

## Electronic Supplementary Information (ESI)

---

### **Reversibility and two state behavior in the thermal unfolding of oligomeric TIM barrel proteins**

**Sergio Romero-Romero<sup>1</sup>, Miguel Costas<sup>2</sup>, Adela Rodríguez-Romero<sup>3</sup> and D. Alejandro Fernández-Velasco\*<sup>1</sup>**

*1 - Laboratorio de Físicoquímica e Ingeniería de Proteínas, Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, México. \**

*E-mail: fdaniel@unam.mx*

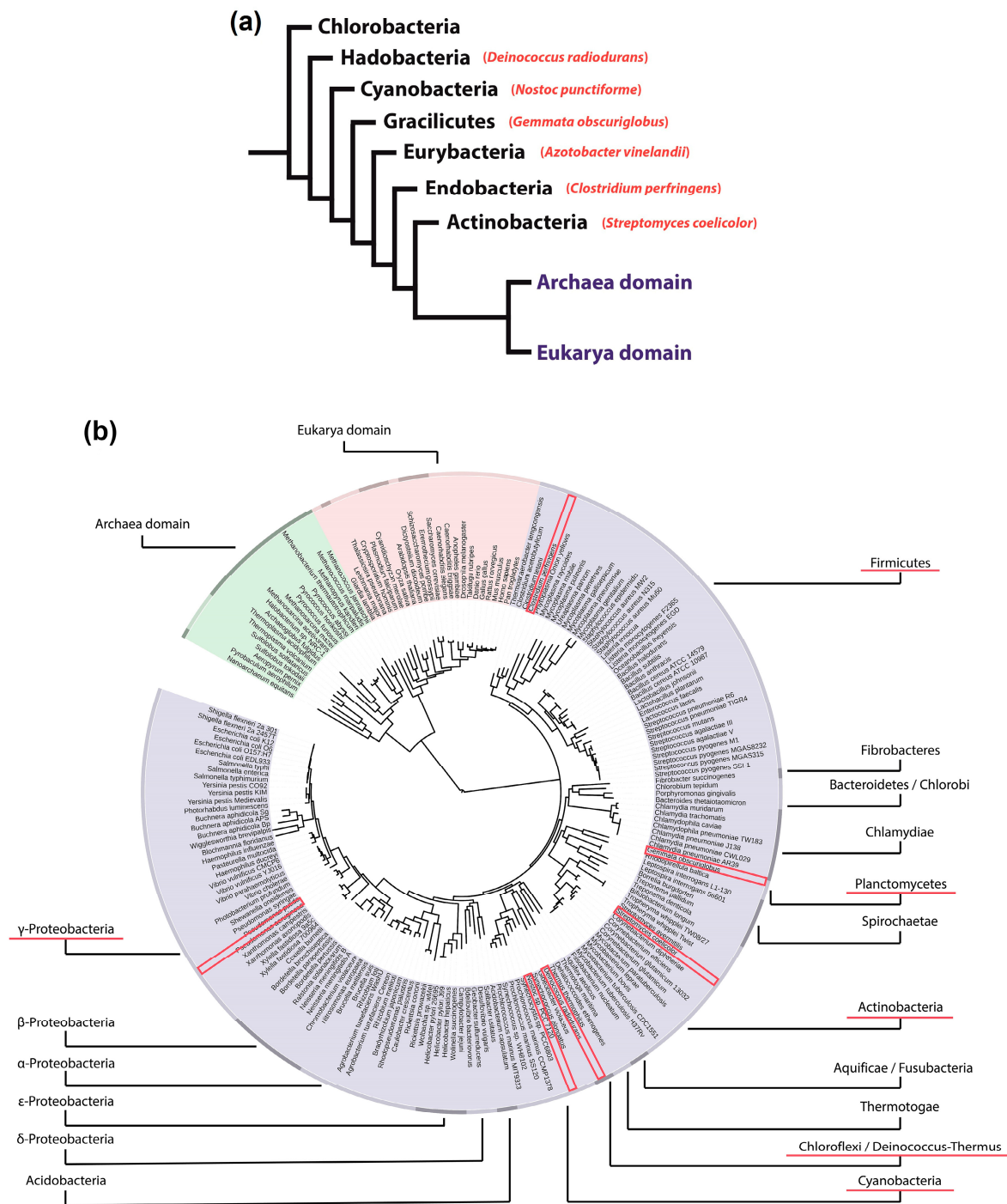
*2 - Laboratorio de Biofísicoquímica, Departamento de Físicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, México.*

*3 - Laboratorio de Química de Biomacromoléculas 3, Departamento de Química de Biomacromoléculas, Instituto de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, México.*

## Table of Contents:

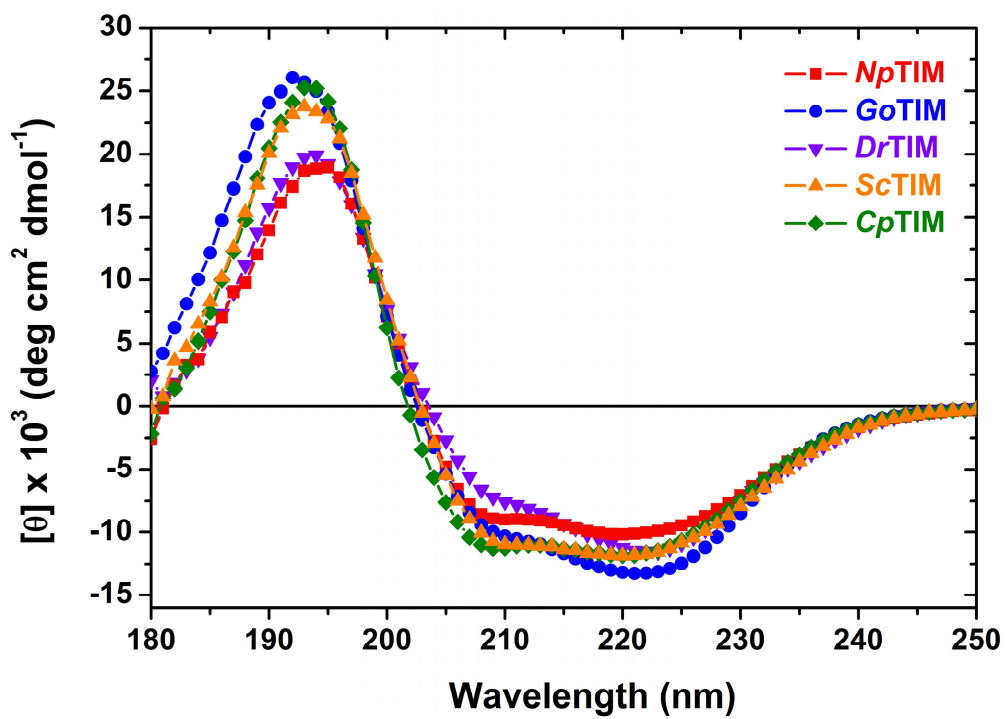
<b>Table of Contents</b> .....	<b>p. S2-S3</b>
<b>Fig. S1:</b> Phylogenetic representations of BacTIMs.....	<b>p. S4</b>
<b>Fig. S2:</b> Far-UV CD spectra of BacTIMs at 25 °C .....	<b>p. S5</b>
<b>Fig. S3:</b> Intrinsic tryptophan fluorescence spectra of BacTIMs .....	<b>p. S6</b>
<b>Fig. S4:</b> Dependence of the $T_m$ with respect to scan rate for BacTIMs .....	<b>p. S7</b>
<b>Fig. S5:</b> Far-UV CD spectra of BacTIMs at 80 °C .....	<b>p. S8</b>
<b>Fig. S6:</b> Amino acid sequence alignment of diverse TIMs used for comparison.....	<b>p. S9-S10</b>
<b>Table S1:</b> Data collection and refinement statistics of <i>Go</i> TIM, <i>Dr</i> TIM, <i>Sc</i> TIM and <i>Cp</i> TIM .....	<b>p. S11</b>
<b>Table S2:</b> Dimeric TIM PDB files used in the structural comparison of IrrevTIMs and RevTIMs. ....	<b>p. S12</b>
<b>Table S3:</b> Thermodynamic parameters of irreversible thermal unfolding of <i>Np</i> TIM and <i>Go</i> TIM.....	<b>p. S13</b>
<b>Table S4:</b> Matrix with identity sequence percentage and average RMSD values for IrrevTIMs and RevTIMs .....	<b>p. S14</b>

<b>Table S5:</b> Amino acidic composition used in the sequence comparison of IrrevTIMs and RevTIMs.....	<b>p. S15</b>
<b>Table S6:</b> Secondary structure elements and stabilizing interactions for IrrevTIMs and RevTIMs.....	<b>p. S16</b>
<b>References for Electronic Supplementary Information.....</b>	<b>p. S17-S18</b>

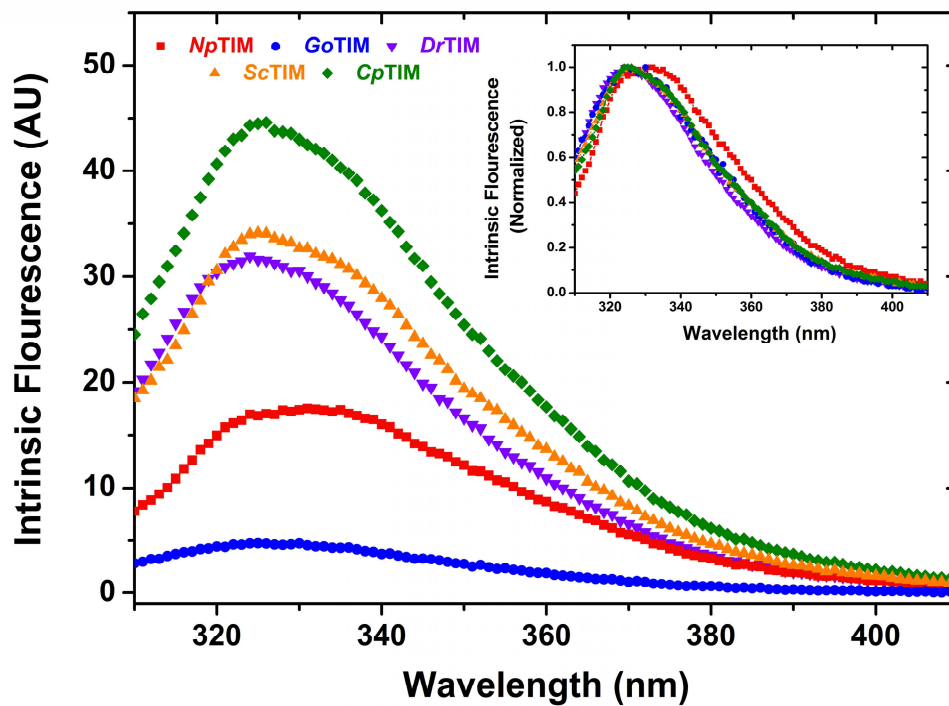


**Fig. S1.** Phylogenetic representations of BacTIMs. **(a)** Cladogram of bacterial supertaxa (according to ref. 1) showing the species from which the indicated TIM sequence was studied. **(b)** Phylogenetic tree showing the TIMs studied in this work (modified from ref. 2). For *Av*TIM phylum ( $\gamma$ -Proteobacteria) only an approximate location in the phylogenetic tree is shown.

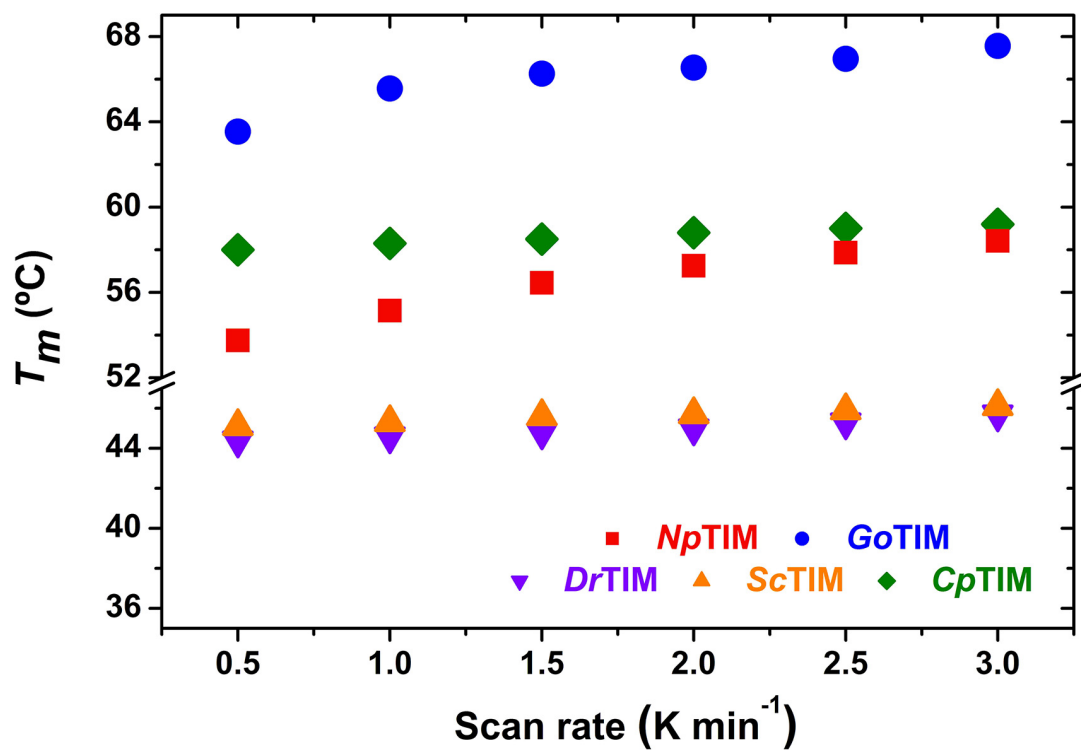




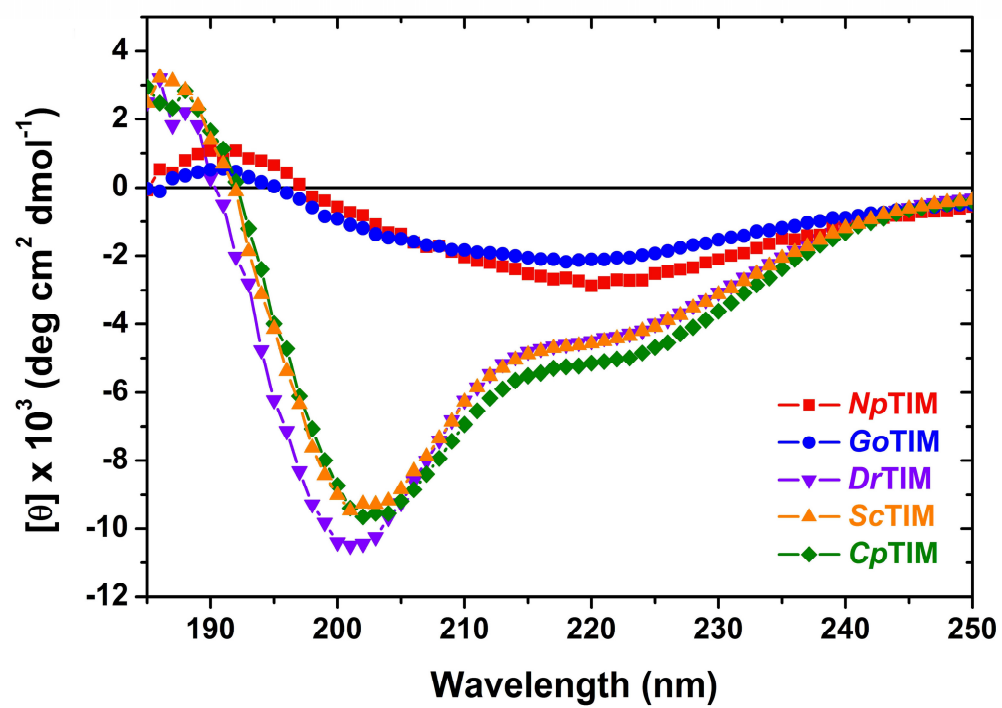
**Fig. S2.** Far-UV CD spectra of BacTIMs at 25 °C. Data shown are the average of five scans recorded using a protein concentration of 15  $\mu$ M in buffer C (10 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8.0).



**Fig. S3.** Intrinsic tryptophan fluorescence spectra of BacTIMs. The spectra shown are the average of five scans recorded using a protein concentration of 15  $\mu$ M at 25  $^{\circ}$ C in buffer C (10 mM  $\text{NaH}_2\text{PO}_4$  pH 8.0). The excitation wavelength was 295 nm. The inset shows normalized data.



**Fig. S4.** Dependence of the  $T_m$  with respect to scan rate for BacTIMs. In all experiments, protein concentration was 15  $\mu\text{M}$ .



**Fig. S5.** Far-UV CD spectra of BacTIMs at 80 °C. Data shown are the average of five scans recorded using a protein concentration of 15  $\mu\text{M}$  in buffer C (10 mM  $\text{NaH}_2\text{PO}_4$  pH 8.0).

```

.....EcTIM..... 1 . . . MRHPLVMGNWKLN . GSRHMVHEL VSNLRK . . . ELAGV . . AGCAVA I APPEMY IDMAKREA . . . EGS HIM 60
.....EhTIM..... 1 . MSARKFVVGGNWKC N . GT LAS IET LTKGVAASVDAELA . . . KKVEV I VGVPF I Y I PKVQQ I LAGEANGAN I L 68
.....GgTIM..... 1 MAP . RKFFVVGGNWKM N . GDKKSLGEL IHTLNG . . . AKLS . . . ADTEVVCGAPS I YLDFARQKL . . . . . DAK I G 60
.....GtTIM..... 1 . MPARRPF I GGNFKCN . GS LDF I KSHVAA IAA . . . HK I P . . . DSVDV V I APSAVHLSTA I AAN . . . . . TSKQLR 61
.....GsTIM..... 1 . . . MRKPI I AGNWKMH . KT LAEAVQFVEDVKG . . . HVPPA . . . DEVI SVVCA PFL FLDRLVQAA . . . . . DGTDLK 60
.....HsTIM..... 1 MAPSRKFFVVGGNWKM N . GRKQSLGEL I GTLNA . . . AKVP . . . ADTEVVCAPPTAY I D FARQKL . . . . . DPK I A 61
.....LmTIM..... 1 MSAKPQP I A AANWKC N . GTTAS IEKLVQVFNE . . . HT I S . . . HDVQC VVAPT FVH I PLVQAKL . . . . . RNPKYV 62
.....MmTIM..... 1 . . . MRHPVVMGNWKL N . GSKEMVVD LNLGNA . . . ELEGV . . . TGVDVAVAPPAL FVDLAERT LT . . . EAGSA I I 62
.....MtTIM..... 1 . . . MSRKPL I AGNWKMN . LNHYEA IALVQK IAF . . . SLPDKYYDRVDVAV I PPF TDLRSVQTLV . . . . . DGDKLR 63
.....OcTIM..... 1 MAPSRKFFVVGGNWKM N . GRKKNLGEL I ETLNA . . . AKVP . . . ADTEVVCAPPTAY I D FARQKL . . . . . DPK I A 61
.....PftTIM..... 1 . . . MARKYFVAANWKC N . GTLES I KSLTNSFN N . . . LDFDP . . . SKLDVVVFPVSVHYDHRKLL . . . . . QSKFS 60
.....SsTIM..... 1 MAPARKFFVVGGNWKM N . GRKNNLGEL I INTLNA . . . AKLP . . . ADTEVVCAPPTAY I D FARQKL . . . . . DPK I A 61
.....TbTIM..... 1 . MSKPQP I A AANWKC N . GSQQSLSEL I DLFNS . . . TS I N . . . HDVQC VVAST FVH LAMTKERL . . . . . SHPKFV 61
.....TcTIM..... 1 MASKPQP I A AANWKC N . GSESLVPL I ETLNA . . . ATFD . . . HDVQC VVAPT FLH I PMTKARL . . . . . TNPKFQ 62
.....TsTIM..... 1 . . . MTRKFLVVGGNWKM N . GSYSH INTFFDTLQK . . . ADTD . . . PNAD I V I GVPACYLKYAQDKA . . . . . PKGIK 59
.....TvTIM..... 1 . . . MRTFFVVGGNWKA NPKTVEEAEL I EMLNG . . . AKVE . . . GNVEVVVAAPF I FLPTLQKQL . . . . . RKDWK 59
.....YtTIM..... 1 . . . MARTFFVVGGNFKLN . GSKQS I KE I VERLNT . . . AS I P . . . ENVEV V ICPPATYLDYSVSLV . . . . . KKPQVT 60
.....NpTIM..... 1 . . . MRK I V I AGNWKMF . KTQAETQEF LQGF L P . . . HLEETP . QGREV I LCPPFTDLSVLSKTL . . . . . HGSL I Q 61
.....GoTIM..... 1 . . . MPTRKKFVAGNWKMN . TTLAEAKALGA AVAK . . . GVTD . . . DRVTVAVFPYPWLTA VGEVL . . . . . KGSPVA 61
.....DrTIM..... 1 . . . MQTLLALNWKMN . KPTPEARSWAELTT . . . KYAPA . . . EGVDLAVLAPALDLSALANL . . . . . PAG I A 58
.....ScTIM..... 1 . . . MTRTPTLMAGNWKMN . LNHLEA IAHVQKLA F . . . ALADKDYDAVEVA VLAPFTDLRSVQTLV . . . . . DGDK LK 64
.....CpTIM..... 1 . . . MRTPI I AGNWKMH . YTI DEAVK LVEELKP . . . LVKD . . . AKCEVVVCPT FVCLDAVKKAV . . . . . EGTN I K 59

```

```

Consensus R K F F V A G N W K L N G L L N A V E V V V A P P L D L L L K E
M+PMRKPFVAGNWKMN -GTKAEA+ELVETLNA - -AKI+- -YADVEVVVAPPFVYLDFAQQKL - - - -DGPKI+

```

```

.....EcTIM..... 61 L . . GAQNVLDNL SGAFTGET SAAMLKD IGAQY I I I GHSERRTYHKESEDEL IAKKFAVLKEQGLTFVLCI GETE 131
.....EhTIM..... 69 V . . SAENAW . TKSGAYTGEVHVGM L VDCQVYV I I L GHSERRQ I FHESNEQVAEKVKVA I DAGLKV IAC I GETE 138
.....GgTIM..... 61 V . . AAQNCYKVPKG AFTGE I SPAM I KD IGAAW I L GHSERRH V FGESEDEL I GQKVAHALAEGLGV IAC I GEKL 131
.....GtTIM..... 62 I . . AAQNVYLEGNGAWTGETS VEM LQDMGLKHV I V GHSERR I MGETDEQS AKKAKRAL EKGMTV I FCVGETL 132
.....GsTIM..... 61 I . . GAQTMHFADQGA YTTGEVSPVMLKDLGVTYV I L GHSERRQMF AETDETVNKKVLA AFTRGL I P I ICCGESL 131
.....HsTIM..... 62 V . . AAQNCYKVTNGAFTGE I SPGMIKDCGATWVVLGHSERRH V FGESEDEL I GQKVAHALAEGLGV IAC I GEKL 132
.....LmTIM..... 63 I . . SAENA I . AKSGAFTGEVSPM I LKD IGVHVV I L GHSERRTY YGETDE I VAQKVSEACKQGFV IAC I GETL 132
.....MmTIM..... 63 L . . GAQNTDLNNSGAFTGDMSPAM LKEFGATH I I I GHSERR EYHAESEDFVAKKFAFLKENGLTPVLCI GESD 133
.....MtTIM..... 64 LTYGAQDLSPHDSGAYTGDVSGAFLAKLGC SYVVVGH SERRTYHNEDDALVAAKAAT ALKHGLTP IVC I GEHL 136
.....OcTIM..... 62 V . . AAQNCYKVTNGAFTGE I SPGMIKDCGATWVVLGHSERRH V FGESEDEL I GQKVAHALAEGLGV IAC I GEKL 132
.....PftTIM..... 61 T . . GIQNVSKFGNGSYTGEVSAE IAKDLN I EYV I I GHSERRR YFHETDE D VREK LQASL KNNLKA VVCFGESL 131
.....SsTIM..... 62 V . . AAQNCYKVANGAFTGE I GPGMIKDLGATWVVLGHSERRH V FGESEDEL I GQKVAHALAEGLGV IAC I GEKL 132
.....TbTIM..... 62 I . . AAQNA I . AKSGAFTGEVSLP I LKDFGVNWI V L GHSERRA Y YGETNE I VADKVA AVASGFV IAC I GETL 131
.....TcTIM..... 63 I . . AAQNA I . TRSGAFTGEVSLQ I LKDYGIK WVV L GHSERR L Y YGETNE I VAEKVAQACAGFHV IVCVGETN 132
.....TsTIM..... 60 I . . AAENCYKVGSGAFTGE ISTEM I KDCGCEWV I L GHSERRH I FGESEDEL I GEKVKHALDSGLNV I PC I GELL 130
.....TvTIM..... 60 V . . SAENVFTKPNGAFTGEVTVPM I KSFGEIWT I L GHSERRD I LKEDDEF LAAKAKFAL ENGMK I I YCCGEHL 130
.....YtTIM..... 61 V . . GAQNAYLKASGAFTGENSVDQ I KDVGAKW I L GHSERRS YFHEDDKF IADKTKFALGQGVGV ILC I GETL 131
.....NpTIM..... 62 L . . GAQNIHWEEFGAYTGE I SGPM LTESGVRFV I V GHSERRQ YFGETDATVNLRLRTAQRFG LTP I LCVGETK 132
.....GoTIM..... 62 L . . GAQDVSEKKG AFTGEVSPAM LLETGCKYAL I GHSERRH I I G ESETF I NHKVHTALEEGLSVVLCMGETL 132
.....DrTIM..... 59 F . . GGQDVSAHESGAYTGE I SAAMLKDAGASC VVVGH SERR EYHDES DATVA AKARQAQANGLLP I VCVGENL 129
.....ScTIM..... 65 I KYGAQD I SAHDGGA YTTGE I SGPM LAKLKCTYVAVGH SERRQYH AETDE I VNAKVKAA YKHGLTP I LCVGEEL 137
.....CpTIM..... 60 V . . GAQNMHFEKGAFTGE I APRMLEAMN I D YV I I GHSERR EYFNETDETCNKVKAAFAHNLTP I LCCGETL 130

```

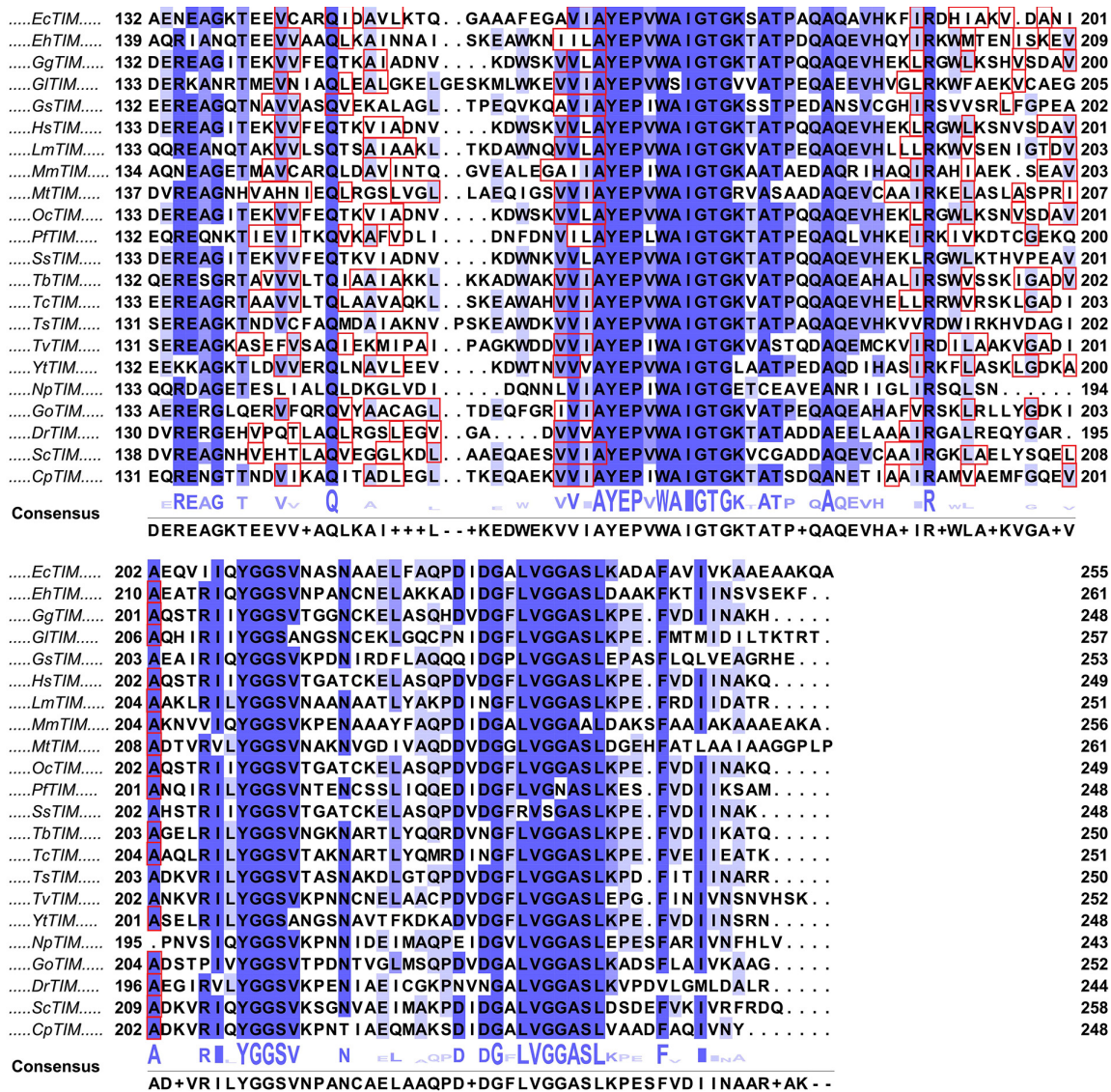
```

Consensus <AQN <GAFTGE S MLKD G <V I L GHSERR Y F G E D E V A K V A A L G L V I C G E L
V -YGAQNVYKVKSGAFTGE+SPAM LKDLGATW I L GHSERRHYFGESEDELVA+KVA+ALAEGLTV I +C I GETL

```

Continue in the next page...





**Fig. S6.** Sequence alignment of diverse TIMs used for comparison. The conserved residues in these TIMs are shown on the *consensus line* and represented from blue light to dark in accordance to increase conservation. In red squares are shown the residues (both internal and boundary) involved in the formation of the biggest cavity in these structures, except for *NpTIM*, *SsTIM* and *TsTIM* for which there is not structural data reported.

**Table S1.** Thermodynamic parameters for the irreversible thermal unfolding of *Np*TIM and *Go*TIM.

Thermodynamic parameter	Bacterial TIMs			Eukaryotic TIMs			
	<i>Np</i> TIM	<i>Go</i> TIM	<i>Gl</i> TIM <sup>a</sup>	<i>Hs</i> TIM <sup>a</sup>	<i>Lm</i> TIM <sup>b</sup>	<i>Tb</i> TIM <sup>b</sup>	<i>Tc</i> TIM <sup>b</sup>
Activation free energy (kJ mol <sup>-1</sup> ) from average of individual fits to DSC thermograms (Eqs. (3) and (4))	368 ± 46	716 ± 63	NR	NR	312 ± 5	397 ± 5	809 ± 7
Activation free energy (kJ mol <sup>-1</sup> ) from Arrhenius plot (Eq. (2))	368 ± 17	716 ± 25	367 ± 7	361 ± 4	315 ± 3	398 ± 5	793 ± 18
Activation free energy (kJ mol <sup>-1</sup> ) from consistency test using several scanning rates (Eq. (5))	343 ± 21	715 ± 88	NR	NR	347 ± 39	420 ± 55	715 ± 137
Activation free energy average (kJ mol <sup>-1</sup> )	360 ± 25	715 ± 46	NR	NR	325 ± 21	406 ± 29	774 ± 75
$m^\ddagger$ (kJ mol <sup>-1</sup> M <sup>-1</sup> )	1.67	8.12	3.94	2.77	1.47	1.69	8.32
$m_{eq}$ (kJ mol <sup>-1</sup> M <sup>-1</sup> )	27.95 <sup>c</sup>	27.77	30.02	30.09	30.80	30.35	30.67
$m^\ddagger / m_{eq}$	0.06	0.29	0.13	0.09	0.05	0.06	0.27
Number of residues unfolded in the transition state <sup>d</sup>	30	148	67	45	24	28	136

<sup>a</sup> From Ref. 18. <sup>b</sup> From Ref. 19. <sup>c</sup> Because the structure of *Np*TIM is not known, the  $\Delta$ ASA value used for the calculation of  $m_{eq}$  is the average of  $\Delta$ ASA from all three-dimensional structures employed for structural comparison ( $57388 \pm 1365$  Å<sup>2</sup>; Table 4). <sup>d</sup> Calculated from:  $N_r(m^\ddagger/m_{eq})$ , where  $N_r$  is the total number of residues in the dimeric protein, and  $m^\ddagger/m_{eq}$  is the degree of unfolding as estimated from the urea  $m$  values. NR: value not reported.

**Table S2.** Data collection and refinement statistics of *Go*TIM, *Dr*TIM, *Sc*TIM and *Cp*TIM.

Data collection <sup>a</sup>				
	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
PDB ID	4Y96	4Y90	4Y9A	4Y8F
Resolution range (Å)	36.36-1.58 (1.64-1.58)	36.13-2.09 (2.17-2.09)	43.03-2.29 (2.37-2.29)	37.59-1.54 (1.59-1.54)
Space group	P 6 <sub>5</sub> 2 2	R 3 2	P 4 <sub>3</sub>	C 1 2 1
Unit cell dimensions				
a, b, c, (Å)	124.8, 127.8, 134.2	169.6, 169.6, 202.3	86.1, 86.1, 134.0	75.3, 49.6, 71.5
α, β, γ, (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0	90.0, 120.0, 90.0
Total reflections	779530 (107888)	261090 (36321)	128420 (19245)	142402 (17348)
Unique reflections	81335 (8192)	64962 (6342)	43475 (4241)	33930 (3130)
Multiplicity	9.6 (9.0)	4.0 (3.9)	3.0 (3.1)	4.2 (3.7)
Completeness (%)	99.3 (97.1)	99.7 (98.6)	99.7 (98.1)	99.2 (94.4)
Mean I / sigma (I)	29.0 (4.8)	11.0 (2.9)	10.0 (3.0)	14.6 (3.1)
R-merge	0.040 (0.487)	0.097 (0.464)	0.071 (0.371)	0.062 (0.448)
Refinement statistics				
	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
R <sub>work</sub> / R <sub>free</sub> (%)	17.8 / 20.1	13.8 / 18.6	23.1 / 28.2	14.8 / 18.1
Average B-value (Å <sup>2</sup> )	24.9	25.9	35.8	16.6
Protein	23.1	24.8	35.9	14.1
Ligand	32.8	49.3	---	31.5
Solvent	35.2	33.2	32.7	29.5
Number of atoms	4446	8069	7943	2362
Protein	3777	7142	7690	1977
Ligand	14	87	---	11
Water	639	832	250	372
Protein residues	500	976	1018	251
RMS (bonds) (Å)	0.007	0.011	0.009	0.010
RMS (angles) (°)	1.11	1.30	1.33	1.24
Ramachandran favored (%)	98.43	97.74	96.15	98.47
Ramachandran allowed (%)	1.18	1.85	3.56	1.15
Ramachandran outliers (%)	0.39	0.41	0.30	0.38
Clashcore	1.83	1.25	5.41	1.76

<sup>a</sup> Statistics for the highest-resolution shell are shown in parentheses.



**Table S3.** Matrix with identity sequence percentage (below) and average RMSD values (top) for IrrevTIMs and RevTIMs.<sup>a</sup>

		RMSD values																						
		IrrevTIMs															RevTIMs							
		<i>Ec</i> TIM	<i>Eh</i> TIM	<i>Gg</i> TIM	<i>Gf</i> TIM	<i>Gs</i> TIM	<i>Hs</i> TIM	<i>Lm</i> TIM	<i>Mm</i> TIM	<i>Mt</i> TIM	<i>Oc</i> TIM	<i>Pf</i> TIM	<i>Ss</i> TIM	<i>Tb</i> TIM	<i>Tc</i> TIM	<i>Ts</i> TIM	<i>Tv</i> TIM	<i>Yf</i> TIM	<i>Np</i> TIM	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	
Identity sequence percentage	IrrevTIMs	<i>Ec</i> TIM	-----	1.51	1.30	1.53	1.48	1.39	1.13	1.24	1.10	1.50	1.27	1.41	1.48		1.24	1.08		1.49	1.37	1.36	1.46	
		<i>Eh</i> TIM	41	-----	0.83	1.50	1.16	0.97	1.09	1.65	1.15	0.93	1.04	0.82	1.07		0.90	1.21		1.51	1.20	0.87	1.07	
		<i>Gg</i> TIM	45	50	-----	1.01	0.92	0.40	0.65	1.20	1.05	0.34	0.85	0.61	0.73		0.62	0.68		1.36	1.00	0.87	1.22	
		<i>Gf</i> TIM	40	43	44	-----	1.21	1.01	1.05	1.09	1.37	0.98	1.30	1.10	0.88		1.04	1.12		0.86	1.49	1.45	1.52	
		<i>Gs</i> TIM	39	41	39	38	-----	0.99	1.49	1.51	1.20	0.94	1.30	1.13	1.30		1.11	1.24		1.33	1.07	0.88	0.78	
		<i>Hs</i> TIM	44	49	90	45	38	-----	0.93	1.26	1.06	0.22	0.96	ND	0.69	0.79	ND	0.64	0.87	ND	1.13	1.29	0.91	1.17
		<i>Lm</i> TIM	42	51	50	46	43	50	-----	1.16	1.13	0.91	0.84	0.46	0.51		1.03	0.92		1.46	1.05	1.23	1.39	
		<i>Mm</i> TIM	65	41	40	35	41	40	38	-----	1.42	1.33	1.36	1.40	1.17		1.17	0.99		1.31	1.39	1.44	1.26	
		<i>Mt</i> TIM	41	36	37	38	43	37	38	41	-----	1.11	0.93	1.11	1.25		1.15	1.10		1.49	0.61	0.70	1.40	
		<i>Oc</i> TIM	44	49	89	45	38	98	50	40	37	-----	1.10	0.76	0.87		0.63	0.86		1.14	1.13	0.92	1.11	
	<i>Pf</i> TIM	39	44	43	40	37	42	45	38	35	42	-----	0.87	0.93		1.07	1.01		1.52	1.10	1.01	1.42		
	<i>Ss</i> TIM	43	48	87	44	37	93	49	39	35	93	42	-----						ND					
	<i>Tb</i> TIM	42	43	51	44	40	52	69	38	38	51	43	39	-----	0.47		0.89	0.98		1.48	1.07	0.96	1.35	
	<i>Tc</i> TIM	42	45	50	46	41	52	68	39	39	51	43	50	74	-----	ND	0.91	0.88	ND	1.38	1.49	1.17	1.48	
	<i>Ts</i> TIM	49	48	61	47	39	59	46	39	38	59	41	58	47	50	-----			ND					
	<i>Tv</i> TIM	40	44	48	41	43	47	43	40	39	47	38	46	42	42	48	-----	1.02		1.15	1.01	0.89	0.95	
	<i>Yf</i> TIM	44	43	53	48	38	52	46	42	39	52	42	51	48	46	49	46	-----	ND	1.49	1.27	1.15	1.47	
	<i>Np</i> TIM	39	38	37	33	47	36	36	43	40	36	33	35	34	32	34	38	35	-----			ND		
	<i>Go</i> TIM	46	41	46	39	45	45	39	43	44	45	40	43	41	40	45	39	40	38	-----	1.25	1.43	1.44	
	<i>Dr</i> TIM	38	34	38	36	41	38	37	39	48	37	33	37	37	39	38	39	37	40	40	-----	0.76	1.35	
<i>Sc</i> TIM	42	40	38	38	48	37	40	40	65	37	34	36	37	38	40	43	37	46	45	49	-----	0.94		
<i>Cp</i> TIM	45	42	41	38	55	40	42	44	41	40	40	40	40	40	41	43	40	48	44	42	52	-----		

<sup>a</sup> ND: average RMSD value not determined because there is no structural data for *Np*TIM, *Ss*TIM and *Ts*TIM.

**Table S4.** Amino acidic composition used in the sequence comparison of IrrevTIMs and RevTIMs.

Structural property	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	Average RevTIMs	Average IrrevTIMs
<b>Amino acid composition (%)</b>					
Alanine (A)	17.0	13.6	12.4	14.3 ± 2.0	12.0 ± 3.1
Cysteine (C)	1.2	1.6	2.4	1.7 ± 0.5	1.5 ± 0.4
Aspartic acid (D)	4.9	7.4	4.8	5.7 ± 1.2	4.4 ± 1.4
Glutamic acid (E)	8.1	7.4	9.6	8.3 ± 0.9	7.2 ± 1.3
Phenylalanine (F)	0.4	1.6	2.8	1.6 ± 1.0	3.2 ± 0.9
Glycine (G)	10.9	8.9	7.2	9.0 ± 1.5	8.8 ± 1.4
Histidine (H)	2.0	3.1	2.0	2.4 ± 0.5	2.3 ± 0.6
Isoleucine (I)	3.2	5.0	8.0	5.4 ± 1.9	6.9 ± 1.2
Lysine (K)	3.6	7.0	7.6	6.1 ± 1.7	6.9 ± 1.8
Leucine (L)	10.1	8.5	4.8	7.8 ± 2.2	7.1 ± 1.3
Methionine (M)	1.6	1.9	3.2	2.2 ± 0.7	1.6 ± 0.7
Asparagine (N)	3.2	2.3	4.8	3.4 ± 1.0	4.0 ± 1.2
Proline (P)	4.0	2.3	2.8	3.1 ± 0.7	3.5 ± 0.9
Glutamine (Q)	2.8	3.9	3.6	3.4 ± 0.4	4.0 ± 1.1
Arginine (R)	4.9	3.5	2.8	3.7 ± 0.9	3.6 ± 0.8
Serine (S)	4.9	3.9	2.4	3.7 ± 1.0	5.1 ± 1.1
Threonine (T)	4.5	4.3	7.2	5.3 ± 1.3	5.1 ± 0.9
Valine (V)	8.9	9.7	8.8	9.1 ± 0.4	9.2 ± 1.1
Tryptophan (W)	1.2	0.8	0.8	0.9 ± 0.2	1.5 ± 0.5
Tyrosine (Y)	2.4	3.5	2.4	2.8 ± 0.5	2.0 ± 0.6
<b>Physicochemical amino acid properties (%)</b>					
Charged (DEHKR)	23.5	28.3	26.7	26.2 ± 2.0	24.4 ± 1.9
Positively charged (HKR)	10.5	13.6	12.4	12.1 ± 1.2	12.7 ± 1.3
Negatively charged (DE)	13.0	14.7	14.3	14.0 ± 0.8	11.6 ± 1.4
Aliphatic (AGILPV)	54.3	48.1	43.8	48.7 ± 4.3	47.5 ± 3.3
Aromatic (FHXY)	6.1	8.9	8.0	7.7 ± 1.2	9.0 ± 0.8
Polar (DEKNQR)	27.5	31.4	33.1	30.7 ± 2.3	30.1 ± 2.4
Neutral, polar (CNQSTY)	19.0	19.4	22.7	20.4 ± 1.7	21.8 ± 2.7
Neutral, non-polar (AFGILMPVW)	57.5	52.3	50.6	53.5 ± 2.9	53.8 ± 2.9
Hydrophobic (CFILMVW)	26.7	29.1	30.7	28.8 ± 1.6	31.0 ± 1.4
Small size (ACGSTV)	47.4	41.9	40.2	43.2 ± 3.1	41.8 ± 3.0
Medium size (DEHILMNQ)	40.1	41.9	43.4	41.8 ± 1.4	41.0 ± 2.9
Large size (FKRWY)	12.6	16.3	16.3	15.1 ± 1.8	17.2 ± 2.1

**Table S5.** Secondary structure elements and stabilizing interactions for IrrevTIMs and RevTIMs.

Structural property	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	Average RevTIMs	Average IrrevTIMs
<b>Secondary structure elements (%)</b>					
$\alpha$ -helix	44	48	43	45.0 $\pm$ 2.2	44.9 $\pm$ 2.1
$\beta$ -strand	21	20	21	20.7 $\pm$ 0.5	21.8 $\pm$ 1.0
Random coil	35	32	36	34.3 $\pm$ 1.7	33.3 $\pm$ 1.8
<b>Stabilizing interactions</b>					
Total H-bonds of the oligomer	490	503	510	501 $\pm$ 8	521 $\pm$ 24
Interface H-bonds	27	19	24	23 $\pm$ 3	28 $\pm$ 5
Total salt bridges of the oligomer	37	50	44	44 $\pm$ 5	34 $\pm$ 7
Interface salt bridges	2	2	12	5 $\pm$ 5	6 $\pm$ 3

**Table S6.** Dimeric TIM PDB files used in the structural comparison of IrrevTIMs and RevTIMs.

Organism	Code	PDB ID	Resolution (Å)	Spatial group	Asymmetric unit <sup>a</sup>	Conformation of active site loops		Reference
						Subunit A	Subunit B	
<i>Clostridium perfringens</i>	<i>Cp</i> TIM	4Y8F	1.54	C 1 2 1	Monomer	Open	Open	This work
<i>Deinococcus radiodurans</i>	<i>Dr</i> TIM	4Y90	2.10	R 3 2	Tetramer	Open	Open	This work
<i>Entamoeba histolytica</i>	<i>Et</i> TIM	1M6J	1.50	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Dimer	Open	Open	3
<i>Escherichia coli</i>	<i>Ec</i> TIM	4K6A	1.80	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Open	Open	4
<i>Gallus gallus</i>	<i>Gg</i> TIM	1TPH	1.80	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Closed	Closed	5
<i>Gemmata obscuriglobus</i>	<i>Go</i> TIM	4Y96	1.58	P 6 <sub>5</sub> 2 2	Dimer	Closed	Open	This work
<i>Geobacillus stearothermophilus</i>	<i>Gs</i> TIM	1BTM	2.80	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Dimer	Closed	Closed	6
<i>Giardia lamblia</i>	<i>Gl</i> TIM	2DP3	2.10	I 2 2 2	Monomer	Closed	Open	7
<i>Homo sapiens</i>	<i>Hs</i> TIM	2JK2	1.70	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Open	Open	8
<i>Leishmania mexicana</i>	<i>Lm</i> TIM	1AMK	1.83	C 1 2 1	Monomer	Closed	Closed	9
<i>Moritella marina</i>	<i>Mm</i> TIM	1AW2	2.65	P 1 2 <sub>1</sub> 1	Dimer	Open	Open	10
<i>Mycobacterium tuberculosis</i>	<i>Mt</i> TIM	3TA6	1.41	C 1 2 1	Dimer	Open	Open	11
<i>Oryctolagus cuniculus</i>	<i>Oc</i> TIM	1R2R	1.50	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Open	Open	12
<i>Plasmodium falciparum</i>	<i>Pf</i> TIM	1YDV	2.20	C 1 2 1	Dimer	Open	Open	13
<i>Saccharomyces cerevisiae</i>	<i>Yt</i> TIM	1NF0	1.60	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Open	Closed	14
<i>Streptomyces coelicolor</i>	<i>Sc</i> TIM	4Y9A	2.30	P 4 <sub>3</sub>	Dimer	Open	Open	This work
<i>Trichomonas vaginalis</i> <sup>b</sup>	<i>Tv</i> TIM	3QSR	2.05	P 2 2 <sub>1</sub> 2 <sub>1</sub>	Monomer	Open	Open	15
<i>Trypanosoma brucei</i>	<i>Tb</i> TIM	5TIM	1.83	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Open	Closed	16
<i>Trypanosoma cruzi</i>	<i>Tc</i> TIM	1TCD	1.83	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Dimer	Closed	Open	17

<sup>a</sup> The number of molecules in asymmetric unit was determined by the Matthews coefficient analysis by solvent content percentage in the crystal.

<sup>b</sup> Structural comparison of *Tv*TIM was realized with the three-dimensional structure of its dimeric Ile45 variant.

## References for Electronic Supplementary Information

1. T. Cavalier-Smith, *Biol. Direct.*, 2006, **1**, 1-19.
2. I. Letunic and P. Bork, *Nucleic Acids Res.*, 2011, **39**, W475-W478.
3. A. Rodríguez-Romero, A. Hernandez-Santoyo, L. del Pozo-Yauner, A. Kornhauser and D. A. Fernández-Velasco, *J. Mol. Biol.*, 2002, **322**: 669-75.
4. G. Minasov, M. Kuhn, A. Halavaty, L. Shuvalova, I. Dubrovskaya, J. Winsor, S. Grimshaw, W. F. Anderson and Center for Structural Genomics of Infectious Diseases, 2013, Forthcoming.
5. Z. Zhang, S. Sugio, E. A. Komives, K. D. Liu, J. R. Knowles, G. A. Petsko and D. Ringe, *Biochemistry*, 1994, **33**, 2830-2837.
6. L. F. Delboni, S. C. Mande, F. Reintier-Delrue, V. Mainfroid, S. Turtley, F. M. D. Vellieux, J. A. Martial and W. G. J. Hol, *Protein Sci.*, 1995, **4**, 2594-2604.
7. H. Reyes-Vivas, A. Díaz, J. Peón, G. Mendoza-Hernández, G. Hernández-Alcántara, I. de la Mora-de la Mora, S. Enríquez-Flores, L. Domínguez-Ramírez and G. López-Velázquez, *J. Mol. Biol.*, 2007, **365**, 752-763.
8. C. Rodríguez-Almazán, R. Arreola, D. Rodríguez-Larrea, B. Aguirre-López, M. T. de Gómez-Puyou, R. Pérez-Montfort, M. Costas, A. Gómez-Puyou and A. Torres-Larios, *J. Biol. Chem.*, 2008, **283**, 23254-23263.
9. J. C. Williams, J. P. Zeelen, G. Neubauer, G. Vriend, J. Backmann, P. A. Michels, A. M. Lambeir and R. K. Wierenga, *Protein Eng.*, 1999, **12**, 243-250.
10. M. Alvarez, J. P. Zeelen, V. Mainfroid, F. Reintier-Delrue, J. A. Martial, L. Wyns, R. K. Wierenga and D. Maes, *J. Biol. Chem.*, 1998, **273**, 2199-2206.
11. S. E. Connor, G. C. Capodagli, M. K. Deaton and S. D. Pegan, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 1017-1022.
12. R. Aparicio, S. T. Ferreira and I. Polikarpov, *J. Mol. Biol.*, 2003, **334**, 1023-1041.
13. S. S. Velanker, S. S. Ray, R. S. Gokhale, S. Suma, H. Balaram, P. Balaram and M. R. N. Murthy, *Structure*, 1997, **5**, 751-761.
14. G. Jögl, S. Rozovsky, A. E. McDermott and L. Tong, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 50-55.

15. S. Lara-González, P. Estrella-Hernández, A. Ochoa-Leyva, M. Del Carmen Portillo-Téllez, L. A. Caro-Gómez, E. E. Figueroa-Angulo, H. Salgado-Lugo, J. F. T. Miranda-Ozuna, J. Ortega-López, R. Arroyo, L. G. Brieba and C. B. Benítez-Cardoza, *Proteins*, 2014, **82**, 22-33.
16. R. K. Wierenga, M. E. M. Noble, G. Vriend and S. Nauche, *J. Mol. Biol.*, 1991, **220**, 995-1015.
17. E. Maldonado, M. Soriano-García, A. Moreno, N. Cabrera, G. Garza-Ramos, M. T. de Gómez-Puyou, A. Gómez-Puyou and R. Pérez-Montfort, *J. Mol. Biol.*, 1998, **283**: 193-203.
18. Y. Aguirre, N. Cabrera, B. Aguirre, R. Pérez-Montfort, A. Hernandez-Santoyo, H. Reyes-Vivas, S. Enríquez-Flores, M. T. de Gómez-Puyou, A. Gómez-Puyou, J. M. Sanchez-Ruiz and M. Costas, *Proteins*, 2014, **82**, 323-335.
19. M. Costas, D. Rodríguez-Larrea, L. De Maria, T. V. Borchert, A. Gómez-Poyou and J. M. Sanchez-Ruiz, *J. Mol. Biol.*, 2009, **385**, 924-937.