



Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 07:18 am BST

PDB ID : 6G7Q
Title : Trichodesmium Tery_3377 (IdiA) (FutA) in complex with iron and citrate ligands.
Authors : Machelett, M.M.; Tews, I.
Deposited on : 2018-04-06
Resolution : 1.20 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

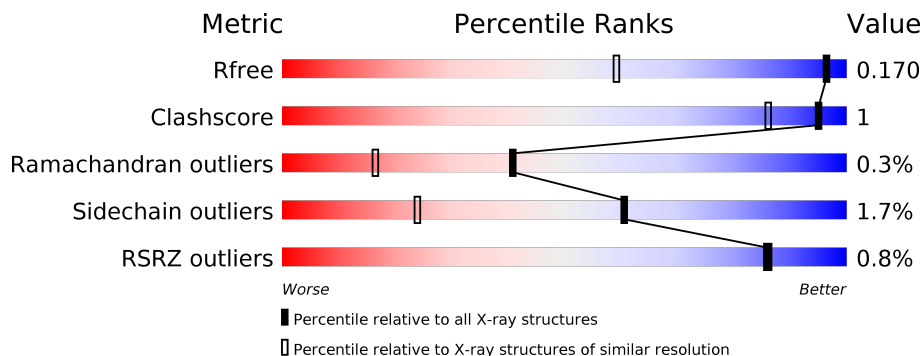
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.20 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)
RSRZ outliers	127900	1200 (1.22-1.18)

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 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 $\leq 5\%$. 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	318	 % 95% 5%
1	B	318	 % 93% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	405	-	X	-	-
3	GOL	B	403	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6050 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 Extracellular solute-binding protein, family 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	316	2599	1660	442	489	8	0	31	0
1	A	318	2535	1614	436	479	6	0	14	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	320	LEU	-	expression tag	UNP Q10Z45
B	321	GLU	-	expression tag	UNP Q10Z45
A	320	LEU	-	expression tag	UNP Q10Z45
A	321	GLU	-	expression tag	UNP Q10Z45

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Cl	0	0
			1	1		
2	A	1	Total	Cl	0	0
			1	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).

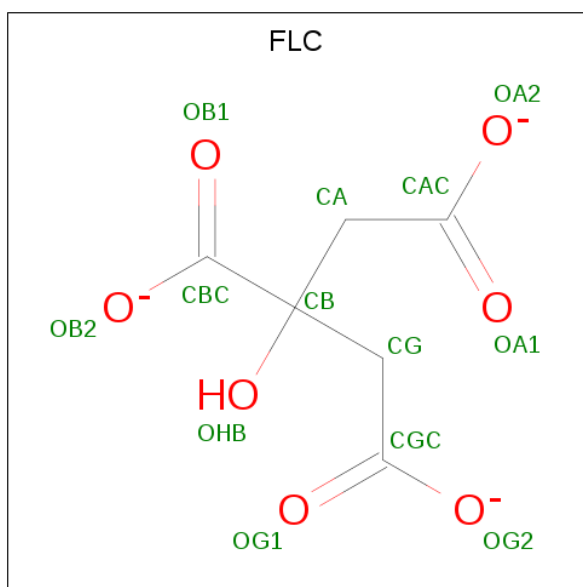


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0

- Molecule 4 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Fe 1 1	0	0

- Molecule 5 is CITRATE ANION (three-letter code: FLC) (formula: C₆H₅O₇).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 13 6 7	0	0
5	A	1	Total C O 13 6 7	0	0

- Molecule 6 is water.

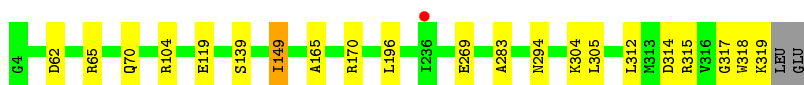
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	508	Total O 515 515	0	7
6	A	344	Total O 348 348	0	4

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 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: Extracellular solute-binding protein, family 1

Chain B: 



- Molecule 1: Extracellular solute-binding protein, family 1

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.33Å 102.01Å 57.67Å 90.00° 112.85° 90.00°	Depositor
Resolution (Å)	30.00 – 1.20 28.64 – 1.20	Depositor EDS
% Data completeness (in resolution range)	98.4 (30.00-1.20) 98.5 (28.64-1.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.72 (at 1.20Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.136 , 0.165 0.143 , 0.170	Depositor DCC
R_{free} test set	1991 reflections (1.08%)	wwPDB-VP
Wilson B-factor (Å ²)	12.4	Xtrriage
Anisotropy	0.520	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 45.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.017 for l,-k,h	Xtrriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	6050	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.12% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, FLC, FE, CL

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	0.81	1/2627 (0.0%)	0.93	7/3545 (0.2%)
1	B	0.78	1/2729 (0.0%)	0.92	6/3682 (0.2%)
All	All	0.80	2/5356 (0.0%)	0.92	13/7227 (0.2%)

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	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	119	GLU	CD-OE2	-5.46	1.19	1.25
1	A	207	ARG	CZ-NH1	5.07	1.39	1.33

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	315	ARG	NE-CZ-NH1	-16.42	112.09	120.30
1	A	170	ARG	NE-CZ-NH1	10.11	125.35	120.30
1	A	170	ARG	NE-CZ-NH2	-9.95	115.33	120.30
1	B	315	ARG	NE-CZ-NH2	7.92	124.26	120.30
1	A	315	ARG	NE-CZ-NH1	6.62	123.61	120.30
1	A	207	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	A	230	ARG	NE-CZ-NH2	-5.41	117.60	120.30
1	B	314	ASP	CB-CG-OD2	5.33	123.09	118.30
1	A	22	ASP	CB-CG-OD2	-5.29	113.54	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	218	ASP	CB-CG-OD2	-5.18	113.64	118.30
1	B	65	ARG	NE-CZ-NH1	5.17	122.89	120.30
1	B	149	ILE	CB-CG1-CD1	5.11	128.20	113.90
1	B	318	TRP	O-C-N	5.08	130.82	122.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	236	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	2535	0	2564	3	0
1	B	2599	0	2646	9	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	6	0	8	1	0
3	B	18	0	23	5	4
4	A	1	0	0	0	0
5	A	26	0	8	0	0
6	A	348	0	0	0	7
6	B	515	0	0	1	3
All	All	6050	0	5249	13	7

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 (13) 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:319:LYS:C	3:B:402:GOL:O3	1.65	1.32
1:B:317:GLY:O	3:B:404:GOL:H2	1.88	0.74

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:170[A]:ARG:HD2	6:B:563[A]:HOH:O	2.10	0.52
1:B:319:LYS:CA	3:B:402:GOL:O3	2.54	0.51
3:B:402:GOL:O1	3:B:403:GOL:H32	2.12	0.50
1:B:304[B]:LYS:HG2	1:B:305[B]:LEU:HD12	1.97	0.45
1:A:86:LYS:NZ	3:A:405:GOL:H11	2.33	0.44
1:A:68:ARG:HB2	1:A:68:ARG:CZ	2.47	0.44
1:A:269:GLU:HG2	1:A:283:ALA:HB2	2.01	0.43
1:B:165:ALA:CB	1:B:312[B]:LEU:HD21	2.49	0.42
1:B:269:GLU:HG2	1:B:283:ALA:HB2	2.01	0.42
1:B:139[A]:SER:OG	3:B:403:GOL:C1	2.69	0.41

All (7) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:B:905:HOH:O	6:A:539:HOH:O[2_454]	0.84	1.36
6:B:974:HOH:O	6:A:808:HOH:O[2_555]	1.01	1.19
3:B:403:GOL:C2	6:A:753:HOH:O[2_454]	1.50	0.70
3:B:403:GOL:O3	6:A:753:HOH:O[2_454]	1.59	0.61
3:B:403:GOL:C3	6:A:753:HOH:O[2_454]	1.69	0.51
6:B:807:HOH:O	6:A:725:HOH:O[2_454]	2.07	0.13
3:B:403:GOL:C1	6:A:753:HOH:O[2_454]	2.15	0.05

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	330/318 (104%)	327 (99%)	2 (1%)	1 (0%)	41 16
1	B	345/318 (108%)	341 (99%)	3 (1%)	1 (0%)	41 16
All	All	675/636 (106%)	668 (99%)	5 (1%)	2 (0%)	41 16

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	62	ASP
1	A	62	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	275/261 (105%)	271 (98%)	4 (2%)	65 29
1	B	284/261 (109%)	278 (98%)	6 (2%)	53 15
All	All	559/522 (107%)	549 (98%)	10 (2%)	60 21

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	70	GLN
1	B	104	ARG
1	B	149	ILE
1	B	196[A]	LEU
1	B	196[B]	LEU
1	B	294	ASN
1	A	70	GLN
1	A	104	ARG
1	A	149	ILE
1	A	294	ASN

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

Mol	Chain	Res	Type
1	A	113	ASN
1	A	116	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 3 are monoatomic - leaving 6 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	GOL	B	404	-	5,5,5	0.73	0	5,5,5	1.41	1 (20%)
3	GOL	B	403	-	5,5,5	0.29	0	5,5,5	1.15	0
5	FLC	A	402	4	3,12,12	1.81	2 (66%)	3,17,17	0.84	0
3	GOL	A	405	-	5,5,5	0.95	0	5,5,5	2.00	3 (60%)
5	FLC	A	403	4	3,12,12	1.07	0	3,17,17	3.23	1 (33%)
3	GOL	B	402	-	5,5,5	1.20	1 (20%)	5,5,5	2.11	2 (40%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	B	404	-	-	2/4/4/4	-
3	GOL	B	403	-	-	4/4/4/4	-
5	FLC	A	402	4	-	0/6/16/16	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	405	-	-	4/4/4/4	-
5	FLC	A	403	4	-	0/6/16/16	-
3	GOL	B	402	-	-	2/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	GOL	O3-C3	2.49	1.52	1.42
5	A	402	FLC	OHB-CB	-2.21	1.39	1.43
5	A	402	FLC	CA-CB	-2.18	1.51	1.54

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	403	FLC	CB-CG-CGC	5.58	123.93	114.98
3	B	402	GOL	O2-C2-C3	3.84	126.04	109.12
3	A	405	GOL	C3-C2-C1	2.73	122.31	111.70
3	B	404	GOL	O3-C3-C2	2.21	120.81	110.20
3	A	405	GOL	O2-C2-C3	-2.10	99.88	109.12
3	B	402	GOL	O1-C1-C2	-2.05	100.36	110.20
3	A	405	GOL	O3-C3-C2	-2.05	100.37	110.20

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	404	GOL	O1-C1-C2-C3
3	A	405	GOL	O1-C1-C2-C3
3	A	405	GOL	C1-C2-C3-O3
3	B	402	GOL	O1-C1-C2-C3
3	A	405	GOL	O2-C2-C3-O3
3	B	403	GOL	O1-C1-C2-C3
3	B	403	GOL	C1-C2-C3-O3
3	A	405	GOL	O1-C1-C2-O2
3	B	402	GOL	O1-C1-C2-O2
3	B	404	GOL	O1-C1-C2-O2
3	B	403	GOL	O1-C1-C2-O2
3	B	403	GOL	O2-C2-C3-O3

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	404	GOL	1	0
3	B	403	GOL	2	4
3	A	405	GOL	1	0
3	B	402	GOL	3	0

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	318/318 (100%)	0.07	4 (1%) 77 77	10, 16, 28, 39	0
1	B	316/318 (99%)	0.03	1 (0%) 94 94	10, 15, 25, 35	0
All	All	634/636 (99%)	0.05	5 (0%) 86 86	10, 15, 27, 39	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	209	SER	4.6
1	A	236	ILE	2.3
1	A	51	ALA	2.3
1	B	236	ILE	2.2
1	A	321	GLU	2.1

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	GOL	A	405	6/6	0.70	0.25	25,32,42,49	0
5	FLC	A	403	13/13	0.70	0.23	17,25,29,34	13
3	GOL	B	404	6/6	0.76	0.39	29,31,34,37	0
3	GOL	B	403	6/6	0.80	0.26	29,33,40,41	0
3	GOL	B	402	6/6	0.85	0.24	26,32,35,40	0
5	FLC	A	402	13/13	0.88	0.19	20,24,29,31	0
4	FE	A	401	1/1	0.99	0.05	23,23,23,23	1
2	CL	A	404	1/1	1.00	0.05	15,15,15,15	0
2	CL	B	401	1/1	1.00	0.04	15,15,15,15	0

6.5 Other polymers [i](#)

There are no such residues in this entry.