99	0.43
1:G:63:ILE:HD11	1:G:83:SER:HB2	2.00	0.43
1:A:64:HIS:CE1	1:A:86:LEU:HD12	2.53	0.43
1:A:141:ALA:HB2	1:A:185:TRP:CZ2	2.54	0.43
1:C:46:THR:HG21	1:C:130:ALA:HB3	2.02	0.42
1:G:139:ILE:HG13	1:G:140:ILE:H	1.84	0.42
1:G:139:ILE:HD13	1:G:154:TYR:N	2.35	0.42

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	371/425 (87%)	362 (98%)	8 (2%)	1 (0%)	46	85
1	C	370/425 (87%)	353 (95%)	16 (4%)	1 (0%)	46	85
1	E	370/425 (87%)	352 (95%)	16 (4%)	2 (0%)	34	78
1	G	371/425 (87%)	349 (94%)	20 (5%)	2 (0%)	34	78
2	B	65/85 (76%)	59 (91%)	5 (8%)	1 (2%)	13	55
2	D	65/85 (76%)	58 (89%)	5 (8%)	2 (3%)	5	34
2	F	64/85 (75%)	59 (92%)	4 (6%)	1 (2%)	12	54
2	H	64/85 (75%)	59 (92%)	4 (6%)	1 (2%)	12	54
All	All	1740/2040 (85%)	1651 (95%)	78 (4%)	11 (1%)	30	75

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	315	SER
1	E	315	SER
1	C	315	SER
1	G	315	SER
1	E	307	LYS
2	B	532	LYS
2	D	532	LYS
2	F	532	LYS
2	H	532	LYS
2	D	545	PRO
1	G	139	ILE

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	340/375 (91%)	338 (99%)	2 (1%)	90	97
1	C	339/375 (90%)	336 (99%)	3 (1%)	84	95
1	E	339/375 (90%)	335 (99%)	4 (1%)	78	93
1	G	340/375 (91%)	322 (95%)	18 (5%)	28	69
2	B	58/71 (82%)	53 (91%)	5 (9%)	13	46
2	D	59/71 (83%)	53 (90%)	6 (10%)	9	36
2	F	58/71 (82%)	55 (95%)	3 (5%)	29	69
2	H	59/71 (83%)	57 (97%)	2 (3%)	44	80
All	All	1592/1784 (89%)	1549 (97%)	43 (3%)	52	85

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	68	LEU
1	A	126	GLU
2	B	477	LYS
2	B	479	THR
2	B	481	ILE
2	B	508	ILE
2	B	531	ARG
1	C	40	LEU
1	C	63	ILE
1	C	189	LEU
2	D	468	ARG
2	D	479	THR
2	D	481	ILE
2	D	508	ILE
2	D	531	ARG
2	D	546	ARG
1	E	75	GLU
1	E	195	SER
1	E	207	ILE
1	E	307	LYS
2	F	481	ILE
2	F	508	ILE
2	F	531	ARG
1	G	54	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	63	ILE
1	G	74	ASP
1	G	115	ILE
1	G	117	ILE
1	G	167	CYS
1	G	183	LEU
1	G	185	TRP
1	G	187	PRO
1	G	194	LEU
1	G	207	ILE
1	G	210	VAL
1	G	215	LYS
1	G	238	LEU
1	G	240	GLU
1	G	397	ASN
1	G	410	ASN
1	G	411	ASP
2	H	481	ILE
2	H	531	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	250	GLN
1	E	128	ASN
1	G	122	ASN

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)

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	381/425 (89%)	0.19	1 (0%)	94	93	54, 79, 108, 137	0
1	C	380/425 (89%)	0.14	2 (0%)	91	87	57, 84, 118, 135	0
1	E	380/425 (89%)	0.32	10 (2%)	59	45	61, 90, 119, 149	0
1	G	381/425 (89%)	0.39	18 (4%)	35	22	65, 109, 141, 172	0
2	B	69/85 (81%)	0.15	1 (1%)	78	65	63, 94, 128, 134	0
2	D	69/85 (81%)	0.10	1 (1%)	78	65	71, 92, 136, 147	0
2	F	68/85 (80%)	0.21	2 (2%)	55	41	66, 101, 131, 146	0
2	H	68/85 (80%)	0.18	2 (2%)	55	41	66, 85, 133, 148	0
All	All	1796/2040 (88%)	0.24	37 (2%)	67	52	54, 89, 129, 172	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	214	GLY	5.5
1	G	167	CYS	5.0
2	H	531	ARG	4.6
2	D	546	ARG	3.9
2	F	546	ARG	3.7
1	E	79	LEU	3.6
1	E	113	GLY	3.1
1	G	174	ARG	2.9
1	G	150	LEU	2.9
1	C	399	MET	2.8
1	E	173	LEU	2.7
1	G	67	VAL	2.7
1	E	37	THR	2.6
1	A	171	LEU	2.6
1	G	215	LYS	2.5
1	G	162	ASP	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	G	74	ASP	2.4
1	G	216	VAL	2.4
1	G	126	GLU	2.4
1	G	70	THR	2.3
1	G	119	ILE	2.3
1	G	211	PRO	2.3
1	G	308	LEU	2.3
1	G	79	LEU	2.3
1	E	172	ARG	2.2
1	G	193	LEU	2.2
2	B	546	ARG	2.2
1	E	221	THR	2.2
1	G	207	ILE	2.1
1	E	115	ILE	2.1
1	E	103	GLY	2.1
1	C	67	VAL	2.1
1	E	203	CYS	2.1
2	H	530	VAL	2.1
1	G	120	LYS	2.1
2	F	478	LEU	2.0
1	E	68	LEU	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.