



wwPDB X-ray Structure Validation Summary Report

Feb 1, 2016 – 09:11 AM GMT

PDB ID : 3HGZ
Title : Crystal structure of human insulin-degrading enzyme in complex with amylin
Authors : Guo, Q.; Bian, Y.; Tang, W.J.
Deposited on : 2009-05-14
Resolution : 2.91 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
<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 : 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

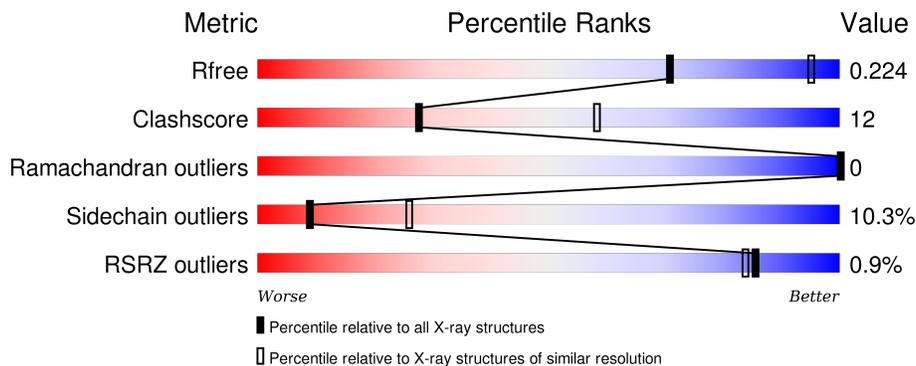
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.91 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	1643 (2.94-2.90)
Clashscore	102246	1871 (2.94-2.90)
Ramachandran outliers	100387	1824 (2.94-2.90)
Sidechain outliers	100360	1826 (2.94-2.90)
RSRZ outliers	91569	1650 (2.94-2.90)

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 ≥ 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	969	 73% 22% . .
1	B	969	 % 70% 25% . .
2	D	37	 3% 14% . 81%
2	E	37	 11% 11% 8% 5% . 73%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 16156 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 Insulin-degrading enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	958	7837	5045	1317	1453	22	0	0	0
1	B	959	7844	5050	1318	1454	22	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	LEU	CYS	ENGINEERED	UNP P14735
A	111	GLN	GLU	ENGINEERED	UNP P14735
A	171	SER	CYS	ENGINEERED	UNP P14735
A	178	ALA	CYS	ENGINEERED	UNP P14735
A	257	VAL	CYS	ENGINEERED	UNP P14735
A	414	LEU	CYS	ENGINEERED	UNP P14735
A	573	ASN	CYS	ENGINEERED	UNP P14735
A	590	SER	CYS	ENGINEERED	UNP P14735
A	789	SER	CYS	ENGINEERED	UNP P14735
A	812	ALA	CYS	ENGINEERED	UNP P14735
A	819	ALA	CYS	ENGINEERED	UNP P14735
A	904	SER	CYS	ENGINEERED	UNP P14735
A	966	ASN	CYS	ENGINEERED	UNP P14735
A	974	ALA	CYS	ENGINEERED	UNP P14735
B	110	LEU	CYS	ENGINEERED	UNP P14735
B	111	GLN	GLU	ENGINEERED	UNP P14735
B	171	SER	CYS	ENGINEERED	UNP P14735
B	178	ALA	CYS	ENGINEERED	UNP P14735
B	257	VAL	CYS	ENGINEERED	UNP P14735
B	414	LEU	CYS	ENGINEERED	UNP P14735
B	573	ASN	CYS	ENGINEERED	UNP P14735
B	590	SER	CYS	ENGINEERED	UNP P14735
B	789	SER	CYS	ENGINEERED	UNP P14735
B	812	ALA	CYS	ENGINEERED	UNP P14735
B	819	ALA	CYS	ENGINEERED	UNP P14735

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	904	SER	CYS	ENGINEERED	UNP P14735
B	966	ASN	CYS	ENGINEERED	UNP P14735
B	974	ALA	CYS	ENGINEERED	UNP P14735

- Molecule 2 is a protein called Islet amyloid polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	7	Total	C	N	O	S	0	0	0
			55	37	9	8	1			
2	E	10	Total	C	N	O	S	0	0	0
			71	46	12	11	2			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

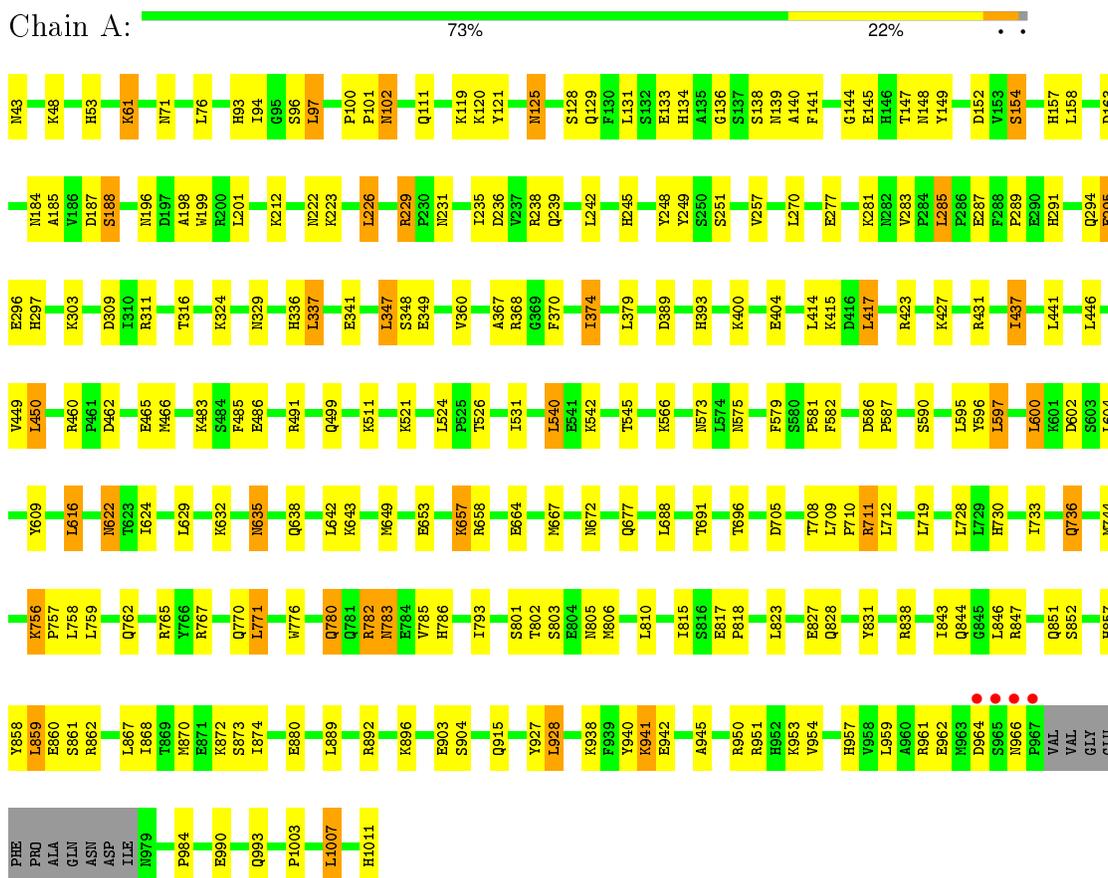
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	123	Total	O	0	0
			123	123		
4	B	221	Total	O	0	0
			221	221		
4	D	2	Total	O	0	0
			2	2		
4	E	1	Total	O	0	0
			1	1		

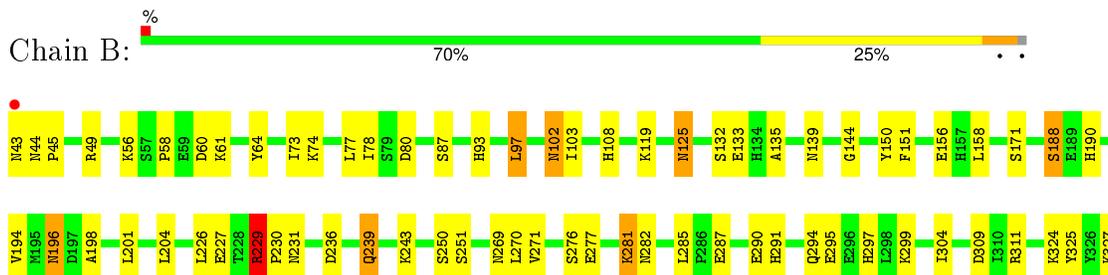
3 Residue-property plots

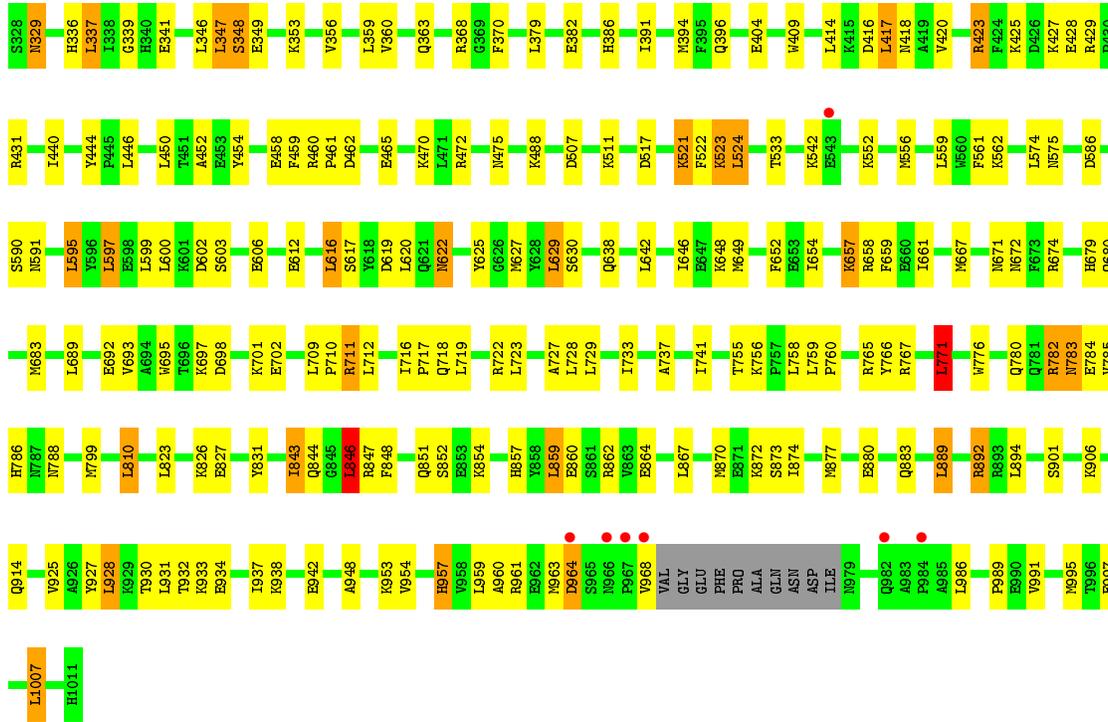
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Insulin-degrading enzyme



- Molecule 1: Insulin-degrading enzyme





• Molecule 2: Islet amyloid polypeptide



• Molecule 2: Islet amyloid polypeptide



4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	262.87Å 262.87Å 90.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.69 – 2.91 49.68 – 2.91	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.69-2.91) 99.8 (49.68-2.91)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.16 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.176 , 0.224 0.183 , 0.224	Depositor DCC
R_{free} test set	3956 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	60.5	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 57.4	EDS
Estimated twinning fraction	0.026 for h,-h-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 79012 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	16156	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.77% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.60	0/8033	0.71	3/10866 (0.0%)
1	B	0.57	0/8040	0.68	4/10876 (0.0%)
2	D	0.98	0/54	1.23	1/69 (1.4%)
2	E	1.05	0/69	1.11	1/88 (1.1%)
All	All	0.59	0/16196	0.70	9/21899 (0.0%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	635	ASN	CB-CA-C	-8.45	93.50	110.40
1	B	964	ASP	CB-CA-C	6.74	123.89	110.40
1	A	226	LEU	CA-CB-CG	5.86	128.78	115.30
1	A	285	LEU	CA-CB-CG	5.85	128.76	115.30
1	B	229	ARG	CG-CD-NE	5.13	122.56	111.80

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	7837	0	7773	156	0
1	B	7844	0	7782	206	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	55	0	61	10	0
2	E	71	0	70	9	0
3	A	1	0	0	1	0
3	B	1	0	0	1	0
4	A	123	0	0	30	0
4	B	221	0	0	58	0
4	D	2	0	0	3	0
4	E	1	0	0	3	0
All	All	16156	0	15686	371	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 12.

The worst 5 of 371 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:892:ARG:HD3	4:A:2:HOH:O	1.41	1.14
1:A:296:GLU:HG3	4:A:1081:HOH:O	1.50	1.11
1:B:784:GLU:HG2	4:B:1054:HOH:O	1.49	1.09
2:D:15:PHE:O	4:D:432:HOH:O	1.67	1.09
2:E:15:PHE:O	4:E:334:HOH:O	1.69	1.08

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	954/969 (98%)	908 (95%)	46 (5%)	0	100 100
1	B	955/969 (99%)	909 (95%)	46 (5%)	0	100 100
2	D	3/37 (8%)	3 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	4/37 (11%)	4 (100%)	0	0	100	100
All	All	1916/2012 (95%)	1824 (95%)	92 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	854/863 (99%)	764 (90%)	90 (10%)	8	25
1	B	855/863 (99%)	774 (90%)	81 (10%)	11	30
2	D	6/31 (19%)	3 (50%)	3 (50%)	0	0
2	E	7/31 (23%)	4 (57%)	3 (43%)	0	0
All	All	1722/1788 (96%)	1545 (90%)	177 (10%)	9	26

5 of 177 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	889	LEU
1	B	201	LEU
1	B	901	SER
1	A	928	LEU
1	A	1007	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 69 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	52	ASN
1	B	196	ASN
1	B	922	ASN
1	B	93	HIS
1	B	129	GLN

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, 95th 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	958/969 (98%)	-0.17	4 (0%) 93 92	38, 52, 68, 102	1 (0%)
1	B	959/969 (98%)	-0.09	8 (0%) 87 86	44, 57, 74, 103	1 (0%)
2	D	7/37 (18%)	1.67	1 (14%) 4 2	62, 103, 104, 104	0
2	E	10/37 (27%)	2.06	4 (40%) 0 0	84, 104, 105, 106	0
All	All	1934/2012 (96%)	-0.11	17 (0%) 85 84	38, 55, 73, 106	2 (0%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	966	ASN	5.7
1	A	967	PRO	4.4
1	B	968	VAL	4.4
1	B	966	ASN	4.3
2	E	16	LEU	3.5

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, 95th 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(\AA^2)	Q<0.9
3	ZN	B	2	1/1	0.98	0.19	-0.56	2,2,2,2	0
3	ZN	A	1	1/1	0.96	0.17	-0.82	2,2,2,2	0

6.5 Other polymers [i](#)

There are no such residues in this entry.