



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 09:05 AM EST

PDB ID : 4FEB
Title : Crystal Structure of Htt36Q3H-EX1-X1-C2(Beta)
Authors : Kim, M.
Deposited on : 2012-05-29
Resolution : 2.80 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

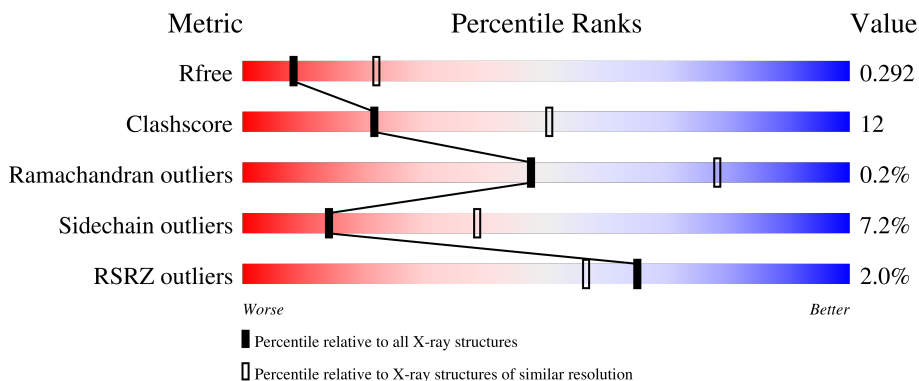
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

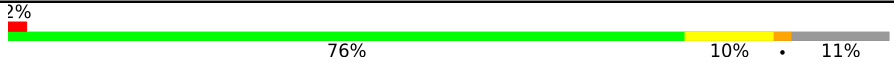
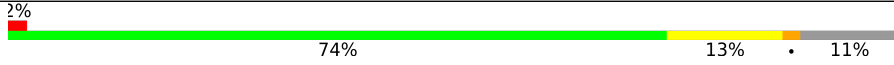

The reported resolution of this entry is 2.80 Å.

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	452	 2% 76% 10% • 11%
1	B	452	 2% 74% 13% • 11%
1	C	452	 2% 74% 12% • 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ZN	A	504	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9631 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 Maltose-binding periplasmic protein,Huntingtin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	402	Total 3155	C 2021	N 522	O 604	S 8	10	3	0
1	B	402	Total 3146	C 2016	N 520	O 602	S 8	14	2	0
1	C	402	Total 3149	C 2020	N 520	O 601	S 8	22	2	0

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	359	ALA	-	linker	UNP P0AEY0
A	360	ALA	-	linker	UNP P0AEY0
A	361	LEU	-	linker	UNP P0AEY0
A	362	ALA	-	linker	UNP P0AEY0
A	363	ALA	-	linker	UNP P0AEY0
A	364	ALA	-	linker	UNP P0AEY0
A	365	GLN	-	linker	UNP P0AEY0
A	366	THR	-	linker	UNP P0AEY0
A	367	ASN	-	linker	UNP P0AEY0
A	368	ALA	-	linker	UNP P0AEY0
A	369	ALA	-	linker	UNP P0AEY0
A	370	ALA	-	linker	UNP P0AEY0
A	388	GLN	-	insertion	UNP P42858
A	389	GLN	-	insertion	UNP P42858
A	390	GLN	-	insertion	UNP P42858
A	391	GLN	-	insertion	UNP P42858
A	392	GLN	-	insertion	UNP P42858
A	393	GLN	-	insertion	UNP P42858
A	394	GLN	-	insertion	UNP P42858
A	395	HIS	-	insertion	UNP P42858
A	396	GLN	-	insertion	UNP P42858
A	397	HIS	-	insertion	UNP P42858
A	398	GLN	-	insertion	UNP P42858

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	399	HIS	-	insertion	UNP P42858
A	400	GLN	-	insertion	UNP P42858
A	401	GLN	-	insertion	UNP P42858
A	402	GLN	-	insertion	UNP P42858
A	403	GLN	-	insertion	UNP P42858
A	404	GLN	-	insertion	UNP P42858
A	405	GLN	-	insertion	UNP P42858
B	359	ALA	-	linker	UNP P0AEY0
B	360	ALA	-	linker	UNP P0AEY0
B	361	LEU	-	linker	UNP P0AEY0
B	362	ALA	-	linker	UNP P0AEY0
B	363	ALA	-	linker	UNP P0AEY0
B	364	ALA	-	linker	UNP P0AEY0
B	365	GLN	-	linker	UNP P0AEY0
B	366	THR	-	linker	UNP P0AEY0
B	367	ASN	-	linker	UNP P0AEY0
B	368	ALA	-	linker	UNP P0AEY0
B	369	ALA	-	linker	UNP P0AEY0
B	370	ALA	-	linker	UNP P0AEY0
B	388	GLN	-	insertion	UNP P42858
B	389	GLN	-	insertion	UNP P42858
B	390	GLN	-	insertion	UNP P42858
B	391	GLN	-	insertion	UNP P42858
B	392	GLN	-	insertion	UNP P42858
B	393	GLN	-	insertion	UNP P42858
B	394	GLN	-	insertion	UNP P42858
B	395	HIS	-	insertion	UNP P42858
B	396	GLN	-	insertion	UNP P42858
B	397	HIS	-	insertion	UNP P42858
B	398	GLN	-	insertion	UNP P42858
B	399	HIS	-	insertion	UNP P42858
B	400	GLN	-	insertion	UNP P42858
B	401	GLN	-	insertion	UNP P42858
B	402	GLN	-	insertion	UNP P42858
B	403	GLN	-	insertion	UNP P42858
B	404	GLN	-	insertion	UNP P42858
B	405	GLN	-	insertion	UNP P42858
C	359	ALA	-	linker	UNP P0AEY0
C	360	ALA	-	linker	UNP P0AEY0
C	361	LEU	-	linker	UNP P0AEY0
C	362	ALA	-	linker	UNP P0AEY0
C	363	ALA	-	linker	UNP P0AEY0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	364	ALA	-	linker	UNP P0AEY0
C	365	GLN	-	linker	UNP P0AEY0
C	366	THR	-	linker	UNP P0AEY0
C	367	ASN	-	linker	UNP P0AEY0
C	368	ALA	-	linker	UNP P0AEY0
C	369	ALA	-	linker	UNP P0AEY0
C	370	ALA	-	linker	UNP P0AEY0
C	388	GLN	-	insertion	UNP P42858
C	389	GLN	-	insertion	UNP P42858
C	390	GLN	-	insertion	UNP P42858
C	391	GLN	-	insertion	UNP P42858
C	392	GLN	-	insertion	UNP P42858
C	393	GLN	-	insertion	UNP P42858
C	394	GLN	-	insertion	UNP P42858
C	395	HIS	-	insertion	UNP P42858
C	396	GLN	-	insertion	UNP P42858
C	397	HIS	-	insertion	UNP P42858
C	398	GLN	-	insertion	UNP P42858
C	399	HIS	-	insertion	UNP P42858
C	400	GLN	-	insertion	UNP P42858
C	401	GLN	-	insertion	UNP P42858
C	402	GLN	-	insertion	UNP P42858
C	403	GLN	-	insertion	UNP P42858
C	404	GLN	-	insertion	UNP P42858
C	405	GLN	-	insertion	UNP P42858

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	8	Total Zn 8 8	0	0
2	B	19	Total Zn 19 19	0	0
2	C	11	Total Zn 11 11	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	5	Total Na 5 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	Na	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	50	Total	O	0	0
			50	50		
4	B	56	Total	O	0	0
			56	56		
4	C	31	Total	O	0	0
			31	31		

GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
GLN
PRO
PRO
PRO
PRO
PRO
PRO
PRO
PRO
PRO
PRO
GLN
LEU
PRO
GLN
PRO
PRO
PRO
GLN
ALA
GLN
PRO
LEU
LEU
PRO
GLN

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	155.05Å 177.28Å 78.87Å 90.00° 109.03° 90.00°	Depositor
Resolution (Å)	35.00 – 2.80 34.46 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.5 (35.00-2.80) 96.4 (34.46-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.30 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.229 , 0.275 0.247 , 0.292	Depositor DCC
R_{free} test set	2375 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	78.1	Xtrriage
Anisotropy	0.058	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 59.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9631	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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, 0.333 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: NA, 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	0.35	1/3229 (0.0%)	0.53	2/4379 (0.0%)
1	B	0.37	0/3220	0.48	2/4366 (0.0%)
1	C	0.38	0/3224	0.47	0/4372
All	All	0.37	1/9673 (0.0%)	0.49	4/13117 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	390	GLN	CG-CD	-5.75	1.37	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	390	GLN	CG-CD-OE1	-12.45	96.69	121.60
1	A	390	GLN	CG-CD-NE2	9.17	138.70	116.70
1	B	4	GLU	N-CA-C	-6.51	93.42	111.00
1	B	374	LEU	CA-CB-CG	5.58	128.13	115.30

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	3155	0	3108	79	0
1	B	3146	0	3102	76	0
1	C	3149	0	3105	74	13
2	A	8	0	0	0	0
2	B	19	0	0	0	0
2	C	11	0	0	0	0
3	A	5	0	0	0	0
3	C	1	0	0	0	0
4	A	50	0	0	0	0
4	B	56	0	0	1	0
4	C	31	0	0	1	0
All	All	9631	0	9315	223	13

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 223 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:B:398:GLN:OE1	1:B:399:HIS:CD2	1.65	1.47
1:B:398:GLN:CD	1:B:399:HIS:HD2	1.30	1.31
1:A:396:GLN:CG	1:A:402:GLN:HA	1.62	1.30
1:A:396:GLN:CG	1:A:402:GLN:CA	2.09	1.30
1:B:398:GLN:CD	1:B:399:HIS:CD2	2.01	1.29

The worst 5 of 13 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:153:GLU:CG	1:C:396:GLN:NE2[2_556]	0.19	2.01
1:C:153:GLU:CD	1:C:396:GLN:CD[2_556]	0.70	1.50
1:C:153:GLU:CD	1:C:396:GLN:OE1[2_556]	0.72	1.48
1:C:153:GLU:OE1	1:C:396:GLN:OE1[2_556]	1.02	1.18
1:C:153:GLU:OE2	1:C:396:GLN:OE1[2_556]	1.19	1.01

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	403/452 (89%)	390 (97%)	13 (3%)	0	100	100
1	B	402/452 (89%)	387 (96%)	14 (4%)	1 (0%)	47	78
1	C	402/452 (89%)	389 (97%)	12 (3%)	1 (0%)	47	78
All	All	1207/1356 (89%)	1166 (97%)	39 (3%)	2 (0%)	47	78

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	3	GLU
1	B	2	ILE

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	324/370 (88%)	302 (93%)	22 (7%)	16	42
1	B	323/370 (87%)	302 (94%)	21 (6%)	17	44
1	C	323/370 (87%)	296 (92%)	27 (8%)	11	31
All	All	970/1110 (87%)	900 (93%)	70 (7%)	14	38

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	344	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	373	THR
1	C	393	GLN
1	B	6	LYS
1	B	3	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	395	HIS
1	C	397	HIS
1	C	402	GLN
1	B	205	ASN
1	B	201	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 44 ligands modelled in this entry, 44 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	402/452 (88%)	-0.14	8 (1%) 65 56	57, 85, 122, 149	18 (4%)
1	B	402/452 (88%)	-0.02	7 (1%) 70 63	56, 89, 131, 157	16 (3%)
1	C	402/452 (88%)	-0.01	9 (2%) 62 52	33, 94, 147, 183	18 (4%)
All	All	1206/1356 (88%)	-0.06	24 (1%) 65 56	33, 89, 133, 183	52 (4%)

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1	LYS	5.0
1	C	398	GLN	4.5
1	A	399	HIS	4.4
1	A	396	GLN	4.0
1	B	34	LYS	3.6

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 monosaccharides 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, 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	B-factors(Å ²)	Q<0.9
2	ZN	A	504	1/1	0.70	0.47	60,60,60,60	1
3	NA	A	513	1/1	0.74	0.22	60,60,60,60	0
2	ZN	A	501	1/1	0.75	0.11	60,60,60,60	1
2	ZN	B	513	1/1	0.80	0.12	60,60,60,60	1
2	ZN	A	505	1/1	0.80	0.33	60,60,60,60	1
2	ZN	A	507	1/1	0.81	0.24	60,60,60,60	1
2	ZN	C	501	1/1	0.81	0.14	60,60,60,60	1
2	ZN	B	510	1/1	0.81	0.10	60,60,60,60	1
3	NA	A	511	1/1	0.83	0.16	60,60,60,60	0
2	ZN	A	502	1/1	0.84	0.08	60,60,60,60	1
2	ZN	B	509	1/1	0.85	0.20	60,60,60,60	1
2	ZN	C	508	1/1	0.85	0.23	60,60,60,60	1
2	ZN	C	510	1/1	0.87	0.22	60,60,60,60	1
2	ZN	B	514	1/1	0.87	0.37	60,60,60,60	1
2	ZN	B	516	1/1	0.87	0.19	60,60,60,60	1
2	ZN	B	511	1/1	0.88	0.13	60,60,60,60	1
2	ZN	B	518	1/1	0.88	0.24	60,60,60,60	1
3	NA	A	512	1/1	0.88	0.39	60,60,60,60	0
2	ZN	C	509	1/1	0.88	0.14	60,60,60,60	1
3	NA	C	512	1/1	0.88	0.32	60,60,60,60	0
2	ZN	C	507	1/1	0.90	0.08	60,60,60,60	1
2	ZN	A	508	1/1	0.91	0.18	60,60,60,60	1
2	ZN	C	505	1/1	0.91	0.23	60,60,60,60	1
2	ZN	A	503	1/1	0.91	0.16	60,60,60,60	1
3	NA	A	509	1/1	0.91	0.15	60,60,60,60	0
2	ZN	C	506	1/1	0.92	0.12	60,60,60,60	1
2	ZN	B	506	1/1	0.92	0.13	60,60,60,60	1
2	ZN	C	511	1/1	0.93	0.36	55,55,55,55	1
2	ZN	C	502	1/1	0.94	0.27	60,60,60,60	1
3	NA	A	510	1/1	0.94	0.15	60,60,60,60	0
2	ZN	B	515	1/1	0.94	0.24	56,56,56,56	1
2	ZN	B	505	1/1	0.95	0.14	60,60,60,60	1
2	ZN	B	504	1/1	0.95	0.13	60,60,60,60	1
2	ZN	B	519	1/1	0.95	0.21	60,60,60,60	1
2	ZN	B	517	1/1	0.96	0.25	45,45,45,45	1
2	ZN	B	508	1/1	0.97	0.06	60,60,60,60	1
2	ZN	C	504	1/1	0.97	0.07	60,60,60,60	1
2	ZN	B	502	1/1	0.97	0.09	60,60,60,60	1
2	ZN	B	501	1/1	0.98	0.12	60,60,60,60	1
2	ZN	A	506	1/1	0.98	0.28	60,60,60,60	1
2	ZN	B	512	1/1	0.99	0.18	60,60,60,60	1
2	ZN	C	503	1/1	0.99	0.23	60,60,60,60	1
2	ZN	B	503	1/1	0.99	0.09	60,60,60,60	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	ZN	B	507	1/1	0.99	0.07	60,60,60,60	1

6.5 Other polymers [i](#)

There are no such residues in this entry.