wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 04:32 am BST

PDB ID	:	4BRP
Title	:	Legionella pneumophila NTPDase1 crystal form V (part-open)
Authors	:	Zebisch, M.; Schaefer, P.; Lauble, P.; Straeter, N.
Deposited on	:	2013-06-04
Resolution	:	2.50 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

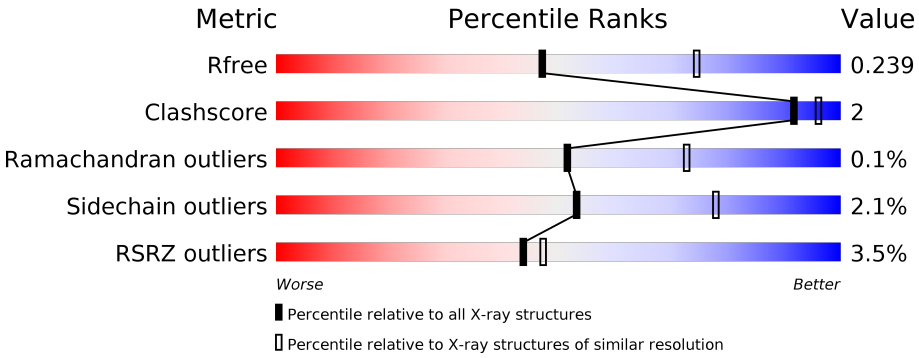
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 <=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	368	3% 90% 6%
1	B	368	4% 90% 6%
1	C	368	4% 90% 7%
1	D	368	3% 89% 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11107 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 ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	355	Total	C	N	O	S	0	0	0
			2804	1788	463	539	14			
1	B	347	Total	C	N	O	S	0	0	0
			2728	1739	449	526	14			
1	C	357	Total	C	N	O	S	0	1	0
			2808	1788	465	541	14			
1	D	353	Total	C	N	O	S	0	2	0
			2746	1750	454	528	14			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	MET	-	expression tag	UNP Q5ZUA2
A	137	ASP	GLU	conflict	UNP Q5ZUA2
A	149	VAL	ALA	conflict	UNP Q5ZUA2
A	394	LEU	-	expression tag	UNP Q5ZUA2
A	395	GLU	-	expression tag	UNP Q5ZUA2
A	396	HIS	-	expression tag	UNP Q5ZUA2
A	397	HIS	-	expression tag	UNP Q5ZUA2
A	398	HIS	-	expression tag	UNP Q5ZUA2
A	399	HIS	-	expression tag	UNP Q5ZUA2
A	400	HIS	-	expression tag	UNP Q5ZUA2
A	401	HIS	-	expression tag	UNP Q5ZUA2
B	34	MET	-	expression tag	UNP Q5ZUA2
B	137	ASP	GLU	conflict	UNP Q5ZUA2
B	149	VAL	ALA	conflict	UNP Q5ZUA2
B	394	LEU	-	expression tag	UNP Q5ZUA2
B	395	GLU	-	expression tag	UNP Q5ZUA2
B	396	HIS	-	expression tag	UNP Q5ZUA2
B	397	HIS	-	expression tag	UNP Q5ZUA2
B	398	HIS	-	expression tag	UNP Q5ZUA2
B	399	HIS	-	expression tag	UNP Q5ZUA2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	400	HIS	-	expression tag	UNP Q5ZUA2
B	401	HIS	-	expression tag	UNP Q5ZUA2
C	34	MET	-	expression tag	UNP Q5ZUA2
C	137	ASP	GLU	conflict	UNP Q5ZUA2
C	149	VAL	ALA	conflict	UNP Q5ZUA2
C	394	LEU	-	expression tag	UNP Q5ZUA2
C	395	GLU	-	expression tag	UNP Q5ZUA2
C	396	HIS	-	expression tag	UNP Q5ZUA2
C	397	HIS	-	expression tag	UNP Q5ZUA2
C	398	HIS	-	expression tag	UNP Q5ZUA2
C	399	HIS	-	expression tag	UNP Q5ZUA2
C	400	HIS	-	expression tag	UNP Q5ZUA2
C	401	HIS	-	expression tag	UNP Q5ZUA2
D	34	MET	-	expression tag	UNP Q5ZUA2
D	137	ASP	GLU	conflict	UNP Q5ZUA2
D	149	VAL	ALA	conflict	UNP Q5ZUA2
D	394	LEU	-	expression tag	UNP Q5ZUA2
D	395	GLU	-	expression tag	UNP Q5ZUA2
D	396	HIS	-	expression tag	UNP Q5ZUA2
D	397	HIS	-	expression tag	UNP Q5ZUA2
D	398	HIS	-	expression tag	UNP Q5ZUA2
D	399	HIS	-	expression tag	UNP Q5ZUA2
D	400	HIS	-	expression tag	UNP Q5ZUA2
D	401	HIS	-	expression tag	UNP Q5ZUA2

- Molecule 2 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Br 2 2	0	0
2	A	2	Total Br 2 2	0	0
2	D	1	Total Br 1 1	0	0
2	C	2	Total Br 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	5	Total O 5 5	0	0

Continued on next page...

Continued from previous page...

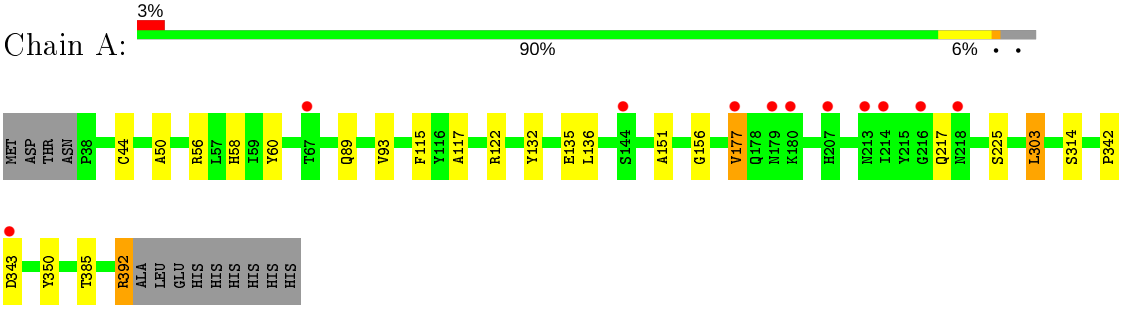
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	3	Total O 3 3	0	0
3	D	6	Total O 6 6	0	0

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.

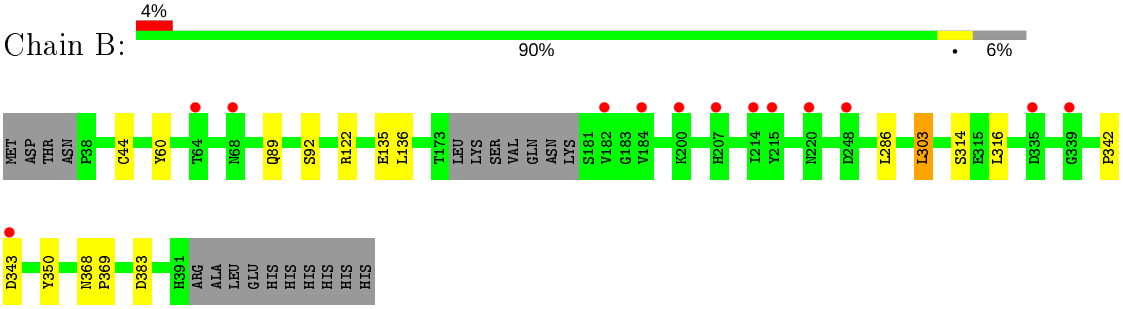
• Molecule 1: ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE I

Chain A:



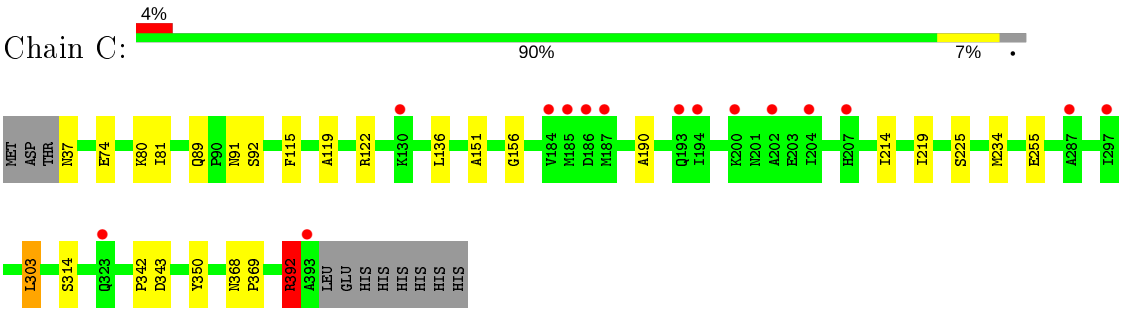
• Molecule 1: ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE I

Chain B:



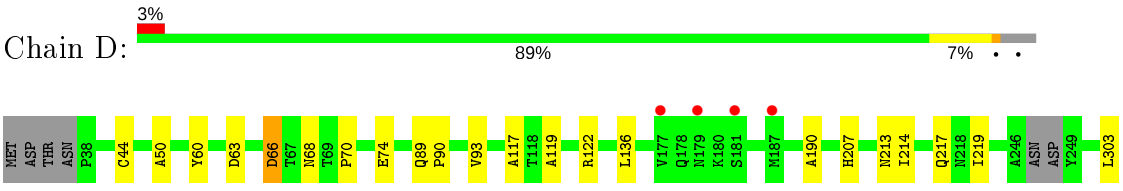
• Molecule 1: ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE I

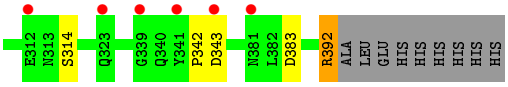
Chain C:



• Molecule 1: ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE I

Chain D:




E312
N313
S314
Q323
G339
Q340
Y341
P342
D343
N381
L382
D383
R392
ALA
LEU
GLU
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	129.61Å 129.61Å 162.70Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.08 – 2.50 29.08 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.08-2.50) 99.3 (29.08-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.23 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.177 , 0.219 0.199 , 0.239	Depositor DCC
R_{free} test set	1397 reflections (2.54%)	wwPDB-VP
Wilson B-factor (Å^2)	54.8	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 35.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtriage
Reported twinning fraction	0.889 for H, K, L 0.111 for -h,-k,l	Depositor
Outliers	0 of 55087 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11107	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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: BR

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	0.94	4/2880 (0.1%)	0.91	4/3931 (0.1%)
1	B	0.90	0/2802	0.90	3/3825 (0.1%)
1	C	0.89	0/2887	1.02	7/3945 (0.2%)
1	D	0.94	0/2827	0.95	5/3863 (0.1%)
All	All	0.92	4/11396 (0.0%)	0.95	19/15564 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	132	TYR	CE1-CZ	-7.02	1.29	1.38
1	A	132	TYR	CG-CD1	-6.94	1.30	1.39
1	A	132	TYR	CE2-CZ	-6.57	1.30	1.38
1	A	132	TYR	CG-CD2	-5.30	1.32	1.39

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	122	ARG	NE-CZ-NH2	-19.65	110.48	120.30
1	C	122	ARG	NE-CZ-NH1	17.34	128.97	120.30
1	D	122	ARG	NE-CZ-NH1	13.15	126.88	120.30
1	A	122	ARG	NE-CZ-NH1	12.60	126.60	120.30
1	D	122	ARG	NE-CZ-NH2	-11.74	114.43	120.30

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	2804	0	2646	10	0
1	B	2728	0	2549	6	1
1	C	2808	0	2625	11	1
1	D	2746	0	2535	10	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	1	0	0	0	0
3	A	5	0	0	0	0
3	C	3	0	0	0	0
3	D	6	0	0	0	0
All	All	11107	0	10355	35	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 2.

The worst 5 of 35 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:D:213:ASN:HA	1:D:217:GLN:O	2.05	0.56
1:A:342:PRO:O	1:A:343:ASP:HB2	2.08	0.54
1:D:214:ILE:HG22	1:D:219:ILE:HD12	1.89	0.53
1:C:368:ASN:OD1	1:C:369:PRO:HD2	2.12	0.50
1:D:342:PRO:O	1:D:343:ASP:CB	2.57	0.49

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)
1:B:135:GLU:OE2	1:C:74:GLU:OE2[5_545]	2.12	0.08

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	353/368 (96%)	343 (97%)	9 (2%)	1 (0%)	41 61
1	B	343/368 (93%)	333 (97%)	10 (3%)	0	100 100
1	C	356/368 (97%)	345 (97%)	11 (3%)	0	100 100
1	D	351/368 (95%)	338 (96%)	12 (3%)	1 (0%)	41 61
All	All	1403/1472 (95%)	1359 (97%)	42 (3%)	2 (0%)	51 73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	177	VAL
1	D	90	PRO

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	306/328 (93%)	300 (98%)	6 (2%)	55 79
1	B	295/328 (90%)	290 (98%)	5 (2%)	60 82
1	C	303/328 (92%)	297 (98%)	6 (2%)	55 79
1	D	291/328 (89%)	283 (97%)	8 (3%)	44 71
All	All	1195/1312 (91%)	1170 (98%)	25 (2%)	53 78

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	89	GLN
1	C	136	LEU
1	D	314	SER
1	C	92	SER
1	C	303	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	292	ASN
1	C	260	ASN
1	B	331	HIS
1	A	280	GLN
1	C	207	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

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	355/368 (96%)	-0.11	11 (3%)	49	52	34, 56, 92, 118	0
1	B	347/368 (94%)	0.03	13 (3%)	41	45	38, 61, 102, 119	0
1	C	357/368 (97%)	-0.00	15 (4%)	36	39	35, 60, 98, 119	0
1	D	353/368 (95%)	-0.02	10 (2%)	53	56	31, 58, 106, 141	1 (0%)
All	All	1412/1472 (95%)	-0.03	49 (3%)	44	47	31, 58, 100, 141	1 (0%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	179	ASN	5.0
1	B	64	THR	4.7
1	A	216	GLY	4.6
1	B	184	VAL	4.5
1	B	343	ASP	3.6

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 carbohydrates 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	BR	B	1393	1/1	0.77	0.22	116,116,116,116	0
2	BR	A	1393	1/1	0.85	0.27	93,93,93,93	0
2	BR	C	1394	1/1	0.92	0.34	116,116,116,116	0
2	BR	D	1393	1/1	0.96	0.06	60,60,60,60	0
2	BR	A	1394	1/1	0.98	0.07	49,49,49,49	0
2	BR	C	1395	1/1	1.00	0.03	44,44,44,44	0
2	BR	B	1392	1/1	1.00	0.04	52,52,52,52	0

6.5 Other polymers (i)

There are no such residues in this entry.