Full wwPDB X-ray Structure Validation Report (i)

Oct 30, 2021 – 01:04 pm BST

PDB ID	:	7O4F
Title	:	The DYW domain of A. thaliana OTP86 in its active state
Authors	:	Weber, G.; Palm, G.J.; Takenaka, M.; Barthel, T.; Feiler, C.; Weiss, M.S.
Deposited on	:	2021-04-06
Resolution	:	1.65 Å(reported)

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

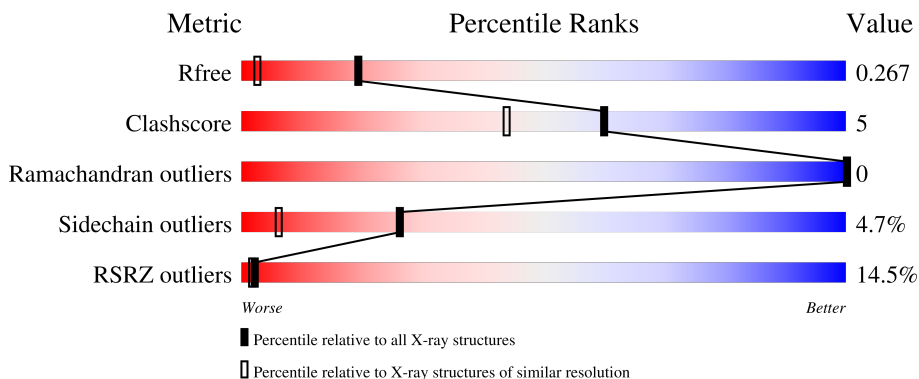|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.65 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	Percentile Ranks	Value
Rfree		0.267
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		4.7%
RSRZ outliers		14.5%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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 <=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	138	14% / 89% / 9%
1	B	138	16% / 81% / 16%
1	D	138	12% / 84% / 11%
1	G	138	14% / 85% / 12%

2　Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4821 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 Pentatricopeptide repeat-containing protein At3g63370, chloroplastic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	136	Total	C	N	O	S	0	0	0
			1109	691	208	199	11			
1	B	136	Total	C	N	O	S	0	1	0
			1112	693	208	199	12			
1	D	135	Total	C	N	O	S	0	2	0
			1109	693	207	198	11			
1	G	136	Total	C	N	O	S	0	0	0
			1109	691	208	199	11			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	823	GLY	-	expression tag	UNP Q9M1V3
A	824	ALA	-	expression tag	UNP Q9M1V3
A	825	MET	-	expression tag	UNP Q9M1V3
B	823	GLY	-	expression tag	UNP Q9M1V3
B	824	ALA	-	expression tag	UNP Q9M1V3
B	825	MET	-	expression tag	UNP Q9M1V3
D	823	GLY	-	expression tag	UNP Q9M1V3
D	824	ALA	-	expression tag	UNP Q9M1V3
D	825	MET	-	expression tag	UNP Q9M1V3
G	823	GLY	-	expression tag	UNP Q9M1V3
G	824	ALA	-	expression tag	UNP Q9M1V3
G	825	MET	-	expression tag	UNP Q9M1V3

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Zn	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total 2	Zn 2	0	0
2	D	2	Total 2	Zn 2	0	0
2	G	2	Total 2	Zn 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	81	Total 81	O 81	0	0
3	B	107	Total 107	O 107	0	0
3	D	99	Total 99	O 99	0	0
3	G	87	Total 87	O 87	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.
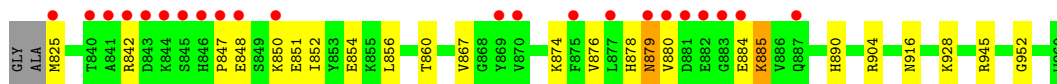
- Molecule 1: Pentatricopeptide repeat-containing protein At3g63370, chloroplastic
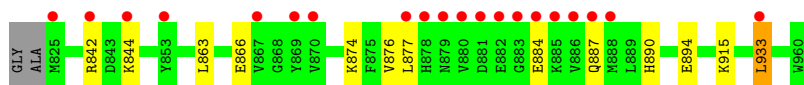
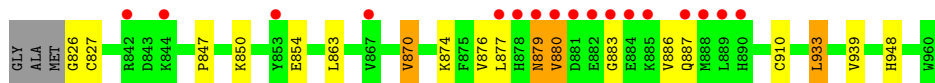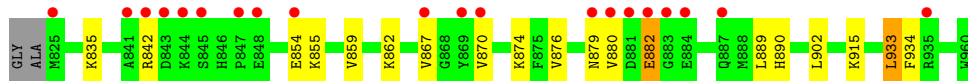- Molecule 1: Pentatricopeptide repeat-containing protein At3g63370, chloroplastic

- Molecule 1: Pentatricopeptide repeat-containing protein At3g63370, chloroplastic

- Molecule 1: Pentatricopeptide repeat-containing protein At3g63370, chloroplastic

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	117.60Å 132.88Å 30.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.03 – 1.65 44.03 – 1.65	Depositor EDS
% Data completeness (in resolution range)	97.3 (44.03-1.65) 97.7 (44.03-1.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 1.65Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.221 , 0.268 0.221 , 0.267	Depositor DCC
R_{free} test set	2100 reflections (3.62%)	wwPDB-VP
Wilson B-factor (Å^2)	18.9	Xtriage
Anisotropy	0.191	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.59$, $< L^2 > = 0.44$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4821	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.67 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.2175e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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: 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	0.57	0/1131	0.72	0/1513
1	B	0.61	0/1137	0.74	0/1521
1	D	0.64	1/1137 (0.1%)	0.71	0/1522
1	G	0.57	0/1131	0.71	0/1513
All	All	0.60	1/4536 (0.0%)	0.72	0/6069

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	827	CYS	CB-SG	-6.52	1.71	1.82

There are no bond angle outliers.

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	1109	0	1095	6	0
1	B	1112	0	1100	16	0
1	D	1109	0	1102	12	0
1	G	1109	0	1095	12	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	2	0	0	0	0
2	G	2	0	0	0	0
3	A	81	0	0	0	1
3	B	107	0	0	1	3
3	D	99	0	0	3	2
3	G	87	0	0	3	2
All	All	4821	0	4392	41	5

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.

All (41) 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:D:847:PRO:O	3:D:1101:HOH:O	1.89	0.90
1:B:847:PRO:O	3:B:1101:HOH:O	1.96	0.83
1:D:826:GLY:N	3:D:1102:HOH:O	2.11	0.82
1:G:876:VAL:HB	3:G:1115:HOH:O	1.89	0.72
1:A:866:GLU:HB2	1:A:933:LEU:HD21	1.77	0.66
1:A:874:LYS:HG2	1:G:876:VAL:HG12	1.77	0.66
1:G:882:GLU:OE2	3:G:1101:HOH:O	2.13	0.65
1:B:879:ASN:HB3	1:B:884:GLU:HG3	1.82	0.62
1:B:825:MET:HG2	1:B:916:ASN:HD21	1.64	0.61
1:D:948:HIS:HD2	3:D:1191:HOH:O	1.85	0.59
1:D:880:VAL:HG22	1:D:883:GLY:H	1.70	0.57
1:B:878:HIS:HB3	1:D:870:VAL:HG12	1.87	0.56
1:G:842:ARG:HB3	1:G:890:HIS:CG	2.42	0.55
1:A:894:GLU:OE2	1:A:915:LYS:HE2	2.07	0.55
1:B:856:LEU:O	1:B:860:THR:HG23	2.08	0.54
1:B:851:GLU:OE2	1:B:904:ARG:NH2	2.28	0.54
1:A:863:LEU:HD23	1:A:933:LEU:HD22	1.92	0.51
1:G:862:LYS:HD3	1:G:934:PHE:CZ	2.46	0.51
1:B:879:ASN:OD1	1:B:879:ASN:N	2.44	0.50
1:B:848:GLU:HB3	1:B:852:ILE:HD12	1.92	0.49
1:D:863:LEU:HD23	1:D:933:LEU:HD12	1.95	0.49
1:A:876:VAL:HG12	1:G:874:LYS:HG2	1.95	0.49
1:B:879:ASN:HA	1:B:884:GLU:HA	1.95	0.47
1:A:842:ARG:HB3	1:A:890:HIS:CG	2.49	0.47
1:B:876:VAL:HG12	1:D:874:LYS:HG2	1.97	0.47
1:B:842:ARG:HB3	1:B:890:HIS:CG	2.51	0.46
1:B:850:LYS:O	1:B:854:GLU:HG3	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:874:LYS:HG3	1:D:876:VAL:HG12	1.99	0.45
1:G:876:VAL:HG13	1:G:889:LEU:HD11	1.98	0.45
1:G:855:LYS:O	1:G:859:VAL:HG23	2.17	0.44
1:B:928:LYS:O	1:B:952:GLY:HA2	2.19	0.43
1:G:867:VAL:HG13	1:G:933:LEU:HD21	2.00	0.43
1:G:902:LEU:O	3:G:1103:HOH:O	2.21	0.43
1:B:885:LYS:HB3	1:B:885:LYS:HE3	1.52	0.43
1:D:879:ASN:OD1	1:D:879:ASN:N	2.52	0.43
1:B:825:MET:HB3	1:B:916:ASN:ND2	2.34	0.42
1:G:835:LYS:HB2	1:G:835:LYS:HE2	1.86	0.42
1:D:850:LYS:NZ	1:D:854:GLU:OE2	2.51	0.42
1:D:910[A]:CYS:SG	1:D:939:VAL:HG23	2.61	0.41
1:D:880:VAL:HG22	1:D:883:GLY:O	2.21	0.41
1:G:915:LYS:HE3	1:G:915:LYS:HB3	1.83	0.40

All (5) 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)
3:A:1116:HOH:O	3:B:1111:HOH:O[2_555]	2.02	0.18
3:D:1170:HOH:O	3:G:1162:HOH:O[2_654]	2.11	0.09
3:B:1125:HOH:O	3:B:1166:HOH:O[2_555]	2.16	0.04
3:B:1147:HOH:O	3:B:1160:HOH:O[2_555]	2.16	0.04
3:D:1170:HOH:O	3:G:1173:HOH:O[2_654]	2.19	0.01

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	134/138 (97%)	131 (98%)	3 (2%)	0	100	100
1	B	135/138 (98%)	129 (96%)	6 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	135/138 (98%)	131 (97%)	4 (3%)	0	100	100
1	G	134/138 (97%)	131 (98%)	3 (2%)	0	100	100
All	All	538/552 (98%)	522 (97%)	16 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	122/122 (100%)	117 (96%)	5 (4%)	30	8
1	B	123/122 (101%)	118 (96%)	5 (4%)	30	8
1	D	123/122 (101%)	116 (94%)	7 (6%)	20	4
1	G	122/122 (100%)	116 (95%)	6 (5%)	25	5
All	All	490/488 (100%)	467 (95%)	23 (5%)	26	6

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	844	LYS
1	A	877	LEU
1	A	884	GLU
1	A	887	GLN
1	A	933	LEU
1	B	867	VAL
1	B	879	ASN
1	B	880	VAL
1	B	885	LYS
1	B	945	ARG
1	D	870	VAL
1	D	877	LEU
1	D	879	ASN
1	D	880	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	886	VAL
1	D	887	GLN
1	D	933	LEU
1	G	854	GLU
1	G	870	VAL
1	G	879	ASN
1	G	880	VAL
1	G	882	GLU
1	G	933	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	948	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 8 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 ⓘ

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, 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	136/138 (98%)	0.78	20 (14%)	2	2	16, 29, 106, 177	0
1	B	136/138 (98%)	0.72	22 (16%)	1	1	14, 25, 92, 125	0
1	D	135/138 (97%)	0.63	17 (12%)	3	3	15, 26, 87, 140	0
1	G	136/138 (98%)	0.67	20 (14%)	2	2	16, 29, 104, 165	0
All	All	543/552 (98%)	0.70	79 (14%)	2	2	14, 27, 96, 177	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	883	GLY	24.7
1	B	880	VAL	13.3
1	G	882	GLU	11.9
1	D	880	VAL	10.2
1	G	883	GLY	10.1
1	A	881	ASP	9.2
1	D	883	GLY	9.0
1	D	881	ASP	6.9
1	A	879	ASN	6.4
1	A	825	MET	6.4
1	D	884	GLU	6.3
1	A	880	VAL	6.2
1	D	882	GLU	6.0
1	B	881	ASP	5.8
1	B	844	LYS	5.6
1	A	878	HIS	5.4
1	A	882	GLU	5.2
1	A	884	GLU	5.1
1	G	884	GLU	4.9
1	G	879	ASN	4.7
1	G	881	ASP	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	842	ARG	4.5
1	B	882	GLU	4.4
1	B	841	ALA	4.4
1	D	879	ASN	4.4
1	G	847	PRO	4.3
1	B	845	SER	4.3
1	D	842	ARG	4.3
1	G	844	LYS	4.1
1	G	870	VAL	4.1
1	B	846	HIS	4.0
1	D	878	HIS	3.7
1	D	887	GLN	3.7
1	B	879	ASN	3.7
1	B	842	ARG	3.6
1	B	887	GLN	3.6
1	A	867	VAL	3.6
1	D	885	LYS	3.5
1	A	877	LEU	3.3
1	A	870	VAL	3.3
1	G	880	VAL	3.3
1	B	825	MET	3.3
1	A	885	LYS	3.2
1	A	853	TYR	3.2
1	D	867	VAL	3.1
1	D	890	HIS	3.1
1	B	883	GLY	3.0
1	B	870	VAL	3.0
1	B	848	GLU	2.9
1	B	884	GLU	2.9
1	B	843	ASP	2.9
1	A	887	GLN	2.9
1	G	841	ALA	2.9
1	B	875	PHE	2.8
1	G	825	MET	2.8
1	A	844	LYS	2.8
1	B	840	THR	2.7
1	B	850	LYS	2.7
1	D	889	LEU	2.6
1	A	933	LEU	2.6
1	G	845	SER	2.6
1	D	844	LYS	2.5
1	G	935	ARG	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	877	LEU	2.5
1	G	854	GLU	2.4
1	B	877	LEU	2.4
1	B	869	TYR	2.4
1	D	888	MET	2.3
1	G	842	ARG	2.3
1	G	843	ASP	2.3
1	A	888	MET	2.3
1	D	853	TYR	2.2
1	G	867	VAL	2.2
1	A	869	TYR	2.2
1	B	847	PRO	2.2
1	G	848	GLU	2.1
1	A	886	VAL	2.1
1	G	887	GLN	2.0
1	G	869	TYR	2.0

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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	1002	1/1	0.99	0.06	19,19,19,19	0
2	ZN	B	1002	1/1	0.99	0.06	19,19,19,19	0
2	ZN	D	1002	1/1	0.99	0.04	18,18,18,18	0
2	ZN	G	1002	1/1	0.99	0.06	21,21,21,21	0
2	ZN	D	1001	1/1	1.00	0.07	15,15,15,15	0
2	ZN	B	1001	1/1	1.00	0.07	14,14,14,14	0
2	ZN	G	1001	1/1	1.00	0.06	15,15,15,15	0

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
2	ZN	A	1001	1/1	1.00	0.05	15,15,15,15	0

6.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK