



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 03:24 pm GMT

PDB ID : 3N1G  
Title : Crystal structure of DhhN bound to BOCFn3  
Authors : Kavran, J.M.; Leahy, D.J.  
Deposited on : 2010-05-15  
Resolution : 1.90 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

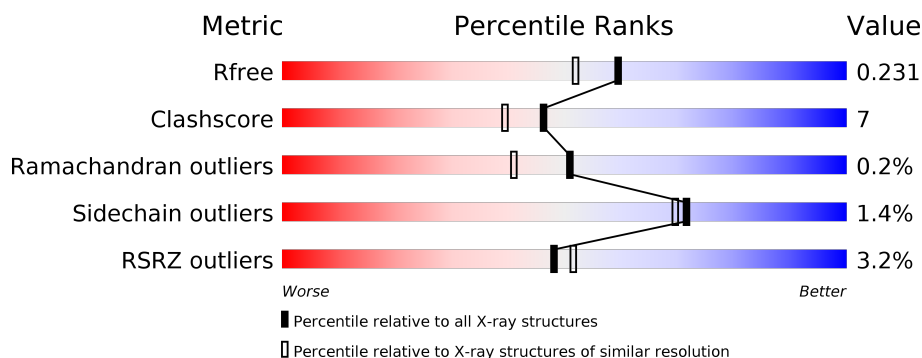
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	100719	5047 (1.90-1.90)
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)
RSRZ outliers	101464	5100 (1.90-1.90)

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. 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	170	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -10px; left: 0;">%</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>79%</span> <span>9%</span> <span>11%</span> </div> </div>
1	B	170	<div> <div style="width: 100%; height: 10px; background-color: green; position: relative;"> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>77%</span> <span>12%</span> <span>11%</span> </div> </div>
2	C	111	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -10px; left: 0;">7%</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>68%</span> <span>25%</span> <span>• 6%</span> </div> </div>
2	D	111	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -10px; left: 0;">5%</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>70%</span> <span>17%</span> <span>•• 11%</span> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4816 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 Desert hedgehog protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	151	Total	C	N	O	S	0	1	0
			1214	757	224	229	4			
1	A	151	Total	C	N	O	S	0	0	0
			1204	751	223	226	4			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	GLY	-	EXPRESSION TAG	UNP O43323
B	21	SER	-	EXPRESSION TAG	UNP O43323
B	22	GLY	-	EXPRESSION TAG	UNP O43323
B	23	PRO	-	EXPRESSION TAG	UNP O43323
A	20	GLY	-	EXPRESSION TAG	UNP O43323
A	21	SER	-	EXPRESSION TAG	UNP O43323
A	22	GLY	-	EXPRESSION TAG	UNP O43323
A	23	PRO	-	EXPRESSION TAG	UNP O43323

- Molecule 2 is a protein called Brother of CDO.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	104	Total	C	N	O	S	0	3	0
			861	545	137	171	8			
2	D	99	Total	C	N	O	S	0	1	0
			820	522	130	161	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	707	GLY	-	EXPRESSION TAG	UNP Q9BWV1
C	708	SER	-	EXPRESSION TAG	UNP Q9BWV1
C	709	THR	-	EXPRESSION TAG	UNP Q9BWV1
D	707	GLY	-	EXPRESSION TAG	UNP Q9BWV1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	708	SER	-	EXPRESSION TAG	UNP Q9BWV1
D	709	THR	-	EXPRESSION TAG	UNP Q9BWV1

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Zn 1 1	0	0
4	A	1	Total Zn 1 1	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mg 1 1	0	0


- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	252	Total O 252 252	0	0
6	C	99	Total O 99 99	0	0
6	A	239	Total O 239 239	0	0
6	D	120	Total O 120 120	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.

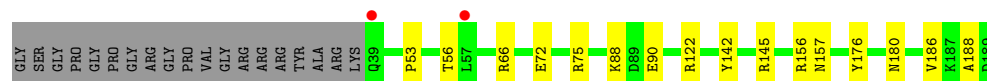
- Molecule 1: Desert hedgehog protein

Chain B: 



- Molecule 1: Desert hedgehog protein

Chain A: 



- Molecule 2: Brother of CDO

Chain C: 



- Molecule 2: Brother of CDO

Chain D: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.21Å 92.98Å 74.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.65 – 1.90 24.65 – 1.90	Depositor EDS
% Data completeness (in resolution range)	86.5 (24.65-1.90) 96.3 (24.65-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 1.90Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.182 , 0.222 0.194 , 0.231	Depositor DCC
$R_{free}$ test set	2365 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.6	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 57.7	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	0.94	EDS
Total number of atoms	4816	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% 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: MG, CA, ZN

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	0.38	0/1231	0.60	0/1668
1	B	0.39	0/1244	0.61	1/1685 (0.1%)
2	C	0.40	0/894	0.50	0/1211
2	D	0.34	0/846	0.55	1/1145 (0.1%)
All	All	0.38	0/4215	0.58	2/5709 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	731	LEU	CA-CB-CG	5.18	127.21	115.30
1	B	189	ASP	CB-CG-OD2	5.11	122.90	118.30

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	1204	0	1159	11	0
1	B	1214	0	1171	12	0
2	C	861	0	809	23	0
2	D	820	0	767	14	0
3	A	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	2	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	1	0	0	0	0
6	A	239	0	0	3	0
6	B	252	0	0	3	0
6	C	99	0	0	1	0
6	D	120	0	0	2	0
All	All	4816	0	3906	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 7.

All (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)
2:D:724:VAL:HG22	2:D:728:THR:HB	1.64	0.80
2:D:713:VAL:HG23	2:D:714:ALA:H	1.55	0.70
2:C:740:ASN:HB3	2:C:743:THR:HG22	1.76	0.66
2:D:794:ASN:HB2	6:D:496:HOH:O	1.97	0.63
2:C:724:VAL:HG22	2:C:728:THR:HB	1.81	0.63
2:C:730[A]:MET:SD	2:C:774:TRP:CE3	2.95	0.59
2:C:740:ASN:HB3	2:C:743:THR:CG2	2.31	0.59
2:C:734:MET:HE2	6:C:255:HOH:O	2.04	0.58
2:C:713:VAL:O	2:C:800:GLU:HG2	2.05	0.57
1:A:145:ARG:HD2	1:A:186:VAL:O	2.05	0.56
2:D:735:TYR:O	2:D:737:PRO:HD3	2.05	0.56
2:C:812:ARG:O	2:C:813:LYS:C	2.45	0.55
2:C:724:VAL:CG2	2:C:728:THR:HB	2.39	0.53
1:B:115:MET:HE3	6:B:400:HOH:O	2.08	0.53
2:C:813:LYS:O	2:C:813:LYS:HG2	2.09	0.53
2:C:730[A]:MET:SD	2:C:774:TRP:HE3	2.32	0.52
1:B:156:ARG:HH21	1:B:180:ASN:ND2	2.07	0.52
1:B:104:LYS:HE3	1:B:105[B]:GLU:OE2	2.09	0.52
1:A:156:ARG:HH21	1:A:180:ASN:ND2	2.08	0.51
1:A:145:ARG:HD3	1:A:188:ALA:HA	1.93	0.51
1:A:66:ARG:HG3	1:A:142:TYR:HB3	1.94	0.50
1:A:157:ASN:ND2	6:A:246:HOH:O	2.43	0.50
1:B:137:GLN:O	1:B:138:ASP:HB2	2.11	0.50
2:C:726:GLU:HG2	2:C:727:THR:HG23	1.94	0.49
1:B:134:HIS:HB3	2:C:805:MET:CE	2.42	0.48

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:746:HIS:HB2	2:C:793:PHE:CE1	2.48	0.48
2:D:716:PRO:HG2	2:D:790:MET:HG3	1.98	0.45
2:D:723:ALA:HB2	2:D:808:GLU:O	2.16	0.45
1:A:156:ARG:HH21	1:A:180:ASN:HD22	1.66	0.44
1:A:72:GLU:OE1	1:A:75:ARG:HD2	2.18	0.44
1:B:118:TRP:CZ2	1:B:161:LEU:HD23	2.51	0.44
2:D:743:THR:HA	2:D:744:PRO:HD3	1.88	0.44
2:D:749:TYR:CG	2:D:765:LYS:HD2	2.53	0.44
1:A:156:ARG:NH1	6:A:289:HOH:O	2.50	0.44
1:A:53:PRO:HB2	1:A:56:THR:HG23	2.00	0.44
2:D:712:PRO:HB2	2:D:799:SER:HB3	1.99	0.44
1:A:88:LYS:HE3	1:A:90:GLU:HB3	1.99	0.43
1:B:55:ARG:NH1	6:B:645:HOH:O	2.51	0.43
2:C:713:VAL:HG13	2:C:714:ALA:H	1.83	0.43
2:C:723:ALA:HB2	2:C:808:GLU:O	2.18	0.43
1:B:187:LYS:HE2	6:B:640:HOH:O	2.18	0.43
2:C:720:PHE:CZ	2:C:722:ASP:HB2	2.53	0.43
2:D:721:THR:HG22	2:D:731:LEU:HB3	2.01	0.43
1:B:88:LYS:HE2	2:C:755:THR:HG23	2.00	0.43
1:B:45:TYR:O	1:B:46:LYS:HB2	2.19	0.42
2:C:740:ASN:CB	2:C:743:THR:HG22	2.44	0.42
2:C:712:PRO:HD3	2:C:794:ASN:ND2	2.35	0.42
2:C:782:PRO:HG2	2:C:812:ARG:NH2	2.35	0.42
1:A:122:ARG:NH2	6:A:555:HOH:O	2.50	0.41
2:C:746:HIS:CD2	2:C:794:ASN:HA	2.55	0.41
2:D:749:TYR:CD2	2:D:765:LYS:HD2	2.55	0.41
2:C:753:ARG:HG2	2:C:754:PRO:O	2.21	0.41
1:B:156:ARG:HE	1:B:180:ASN:ND2	2.19	0.41
2:D:743:THR:HG22	6:D:496:HOH:O	2.20	0.41
2:C:710:GLU:HG3	2:C:711:ARG:N	2.36	0.41
1:B:57:LEU:HD23	1:B:57:LEU:HA	1.86	0.40
2:D:726:GLU:HG2	2:D:727:THR:HG23	2.03	0.40
2:D:724:VAL:CG2	2:D:728:THR:HB	2.43	0.40

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	149/170 (88%)	147 (99%)	2 (1%)	0	100	100
1	B	150/170 (88%)	148 (99%)	2 (1%)	0	100	100
2	C	105/111 (95%)	99 (94%)	5 (5%)	1 (1%)	18	7
2	D	96/111 (86%)	93 (97%)	3 (3%)	0	100	100
All	All	500/562 (89%)	487 (97%)	12 (2%)	1 (0%)	51	41

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	812	ARG

### 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	125/138 (91%)	124 (99%)	1 (1%)	85	85
1	B	127/138 (92%)	124 (98%)	3 (2%)	54	47
2	C	98/100 (98%)	98 (100%)	0	100	100
2	D	92/100 (92%)	90 (98%)	2 (2%)	57	51
All	All	442/476 (93%)	436 (99%)	6 (1%)	71	69

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	39	GLN
1	B	157	ASN
1	B	176	TYR
1	A	176	TYR
2	D	724	VAL
2	D	731	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	180	ASN
2	C	791	GLN
1	A	180	ASN
2	D	779	HIS

### 5.3.3 RNA [i](#)

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](#)

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<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	151/170 (88%)	-0.10	2 (1%) 77 80	14, 21, 40, 78	0
1	B	151/170 (88%)	-0.10	0 100 100	14, 20, 35, 54	0
2	C	104/111 (93%)	0.64	8 (7%) 14 16	18, 33, 59, 89	0
2	D	99/111 (89%)	0.32	6 (6%) 22 25	17, 29, 57, 65	0
All	All	505/562 (89%)	0.14	16 (3%) 48 51	14, 25, 53, 89	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	711	ARG	8.6
2	D	813	LYS	7.8
2	C	812	ARG	6.4
2	C	710	GLU	6.1
2	C	741	ASN	5.9
2	C	813	LYS	5.9
2	D	812	ARG	5.3
2	C	795	GLU	3.7
1	A	39	GLN	3.5
2	D	773	TYR	3.0
2	D	743	THR	3.0
2	D	736	ILE	2.5
2	C	773	TYR	2.5
2	C	767	MET	2.4
1	A	57	LEU	2.2
2	D	744	PRO	2.1

### 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 [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<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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	MG	A	193	1/1	0.97	0.17	1.67	30,30,30,30	0
3	CA	B	190	1/1	1.00	0.10	-0.13	19,19,19,19	0
4	ZN	B	192	1/1	0.91	0.10	-0.59	41,41,41,41	0
3	CA	B	191	1/1	1.00	0.07	-1.40	17,17,17,17	0
3	CA	A	192	1/1	1.00	0.06	-1.61	16,16,16,16	0
3	CA	A	190	1/1	1.00	0.07	-1.91	20,20,20,20	0
4	ZN	A	191	1/1	0.95	0.08	-2.36	39,39,39,39	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.