Supporting Information

Song et al. 10.1073/pnas.1218581109



Fig. S1. TSR domain is present in pfTRAP (26-299) crystals, but missing in electron density. (A) The structure of pfTRAP (26-299) VWA domain in the same orientation as in Fig. 1 *B* and *C*. (*B*) SDS/PAGE of dissolved crystals and the TRAP protein that was subjected to crystallization. (C) Packing of pfTRAP (26-299) VWA domain in crystal lattice. The chains are shown as ribbons and the C α of terminal residues K240 are shown as spheres.

		Ε-βΑ	β1	α1		β	2 β3	
pvTRAP	DEKVV D EVI	KYSEEVCNES	VDLYLLVDGSGS	IGYPNWITKV	IPMLNGLINSLS	SLSRDTINLY	MNLFGNYTTELIR	l 99
PfTRAP	QNNIVDEI	KYRE evcnDE	VDLYLLMDGSGS	IRRHNWVNHA	VPLAMKLIQQLN	ILNDNAIHLY	ASVFSNNAREIIR	L 103
MIC2	DVIQSDSA	IGAAEG ctNQ	LDICFLIDSSGS	I-GIQNFRLV	KQFLHTFLMVL	PIGPEEVNNA	VVTYSTDVHLQWD	L 136
CMG2		- <u>T</u> scrRA	FDLYFVLDKSGS	V-AN-NWIEI	YNFVQQLAERF\	SPEMRLS	FIVFSSQATIILP	L 95
Mac1		D	SDIAFLIDGSGS	I-IPHDFRRM	KEFVSTVMEQL	KKSKTLFS	LMQYSEEFRIHFT	F 186
VWFA3		csQP	LDVILLLDGSSS	F-PASYFDEM	KSFAKAFISKAN	IIGPRLTQVS	VLQYGSITTIDVP	W 1745
	α2	α3		α4		β4	α5	
pvTRAP	GSGQSIDK	RQALSKVTEL	RKTYTPYGTTNM	TAALDEVQKH	LNDRVNREKA	IQLVILMTD	GVPNSKYRALE	v 170
pfTRAP	HSDASKNKE	EKALIIIKSL	LSTNLPYGKTSL	TDALLQVRKH	LNDRINRENA	NQLVVILTD	GIPDSIQDSLK	E 174
MIC2	QSPNAVDKQ	QLAAHAVLDM	PYKKGSTNT	SDGLKACKQI	LFTGsrpgREH\	PKLVIGMTD	gesd <mark>sdfrtvR</mark>	A 200
CMG2	TGDRO	GKISKGLEDL	kR-VSPVGETYI	HEGLKLANEÇ	IQKagg-lk1	SSIIIALTD	GKLDglvpsyAEK	E 162
Mac1	K-efqn-nH	PNPRSLVKPI	TQLLGRTHT	ATGIRKVVRE	LFNItngaRKNA	FKILVVITD	GEKFg-dplgYED	V 255
VWFA3	N-vv-pEKA	AHLLSLVDVM	QREGGPSQI	GDALGFAVRY	LTSemhgaRPG <i>I</i>	SKAVVILVT	DVSVdsVDA	A 1811
		β5	α6		β6	α7		
pvTRAP	ANKLKORN	/SLAVIGIGQ	GINHQFNR	LIAGCRPR	-EPNCKFYSYA-	-DWNEAVAL	IKPFIAKVCTE	233
PfTRAP	SRKLSDRG	VKIAVFGIGQ	GINVAFNR	FLVGCHPS	-DGKCNLYADS-	-AWENVKNV	IGPFMKAVCVE	237
MIC2	AKEIRELGO	GIVTVLAVGH	YVKHSECR	SMCGCSGTSI	DDSPCPLYLRA-	-DWGQLATA	IKPMLKEVCKT	266
CMG2	AKISRSLGA	ASVYCVGVLD	-FEQAQLE	RIADsk	EQVFPVk	ggFQALKGI	INSILAqsc	218
Mac1	I PEADREG	VIRYVIGVGD	AFrsekSRQELN	TIASkpp	rDHVFQVr	NFEALKTI	QNQLREKIFA-	318
VWFA3	ADAARSNRV	VTVFPIGIGD	RYDAAQLR	ILAGpag	dSNVVK1	RIEDLPTMV	tlgNSFLHKLc	1872
	Ε-βΒ	TSR-1	TSR-β2			TSR-	β3	
pvTRAP	VERVANCG	PWDPWTACSV	TCGRGTHSRSRP	SLH	EF	CTTHM	VSECEEGECP	283
pfTRAP	VEKTASCG	VWDEWSPCSV	TCGKGTRSRKRE	ILH	E(GCTSEL	<i>QEQCEEERCL</i>	287
MIC2	LPQDAICSI	DWSAWSPCSV	SCGDGSQIRTRT	EVSAPQPGTI	TCPDCPAPMGR	CVEQGGLEE.	<i>IRECSAGVCA</i>	337

Fig. S2. Structure-based alignment of TRAP with representative VWA domains. VWA domain structures were aligned using secondary-structure matching (1). Some gaps were closed up manually; resulting structurally nonequivalent aligned residues are in lowercase. Residues not present in crystallized proteins or missing in density are in italics. Aligned structures are MIC (2), CMG2 (3), Mac-1 (4), and VWF A3 (5). The MPP2 cleavage site in MIC2 is indicated by an arrow.

- 1. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr 60(Pt 12 Pt 1): 2256–2268.
- 2. Tonkin ML, Grujic O, Pearce M, Crawford J, Boulanger MJ (2010) Structure of the micronemal protein 2 A/I domain from Toxoplasma gondii. Protein Sci 19(10):1985–1990.
- 3. Lacy DB, Wigelsworth DJ, Scobie HM, Young JA, Collier RJ (2004) Crystal structure of the von Willebrand factor A domain of human capillary morphogenesis protein 2: An anthrax toxin receptor. Proc Natl Acad Sci USA 101(17):6367–6372.
- 4. Lee JO, Rieu P, Arnaout MA, Liddington R (1995) Crystal structure of the A domain from the α subunit of integrin CR3 (CD11b/CD18). Cell 80(4):631–638.
- 5. Bienkowska J, Cruz M, Atiemo A, Handin R, Liddington R (1997) The von willebrand factor A3 domain does not contain a metal ion-dependent adhesion site motif. J Biol Chem 272(40): 25162–25167.

SANG SAN

			• •	• • E-	·βA		β1	**	*		α1			β	2	
P.vivax	25 -	D-EK	<mark>vv</mark> bev	<mark>/KY</mark> SE	E-VCI	JESVI	LYLL	<mark>/ĎGS</mark> G	SI <mark>G</mark> YP	NW ⁱ I T K	VI <mark>PML</mark> N	Ğ<mark>lı</mark>ns	LS <mark>LS</mark> R	D T INL Y	MNLFGNY	T 93
P.falciparum	26 R	DVQNN-	<mark>IVDE</mark>	<mark>KY</mark> RÉ	E- <mark>V</mark> CI	VD <mark>EV</mark> E	LYLL	IDC <mark>S</mark> G	SI RRH	NWVNH	AV <mark>P</mark> LAM	IK <mark>LI</mark> QQ	L <mark>NL</mark> ND	NA <mark>T</mark> HLY	ASVFSNN	A 97
P.knowlesi	25 -	- <mark>D-Q</mark> K	<mark>IV</mark> DEV	<mark>/KY</mark> NE	E- <mark>V</mark> CI	JEK <mark>V</mark> I	LYLL1	/D <mark>GS</mark> G	SI <mark>G</mark> YAI	NWI TR	VI <mark>P</mark> MLT	'G <mark>LI</mark> EN	L <mark>NLS</mark> K	DS INLY	MSLFASH	T 93
P.berghei	25 -	· <mark>QE</mark>	ILDE	IKYSE	E- <mark>VC</mark> I	NEQIC	LHILI	L <mark>DGS</mark> G	SIGHS	NWISH	VI <mark>P</mark> MLT	TLVDN	LNISR	R DEINI S	MTLFSTY	A 92
P.yoelii	25 -	QE	TLDE	KYSE	E-VC1	re <mark>qi</mark> t	THILI	DGSG	SIGYS	NWKAH	VIPMLN	TLVDN.	LNISN	DEINVS	LTLFSTN	S 92
P.gallinaceum	25-	·A-DQ		TYNE			LYLLN		SIGYY		AVPLVE	EIVQN			LSVFTHI	L 93
B.DOVIS T gondii	520		TOCDCA				TCEL	DESA		QWEGQ					UVTV CTD	1 94 V 124
N caninum	54 9		MK GCCT				TCFL		GTGFAI	HVR-F						V 124 V 125
E tenella	27 -	-FSS-I				rst.t.t	VMT.VA	DESC	STOTS	NFR-K	VROFIE				TTTEATR	Q 98
2100110110																
		β3	α2	<u> </u>	α3			*		α4				β4	*	
P.vivax	94 T	ELIRL	Ġ <mark>s</mark> gqsıı	0 <mark>K</mark> RQ <mark>Å</mark>	LS-K	/TE <mark>L</mark> F	RKT <mark>Y</mark> TI	YGTT	NMTAA	LDE <mark>V</mark> -	QK <mark>HĽ</mark> ND	- <mark>RV</mark> NRI	EK <mark>AİQ</mark>	LVILMI	DG <mark>VPN</mark> SK	Y 165
P.falciparum	98 <mark>R</mark>	EIIRL	H <mark>S</mark> DA <mark>S</mark> KI	N <mark>K</mark> EK <mark>A</mark>	LI-II	KŠLI	STNL	YGKT	NLTDA	LLQ <mark>V</mark> -	RK <mark>HL</mark> ND	-RINR	ENANQ	LVVILI	DG <mark>IPD</mark> SI	Q 169
P.knowlesi	94 T	ELIRL	G <mark>S</mark> GP <mark>S</mark> MI	0 <mark>K</mark> KQ <mark>A</mark>	LN-V	/RDLF	RKG <mark>Y</mark> E <mark>I</mark>	YGNT	S <mark>MS</mark> SAI	LSE <mark>V</mark> -	EM <mark>HL</mark> KD	– <mark>RV</mark> NRI	PNAIQ	LVILMI	DG <mark>IPN</mark> NK	Y 165
P.berghei	93 <mark>R</mark>	ELVRLE	KRYG <mark>ST</mark> S	S <mark>K</mark> A- <mark>S</mark>	LRFI]	LAQLC	<u>NNY</u> S <mark>I</mark>	HGTT	NLTSA	LLNV-	DN <mark>LI</mark> QK	- <mark>KM</mark> NR	P <mark>NAI</mark> Q	LVIILI	'DG <mark>IPN</mark> NL	к 164
P.yoelii	93 <mark>R</mark>	ELIKL!	KG <mark>Y</mark> G <mark>ST</mark> S	SKD-S	LRFII	AHLÇ	onn <mark>y</mark> s <mark>i</mark>	PNGNT	NLTSAI	LLV <mark>V</mark> -	DT <mark>LI</mark> NE	-RMYR	PDAIQ	LAIILI	DG <mark>IPN</mark> DL	P 164
P.gallinaceum	94 K	EYIPLI	NSIFSTI	RDFA		RSLF	RTK <mark>Y</mark> SÇ	NG <mark>ST</mark>	NLTLA	LSRVL	KN YF LT	-KGSRI	EDAVQ	LVIIFI	DGS <mark>PD</mark> NK	E 166
B.bovis	95 R	QIFTFI	LD <mark>A</mark> AA <mark>S</mark> S	STRLA	LT-KI	DWMA	GTKAI	RSGMT	Y T GRA	LNYVR	KAILPY	GRI	K <mark>NV</mark> PK	ALLLIT	DGVSSDG	S 165
T.gondii	125 H				AH-AL		-MPYP		NTSDG		QILFTG	SRPGR	CHVPK	LVIGMI	DGESDSD	F 195
N.Caninum	126 H		RANNAS	JKETA		L 1	- IPY	IGGTT	VIIIVO			PREER	D.L.A.BR			F 196
E.Cenerra			SDPKAT	NF SLA	T2-M	Ra	-1913		I I HIG	LQDAR		NAGAR				5 109
		α5		β	5		α6				β6		α7		E-βE	3
P.vivax	166 <mark>R</mark>	ALEVAI	NK <mark>LK</mark> QRI	vvs LA	VIGIC	GQ <mark>GİN</mark>	HQFN	LIAĠ	CR <mark>P</mark>	- <mark>RE</mark> PN	C K FYS Y	ADWNE	A V̈́A LI	KPFIAK	VCTEVER	v 237
P.falciparum	170	SLKES			VFGT		VAFNE	FLVG		-SDGK		SAWEN		GPEMKA	VCVEVER	т 2.41
P.knowlesi	166 8	ALELS	RALKERI	VKLA	VIGIO	GOGIN	HOYN	LMAG		-RERS	CKFYSS		AISLI	KPFIAK	VCTEVER	I 237
P.berghei	165 K	STTVVI	NO <mark>lk</mark> kk		IIGV	GAGVN	NMFNE	ILVG	CG	-KLGP	CPYYSY	GSWDO	AOTMI	KPFLSK		V 235
P.yoelii	165 <mark>R</mark>	STAVV	HQ <mark>LK</mark> RKI	IVNVA	II <mark>G</mark> V	GA <mark>GVN</mark>	NEYNF	RI <mark>LV</mark> G	CD	- <mark>R</mark> YAP	CPYY <mark>S</mark> S	GSWNE	AQNMI	KPFL TK	VCQEVER	I 235
P.gallinaceum	167 <mark>S</mark>	SA <mark>M</mark> KE <mark>V</mark> I	NK <mark>LK</mark> KMI	K <mark>akfa</mark>	VI <mark>G</mark> V	GM <mark>GI</mark> N	KE FNF	(S <mark>LV</mark> G	C- <mark>P</mark> L-	– <mark>KE</mark> KK	CDLYSE	AS <mark>WNE</mark>	VQNVI	A <mark>PFL</mark> KE	VCIEVEK	V 238
B.bovis	166 Y	TAQVA/	AM <mark>LRD</mark> E(SVNVM	VI <mark>G</mark> VG	3D- <mark>V</mark> N	VAEC	RG <mark>IV</mark> G	CD	-GVMD	CPMF KH	TNWKD	IMGLF	'NS <mark>LM</mark> KE	: <mark>VC</mark> DILPQ	D 235
T.gondii	196 <mark>R</mark>	TVRAA	KE <mark>IRE</mark> LO	GIVT	VLAV	HY VK	HSEC	RS <mark>MC</mark> G	CSGTS	DDDSP	CPLY LR	ADWG Q1	LATAI	KPMLKE	VCKTLPQ	D 270
N.caninum	197 H	TVNEA	KVIRER	GIIT	VLSV	SMYVN	HNECH	RSMCG	CR	DSSP	CPLYLQ	TEWSQ	LLPSI	S <mark>PILK</mark> E	VC KTLPK	D 271
E.tenella	1700	TRSSA	AA <mark>LRD</mark> AC	FAIVV	VL <mark>G</mark> V(3S <mark>GVN</mark>	SSEC	KS <mark>IA</mark> G	cs	TSN	CPRYLQ	SNWSN	V TQQ V	NG <mark>IIK</mark> A	ACKDLAK	D 239
	•	•	TSR-	1		TSR-	-β2 💊	٠			_		TS	SR-β3		
P.vivax	238	N <mark>ĊG</mark> P-	WDPWTA	C <mark>ŠVT</mark> C	G-RG	rh <mark>sr</mark> -	<mark>S</mark> RI	S <mark>LH</mark> -				EKCTT	н- <mark>м</mark>	VSE	CEEG <mark>ECE</mark>	283
P.falciparum	242 🗛	SCGV-	WDEWSPO	CSVTC	G-KG	rrsr-		ILH-				EGCTS	E - <mark>I.</mark>	QE0		287
P.knowlesi	238 <mark>A</mark>	KCGP-I	WDDWTP	CSVTC	G-KG	rh <mark>s</mark> r-	<mark>S</mark> RI	PLLH-				AGCTT	н- <mark>м</mark>	VKE	CEMDEC F	283
P.berghei	236 🗛	LCGK-T	WEEWSE(CSTTC	D-NG	rkir-	K <mark>R</mark> I	K <mark>VLH</mark> -				PNCAG	E- <mark>M</mark>	<mark>T</mark> AI	CK <mark>V</mark> RDC	281
P.yoelii	236 <mark>A</mark>	HCGK-I	WEEWSE(CSTTC	D-EG	R <mark>KIR</mark> -	R <mark>R</mark> Ç	2 <mark>ILH</mark> -				PGCVS	E – <mark>M</mark> – –	· <mark>T</mark> TI	CK <mark>V</mark> RDCF	281
P.gallinaceum	239 <mark>A</mark>	HCGS-T	WG <mark>EWS</mark> PO	CSVTC	G-EG	VRTR-	R <mark>R</mark> I	EVLH-				KGCTD	H- <mark>M</mark>	<mark>T</mark> VI	CEKPNCE	284
B.bovis	236 A	VCEPVI	WAEWSS	CKGEC	GVPG	TRTRA	LLDLI	RMIEK	PVSGS	NGQP-	G	KSCED	QKM NF	TLPQSE1	TCT <mark>I</mark> -ECN	302
T.gondii	271 4	ICSD-	WSAWSPO	svsc	G-DG	SQIR-	T R1	revsa	PQPGT	PTCPD	CPAPMC	RTCVE	Q-G-G	LEEIRE	CSAGVCA	337
N.caninum	272	VCSE-		SATC	G-VG	rQG <mark>R</mark> -	<mark>TR</mark> Ç	2QLSP	PAPGT	PTCPD		RSCEE	Q-G-C	VKENRS	CDAGTCS	338
L.CENEIIA	240 <mark>A</mark>	VCSE-	WSEYGPO	-VGEC	GKEG	vors-		/EISP	QK <mark>P</mark> G <mark>S</mark>	PPCPT			Q-PPC	TRTQI	. T.W. P. A. C. R	. 308
		L														

Fig. S3. Sequence alignment of the N-terminal TRAP segment containing the VWA and TSR domains with orthologs in other apicomplexan species. Sequences have GenBank accession numbers AAC97484 (*P. vivax*), AAA29775 (*P. falciparum*), AAG24613 (*Plasmodium knowlesi*), AAB63302 (*Plasmodium berghei*), AAA29768 (*Plasmodium yoelii*), AAC47461 (*Plasmodium gallinaceum*), ACM44016 (*Babesia bovis*), AAB63303 (*Toxoplasma gondii*), AAF01565 (*Neospora caninum*), and AAD03350 (*Eimeria tenella*). Secondary structures are marked above the sequences of *P. vivax* and *P. falciparum*. TRAP constructs. MIDAS residues are marked with asterisks. Disulfide-bonded residues are linked by black lines for *P. vivax* and *P. falciparum*. Conserved residues within the interface between the extensible β ribbon and TSR domain are marked with blue diamonds. Black dots mark decadal residues. In MIC2, E-βA differs. The long loop between TSR-β2 and TSR-β3 in MIC2 with its additional two cysteines is predicted to extend the interface between the TSR domain and extensible β ribbon and compensate for the different character of extensible β ribbon E-βA in MIC2.

< <



Fig. 54. Superposition of TRAP TSR domains from *P. vivax* crystal structure and *P. falciparum* NMR structures. Backbones are magenta for *P. vivax* and green for *P. falciparum* (model 1 of PDB ID 2BBX). The TSR layer residues and the carbohydrate (present only in the crystal structure) are shown as sticks. Residue numbering is for *P. vivax*; His258 is Arg in *P. falciparum*.



Fig. S5. *N*-glycosylation of back-mutated constructs in 293T cells. The crystallization constructs with all potential *N*-linked sites mutated and indicated backmutations to wild-type sequence are compared. Culture supernatants from transient transfections were subjected to reducing SDS 12.5% PAGE and anti-His Western blotting.

AC DNAS

Table S1. Data collection and refinement statistics

pfTRAP (26-299)		pfTRAP (41-240)	pvTRAP	pvTRAP (Mg)	pvTRAP (Mn)	
Data						
Space group	14	P42 ₁ 2	P2 ₁ 2 ₁ 2 ₁	P212121	P212121	
Cell dimensions						
a, b, c (Å)	110.2, 110.2, 47.0	117.7, 117.7, 65.5	56.3, 100.5, 158.6	59.6, 98.0, 159.2	59.6, 98.6,159.5	
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	
Resolution (Å)	43.27–2.2 (2.26–2.2)	41.03–2.25 (2.29–2.25)	42.43–2.2 (2.24–2.2)	41.72–2.24 (2.28–2.24)	39.88–2.2 (2.24–2.2)	
R _{svm} *	0.165 (1.123)	0.157 (1.000)	0.071 (0.315)	0.085 (0.353)	0.132 (0.948)	
l/σl	11.26 (1.69)	10.22 (1.28)	16.72 (3.69)	13.73 (4.36)	10.61 (1.49)	
Completeness (%)	98.1 (97.5)	99.9 (99.6)	98.3 (88.0)	99.6 (99.1)	99.7 (99.4)	
Redundancy	2.89 (2.80)	6.8 (5.7)	3.8 (3.5)	4.0 (3.8)	4.0 (3.8)	
Refinement						
Resolution (Å)	43.27-2.2	41.03-2.25	42.43-2.2	41.72-2.24	39.88-2.2	
No. reflections	14,268	22,394	46,091	45,385	47,461	
$R_{\rm work}/R_{\rm free}^{\dagger}$	0.17/0.22	0.19/0.24	0.16/0.20	0.16/0.20	0.17/0.21	
rms deviations						
Bond lengths (Å)	0.008	0.003	0.009	0.008	0.007	
Bond angles (°)	1.10	0.70	0.99	1.00	0.99	
Residue range	41–240	41–240	25–283	28–283	28–283	
Ramachandran (%) [‡]	98.0/2.0/0	96.7/3.3/0	96.3/3.7/0	97.4/2.6/0	96.1/3.9/0	
PDB code	4HQF	4HQK	4HQO	4HQL	4HQN	

Values for highest resolution shells are in parentheses.

* $R_{sym} = \Sigma i, h | I(i,h) - \langle I(h) \rangle | \Sigma i, h | I(i,h) |$ where I(i,h) and $\langle I(h) \rangle$ are the *i*th and mean measurement of intensity of reflection *h*. [†] R_{free} was calculated using 5% of the data. [‡]Residues in favored, accepted, and outlier regions of the Ramachandran plot as reported by MOLPROBITY.

PNAS PNAS