



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 23, 2020 – 10:19 PM BST

PDB ID : 6ZEI
Title : Structure of PP1-IRSp53 S455E chimera [PP1(7-304) + linker (G/S)x9 + IRSp53(449-465)] bound to Phactr1 (516-580)
Authors : Mouilleron, S.; Treisman, R.; Fedoryshchak, R.; Lee, R.; Butler, A.M.; Prechova, M.
Deposited on : 2020-06-16
Resolution : 1.39 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	2.14.6
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)	:	2.14.6

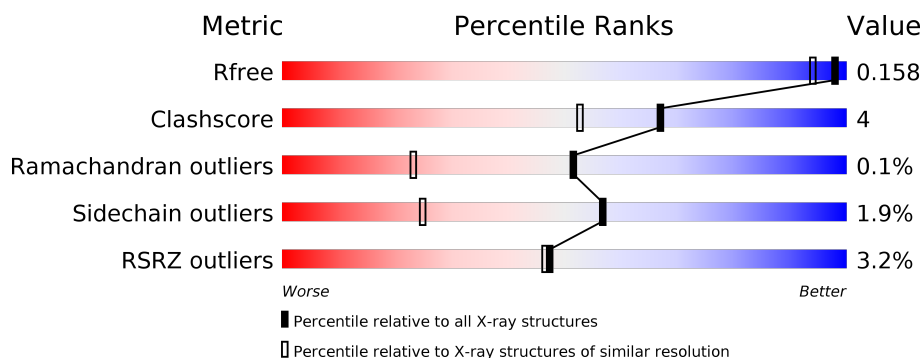
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.39 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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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	338	<div> <div>3%</div> <div> <div></div> <div>83%</div> <div>6%</div> <div>10%</div> </div> </div>
1	B	338	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>8%</div> <div>9%</div> </div> </div>
2	C	70	<div> <div></div> <div> <div></div> <div>86%</div> <div>7%</div> <div>7%</div> </div> </div>
2	D	70	<div> <div>10%</div> <div> <div></div> <div>79%</div> <div>11%</div> <div>7%</div> </div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12842 atoms, of which 6020 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 Serine/threonine-protein phosphatase PP1-alpha catalytic subunit, Brain-specific angiogenesis inhibitor 1-associated protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	305	Total	C	H	N	O	S	0	7	0
			4879	1579	2418	410	452	20			
1	B	307	Total	C	H	N	O	S	0	12	0
			4942	1597	2443	417	465	20			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	-	expression tag	UNP P62136
A	3	HIS	-	expression tag	UNP P62136
A	4	MET	-	expression tag	UNP P62136
A	5	GLY	-	expression tag	UNP P62136
A	6	SER	-	expression tag	UNP P62136
A	305	SER	-	linker	UNP P62136
A	306	GLY	-	linker	UNP P62136
A	307	SER	-	linker	UNP P62136
A	308	GLY	-	linker	UNP P62136
A	309	SER	-	linker	UNP P62136
A	310	GLY	-	linker	UNP P62136
A	311	SER	-	linker	UNP P62136
A	312	GLY	-	linker	UNP P62136
A	313	SER	-	linker	UNP P62136
A	314	GLY	-	linker	UNP P62136
A	315	SER	-	linker	UNP P62136
A	316	GLY	-	linker	UNP P62136
A	317	SER	-	linker	UNP P62136
A	318	GLY	-	linker	UNP P62136
A	319	SER	-	linker	UNP P62136
A	320	GLY	-	linker	UNP P62136
A	321	SER	-	linker	UNP P62136
A	322	GLY	-	linker	UNP P62136
A	329	GLU	SER	conflict	UNP Q9UQB8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	GLY	-	expression tag	UNP P62136
B	3	HIS	-	expression tag	UNP P62136
B	4	MET	-	expression tag	UNP P62136
B	5	GLY	-	expression tag	UNP P62136
B	6	SER	-	expression tag	UNP P62136
B	305	SER	-	linker	UNP P62136
B	306	GLY	-	linker	UNP P62136
B	307	SER	-	linker	UNP P62136
B	308	GLY	-	linker	UNP P62136
B	309	SER	-	linker	UNP P62136
B	310	GLY	-	linker	UNP P62136
B	311	SER	-	linker	UNP P62136
B	312	GLY	-	linker	UNP P62136
B	313	SER	-	linker	UNP P62136
B	314	GLY	-	linker	UNP P62136
B	315	SER	-	linker	UNP P62136
B	316	GLY	-	linker	UNP P62136
B	317	SER	-	linker	UNP P62136
B	318	GLY	-	linker	UNP P62136
B	319	SER	-	linker	UNP P62136
B	320	GLY	-	linker	UNP P62136
B	321	SER	-	linker	UNP P62136
B	322	GLY	-	linker	UNP P62136
B	329	GLU	SER	conflict	UNP Q9UQB8

- Molecule 2 is a protein called Phosphatase and actin regulator.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	65	Total	C	H	N	O	S	0	2	0
			1125	353	562	106	103	1			
2	D	65	Total	C	H	N	O	S	0	4	0
			1150	363	574	107	105	1			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	511	GLY	-	expression tag	UNP Q4VY12
C	512	PRO	-	expression tag	UNP Q4VY12
C	513	LEU	-	expression tag	UNP Q4VY12
C	514	GLY	-	expression tag	UNP Q4VY12
C	515	SER	-	expression tag	UNP Q4VY12
D	511	GLY	-	expression tag	UNP Q4VY12

Continued on next page...

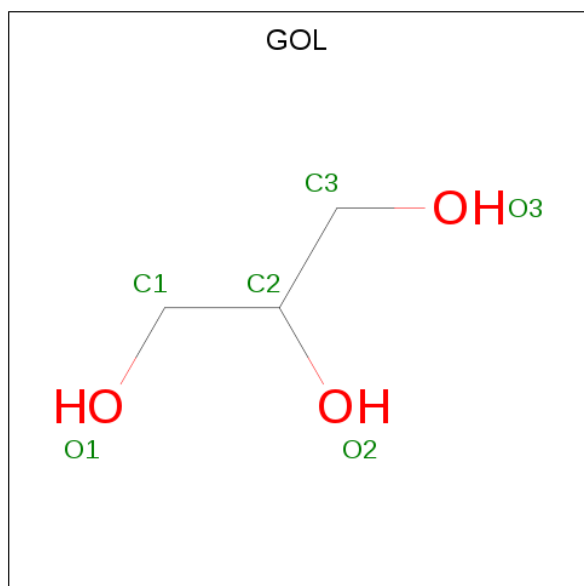
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	512	PRO	-	expression tag	UNP Q4VY12
D	513	LEU	-	expression tag	UNP Q4VY12
D	514	GLY	-	expression tag	UNP Q4VY12
D	515	SER	-	expression tag	UNP Q4VY12

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

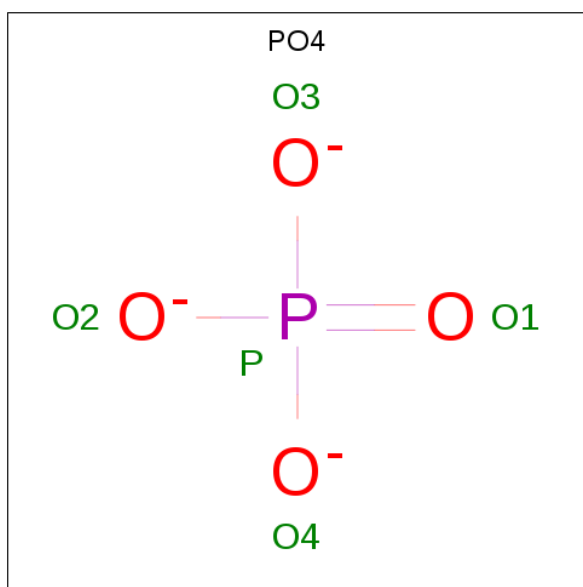
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total Mn 2 2	0	0
3	A	2	Total Mn 2 2	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 14 3 8 3	0	0
4	A	1	Total C H O 13 3 7 3	0	0
4	B	1	Total C H O 14 3 8 3	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	287	Total	O	0	1
			287	287		
6	C	87	Total	O	0	0
			87	87		
6	B	254	Total	O	0	1
			254	254		
6	D	63	Total	O	0	1
			63	63		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	48.80Å 122.33Å 69.42Å 90.00° 92.27° 90.00°	Depositor
Resolution (Å)	69.37 – 1.39 69.37 – 1.39	Depositor EDS
% Data completeness (in resolution range)	99.6 (69.37-1.39) 99.6 (69.37-1.39)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 1.39Å)	Xtriage
Refinement program	PHENIX 1.18_3845, PHENIX 1.18_3845	Depositor
R, R_{free}	0.130 , 0.158 0.130 , 0.158	Depositor DCC
R_{free} test set	1999 reflections (1.23%)	wwPDB-VP
Wilson B-factor (Å ²)	16.4	Xtriage
Anisotropy	0.217	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.44 , 56.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	12842	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PO4, MN

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	0.53	0/2536	0.79	2/3427 (0.1%)
1	B	0.53	0/2590	0.78	2/3499 (0.1%)
2	C	0.57	1/580 (0.2%)	0.72	1/778 (0.1%)
2	D	0.59	0/601	0.74	0/807
All	All	0.54	1/6307 (0.0%)	0.78	5/8511 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	556	GLU	CB-CG	5.19	1.62	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	299	ALA	N-CA-C	5.77	126.57	111.00
1	A	260	LYS	CD-CE-NZ	5.53	124.43	111.70
2	C	521	ARG	NE-CZ-NH1	-5.18	117.71	120.30
1	B	143[A]	ARG	NE-CZ-NH2	-5.09	117.75	120.30
1	B	143[B]	ARG	NE-CZ-NH2	-5.09	117.75	120.30

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	2461	2418	2420	18	0
1	B	2499	2443	2426	22	0
2	C	563	562	562	1	0
2	D	576	574	564	9	1
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	12	15	15	1	0
4	B	6	8	8	1	0
5	A	5	0	0	0	0
5	B	5	0	0	0	0
6	A	287	0	0	7	0
6	B	254	0	0	3	1
6	C	87	0	0	0	0
6	D	63	0	0	0	0
All	All	6822	6020	5995	46	1

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.

The worst 5 of 46 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:127:CYS:HB3	6:A:821:HOH:O	1.78	0.83
1:A:195:VAL:HG21	6:A:821:HOH:O	1.78	0.83
1:B:195:VAL:HG21	6:B:784:HOH:O	1.81	0.78
1:A:260:LYS:NZ	6:A:601:HOH:O	2.19	0.76
1:B:101:LEU:HD12	1:B:143[B]:ARG:HG3	1.80	0.64

All (1) 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)
2:D:524:ASP:OD2	6:B:710:HOH:O[1_655]	2.19	0.01

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	308/338 (91%)	300 (97%)	8 (3%)	0	100	100
1	B	315/338 (93%)	302 (96%)	13 (4%)	0	100	100
2	C	65/70 (93%)	64 (98%)	1 (2%)	0	100	100
2	D	67/70 (96%)	65 (97%)	1 (2%)	1 (2%)	10	1
All	All	755/816 (92%)	731 (97%)	23 (3%)	1 (0%)	51	23

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	524	ASP

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	270/289 (93%)	263 (97%)	7 (3%)	46	13
1	B	276/289 (96%)	274 (99%)	2 (1%)	84	66
2	C	60/61 (98%)	59 (98%)	1 (2%)	60	31
2	D	62/61 (102%)	57 (92%)	5 (8%)	11	1
All	All	668/700 (95%)	653 (98%)	15 (2%)	57	19

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	290[B]	MET
2	C	576	ARG
2	D	520[B]	ILE
1	A	290[A]	MET
2	D	520[A]	ILE

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

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 4 are monoatomic - leaving 5 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	$\# Z > 2$
5	PO4	A	505	3	4,4,4	1.02	0	6,6,6	1.20	0
5	PO4	B	504	3	4,4,4	1.13	0	6,6,6	0.90	0
4	GOL	A	503	-	5,5,5	1.08	0	5,5,5	0.83	0
4	GOL	B	503	-	5,5,5	0.73	0	5,5,5	0.89	0
4	GOL	A	504	-	5,5,5	1.31	1 (20%)	5,5,5	0.63	0

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
4	GOL	A	504	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	503	-	-	0/4/4/4	-
4	GOL	B	503	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	504	GOL	O2-C2	-2.20	1.36	1.43

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	504	GOL	C1-C2-C3-O3
4	A	504	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	503	GOL	1	0
4	A	504	GOL	1	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	305/338 (90%)	-0.16	10 (3%) 46 46	11, 16, 30, 49	0
1	B	307/338 (90%)	-0.17	7 (2%) 60 60	11, 17, 34, 57	0
2	C	65/70 (92%)	-0.15	0 100 100	14, 24, 31, 40	0
2	D	65/70 (92%)	0.36	7 (10%) 5 5	15, 26, 42, 43	0
All	All	742/816 (90%)	-0.12	24 (3%) 47 46	11, 18, 37, 57	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	525[A]	TYR	4.5
2	D	546	THR	3.6
1	B	299	ALA	3.5
2	D	580	PRO	3.4
1	B	24	PRO	3.3

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](#)

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
4	GOL	A	504	6/6	0.86	0.16	30,40,48,50	0
4	GOL	A	503	6/6	0.91	0.15	37,44,48,49	0
4	GOL	B	503	6/6	0.98	0.07	18,22,26,27	0
5	PO4	B	504	5/5	0.99	0.07	13,13,14,14	0
5	PO4	A	505	5/5	0.99	0.09	11,12,13,14	0
3	MN	B	501	1/1	1.00	0.10	12,12,12,12	0
3	MN	B	502	1/1	1.00	0.10	12,12,12,12	0
3	MN	A	501	1/1	1.00	0.11	11,11,11,11	0
3	MN	A	502	1/1	1.00	0.10	11,11,11,11	0

6.5 Other polymers [i](#)

There are no such residues in this entry.