



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:33 PM GMT

PDB ID : 4P8C
Title : Crystal structure of M. tuberculosis DprE1 in complex with the non-covalent inhibitor QN127
Authors : Neres, J.; Pojer, F.; Cole, S.T.
Deposited on : 2014-03-31
Resolution : 1.95 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

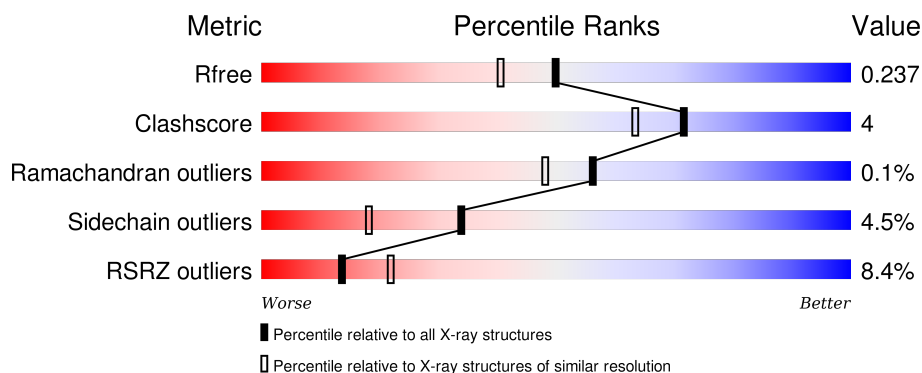
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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}	91344	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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 $\leq 5\%$. 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	480	<div> <div>8%</div> <div>83% 9% • 7%</div> </div>
1	B	480	<div> <div>8%</div> <div>74% 12% • 13%</div> </div>

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 7087 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 Probable decaprenylphosphoryl-beta-D-ribose oxidase.

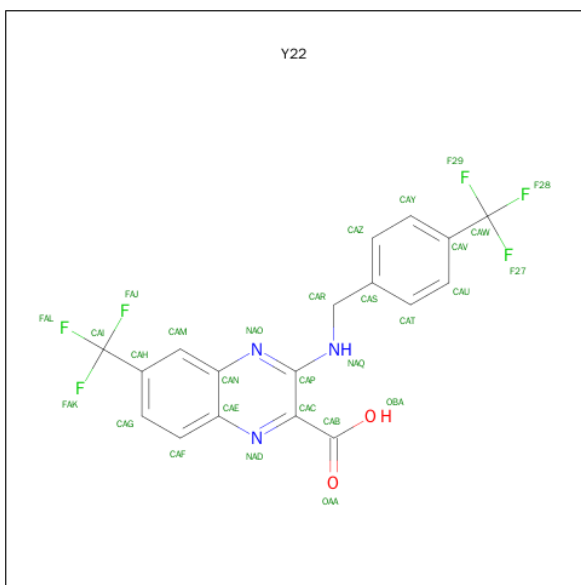
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	448	Total	C	N	O	S	0	0	0
			3447	2194	604	638	11			
1	B	417	Total	C	N	O	S	0	2	0
			3222	2044	574	594	10			

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



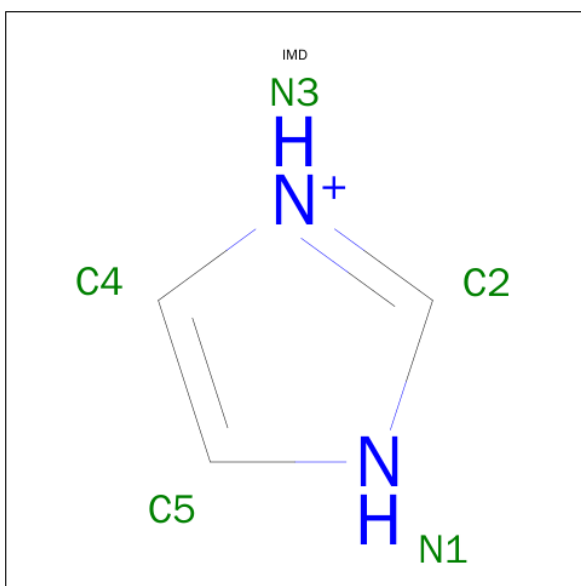
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is 6-(trifluoromethyl)-3-[[4-(trifluoromethyl)benzyl]amino}quinoxaline-2-carboxylic acid (three-letter code: Y22) (formula: $C_{18}H_{11}F_6N_3O_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	0	0
			29	18	6	3	2		
3	B	1	Total	C	F	N	O	0	0
			29	18	6	3	2		

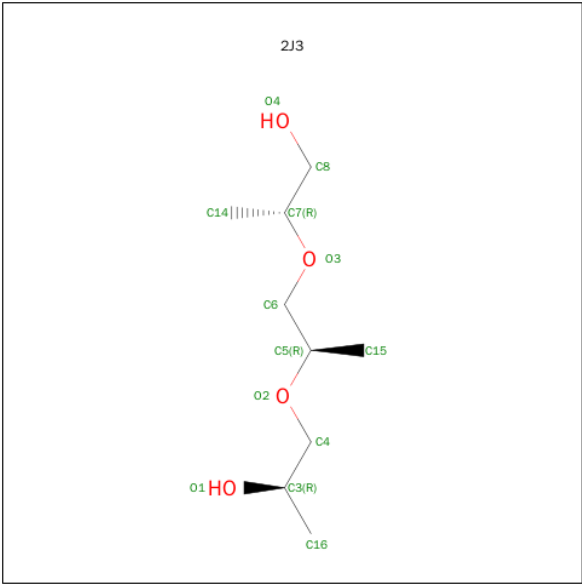
- Molecule 4 is IMIDAZOLE (three-letter code: IMD) (formula: $C_3H_5N_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	N	0	0
			5	3	2		

- Molecule 5 is (2R)-2-([(2R)-2-([(2R)-2-hydroxypropyl]oxy)propyl]oxy)propan-1-ol

(three-letter code: 2J3) (formula: C₉H₂₀O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			13	9	4		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	144	Total	O	0	0
			144	144		
6	B	92	Total	O	0	0
			92	92		

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 ($\text{RSRZ} > 2$). 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.

- Chain A:
-
- 8% 83% 9% 7%
- GLY SER SER HIS HIS HIS HIS HIS SER SER GLY LEU LEU PRO ARG GLY SER HIS MET LEU LEU SER VAL GLY ALA T7 T8 T9 A10 T11 R12 R13 A20 P21 S22 V23 A24 N25 V26 L27 E33 R41 G46 G47 Y60 I72 H81 I92 D93 P116 C117
- T118 R119 H132 D167 W172 G182 I183 L209 D210 S228 A229 W230 F231 G241 E254 Q255 L256 L260 E263 F267 D268 A1A P10 G1N LEU L274 L275 P276 D277 V278 F279 N281 L283 L295 W296 Y297 L317 D318 M319 F320 G321 E322 G323 A324
- ARG A326 Y327 G328 P329 A330 Q334 Y335 Q336 E346 F347 K348 K349 I350 F362 L363 N364 K367 A375 S378 R406 E409 F410 R413 L414 R421 R448 V449 F450 L461

- Chain B:
-
- 8% 74% 12% 13%
- ASP VAL PHE PRO ASN GLY LEU ASN LYS TYR THR PHE GLY PRO ILE GLY ASP MET PHE GLY TRP ASN ARG ALA GLY PRO ALA G331 Q334 Y335 Q336 D345 E346 F347 K348 K349 I350 Q355 F362 L363 N364 V367
- R372 N373 Q374 A375 S378 I392 R413 K418 R421 T422 T423 A424 H428 D435 V444 L447 A455 L461
- GLY SER SER HIS HIS HIS HIS HIS SER SER SER GLY LEU VAL PRO ARG GLY SER HIS MET LEU SER VAL GLY THR T9 A10 T11 R12 L13 W16 G17 R18 T19 A20 P21 S22 V23 A24 N25 V26 L27 R28 T29 E33 M34 K37 A40 R41 V42 A43 E44 THR S45 Q46

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	77.68Å 84.16Å 80.91Å 90.00° 103.17° 90.00°	Depositor
Resolution (Å)	49.24 – 1.95 49.24 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.9 (49.24-1.95) 97.9 (49.24-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.23 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.202 , 0.239 0.201 , 0.237	Depositor DCC
R_{free} test set	3640 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	35.0	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 37.0	EDS
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 72221 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7087	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 4.58% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: Y22, 2J3, IMD, FAD

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	3/3530 (0.1%)	0.61	0/4794
1	B	0.51	5/3302 (0.2%)	0.57	0/4480
All	All	0.52	8/6832 (0.1%)	0.59	0/9274

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	296	TRP	CD2-CE2	5.30	1.47	1.41
1	B	16	TRP	CD2-CE2	5.28	1.47	1.41
1	A	172	TRP	CD2-CE2	5.22	1.47	1.41
1	B	230	TRP	CD2-CE2	5.21	1.47	1.41
1	A	230	TRP	CD2-CE2	5.18	1.47	1.41

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	3447	0	3406	31	0
1	B	3222	0	3203	27	0
2	A	53	0	31	0	0
2	B	53	0	31	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	29	0	10	1	0
3	B	29	0	10	0	0
4	A	5	0	5	0	0
5	A	13	0	20	1	0
6	A	144	0	0	1	0
6	B	92	0	0	2	0
All	All	7087	0	6716	58	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 4.

The worst 5 of 58 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:406:ARG:HH11	1:A:406:ARG:HG2	1.49	0.75
1:B:8:THR:HG23	1:B:25:ASN:HB3	1.68	0.74
1:A:119:ARG:HB2	1:A:119:ARG:HH11	1.55	0.71
1:B:81:HIS:HE1	1:B:93:ASP:OD1	1.76	0.68
1:B:132:HIS:HD2	6:B:663:HOH:O	1.76	0.68

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	442/480 (92%)	433 (98%)	8 (2%)	1 (0%)	52	43
1	B	413/480 (86%)	410 (99%)	3 (1%)	0	100	100
All	All	855/960 (89%)	843 (99%)	11 (1%)	1 (0%)	56	48

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	PRO

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	358/384 (93%)	343 (96%)	15 (4%)	36	21
1	B	336/384 (88%)	319 (95%)	17 (5%)	29	13
All	All	694/768 (90%)	662 (95%)	32 (5%)	34	17

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	448	ARG
1	B	92	ILE
1	B	444	VAL
1	B	41	ARG
1	B	119	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	364	ASN
1	B	364	ASN
1	B	132	HIS
1	A	132	HIS
1	B	81	HIS

5.3.3 RNA ⓘ

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](#)

6 ligands are modelled in this entry.

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	$\# Z > 2$
2	FAD	A	501	-	48,58,58	1.30	6 (12%)	54,89,89	2.18	10 (18%)
3	Y22	A	502	-	27,31,31	1.15	2 (7%)	42,47,47	1.16	2 (4%)
4	IMD	A	503	-	3,5,5	0.42	0	4,5,5	0.52	0
5	2J3	A	504	-	12,12,12	0.75	0	11,14,14	0.85	1 (9%)
2	FAD	B	501	-	48,58,58	1.40	7 (14%)	54,89,89	2.38	13 (24%)
3	Y22	B	502	-	27,31,31	1.13	2 (7%)	42,47,47	1.15	2 (4%)

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
2	FAD	A	501	-	-	0/30/50/50	0/6/6/6
3	Y22	A	502	-	-	0/17/21/21	0/3/3/3
4	IMD	A	503	-	-	0/0/0/0	0/1/1/1
5	2J3	A	504	-	-	0/12/12/12	0/0/0/0
2	FAD	B	501	-	-	0/30/50/50	0/6/6/6
3	Y22	B	502	-	-	0/17/21/21	0/3/3/3

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	FAD	C4X-N5	2.12	1.36	1.33
3	B	502	Y22	CAC-NAD	2.26	1.35	1.32
2	B	501	FAD	C9A-N10	2.28	1.41	1.38
2	A	501	FAD	C9A-N10	2.32	1.42	1.38
3	A	502	Y22	CAC-NAD	2.51	1.36	1.32

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	FAD	N3A-C2A-N1A	-8.55	122.35	128.89
2	A	501	FAD	N3A-C2A-N1A	-8.05	122.73	128.89
2	B	501	FAD	C4-C4X-C10	-6.86	115.55	119.94
2	A	501	FAD	C4-C4X-C10	-5.82	116.22	119.94
3	A	502	Y22	CAC-CAP-NAO	-4.03	118.71	122.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	Y22	1	0
5	A	504	2J3	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	448/480 (93%)	0.41	36 (8%) 15 24	22, 35, 76, 104	0
1	B	417/480 (86%)	0.41	37 (8%) 12 19	24, 41, 71, 99	0
All	All	865/960 (90%)	0.41	73 (8%) 14 22	22, 37, 73, 104	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	47	GLY	9.6
1	A	283	LEU	7.4
1	A	281	ASN	7.3
1	A	323	TRP	7.0
1	A	279	PHE	6.9

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	Y22	B	502	29/29	0.93	0.13	0.30	40,44,58,60	0
2	FAD	A	501	53/53	0.97	0.12	0.22	24,26,29,30	0
5	2J3	A	504	13/13	0.81	0.13	0.20	55,57,62,62	0
2	FAD	B	501	53/53	0.97	0.10	-0.34	28,33,35,38	0
3	Y22	A	502	29/29	0.92	0.10	-0.60	34,38,53,54	0
4	IMD	A	503	5/5	0.97	0.11	-	41,42,42,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.