



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 10:20 pm BST

PDB ID : 2II1  
Title : Crystal structure of Acetamidase (10172637) from Bacillus Halodurans at 1.95 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2006-09-27  
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](mailto: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 : 2.11  
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

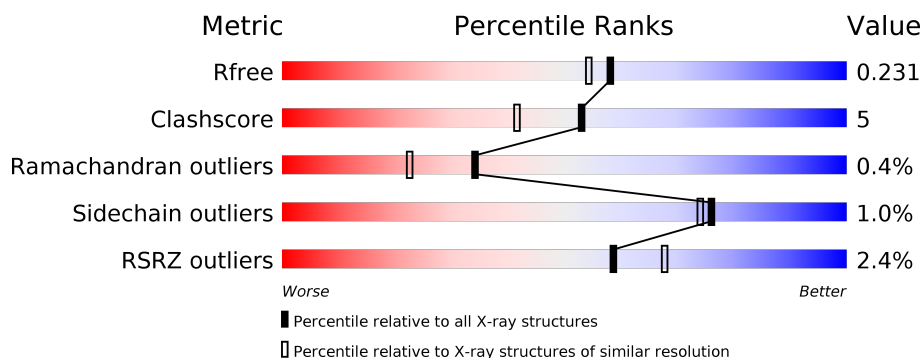
# 1 Overall quality at a glance

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))
$R_{free}$	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (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  $\geq 3$ , 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	301	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>13%</div> <div></div> </div> <div></div> </div>
1	B	301	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>13%</div> <div></div> </div> <div></div> </div>
1	C	301	<div> <div>4%</div> <div> <div></div> <div>84%</div> <div>14%</div> <div></div> </div> <div></div> </div>
1	D	301	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>10%</div> <div></div> </div> <div></div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9233 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 Acetamidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	Se	0	2	0
			2198	1395	363	428	3	9			
1	B	296	Total	C	N	O	S	Se	0	3	1
			2192	1393	358	429	3	9			
1	C	296	Total	C	N	O	S	Se	0	1	0
			2196	1392	360	432	3	9			
1	D	297	Total	C	N	O	S	Se	0	0	0
			2189	1388	357	431	3	10			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	LEADER SEQUENCE	UNP Q9KGN3
A	1	MSE	MET	modified residue	UNP Q9KGN3
A	13	MSE	MET	modified residue	UNP Q9KGN3
A	73	MSE	MET	modified residue	UNP Q9KGN3
A	89	MSE	MET	modified residue	UNP Q9KGN3
A	130	MSE	MET	modified residue	UNP Q9KGN3
A	186	MSE	MET	modified residue	UNP Q9KGN3
A	213	MSE	MET	modified residue	UNP Q9KGN3
A	226	MSE	MET	modified residue	UNP Q9KGN3
A	245	MSE	MET	modified residue	UNP Q9KGN3
A	262	MSE	MET	modified residue	UNP Q9KGN3
B	0	GLY	-	LEADER SEQUENCE	UNP Q9KGN3
B	1	MSE	MET	modified residue	UNP Q9KGN3
B	13	MSE	MET	modified residue	UNP Q9KGN3
B	73	MSE	MET	modified residue	UNP Q9KGN3
B	89	MSE	MET	modified residue	UNP Q9KGN3
B	130	MSE	MET	modified residue	UNP Q9KGN3
B	186	MSE	MET	modified residue	UNP Q9KGN3
B	213	MSE	MET	modified residue	UNP Q9KGN3
B	226	MSE	MET	modified residue	UNP Q9KGN3
B	245	MSE	MET	modified residue	UNP Q9KGN3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	262	MSE	MET	modified residue	UNP Q9KGN3
C	0	GLY	-	LEADER SEQUENCE	UNP Q9KGN3
C	1	MSE	MET	modified residue	UNP Q9KGN3
C	13	MSE	MET	modified residue	UNP Q9KGN3
C	73	MSE	MET	modified residue	UNP Q9KGN3
C	89	MSE	MET	modified residue	UNP Q9KGN3
C	130	MSE	MET	modified residue	UNP Q9KGN3
C	186	MSE	MET	modified residue	UNP Q9KGN3
C	213	MSE	MET	modified residue	UNP Q9KGN3
C	226	MSE	MET	modified residue	UNP Q9KGN3
C	245	MSE	MET	modified residue	UNP Q9KGN3
C	262	MSE	MET	modified residue	UNP Q9KGN3
D	0	GLY	-	LEADER SEQUENCE	UNP Q9KGN3
D	1	MSE	MET	modified residue	UNP Q9KGN3
D	13	MSE	MET	modified residue	UNP Q9KGN3
D	73	MSE	MET	modified residue	UNP Q9KGN3
D	89	MSE	MET	modified residue	UNP Q9KGN3
D	130	MSE	MET	modified residue	UNP Q9KGN3
D	186	MSE	MET	modified residue	UNP Q9KGN3
D	213	MSE	MET	modified residue	UNP Q9KGN3
D	226	MSE	MET	modified residue	UNP Q9KGN3
D	245	MSE	MET	modified residue	UNP Q9KGN3
D	262	MSE	MET	modified residue	UNP Q9KGN3

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	5	Total Ca 5 5	0	0
2	A	2	Total Ca 2 2	0	0
2	D	3	Total Ca 3 3	0	0
2	C	5	Total Ca 5 5	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	115	Total O 115 115	0	0

*Continued on next page...*

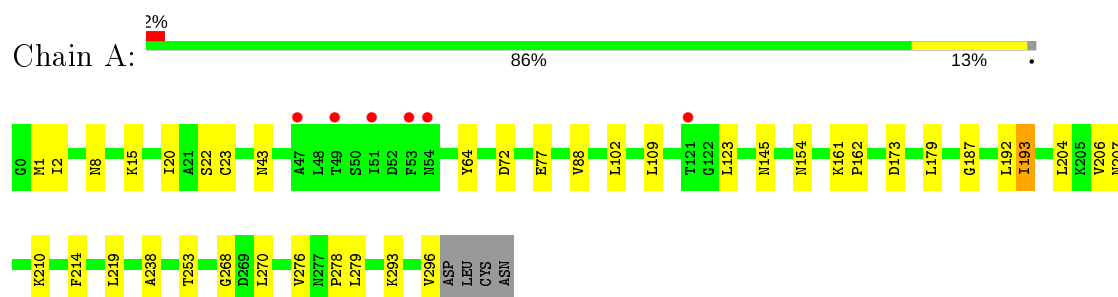
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	104	Total 104	O 104	0	0
3	C	106	Total 106	O 106	0	0
3	D	118	Total 118	O 118	0	0

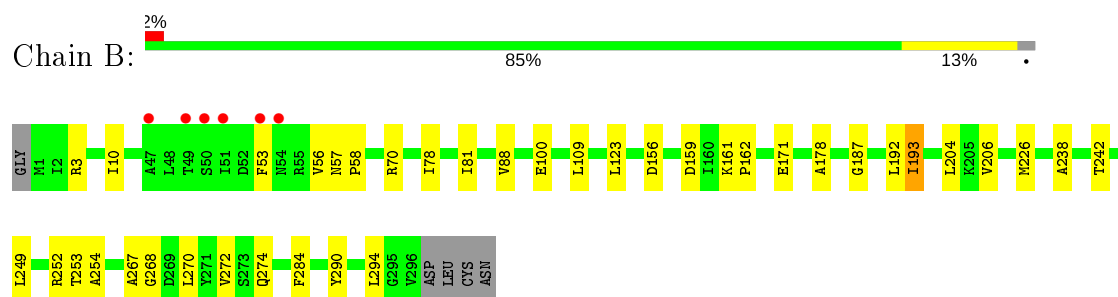
### 3 Residue-property plots [i](#)

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.

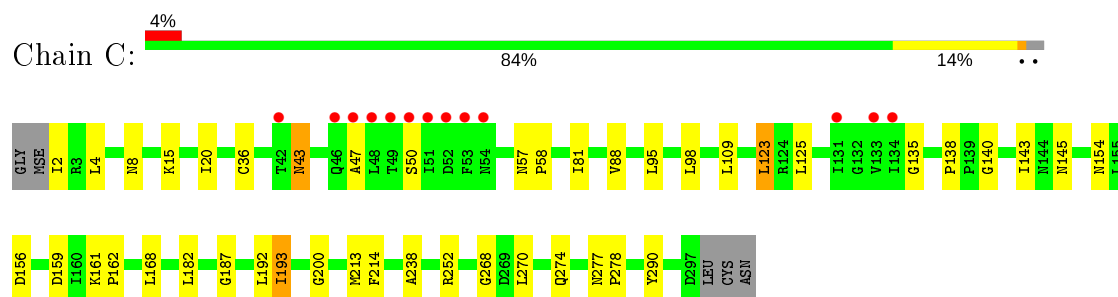
#### • Molecule 1: Acetamidase



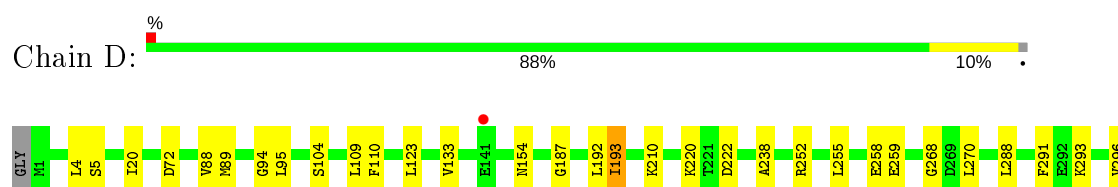
#### • Molecule 1: Acetamidase



#### • Molecule 1: Acetamidase



#### • Molecule 1: Acetamidase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.76Å 69.61Å 94.98Å 73.95° 88.91° 86.63°	Depositor
Resolution (Å)	28.89 – 1.95 28.89 – 1.95	Depositor EDS
% Data completeness (in resolution range)	90.2 (28.89-1.95) 90.2 (28.89-1.95)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.175 , 0.228 0.180 , 0.231	Depositor DCC
$R_{free}$ test set	3594 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.7	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 52.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9233	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.88% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.65	0/2231	0.73	1/3019 (0.0%)
1	B	0.64	0/2229	0.71	1/3018 (0.0%)
1	C	0.62	0/2227	0.69	0/3016
1	D	0.66	0/2217	0.72	0/3002
All	All	0.64	0/8904	0.71	2/12055 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	173	ASP	CB-CG-OD1	5.72	123.44	118.30
1	B	249	LEU	CA-CB-CG	5.27	127.42	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1	MSE	CA

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	2198	0	2224	30	0
1	B	2192	0	2214	27	0
1	C	2196	0	2205	31	0
1	D	2189	0	2198	21	0
2	A	2	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	3	0	0	0	0
3	A	115	0	0	0	0
3	B	104	0	0	1	0
3	C	106	0	0	2	0
3	D	118	0	0	1	0
All	All	9233	0	8841	95	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 5.

The worst 5 of 95 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:88:VAL:HG22	1:A:109:LEU:CD2	1.91	1.00
1:A:88:VAL:HG22	1:A:109:LEU:HD23	1.44	0.96
1:D:88:VAL:HG22	1:D:109:LEU:HD23	1.60	0.82
1:C:4:LEU:HD23	1:C:20:ILE:HD11	1.65	0.77
1:A:123:LEU:HD23	1:D:293:LYS:HD3	1.65	0.77

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	297/301 (99%)	287 (97%)	8 (3%)	2 (1%)	22	11
1	B	297/301 (99%)	285 (96%)	11 (4%)	1 (0%)	41	30
1	C	295/301 (98%)	285 (97%)	9 (3%)	1 (0%)	41	30
1	D	295/301 (98%)	285 (97%)	9 (3%)	1 (0%)	41	30
All	All	1184/1204 (98%)	1142 (96%)	37 (3%)	5 (0%)	34	22

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	ILE
1	B	193	ILE
1	C	193	ILE
1	D	193	ILE
1	A	1	MSE

### 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	237/234 (101%)	234 (99%)	3 (1%)	69	65
1	B	237/234 (101%)	236 (100%)	1 (0%)	91	90
1	C	237/234 (101%)	234 (99%)	3 (1%)	69	65
1	D	236/234 (101%)	234 (99%)	2 (1%)	81	80
All	All	947/936 (101%)	938 (99%)	9 (1%)	76	74

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	43	ASN
1	D	222	ASP
1	C	214	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	214	PHE
1	C	123	LEU

Some 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 15 ligands modelled in this entry, 15 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 ⓘ

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	287/301 (95%)	-0.14	6 (2%) 63 72	6, 14, 25, 35	0
1	B	286/301 (95%)	-0.12	6 (2%) 63 72	9, 16, 25, 33	0
1	C	287/301 (95%)	-0.02	13 (4%) 33 43	8, 16, 29, 39	0
1	D	287/301 (95%)	-0.27	2 (0%) 87 92	6, 13, 25, 41	0
All	All	1147/1204 (95%)	-0.14	27 (2%) 59 68	6, 15, 26, 41	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	51	ILE	5.5
1	A	47	ALA	4.7
1	C	47	ALA	4.2
1	C	48	LEU	4.2
1	C	53	PHE	4.1

### 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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	CA	C	403	1/1	0.96	0.08	31,31,31,31	0
2	CA	D	402	1/1	0.97	0.04	19,19,19,19	0
2	CA	B	402	1/1	0.98	0.06	30,30,30,30	0
2	CA	C	400	1/1	0.99	0.09	21,21,21,21	0
2	CA	C	402	1/1	0.99	0.05	16,16,16,16	0
2	CA	D	401	1/1	0.99	0.11	13,13,13,13	0
2	CA	C	404	1/1	0.99	0.04	28,28,28,28	0
2	CA	A	400	1/1	0.99	0.06	15,15,15,15	0
2	CA	B	403	1/1	0.99	0.04	19,19,19,19	0
2	CA	B	401	1/1	0.99	0.10	15,15,15,15	0
2	CA	C	401	1/1	0.99	0.08	18,18,18,18	0
2	CA	B	400	1/1	1.00	0.07	17,17,17,17	0
2	CA	B	404	1/1	1.00	0.02	19,19,19,19	0
2	CA	A	401	1/1	1.00	0.09	13,13,13,13	0
2	CA	D	400	1/1	1.00	0.09	16,16,16,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.