



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 08:55 PM GMT

PDB ID : 4Z8L  
Title : Crystal structure of DCAF1/SIV-MND VPX/MND SAMHD1 NTD ternary complex  
Authors : Koharudin, L.M.; Wu, Y.; Calero, G.; Ahn, J.; Gronenborn, A.M.  
Deposited on : 2015-04-09  
Resolution : 2.60 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  
<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  
Mogul : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

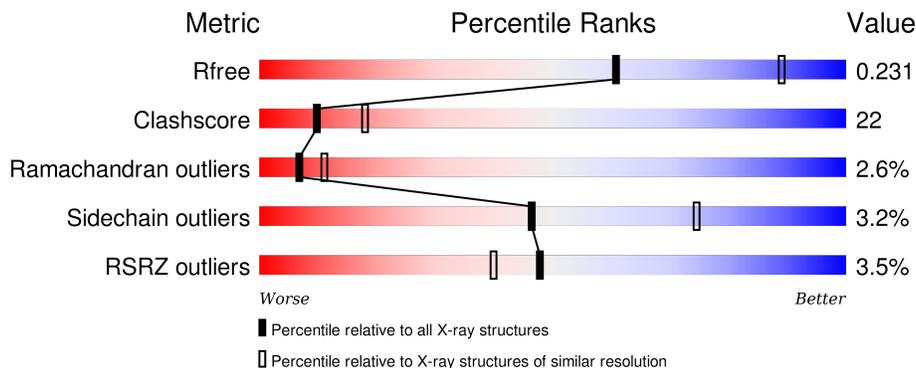
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.



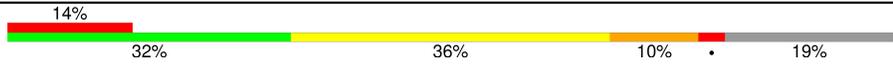
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	349	
1	D	349	
2	B	100	
2	E	100	
3	C	118	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	118	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into five segments: a red segment (14%), a green segment (32%), a yellow segment (36%), an orange segment (10%), and a grey segment (19%). A small black dot is located on the orange segment.</p>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7828 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 Protein VPRBP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	305	2432	1538	419	458	17	0	0	0
1	D	305	2432	1538	419	458	17	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1056	ALA	-	expression tag	UNP Q9Y4B6
A	1397	LEU	-	expression tag	UNP Q9Y4B6
A	1398	GLU	-	expression tag	UNP Q9Y4B6
A	1399	HIS	-	expression tag	UNP Q9Y4B6
A	1400	HIS	-	expression tag	UNP Q9Y4B6
A	1401	HIS	-	expression tag	UNP Q9Y4B6
A	1402	HIS	-	expression tag	UNP Q9Y4B6
A	1403	HIS	-	expression tag	UNP Q9Y4B6
A	1404	HIS	-	expression tag	UNP Q9Y4B6
D	1056	ALA	-	expression tag	UNP Q9Y4B6
D	1397	LEU	-	expression tag	UNP Q9Y4B6
D	1398	GLU	-	expression tag	UNP Q9Y4B6
D	1399	HIS	-	expression tag	UNP Q9Y4B6
D	1400	HIS	-	expression tag	UNP Q9Y4B6
D	1401	HIS	-	expression tag	UNP Q9Y4B6
D	1402	HIS	-	expression tag	UNP Q9Y4B6
D	1403	HIS	-	expression tag	UNP Q9Y4B6
D	1404	HIS	-	expression tag	UNP Q9Y4B6

- Molecule 2 is a protein called Vpx protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	79	676	429	125	116	6	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	E	79	676	429	125	116	6	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	SER	-	expression tag	UNP Q7ZB17
B	2	GLU	-	expression tag	UNP Q7ZB17
B	3	PHE	-	expression tag	UNP Q7ZB17
B	93	LEU	-	expression tag	UNP Q7ZB17
B	94	GLU	-	expression tag	UNP Q7ZB17
B	95	HIS	-	expression tag	UNP Q7ZB17
B	96	HIS	-	expression tag	UNP Q7ZB17
B	97	HIS	-	expression tag	UNP Q7ZB17
B	98	HIS	-	expression tag	UNP Q7ZB17
B	99	HIS	-	expression tag	UNP Q7ZB17
B	100	HIS	-	expression tag	UNP Q7ZB17
E	1	SER	-	expression tag	UNP Q7ZB17
E	2	GLU	-	expression tag	UNP Q7ZB17
E	3	PHE	-	expression tag	UNP Q7ZB17
E	93	LEU	-	expression tag	UNP Q7ZB17
E	94	GLU	-	expression tag	UNP Q7ZB17
E	95	HIS	-	expression tag	UNP Q7ZB17
E	96	HIS	-	expression tag	UNP Q7ZB17
E	97	HIS	-	expression tag	UNP Q7ZB17
E	98	HIS	-	expression tag	UNP Q7ZB17
E	99	HIS	-	expression tag	UNP Q7ZB17
E	100	HIS	-	expression tag	UNP Q7ZB17

- Molecule 3 is a protein called SAM domain and HD domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	92	738	463	135	138	2	0	0	0
3	F	95	758	474	139	143	2	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	SER	-	expression tag	UNP H6WEA4
C	-1	GLU	-	expression tag	UNP H6WEA4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	PHE	-	expression tag	UNP H6WEA4
F	-2	SER	-	expression tag	UNP H6WEA4
F	-1	GLU	-	expression tag	UNP H6WEA4
F	0	PHE	-	expression tag	UNP H6WEA4

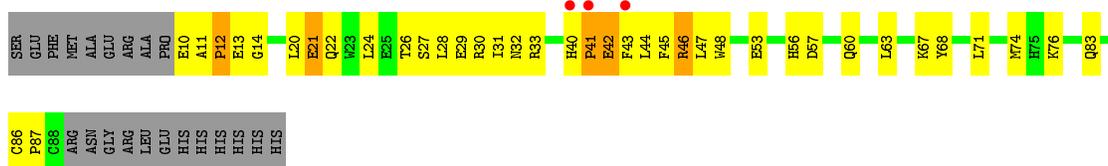
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

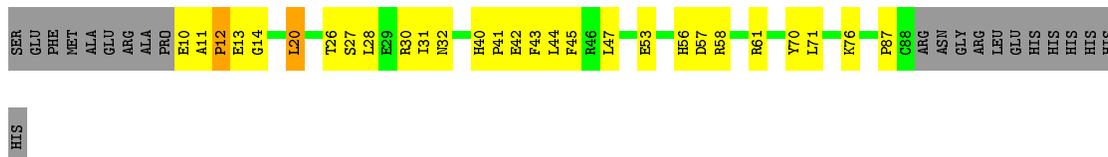
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	29	Total O 29 29	0	0
5	B	12	Total O 12 12	0	0
5	C	11	Total O 11 11	0	0
5	D	45	Total O 45 45	0	0
5	E	9	Total O 9 9	0	0
5	F	8	Total O 8 8	0	0

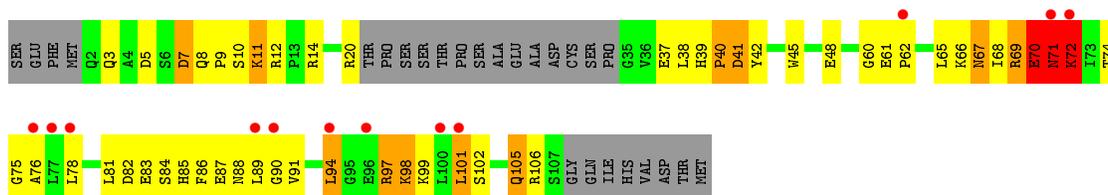




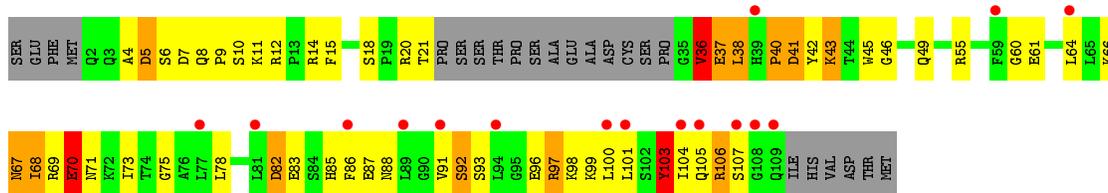
• Molecule 2: Vpx protein



• Molecule 3: SAM domain and HD domain-containing protein



• Molecule 3: SAM domain and HD domain-containing protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.49Å 74.49Å 178.18Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.70 – 2.60 43.69 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.3 (43.70-2.60) 98.3 (43.69-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.199 , 0.228 0.201 , 0.231	Depositor DCC
$R_{free}$ test set	1003 reflections (3.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.7	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 35.3	EDS
Estimated twinning fraction	0.068 for -h,-k,l 0.467 for h,-h-k,-l 0.068 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 33447 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7828	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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, 0.375 respectively for untwinned datasets, and 0.333, 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:  
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	0.26	0/2489	0.48	0/3373
1	D	0.27	0/2489	0.52	1/3373 (0.0%)
2	B	0.38	0/698	0.71	0/943
2	E	0.30	0/698	0.57	0/943
3	C	0.58	0/754	1.16	7/1013 (0.7%)
3	F	0.54	0/774	1.10	6/1040 (0.6%)
All	All	0.36	0/7902	0.69	14/10685 (0.1%)

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
3	C	0	6
3	F	0	3
All	All	0	9

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	94	LEU	CA-CB-CG	8.40	134.61	115.30
3	C	72	LYS	N-CA-C	8.10	132.86	111.00
3	F	67	ASN	N-CA-C	6.99	129.88	111.00
3	C	101	LEU	CA-CB-CG	6.58	130.42	115.30
3	F	67	ASN	CA-C-N	-6.49	102.92	117.20

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	11	LYS	Peptide
3	C	67	ASN	Peptide
3	C	69	ARG	Peptide
3	C	71	ASN	Peptide
3	C	72	LYS	Peptide

## 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	2432	0	2339	74	0
1	D	2432	0	2339	88	0
2	B	676	0	634	49	0
2	E	676	0	634	28	0
3	C	738	0	724	69	0
3	F	758	0	742	73	0
4	B	1	0	0	0	0
4	E	1	0	0	0	0
5	A	29	0	0	7	0
5	B	12	0	0	3	0
5	C	11	0	0	2	0
5	D	45	0	0	14	0
5	E	9	0	0	3	0
5	F	8	0	0	2	0
All	All	7828	0	7412	335	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 22.

The worst 5 of 335 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)
3:F:67:ASN:HB2	3:F:70:GLU:H	1.08	1.12
2:B:46:ARG:HG2	3:C:48:GLU:HG3	1.35	1.04
1:A:1215:PHE:O	5:A:1502:HOH:O	1.84	0.94
1:D:1135:ASN:ND2	3:F:6:SER:O	2.00	0.93
2:B:42:GLU:HB3	2:B:43:PHE:HA	1.53	0.90

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	301/349 (86%)	283 (94%)	16 (5%)	2 (1%)	26	51
1	D	301/349 (86%)	284 (94%)	15 (5%)	2 (1%)	26	51
2	B	77/100 (77%)	65 (84%)	8 (10%)	4 (5%)	2	3
2	E	77/100 (77%)	67 (87%)	8 (10%)	2 (3%)	7	11
3	C	88/118 (75%)	63 (72%)	19 (22%)	6 (7%)	1	1
3	F	91/118 (77%)	72 (79%)	11 (12%)	8 (9%)	1	1
All	All	935/1134 (82%)	834 (89%)	77 (8%)	24 (3%)	7	11

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1167	LYS
2	B	12	PRO
2	B	21	GLU
2	B	41	PRO
3	C	70	GLU

### 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	270/306 (88%)	266 (98%)	4 (2%)	72	90

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	270/306 (88%)	262 (97%)	8 (3%)	48	76
2	B	70/88 (80%)	69 (99%)	1 (1%)	74	90
2	E	70/88 (80%)	69 (99%)	1 (1%)	74	90
3	C	81/104 (78%)	74 (91%)	7 (9%)	13	25
3	F	83/104 (80%)	77 (93%)	6 (7%)	18	35
All	All	844/996 (85%)	817 (97%)	27 (3%)	46	74

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	1073	ARG
1	D	1201	HIS
3	F	71	ASN
1	D	1126	GLN
2	B	46	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	1174	HIS
3	F	105	GLN
3	F	49	GLN
2	B	59	HIS
1	D	1201	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 2 ligands modelled in this entry, 2 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 [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	305/349 (87%)	-0.41	1 (0%) 94 93	24, 42, 61, 87	0
1	D	305/349 (87%)	-0.42	1 (0%) 94 93	23, 39, 60, 109	0
2	B	79/100 (79%)	-0.00	3 (3%) 44 36	26, 44, 83, 116	0
2	E	79/100 (79%)	-0.22	0 100 100	30, 45, 65, 78	0
3	C	92/118 (77%)	0.76	12 (13%) 5 3	38, 81, 97, 116	0
3	F	95/118 (80%)	0.83	16 (16%) 2 1	40, 74, 102, 124	0
All	All	955/1134 (84%)	-0.13	33 (3%) 48 40	23, 44, 90, 124	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1390	GLY	10.4
3	F	108	GLY	7.4
3	F	104	ILE	6.4
3	C	90	GLY	5.0
3	C	77	LEU	4.5

### 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<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(Å <sup>2</sup> )	Q<0.9
4	ZN	B	201	1/1	0.95	0.16	0.36	68,68,68,68	0
4	ZN	E	201	1/1	0.97	0.09	-1.53	56,56,56,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.