Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 10:42 PM GMT

PDB ID	:	1UW4
Title	:	THE STRUCTURAL BASIS OF THE INTERACTION BETWEEN NONSENSE MEDIATED DECAY FACTORS UPF2 AND UPF3
Authors	:	Kadlec, J.; Izaurralde, E.; Cusack, S.
Deposited on	:	2004-01-29
Resolution	:	1.95 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
http://wwpdb.org/validation/2016/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:

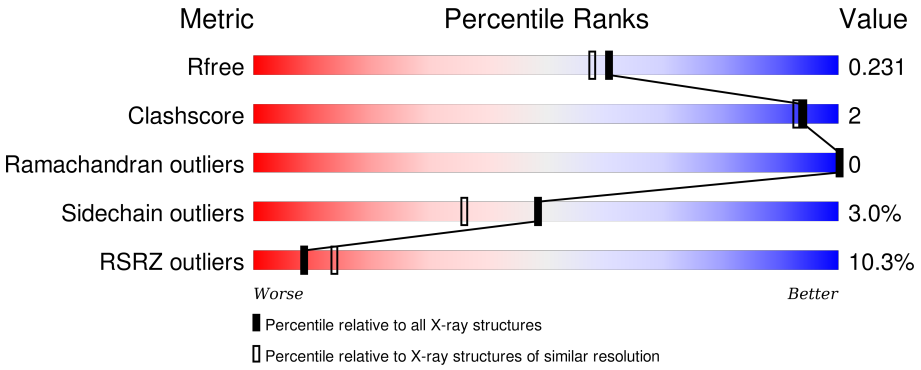
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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 <=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	91	27% 87% 13%
1	C	91	13% 90% 8%
2	B	248	8% 93% 6%
2	D	248	6% 96%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5996 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.

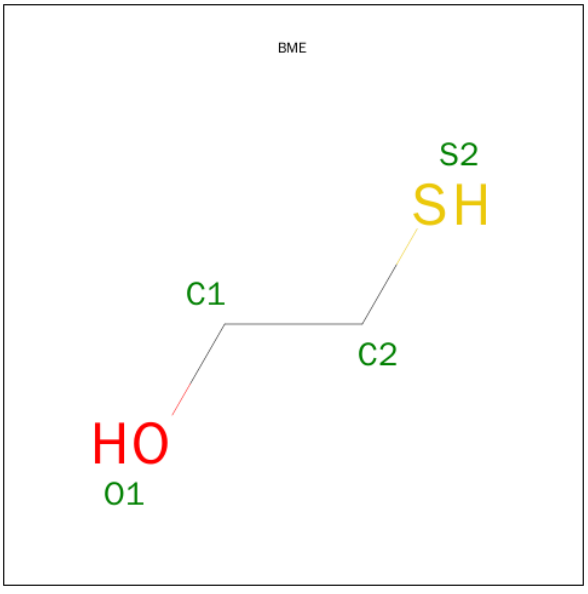
- Molecule 1 is a protein called UPF3X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	91	Total	C	N	O	S	0	0	0
			771	504	127	138	2			
1	C	91	Total	C	N	O	S	0	0	0
			772	504	127	139	2			

- Molecule 2 is a protein called REGULATOR OF NONSENSE TRANSCRIPTS 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	248	Total	C	N	O	S	0	0	0
			2057	1329	350	366	12			
2	D	248	Total	C	N	O	S	0	3	0
			2080	1341	356	371	12			

- Molecule 3 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total S 1 1	0	0
3	D	1	Total S 1 1	0	0

- Molecule 4 is water.

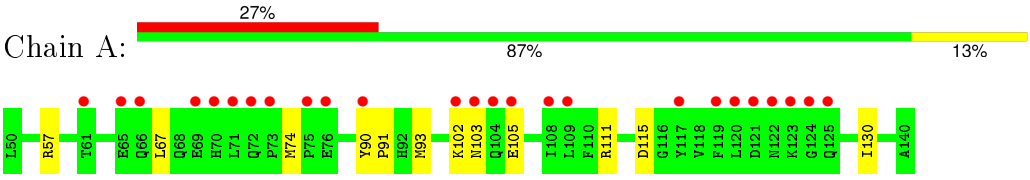
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	18	Total O 18 18	0	0
4	B	121	Total O 121 121	0	0
4	C	24	Total O 24 24	0	0
4	D	151	Total O 151 151	0	0

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 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 (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.

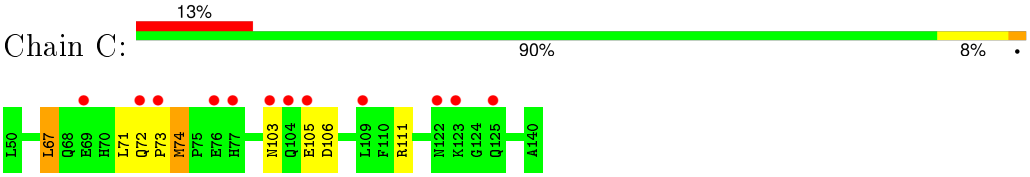
- Molecule 1: UPF3X

Chain A:



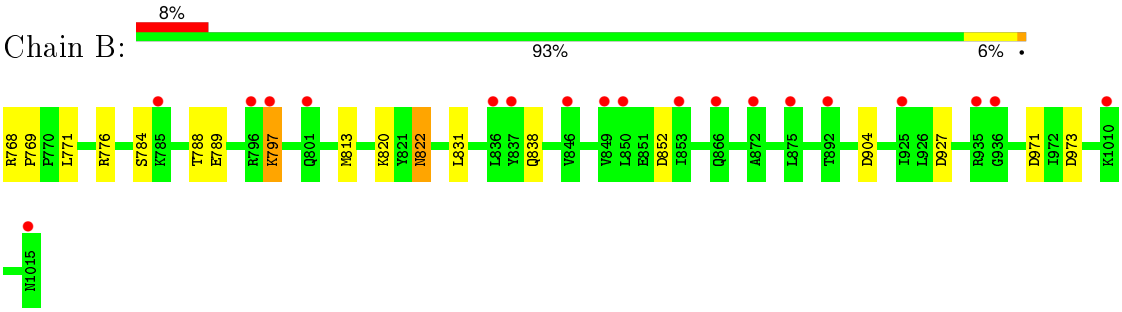
- Molecule 1: UPF3X

Chain C:



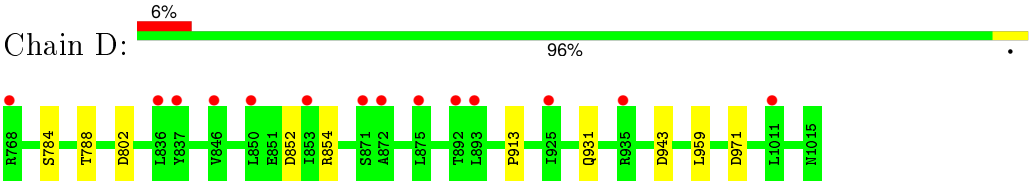
- Molecule 2: REGULATOR OF NONSENSE TRANSCRIPTS 2

Chain B:



- Molecule 2: REGULATOR OF NONSENSE TRANSCRIPTS 2

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	70.31Å 100.18Å 153.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.95 33.68 – 1.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-1.95) 99.5 (33.68-1.95)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.55 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.204 , 0.227 0.211 , 0.231	Depositor DCC
R_{free} test set	2379 reflections (3.09%)	DCC
Wilson B-factor (Å^2)	30.8	Xtriage
Anisotropy	0.324	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 52.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Outliers	0 of 79316 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5996	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.34 % of the origin peak, indicating pseudo translational symmetry.The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.6812e-03.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: BME

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	0.43	0/795	0.70	1/1077 (0.1%)
1	C	0.46	0/796	0.73	1/1078 (0.1%)
2	B	0.48	0/2104	0.70	4/2847 (0.1%)
2	D	0.50	0/2127	0.70	3/2877 (0.1%)
All	All	0.48	0/5822	0.70	9/7879 (0.1%)

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	802	ASP	CB-CG-OD2	6.26	123.93	118.30
2	B	971	ASP	CB-CG-OD2	6.08	123.77	118.30
1	A	115	ASP	CB-CG-OD2	5.78	123.50	118.30
2	D	971	ASP	CB-CG-OD2	5.71	123.44	118.30
1	C	106	ASP	CB-CG-OD2	5.32	123.09	118.30
2	B	904	ASP	CB-CG-OD2	5.09	122.89	118.30
2	D	943	ASP	CB-CG-OD2	5.09	122.88	118.30
2	B	927	ASP	CB-CG-OD2	5.07	122.86	118.30
2	B	973	ASP	CB-CG-OD2	5.04	122.83	118.30

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	771	0	741	6	0
1	C	772	0	741	3	0
2	B	2057	0	2085	7	0
2	D	2080	0	2105	4	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
4	A	18	0	0	1	0
4	B	121	0	0	0	0
4	C	24	0	0	0	0
4	D	151	0	0	2	0
All	All	5996	0	5672	20	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 2.

All (20) 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:90:TYR:HA	1:A:93:MET:CE	2.18	0.71
2:B:797:LYS:C	2:B:797:LYS:HD2	2.25	0.57
1:A:90:TYR:HA	1:A:93:MET:HE3	1.88	0.55
2:B:822:ASN:HD22	2:B:822:ASN:H	1.56	0.54
1:C:72:GLN:HA	1:C:73:PRO:C	2.32	0.50
2:D:784:SER:O	2:D:788:THR:HG23	2.12	0.50
2:B:820:LYS:HE2	2:B:822:ASN:HD21	1.77	0.48
2:B:784:SER:O	2:B:788:THR:HG23	2.13	0.48
1:A:57:ARG:HD3	1:A:130:ILE:HD12	1.95	0.48
1:A:91:PRO:HA	1:A:93:MET:HE3	1.97	0.47
2:B:813:MET:SD	2:B:831:LEU:HD13	2.55	0.46
1:C:67:LEU:HD22	1:C:71:LEU:HG	1.97	0.46
2:B:797:LYS:C	2:B:797:LYS:CD	2.84	0.45
1:A:91:PRO:HD3	4:A:2006:HOH:O	2.17	0.44
1:C:72:GLN:HA	1:C:74:MET:N	2.33	0.43
2:D:931:GLN:HG3	4:D:2090:HOH:O	2.19	0.42
2:D:854[B]:ARG:NH2	4:D:2030:HOH:O	2.52	0.42
1:A:90:TYR:HA	1:A:93:MET:HE1	2.01	0.41
2:B:768:ARG:HA	2:B:769:PRO:HD3	1.97	0.41
2:D:913:PRO:HA	2:D:959:LEU:HD13	2.02	0.41

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	89/91 (98%)	88 (99%)	1 (1%)	0	100	100
1	C	89/91 (98%)	88 (99%)	1 (1%)	0	100	100
2	B	246/248 (99%)	241 (98%)	5 (2%)	0	100	100
2	D	249/248 (100%)	246 (99%)	3 (1%)	0	100	100
All	All	673/678 (99%)	663 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	83/83 (100%)	77 (93%)	6 (7%)	18	6
1	C	83/83 (100%)	78 (94%)	5 (6%)	24	10
2	B	234/234 (100%)	227 (97%)	7 (3%)	48	36
2	D	237/234 (101%)	236 (100%)	1 (0%)	93	93
All	All	637/634 (100%)	618 (97%)	19 (3%)	48	36

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	67	LEU
1	A	74	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	102	LYS
1	A	103	ASN
1	A	105	GLU
1	A	111	ARG
2	B	771	LEU
2	B	776	ARG
2	B	789	GLU
2	B	797	LYS
2	B	822	ASN
2	B	838	GLN
2	B	852	ASP
1	C	67	LEU
1	C	74	MET
1	C	103	ASN
1	C	105	GLU
1	C	111	ARG
2	D	852	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	100	ASN
2	B	772	GLN
2	B	822	ASN
2	B	861	GLN
2	B	865	ASN
1	C	72	GLN
1	C	100	ASN
1	C	103	ASN
1	C	104	GLN
2	D	772	GLN
2	D	794	GLN
2	D	838	GLN
2	D	861	GLN
2	D	865	ASN

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 2 ligands modelled in this entry, 2 are modelled with single atom - 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

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	91/91 (100%)	1.17	25 (27%) 1 0	25, 31, 38, 39	0
1	C	91/91 (100%)	0.83	12 (13%) 4 8	21, 29, 37, 40	0
2	B	248/248 (100%)	0.41	19 (7%) 16 25	20, 27, 34, 39	0
2	D	248/248 (100%)	0.39	14 (5%) 28 39	19, 25, 33, 39	0
All	All	678/678 (100%)	0.56	70 (10%) 9 14	19, 27, 36, 40	0

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	122	ASN	5.1
1	A	120	LEU	4.8
1	C	122	ASN	4.5
2	D	853	ILE	4.3
2	D	837	TYR	3.9
2	B	853	ILE	3.9
2	B	837	TYR	3.9
1	A	72	GLN	3.9
1	A	104	GLN	3.9
2	B	836	LEU	3.8
1	A	124	GLY	3.8
1	A	102	LYS	3.7
1	A	125	GLN	3.6
1	C	105	GLU	3.6
2	D	935	ARG	3.5
1	A	105	GLU	3.5
2	D	850	LEU	3.5
2	B	935	ARG	3.4
1	A	109	LEU	3.4
1	C	104	GLN	3.4
1	C	73	PRO	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	875	LEU	3.1
2	B	850	LEU	3.1
1	A	70	HIS	3.0
1	C	125	GLN	3.0
1	C	103	ASN	2.9
1	A	117	TYR	2.8
1	A	71	LEU	2.8
1	A	119	PHE	2.8
1	A	123	LYS	2.7
1	A	75	PRO	2.7
1	A	73	PRO	2.6
1	A	108	ILE	2.6
1	A	66	GLN	2.6
1	A	121	ASP	2.6
1	C	109	LEU	2.5
1	A	69	GLU	2.5
1	A	76	GLU	2.5
2	B	1015	ASN	2.4
1	C	76	GLU	2.4
1	C	77	HIS	2.4
2	D	925	ILE	2.4
2	B	801	GLN	2.4
2	D	836	LEU	2.3
2	B	872	ALA	2.3
1	C	69	GLU	2.3
1	A	90	TYR	2.3
1	A	61	THR	2.3
1	C	72	GLN	2.3
2	B	785	LYS	2.3
2	D	875	LEU	2.3
2	B	849	VAL	2.3
2	D	846	VAL	2.3
1	A	65	GLU	2.3
2	D	892	THR	2.3
2	D	871	SER	2.2
2	D	893	LEU	2.2
2	B	866	GLN	2.2
1	A	103	ASN	2.2
2	B	797	LYS	2.1
2	B	925	ILE	2.1
2	D	768	ARG	2.1
2	D	872	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	123	LYS	2.1
2	B	796	ARG	2.0
2	D	1011	LEU	2.0
2	B	846	VAL	2.0
2	B	892	THR	2.0
2	B	1010	LYS	2.0
2	B	936	GLY	2.0

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, 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	LLDF	B-factors(Å^2)	Q<0.9
3	BME	D	3016	1/4	0.84	0.11	-	46,46,46,46	0
3	BME	B	3016	1/4	0.93	0.13	-	48,48,48,48	0

6.5 Other polymers (i)

There are no such residues in this entry.