Full wwPDB X-ray Structure Validation Report ⓘ

Jul 7, 2022 – 04:12 pm BST

PDB ID	:	6HV6
Title	:	Crystal structure of PatoxP, a cysteine protease-like domain of Photorhabdus asymbiotica toxin PaTox
Authors	:	Bogdanovic, X.; Wirth, C.; Hunte, C.
Deposited on	:	2018-10-10
Resolution	:	2.00 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 types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

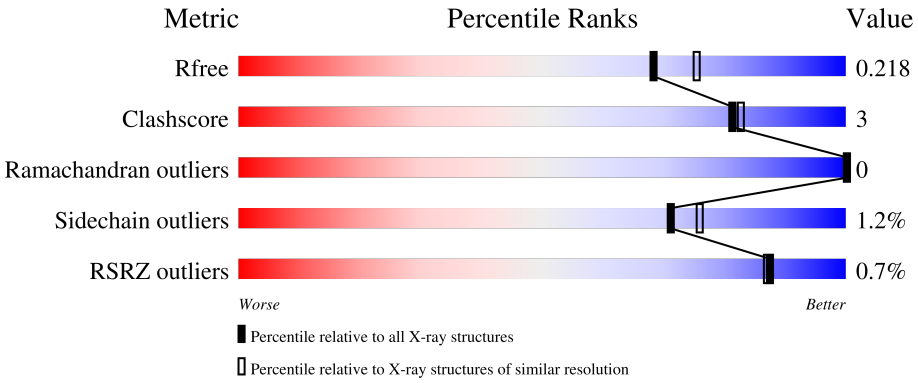
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	463	60% · 37%

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
2	EDO	A	2205	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2555 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 Toxin PAU_02230.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	293	2332	1478	390	454	10	8	0	0

There are 50 discrepancies between the modelled and reference sequences:

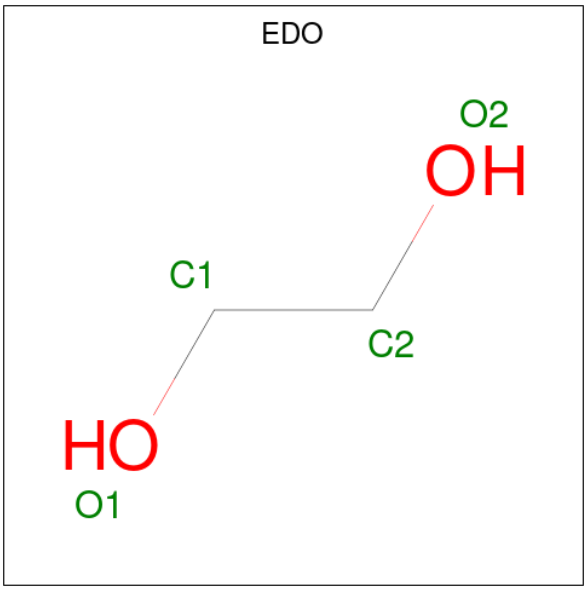
Chain	Residue	Modelled	Actual	Comment	Reference
A	1667	MET	-	initiating methionine	UNP C7BKP9
A	1668	GLY	-	expression tag	UNP C7BKP9
A	1669	SER	-	expression tag	UNP C7BKP9
A	1670	SER	-	expression tag	UNP C7BKP9
A	1671	HIS	-	expression tag	UNP C7BKP9
A	1672	HIS	-	expression tag	UNP C7BKP9
A	1673	HIS	-	expression tag	UNP C7BKP9
A	1674	HIS	-	expression tag	UNP C7BKP9
A	1675	HIS	-	expression tag	UNP C7BKP9
A	1676	HIS	-	expression tag	UNP C7BKP9
A	1677	SER	-	expression tag	UNP C7BKP9
A	1678	SER	-	expression tag	UNP C7BKP9
A	1679	GLY	-	expression tag	UNP C7BKP9
A	1680	LEU	-	expression tag	UNP C7BKP9
A	1681	VAL	-	expression tag	UNP C7BKP9
A	1682	PRO	-	expression tag	UNP C7BKP9
A	1683	ARG	-	expression tag	UNP C7BKP9
A	1684	GLY	-	expression tag	UNP C7BKP9
A	1685	SER	-	expression tag	UNP C7BKP9
A	1686	HIS	-	expression tag	UNP C7BKP9
A	1687	MET	-	expression tag	UNP C7BKP9
A	1688	ALA	-	expression tag	UNP C7BKP9
A	1689	SER	-	expression tag	UNP C7BKP9
A	1690	MET	-	expression tag	UNP C7BKP9
A	1691	THR	-	expression tag	UNP C7BKP9
A	1692	GLY	-	expression tag	UNP C7BKP9
A	1693	GLY	-	expression tag	UNP C7BKP9

Continued on next page...

Continued from previous page...

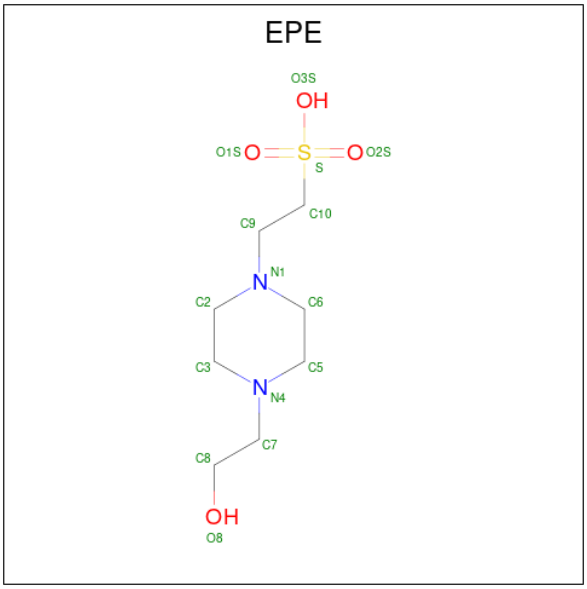
Chain	Residue	Modelled	Actual	Comment	Reference
A	1694	GLN	-	expression tag	UNP C7BKP9
A	1695	GLN	-	expression tag	UNP C7BKP9
A	1696	MET	-	expression tag	UNP C7BKP9
A	1697	GLY	-	expression tag	UNP C7BKP9
A	1698	ARG	-	expression tag	UNP C7BKP9
A	1699	GLY	-	expression tag	UNP C7BKP9
A	1700	SER	-	expression tag	UNP C7BKP9
A	1865	ALA	CYS	engineered mutation	UNP C7BKP9
A	2115	VAL	-	expression tag	UNP C7BKP9
A	2116	ASP	-	expression tag	UNP C7BKP9
A	2117	LYS	-	expression tag	UNP C7BKP9
A	2118	LEU	-	expression tag	UNP C7BKP9
A	2119	ALA	-	expression tag	UNP C7BKP9
A	2120	ALA	-	expression tag	UNP C7BKP9
A	2121	ALA	-	expression tag	UNP C7BKP9
A	2122	LEU	-	expression tag	UNP C7BKP9
A	2123	GLU	-	expression tag	UNP C7BKP9
A	2124	HIS	-	expression tag	UNP C7BKP9
A	2125	HIS	-	expression tag	UNP C7BKP9
A	2126	HIS	-	expression tag	UNP C7BKP9
A	2127	HIS	-	expression tag	UNP C7BKP9
A	2128	HIS	-	expression tag	UNP C7BKP9
A	2129	HIS	-	expression tag	UNP C7BKP9

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O S 15 8 2 4 1	8	0

- Molecule 4 is water.

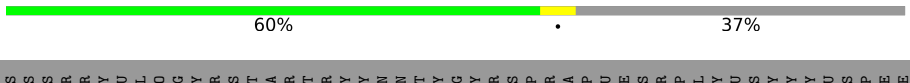
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	184	Total O 184 184	0	0

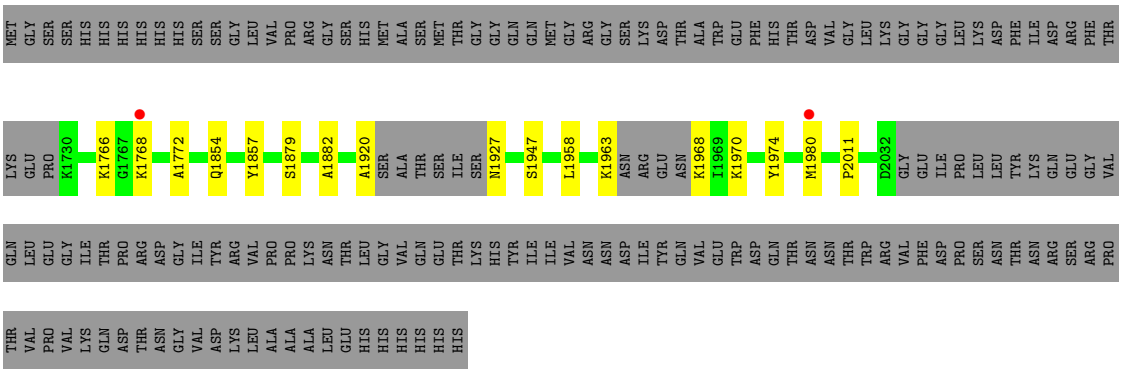
3 Residue-property plots (i)

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: Toxin PAU_02230

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.40Å 76.03Å 76.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.86 – 2.00 53.86 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.3 (53.86-2.00) 98.3 (53.86-2.00)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.175 , 0.214 0.180 , 0.218	Depositor DCC
R_{free} test set	690 reflections (2.75%)	wwPDB-VP
Wilson B-factor (Å^2)	31.8	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.008 for -h,l,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2555	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 6.89% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EPE, EDO

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	0/2378	0.48	0/3200

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

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	2332	0	2310	12	0
2	A	24	0	36	5	0
3	A	15	0	17	0	0
4	A	184	0	0	0	0
All	All	2555	0	2363	12	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 3.

All (12) 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:1882:ALA:H	2:A:2205:EDO:H11	1.30	0.94
1:A:1768:LYS:H	1:A:1768:LYS:HD2	1.68	0.58
1:A:1882:ALA:N	2:A:2205:EDO:H11	2.12	0.50
1:A:1882:ALA:HB3	2:A:2205:EDO:H22	1.97	0.47
1:A:1947:SER:HB3	1:A:1958:LEU:HD11	1.96	0.47
1:A:1854:GLN:HA	1:A:1857:TYR:CE1	2.51	0.46
1:A:1974:TYR:HB2	1:A:1980:MET:HG2	2.00	0.44
1:A:1968:LYS:HG2	1:A:1970:LYS:HE3	1.99	0.43
1:A:1879:SER:C	2:A:2205:EDO:H12	2.38	0.43
1:A:1772:ALA:HB3	2:A:2203:EDO:H22	2.02	0.42
1:A:1920:ALA:HB3	1:A:2011:PRO:HB2	2.02	0.42
1:A:1970:LYS:HD3	1:A:1970:LYS:HA	1.85	0.41

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	287/463 (62%)	283 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	257/404 (64%)	254 (99%)	3 (1%)	71 76

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1766	LYS
1	A	1927	ASN
1	A	1963	LYS

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

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

7 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	EDO	A	2201	-	3,3,3	0.48	0	2,2,2	0.31	0
2	EDO	A	2205	-	3,3,3	0.49	0	2,2,2	0.18	0
3	EPE	A	2207	-	15,15,15	0.62	1 (6%)	18,20,20	0.61	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	EDO	A	2203	-	3,3,3	0.37	0	2,2,2	0.53	0
2	EDO	A	2204	-	3,3,3	0.45	0	2,2,2	0.42	0
2	EDO	A	2206	-	3,3,3	0.44	0	2,2,2	0.38	0
2	EDO	A	2202	-	3,3,3	0.44	0	2,2,2	0.36	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
2	EDO	A	2201	-	-	0/1/1/1	-
2	EDO	A	2205	-	-	1/1/1/1	-
3	EPE	A	2207	-	-	4/9/19/19	0/1/1/1
2	EDO	A	2203	-	-	0/1/1/1	-
2	EDO	A	2204	-	-	0/1/1/1	-
2	EDO	A	2206	-	-	0/1/1/1	-
2	EDO	A	2202	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2207	EPE	O1S-S	2.25	1.51	1.45

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2207	EPE	C8-C7-N4-C5
3	A	2207	EPE	C10-C9-N1-C2
3	A	2207	EPE	C8-C7-N4-C3
3	A	2207	EPE	C10-C9-N1-C6
2	A	2205	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2205	EDO	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2203	EDO	1	0

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, 95^{th} 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	293/463 (63%)	-0.45	2 (0%)	87	87	20, 37, 65, 86	4 (1%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1980	MET	2.7
1	A	1768	LYS	2.0

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

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, 95^{th} 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(Å^2)	Q<0.9
2	EDO	A	2202	4/4	0.82	0.27	62,68,70,76	0
2	EDO	A	2206	4/4	0.83	0.18	53,57,59,63	0
2	EDO	A	2203	4/4	0.90	0.15	62,64,66,69	0
2	EDO	A	2201	4/4	0.91	0.13	44,48,54,64	0
2	EDO	A	2204	4/4	0.92	0.15	57,59,65,70	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	EDO	A	2205	4/4	0.95	0.17	39,42,48,55	0
3	EPE	A	2207	15/15	0.97	0.10	25,40,51,57	8

6.5 Other polymers (i)

There are no such residues in this entry.