wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 2, 2024 – 11:33 pm GMT

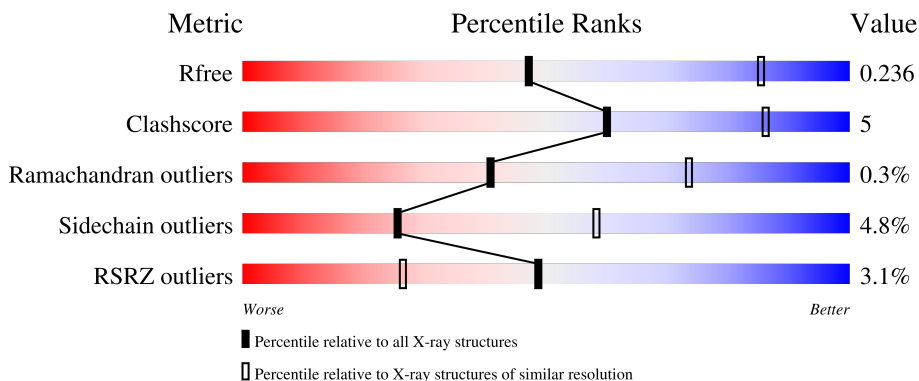|		
PDB ID	:	5FU7
Title	:	drosophila nanos NBR peptide bound to the NOT module of the human CCR4-NOT complex
Authors	:	Raisch, T.; Bhandari, D.; Sabath, K.; Helms, S.; Valkov, E.; Weichenrieder, O.; Izaurralde, E.
Deposited on	:	2016-01-21
Resolution	:	3.10 Å(reported)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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.
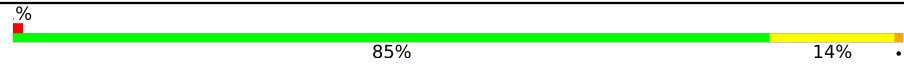
Metric	Percentile Ranks	Value
Rfree		0.236
Clashscore		5
Ramachandran outliers		0.3%
Sidechain outliers		4.8%
RSRZ outliers		3.1%

Worse *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
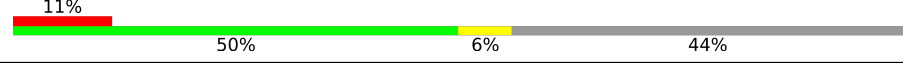
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	535	% / 80% 16% · ·
1	E	535	2% / 80% 14% · ·
2	B	197	10% / 82% 15% ·
2	F	197	3% / 85% 12% · ·
3	C	148	% / 82% 18% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	148	% ■ 85% 14% ▪
4	D	54	4% 50% 11% 39%
4	H	54	11% 50% 6% 44%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 14548 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 CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	519	Total 4213	C 2724	N 722	O 744	S 23	0	0	0
1	E	514	Total 4178	C 2703	N 717	O 735	S 23	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1827	GLY	-	expression tag	UNP A5YKK6
A	1828	PRO	-	expression tag	UNP A5YKK6
A	1829	HIS	-	expression tag	UNP A5YKK6
A	1830	MET	-	expression tag	UNP A5YKK6
A	1831	LEU	-	expression tag	UNP A5YKK6
A	1832	GLU	-	expression tag	UNP A5YKK6
A	2344	GLU	HIS	engineered mutation	UNP A5YKK6
A	2345	GLU	CYS	engineered mutation	UNP A5YKK6
A	2346	GLU	ALA	engineered mutation	UNP A5YKK6
E	1827	GLY	-	expression tag	UNP A5YKK6
E	1828	PRO	-	expression tag	UNP A5YKK6
E	1829	HIS	-	expression tag	UNP A5YKK6
E	1830	MET	-	expression tag	UNP A5YKK6
E	1831	LEU	-	expression tag	UNP A5YKK6
E	1832	GLU	-	expression tag	UNP A5YKK6
E	2344	GLU	HIS	engineered mutation	UNP A5YKK6
E	2345	GLU	CYS	engineered mutation	UNP A5YKK6
E	2346	GLU	ALA	engineered mutation	UNP A5YKK6

- Molecule 2 is a protein called CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	191	Total 1570	C 1012	N 263	O 287	S 8	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	F	191	1570	1012	263	287	8	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	344	GLY	-	expression tag	UNP Q9NZN8
B	345	PRO	-	expression tag	UNP Q9NZN8
B	346	HIS	-	expression tag	UNP Q9NZN8
B	347	MET	-	expression tag	UNP Q9NZN8
B	348	LEU	-	expression tag	UNP Q9NZN8
B	349	GLU	-	expression tag	UNP Q9NZN8
F	344	GLY	-	expression tag	UNP Q9NZN8
F	345	PRO	-	expression tag	UNP Q9NZN8
F	346	HIS	-	expression tag	UNP Q9NZN8
F	347	MET	-	expression tag	UNP Q9NZN8
F	348	LEU	-	expression tag	UNP Q9NZN8
F	349	GLU	-	expression tag	UNP Q9NZN8

- Molecule 3 is a protein called CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	148	1292	840	216	229	7	0	0	0
3	G	148	1293	840	216	230	7	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	601	GLY	-	expression tag	UNP O75175
C	602	PRO	-	expression tag	UNP O75175
C	603	HIS	-	expression tag	UNP O75175
C	604	MET	-	expression tag	UNP O75175
C	605	LEU	-	expression tag	UNP O75175
C	606	GLU	-	expression tag	UNP O75175
G	601	GLY	-	expression tag	UNP O75175
G	602	PRO	-	expression tag	UNP O75175
G	603	HIS	-	expression tag	UNP O75175
G	604	MET	-	expression tag	UNP O75175
G	605	LEU	-	expression tag	UNP O75175
G	606	GLU	-	expression tag	UNP O75175

• Molecule 4 is a protein called NANOS, ISOFORM B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	33	Total 238	C 148	N 35	O 53	S 2	0	0	0
4	H	30	Total 194	C 120	N 31	O 42	S 1	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	110	GLY	-	expression tag	UNP A0A0B4KGY
D	111	PRO	-	expression tag	UNP A0A0B4KGY
D	112	HIS	-	expression tag	UNP A0A0B4KGY
D	113	MET	-	expression tag	UNP A0A0B4KGY
D	114	LEU	-	expression tag	UNP A0A0B4KGY
D	115	GLU	-	expression tag	UNP A0A0B4KGY
H	110	GLY	-	expression tag	UNP A0A0B4KGY
H	111	PRO	-	expression tag	UNP A0A0B4KGY
H	112	HIS	-	expression tag	UNP A0A0B4KGY
H	113	MET	-	expression tag	UNP A0A0B4KGY
H	114	LEU	-	expression tag	UNP A0A0B4KGY
H	115	GLU	-	expression tag	UNP A0A0B4KGY

3 Residue-property plots ⓘ

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: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1
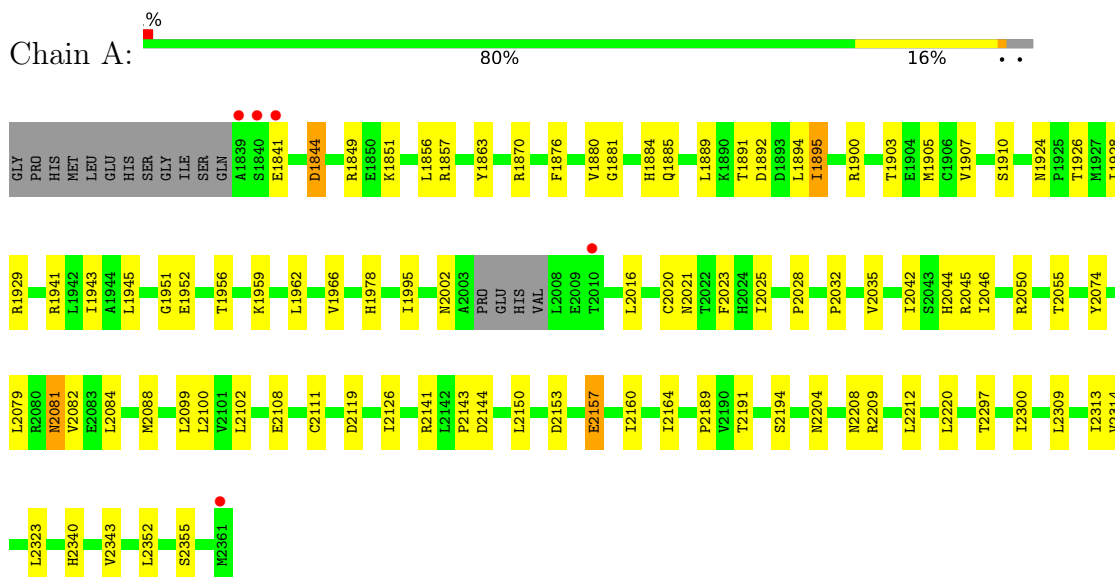
• Molecule 1: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1

• Molecule 2: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2

Chain B:

10%
82%
15%

GLY PRO HIS MET LEU E349 I364 L382 A400 W403 T423 N424 R428 I434 Y439 L443 Y446 M450 N451 Q457 W469 H472 K473 E474 W478 A482 M485 E486 P487 T488 M489 K490 T491 N492 T493 Y494 E495 R496 G497 T498 K512 E513

F514 H515 L516 E517 K520 L521 R524 P525 H526 L527 P528 N532 Y533 N534 Q537 Q538 A539 PHE

• Molecule 2: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2

Chain F:

3%
85%
12%

GLY PRO HIS MET LEU E349 L361 I364 P371 L382 A400 W403 N424 K430 I434 Y439 Y446 M450 V455 W469 E474 A482 M485 M489 T498 Y499 Y500 F501 K509 F514 R524 P525 H526 L527 P528 S529 T530 F531

N534 Q537 Q538 A539 PHE

• Molecule 3: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3

Chain C:

%
82%
18%

G601 E606 E610 E618 A621 H624 S631 L638 C643 P644 T645 P646 H649 F672 L678 R697 M703 M704 W705 R708 P712 D717 E718 Q721 G722 T723 Y726 E738 G739 F740 T741 F742 E743 Y746 L747 E748

• Molecule 3: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3

Chain G:

%
85%
14%

G601 E606 E610 E618 A621 H624 L638 N641 P642 C643 P646 P647 Y648 H649 E661 Q664 L678 M703 R708 P712 T723 E730 E738 F742 L747 E748

• Molecule 4: NANOS, ISOFORM B

Chain D:

4%
50%
11%
39%

GLY PRO HIS MET LEU GLU SER HIS GLN GLN T120 D121 E122 I123 I129 F130 T134 THR GLY ALA ALA GLU ASN ALA GLY S144 D147 D161 LEU GLY

• Molecule 4: NANOS, ISOFORM B

Chain H:

11%
50%
6%
44%

GLY PRO HIS MET LEU GLU SER HIS GLN GLN T120 D121 E122 I123 S126 L127 K128 I129 F130 A131 GLN VAL THR THR GLY ALA ALA GLU ASN ALA ALA G143 S144 D147 D160 ASP LEU GLY

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	75.95Å 135.67Å 104.97Å 90.00° 107.98° 90.00°	Depositor
Resolution (Å)	48.58 – 3.10 48.58 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.58-3.10) 99.7 (48.58-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 3.12Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.165 , 0.227 0.180 , 0.236	Depositor DCC
R_{free} test set	1892 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	82.0	Xtriage
Anisotropy	0.595	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 73.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14548	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.31% 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	0.54	0/4320	0.70	0/5874
1	E	0.53	0/4285	0.69	0/5826
2	B	0.49	0/1617	0.70	0/2198
2	F	0.49	0/1617	0.70	0/2198
3	C	0.49	0/1343	0.64	0/1819
3	G	0.50	0/1344	0.66	0/1819
4	D	0.51	0/239	0.71	0/322
4	H	0.49	0/194	0.67	0/262
All	All	0.52	0/14959	0.69	0/20318

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 (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	4213	0	4224	49	0
1	E	4178	0	4190	44	0
2	B	1570	0	1512	13	0
2	F	1570	0	1512	12	0
3	C	1292	0	1207	15	0
3	G	1293	0	1207	12	0
4	D	238	0	208	5	0
4	H	194	0	158	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	14548	0	14218	133	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.

The worst 5 of 133 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:E:2184:LEU:HD22	1:E:2247:HIS:HD2	1.35	0.92
1:E:1978:HIS:HE1	1:E:2033:GLY:H	1.31	0.79
1:A:1952:GLU:HB2	4:H:129:ILE:HG12	1.72	0.71
2:B:498:THR:HG22	2:B:513:GLU:HB3	1.72	0.71
1:E:2021:ASN:O	1:E:2025:ILE:HG12	1.91	0.71

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	515/535 (96%)	499 (97%)	15 (3%)	1 (0%)	47	79
1	E	510/535 (95%)	488 (96%)	20 (4%)	2 (0%)	34	69
2	B	189/197 (96%)	174 (92%)	13 (7%)	2 (1%)	14	46
2	F	189/197 (96%)	175 (93%)	13 (7%)	1 (0%)	29	64
3	C	146/148 (99%)	138 (94%)	8 (6%)	0	100	100
3	G	146/148 (99%)	139 (95%)	7 (5%)	0	100	100
4	D	29/54 (54%)	29 (100%)	0	0	100	100
4	H	26/54 (48%)	25 (96%)	1 (4%)	0	100	100
All	All	1750/1868 (94%)	1667 (95%)	77 (4%)	6 (0%)	41	73

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	428	ARG
2	F	400	ALA
2	B	400	ALA
1	E	2202	VAL
1	A	2082	VAL

5.3.2 Protein sidechains ⓘ

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	464/479 (97%)	439 (95%)	25 (5%)	22	53
1	E	460/479 (96%)	434 (94%)	26 (6%)	20	52
2	B	167/172 (97%)	161 (96%)	6 (4%)	35	67
2	F	167/172 (97%)	160 (96%)	7 (4%)	30	62
3	C	136/136 (100%)	130 (96%)	6 (4%)	28	61
3	G	136/136 (100%)	132 (97%)	4 (3%)	42	72
4	D	24/41 (58%)	23 (96%)	1 (4%)	30	62
4	H	16/41 (39%)	16 (100%)	0	100	100
All	All	1570/1656 (95%)	1495 (95%)	75 (5%)	25	58

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	2202	VAL
3	G	610	GLU
1	E	2212	LEU
2	F	371	PRO
1	A	2343	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	2081	ASN
1	E	2232	ASN
3	G	694	GLN
1	E	2247	HIS
1	E	2229	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 monosaccharides 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, 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	519/535 (97%)	-0.25	5 (0%)	82	67	49, 81, 137, 163	0
1	E	514/535 (96%)	-0.13	13 (2%)	57	34	55, 92, 142, 185	0
2	B	191/197 (96%)	0.20	20 (10%)	6	2	67, 109, 205, 217	0
2	F	191/197 (96%)	-0.04	6 (3%)	49	26	73, 102, 174, 190	0
3	C	148/148 (100%)	-0.18	2 (1%)	75	56	50, 101, 133, 157	0
3	G	148/148 (100%)	-0.10	1 (0%)	87	75	51, 89, 123, 136	0
4	D	33/54 (61%)	0.55	2 (6%)	21	9	109, 137, 158, 164	0
4	H	30/54 (55%)	0.68	6 (20%)	1	0	120, 136, 165, 181	0
All	All	1774/1868 (94%)	-0.10	55 (3%)	49	26	49, 95, 157, 217	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	488	THR	4.2
2	B	496	ARG	4.1
1	E	2361	MET	3.9
1	E	1841	GLU	3.8
2	B	516	LEU	3.5

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.