A novel galectin-like domain from Toxoplasma gondii micronemal protein 1 assists the folding, assembly, and transport of a cell adhesion complex

SAOUROS, Savvas, et al.

Abstract

Immediately prior to invasion Toxoplasma gondii tachyzoites release a large number of micronemal proteins (TgMICs) that participate in host cell attachment and penetration. The TgMIC4-MIC1-MIC6 complex was the first to be identified in T. gondii and has been recently shown to be critical in invasion. This study establishes that the N-terminal thrombospondin type I repeat-like domains (TSR1-like) from TgMIC1 function as an independent adhesin as well as promoting association with TgMIC4. Using the newly solved three-dimensional structure of the C-terminal domain of TgMIC1 we have identified a novel Galectin-like fold that does not possess carbohydrate binding properties and redefines the architecture of TgMIC1. Instead, the TgMIC1 Galectin-like domain interacts and stabilizes TgMIC6, which provides the basis for a highly specific quality control mechanism for successful exit from the early secretory compartments and for subsequent trafficking of the complex to the micronemes.

Reference

SAOUROS, Savvas, et al. A novel galectin-like domain from Toxoplasma gondii micronemal protein 1 assists the folding, assembly, and transport of a cell adhesion complex. Journal of Biological Chemistry, 2005, vol. 280, no. 46, p. 38583-91

DOI: 10.1074/jbc.C500365200
PMID: 16166092

Available at:
http://archive-ouverte.unige.ch/unige:3971

Disclaimer: layout of this document may differ from the published version.
A Novel Galectin-like Domain from Toxoplasma gondii Micronemal Protein 1 Assists the Folding, Assembly, and Transport of a Cell Adhesion Complex*

Received for publication, August 31, 2005, and in revised form, September 12, 2005. Published, JBC Papers in Press, September 15, 2005, DOI 10.1074/jbc.C500365200

Savvas Saouros†§1, Bryn Edwards-Jones†, Matthias Reiss†, Kovilen Samwynaden‡§, Peter Simpson‡§, Timothy J. Dowse†, Ursula Jäkle‡, Stephanie Ramboarina†, Tara Shiv rattan‡§, Stephen Matthews†§2, and Dominique Soldati-Favre†**3

From the †Department of Biological Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom, the ‡Centre for Structural Biology, Imperial College London, South Kensington, London SW7 2AZ, United Kingdom, the **Department of Microbiology and Genetics, Faculty of Medicine, University of Geneva CMU, 1 rue Michel-Servet, 1211 Geneva 4, Switzerland, the †Zentrum für Molekularbiologie Heidelberg, INF 282, 69120 Heidelberg, Germany, and the ‡Department of Molecular Virology, University Heidelberg, INF 345, 69120 Heidelberg, Germany

Immediately prior to invasion Toxoplasma gondii tachyzoites release a large number of micronemal proteins (TgMICs) that participate in host cell attachment and penetration. The TgMIC4-MIC1-MIC6 complex was the first to be identified in T. gondii and has been recently shown to be critical in invasion. This study establishes that the N-terminal thrombospondin type I repeat-like domains (TSR1-like) from TgMIC1 function as an independent adhesin as well as promoting association with TgMIC4. Using the newly solved three-dimensional structure of the C-terminal domain of TgMIC1 we have identified a novel Galectin-like fold that does not possess carbohydrate binding properties and redefines the architecture of TgMIC1. Instead, the TgMIC1Galectin-like domain interacts and stabilizes TgMIC6, which provides the basis for a highly specific quality control mechanism for successful exit from the early secretory compartments and for subsequent trafficking of the complex to the micronemes.

Toxoplasma gondii is a protozoan parasite of the phylum Apicomplexa, which infects virtually all warm-blooded animals and invades almost any cell type. Host cell invasion by this obligate intracellular parasite is an active process initiated by the formation of a tight association/junction with the host cell plasma membrane and leading to the creation of a parasitophorous vacuole. Contact with the host cell results in an increase in parasite intracellular calcium ions, which trigger apical organelles called micronemes to discharge their contents (1). Several micronemal proteins act as ligands for host cell receptors (2), while TgMIC2 and other transmembrane proteins establish a connection with the parasite actinomyosin system via their cytoplasmic tail (3), thus providing the motive force for penetration (4). It is becoming increasingly apparent that many microneme proteins are found in stable adhesin complexes as well as promoting association with TgMIC4. Using the newly solved three-dimensional structure of the C-terminal domain of TgMIC1 we have identified a novel Galectin-like fold that does not possess carbohydrate binding properties and redefines the architecture of TgMIC1. Instead, the TgMIC1Galectin-like domain interacts and stabilizes TgMIC6, which provides the basis for a highly specific quality control mechanism for successful exit from the early secretory compartments and for subsequent trafficking of the complex to the micronemes.

© 2005 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 46, pp. 38583–38591, November 18, 2005

This work was supported by The Wellcome Trust (Research Leave Award to S. M. J.) and Programme Grant (to D. S.), The Medical Research Council Project Grant (to S. M. J.), and The Deutsche Forschungsgemeinschaft (DFG Grant SO 366/1-3 (to D. S.)). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2bvb) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 These authors contributed equally to this work.
2 To whom correspondence may be addressed. E-mail: s.j.matthews@imperial.ac.uk.
3 An International Scholar of the Howard Hughes Medical Institute. To whom correspondence may be addressed. E-mail: dominique.soldati-favre@medecine.unige.ch.

normally comprise an escorter protein, which is responsible for correct micronemal targeting, and one or more soluble effector proteins. The first such complex to be discovered in T. gondii was TgMIC4-MIC1-MIC6, in which TgMIC6 fulfils the role of the escorter protein, whereas TgMIC1 and TgMIC4 function as adhesins (5). Although TgMIC4-MIC1-MIC6 and the recently identified micronemal complex, TgMIC3-MIC8 (5, 6), are individually dispensable, the generation of double knock-outs for TgMIC1 and TgMIC3 renders the parasites avirulent in vivo, demonstrating functional synergy between these complexes (7). Deletion of the mic1 gene in T. gondii also confirmed the specific and critical role played by TgMIC1 in host cell attachment and invasion in vitro.

Micronemal proteins have a modular structure with common themes in domain organization, for example many possess thrombospondin type-1 repeat domains (TSR1),4 apple (or PAN) domains, and epidermal growth factor-like (EGF) domains (8). A schematic representation of the organization within the TgMIC4-MIC1-MIC6 complex is depicted in Fig. 1. TgMIC6 is a 34-kDa transmembrane protein possessing three EGF domains together with an acidic region and the targeting information for transporting the whole complex to the microneme (5). TgMIC4 is a soluble, adhesive 61-kDa protein containing six conserved apple domains, which is processed extracellularly, releasing the sixth apple domain (9). TgMIC1 is a 49-kDa, soluble protein and forms the core of the complex by simultaneously interacting with TgMIC4, TgMIC6, and host cells (9, 10). Furthermore, TgMIC1 is required for the successful exit of TgMIC4 and TgMIC6 from early compartments of the secretory pathway, namely the endoplasmic reticulum (ER) and the Golgi apparatus (5, 11).

Although earlier studies have clearly highlighted the importance of TgMIC1, the finer structural details are poorly understood. This is particularly apparent in sequence annotations for TgMIC1 (7); while the occurrence of two degenerate TSR1-like domains within the N terminus of TgMIC1 is well accepted, the architecture of the C-terminal 190 amino acids has remained unclassified and disregarded. This study addresses many of these issues and uncovers new features of the TgMIC4-MIC1-MIC6 complex. We have identified the regions within TgMIC1, as well as TgMIC4 and TgMIC6 that contribute to complex

4 The abbreviations used are: TSR1, thrombospondin type I repeat; EGF, epidermal growth factor; ER, endoplasmic reticulum; HFF, human foreskin fibroblast; PBS, phosphate-buffered saline; ESA, excreted secreted antigen; BSA, bovine serum albumin; CBF, cell-bound fraction; NOE, nuclear Overhauser effect; HSQC, heteronuclear single quantum correlation; r.m.s.d., root mean square deviation.
A Novel Galectin-like Domain from *T. Gondii* MIC1

**A**

![Diagram A](image1)

**B**

![Diagram B](image2)

**FIGURE 1. Domain organization of TgMIC1, TgMIC4, and TgMIC6.**

- **A**, schematic representation of the domain structure of TgMIC1 (top), TgMIC4 (middle), and TgMIC6 (bottom). Apple (A), TSR1-like, EGF, trans-membrane (TM), and acidic domains are shown. The C-terminal domain in TgMIC1 (CT) is also indicated. The amino acid positions at which the domains begin and end are indicated.
- **B**, schematic representation of the architecture of TgMIC4-MIC1-MIC6 complex showing key interaction sites (this study). The key cleavage points are also indicated.

formation and binding to host cells. Combined NMR and biochemical experiments reveal a novel interaction surface in the previously unidentified Galectin-like fold of TgMIC1, which interacts with TgMIC6. A novel feature is attributed to the Galectin-like fold of TgMIC1, in that it drives the correct folding and subsequent exit of TgMIC6 from the ER and Golgi, revealing a highly specialized type of quality control mechanism.

**EXPERIMENTAL PROCEDURES**

Reagents and Antibodies—Unless otherwise stated all chemical reagents were obtained from Sigma, and all restriction enzymes were purchased from New England Biolabs. Polyclonal rabbit serum reacting to TgMIC4, TgMIC6, and the monoclonal antibody 5B1 anti-TgMIC4 were described previously (5, 9). TgMIC1 was detected with either a polyclonal rabbit serum (10) or using the anti-myc tag monoclonal antibody 9E10.

**Pichia pastoris Cloning and Expression**—The coding sequence corresponding to the amino acids 17–456 of TgMIC1 was amplified from cDNA by PCR. The PCR product was digested with EcoRI and NotI then ligated into the corresponding restriction sites within the multiple cloning site of pPICZα. pPICZ-TgMIC1 was used as a template from which all other TgMIC1 fragments were amplified. The coding sequence corresponding to amino acids 26–580 of TgMIC4 were amplified from genomic DNA. The coding sequence corresponding to amino acids 29–309 of TgMIC6 were also amplified from genomic DNA. All fragments were cloned directly into pPIC9K or via pGEM-T.

*P. pastoris* transformation and expression was performed using the *Pichia* expression kit (Invitrogen), according to the supplied protocols, and all yeast strains were maintained on yeast extract-peptone-dextrose (YPD)-rich media. Transformation of the supplied host strain GS115 was performed by electroporation following linearization of the plasmids with Sall or BglII for pPICZα- and pPIC9K-based vectors, respectively. Selection of transformants was then performed on YPD Zeocin (100 mg/ml) or on minimal media lacking histidine in the case of pPIC9K. Expression was performed using BMGY and BMMY media according to the manufacturer’s instructions.

*Escherichia coli* Cloning and Expression—TgMIC1-CT, spanning residues 320–456 in TgMIC1, was expressed using the pET 21b plasmid (Novagen) in the BL21 (DE3) *E. coli* strain (Stratagene) as described previously (12, 13). TgMIC6-EGF2acid and TgMIC6-EGF3acid were expressed using the pET 32 Xa/LIC plasmid (Novagen) in the Origami (DE3) *E. coli* strain (Novagen). 15N-Labeled samples of TgMIC6-EGF3acid were produced in minimal media, containing 0.07% 15NH4Cl and 0.2% glucose, supplemented with 50 μg/ml carbenicillin. Protein expression was induced by the addition of 500 μM isopropyl β-D-thiogalactopyranoside. TgMIC6-EGF3acid was purified using the binding of the internal hexahistidine tag to the nickel-nitrilotriacetic acid HISBind resin (Novagen) and cleaved from the thioredoxin fusion protein using thrombin. Samples were dialyzed into 20 mM sodium phosphate buffer, pH 7.0, 150 mM NaCl and concentrated to ~0.5 mM for NMR. In all cases only soluble fractions were used for subsequent purification.

**Host Cells and Parasite Cultures and Transfection**—Tachyzoites from *R. haptotroph* and derived mutant strains were propagated in African green monkey kidney (Vero) cells or human foreskin fibroblasts (HFFs) monolayers grown in Dulbecco’s modified Eagle’s medium (Invitrogen) containing 10% fetal calf serum (Invitrogen), 1 mM glutamine, 10 μg/ml gentamicin. The *mic1ko* mutant used in this study corresponds to the *RHxprt*-strain in which the TgMIC1 gene was disrupted by homologous recombination (5). The transient transfection experiments were performed by electroporation as previously described (14).

**Preparation of Excreted Secreted Antigen**—For the preparation of excretory/secretory products, ~2 × 10^6^ tachyzoites were washed in HBBS, 10 mM HEPES, 0.1 mM EGTA (HHE) and resuspended in PBS, pH 7.4, 0.1% fetal calf serum, 1% ethanol to a cell density of 1 × 10^6^ cells/ml and incubated at 37 °C as described previously (2). Cells were removed by centrifugation and the excreted secreted antigen (ESA) product containing supernatant retained.

Co-immunoprecipitation—200 µl of *P. pastoris* supernatant (0.002–0.01 mg/ml protein) or 100 µl of ESA products (equivalent to 1 × 10^6^ tachyzoites) were incubated with an appropriate antibody overnight at 4 °C with agitation. Supernatant or ESA products in the absence of antibodies, and antibodies in the absence of supernatant or ESA products, were included as controls. 200 µl of 10% Protein A-Sepharose CL4B (Amersham Biosciences) slurry in PBS containing 1% BSA was added to each sample and incubated for 3 h at 4 °C with agitation. The beads were washed five times for 10 min in PBS and boiled for 10 min in
50 µl of SDS-PAGE sample loading buffer containing 0.1 M dithiothreitol. The samples were then analyzed by Western blot.

**Cell Binding Assays**—These were performed as described previously (9). Briefly, confluent monolayers of HFFs, Veros, or Chinese hamster ovary cells, grown in 6- or 12-well plates, were blocked for 30 min at 4 °C with 1% BSA in PBS containing 1 mM CaCl2 and 0.5 mM MgCl2 (CM-PBS). Excess BSA was removed by two 5-min washes with ice-cold CM-PBS, after which the proteins to be assayed were then added either in the form of *P. pastoris* culture supernatant (∼0.25 µg) or ESA products (equivalent to 2 × 10^7 tachyzoites) diluted in CM-PBS to a total volume of 500 µl. After incubation at 4 °C for 1 h the unbound fraction was removed, and the cells were washed four times for 5 min with ice-cold CM-PBS. The cell-bound fraction (CBF) was collected either by the direct addition of 50 µl 1× SDS-PAGE loading buffer or by lysing the cells in 1 ml of RIPA (50 mM Tris, pH 7.5, 1% Triton X-100, 0.5% sodium deoxycholate, 0.2% SDS, 100 mM NaCl, 5 mM EDTA) prior to acetic acid precipitation and re-suspension in SDS-PAGE sample loading buffer with 0.1 M dithiothreitol.

**Indirect Immunofluorescence Assay**—All manipulations were carried out at room temperature. 24–30 h after electroporation, tachyzoite-infected HFF cells on glass coverslips were fixed with 4% paraformaldehyde, 0.05% glutaraldehyde for 20 min, followed by 3-min incubation with 0.1 M glycine in PBS. Fixed cells were permeabilized with 0.2% Triton X-100 in PBS for 20 min and blocked in 2% bovine serum albumin in PBS for 20 min. The cells were then stained with the primary antibodies followed by Alexa 594 goat anti-rabbit or Alexa 488-conjugated goat anti-mouse antibodies (Molecular Probes, Cappel, and Bio-Rad). Previously described rabbit polyclonal antibodies were used for the detection of TgMIC4 (9) and TgMIC6 (6). Confocal images were collected with a Leica laser scanning confocal microscope (TCS-NT DM/IRB) using a 100 × Plan-Apo objective with NA 1.30. Single optical sections were recorded with an optimal pinhole of 1.0 (according to Leica instructions) with 16 times averaging.

**NMR Spectroscopy and Structure Calculation**— Backbone and side-chain assignment were completed using standard double and triple-resonance assignment methodology (12, 13). Hα and Hβ assignments were obtained using HBHA (CBCACO)NH (13). The side-chain assignments were completed using HCCH-total correlation (TOCSY) spectroscopy and (H)CC(CO)NH TOCSY (13). Three-dimensional 1H-15N/13C NOESY-HSQC (mixing time 100 ms at 500 and 800 MHz) experiments provided the distance restraints used in the final structure calculation. Heteronuclear 1H-15N NOE data with minimal water saturation were acquired using the pulse sequence described previously (15).

A total of 42 long range NOEs, providing unambiguous three-dimensional information, were manually assigned from the NOE data. The ARIA protocol (16) was used for completion of the NOE assignment and structure calculation. A total of 3156 NOE-derived distances were assigned from 13C- and 15N-edited spectra, which comprised 2275 unambiguous and 881 ambiguous restraints. Dihedral angle restraints derived from TALOS were also implemented (17). Using NOE data and characteristic side-chain chemical shifts a disulfide bond was imposed between cysteines 35 and 106. Cβ chemical shifts are 52 and 42 ppm, respectively. The frequency window tolerance for assigning NOE was ±0.03 ppm and ±0.04 ppm for direct and indirect proton dimensions and ±0.5 and ±1.2 ppm for nitrogen and carbon dimensions, respectively. The ARIA parameters, p, Tv, and Nv, were set to default values. The 10 lowest energy structures had no NOE violations greater than 0.5 Å and dihedral angle violations greater than 5°. The structural statistics are presented in TABLE ONE.

**Chemical Shift Mapping for the MIC1 and MIC6 Interaction**—For NMR mapping experiments, either 15N TgMIC6-EGF3acid or 15N,13C-labeled TgMIC1-CT were prepared in 20 mM sodium phosphate buffer at pH 7 at ~50 µM in 0.25 ml. Unlabeled TgMIC1-CT or TgMIC6-EGF3acid in the same buffer were introduced at several steps up to a 10-fold molar excess and two-dimensional 1H-15N HSQC spectra were recorded at each stage under identical experimental conditions. Perturbed amide resonances in the 15N-TgMIC1-CT/TgMIC6-EGF3acid complex were assigned to the nearest peaks to the 1H-15N HSQC spectrum of TgMIC1-CT. This is likely to represent an underestimation of the actual chemical shift perturbation.
A Novel Galectin-like Domain from T. Gondii MIC1

RESULTS

TgMIC1 Binds to Host Cells in the Absence of TgMIC4 and TgMIC6—The individual contributions of each protein within the TgMIC4-MIC1-MIC6 complex to cell adhesion are poorly understood, particularly with regard to TgMIC1. As a first step to investigating the role of TgMIC1 and its interdependence within the TgMIC4-MIC1-MIC6 complex, the host cell binding ability of each protein was tested in the absence of the others. To do so, wild-type, mic1ko, mic4ko, and mic6ko mutant parasites were stimulated to secrete their micronemal contents, and the yielded ESAs were incubated with HFF cells and analyzed for the presence of bound TgMIC1 or TgMIC4. When produced in the absence of either TgMIC4 or TgMIC6, TgMIC1 bound HFF cells to the same extent as TgMIC1 released from wild-type parasites (Fig. 2A). The ability of TgMIC4 binding to host cells was not affected by the absence of TgMIC6, which is consistent with the lack of a direct interaction between these two proteins (5). In contrast, the absence of TgMIC1 resulted in a considerable decrease in binding of TgMIC4 to host cells (Fig. 2B). Since the absence of TgMIC1 also causes a significant accumulation of TgMIC4 early in the secretory pathway, it is plausible that the lack of TgMIC4 binding could be explained by its incorrect folding, or alternatively the cell binding properties of TgMIC4 may be predominantly conferred by its association with TgMIC1.

The N-terminal TSR1-like Region from TgMIC1 Performs Dual Functions, Cell Binding and Recruitment of TgMIC4—To dissect the functional regions of TgMIC1, fragments from TgMIC1, TgMIC4, and TgMIC6 were expressed in P. pastoris and subsequently tested for cell binding properties and complex formation. The cell binding observations using parasite-secreted antigens (ESAs) were first confirmed with
recombinant TgMIC1 and TgMIC4, revealing that indeed recombinant TgMIC1 can bind to HFF cells independently of any other T. gondii factors (Fig. 2C). However, unlike TgMIC1, little or no recombinant TgMIC4, when expressed alone, could be detected in the cell binding fraction. This is consistent with TgMIC4 being incorrectly folded or having no independent cell binding property. As expected neither native nor recombinant TgMIC6 exhibit detectable cell binding activity (Fig. 2C).

Co-expression of TgMIC1 and TgMIC4 in P. pastoris followed by co-immunoprecipitation together with Western blot analysis revealed an intact TgMIC1-MIC4 sub-complex (Fig. 3A). There was no corresponding increase in the host cell binding efficiency of TgMIC1 in the presence of TgMIC4 indicating that there is no significant direct contribution of TgMIC4. Furthermore, the cell binding efficiency of TgMIC4 binding was markedly increased when expressed in presence of TgMIC1 (Fig. 3B), confirming findings from miclko parasites.

To define the region of TgMIC1 responsible for cell adhesion, fragments containing both TSR1-like domains (termed TgMIC1-TSR1 hereafter), individual TSR1-like domains, or the C-terminal region (termed TgMIC1-CT hereafter) were expressed in P. pastoris and tested separately. A stretch of ~60 amino acids between these two regions was shown previously to be unstructured (12). Only the fragment containing the entire TSR1-like region bound to host cells, indicating that the cell adhesion properties of TgMIC1 require the presence of both intact TSR1-like domains (Fig. 3C). Further experiments in which TgMIC4 was co-expressed with either TgMIC1-TSR1 or TgMIC1-CT revealed that the tandem TSR1-like domains are also responsible for the interaction between TgMIC1 and TgMIC4 (Fig. 3D).

The C-terminal Domain of TgMIC1 Contains a Novel Galectin-like Domain—Although the C-terminal domain possesses no sequence homology with any other protein, secondary structure prediction algorithms suggest the presence of a folded domain within the C-terminal 137 residues. To investigate the nature of TgMIC1-CT we embarked on a high resolution structure determination. Using a combination of manual and automated NMR assignment methods for analysis (16), a family of high resolution structures for TgMIC1-CT were calculated with excellent agreement with experimental data and structural quality (TABLE ONE). All areas of secondary structure are very well defined (Fig. 4A); the average pairwise root mean square deviation (r.m.s.d.) for the water-refined final structures is 0.47 Å for the backbone atoms and 0.92 Å for the heavy atoms of residues within secondary structure (Fig. 4A).

The final structure of the TgMIC1-CT exhibits a Galectin-like topology (Fig. 4) consisting of an 11-stranded β-barrel, formed from the association of a five-stranded β-sheet (βK, βB, βG, βH, and βI) and a six-stranded β-sheet (βA, β, βC, βD, βE, and βF). The strands are connected by short loops, and with the exception of the C-terminal 5 residues, the entire structure forms a globular, rigid scaffold. Low 1H-15N heteronuclear NOE values (data not shown) for the five C-terminal residues (residues 133–137 in TgMIC1-CT) indicate that this region is highly flexible. The structure of TgMIC1-CT also reveals a short, additional strand at the N terminus (βA) that is not present in archetypal galectin domains (Fig. 4, B and C). Despite sequence identities of only 7%, TgMIC1-CT superimposes with an r.m.s.d. of 2.8 Å over 112 equivalent backbone Ca atoms of the human Galectin-3 (Protein Data Bank code 1a3k) (18, 19).

The Galectin-like Domain of TgMIC1 Promotes Folding of TgMIC6—Co-immunoprecipitation and Western blot analysis performed with TgMIC1 fragments and TgMIC6 (Fig. 5A) revealed a specific interaction between TgMIC1-CT and TgMIC6. The extracellular portion of mature TgMIC6 possesses two EGF domains, namely EGF2 and EGF3, together with an acidic region that extends EGF3 at its C terminus (Fig. 1). To localize the region of TgMIC6 responsible for TgMIC1 binding, we expressed EGF2 and EGF3 together or EGF3 alone plus the C-terminal acidic portions (referred to TgMIC6-EGF3acid and TgMIC6-EGF3acid henceforth) together with TgMIC1 and successfully co-precipitated the complex in both cases, indicating a direct interaction between TgMIC1 and TgMIC6-EGF3acid. It was not possible to produce EGF3 alone in a soluble form, which suggests the acid tail may be, at least in part, an elaboration to the canonical EGF domain or confer enhanced solubility.

High resolution detail for any micrometral protein interaction has yet to be reported. The availability of the high resolution structure for TgMIC1-CT provides the opportunity to perform NMR titration experiments as a means to investigate these interacting surfaces. An analysis of amide line widths and chemical shift changes for TgMIC1-CT in the presence of TgMIC6-EGF3acid hence forth together with TgMIC1 and successfully co-precipitated the complex in both cases, indicating a direct interaction between TgMIC1 and TgMIC6-EGF3acid. It was not possible to produce EGF3 alone in a soluble form, which suggests the acid tail may be, at least in part, an elaboration to the canonical EGF domain or confer enhanced solubility.

### TABLE ONE

| Structural statistics for TgMIC1-CT solution structure calculation |
|---------------------------------------------------------------|
| **Solution structure**                                         |
| TgMIC1-CT (Protein Data Bank code 2bvb)                       |
| Number of experimental restraints                             | 3242 |
| Total NOE-derived                                             | 3156 |
| Ambiguous                                                     | 881  |
| Unambiguous                                                   | 2275 |
| Intrarresidue                                                 | 867  |
| Sequential                                                   | 519  |
| Medium range (i−j ≤ 4)                                        | 123  |
| Long range (i−j > 4)                                          | 766  |
| Talos (ψ/φ)                                                   | 86   |
| r.m.s.d. from experimental restraints                         |       |
| Distance (Å)                                                  | 0.018 ± 0.001 |
| Dihedral angle (degree)                                       | 0.6 ± 0.06 |
| r.m.s.d. from idealized covalent geometry                     |       |
| Bonds (Å)                                                     | 0.0041 ± 0.00008 |
| Angles (degree)                                               | 0.55 ± 0.02 |
| Energies (kcal mol−1)                                         |       |
| ENOE                                                          | 51.1 ± 7 |
| Ebond                                                         | 32.5 ± 1 |
| Eangle                                                        | 165.4 ± 12 |
| Evdw                                                          | −1059 ± 64 |
| Coordinate r.m.s.d. (Å)                                       |       |
| Backbone atoms in secondary structure                         | 0.47 ± 0.07 |
| Heavy atoms in secondary structure                             | 0.92 ± 0.1 |
| Ramachandran plot                                             |       |
| Residues in most favored regions (%)                          | 86   |
| Residues in allowed regions (%)                               | 10   |
| Residues in disallowed regions (%)                            | 4    |

The two-dimensional 1H-15N NMR spectrum of TgMIC6-EGF3acid in the absence of TgMIC1-CT is indicative of an unfolded protein (Fig. 5D, left panel). EGF domains often contain a Ca2+ binding site at their N termini that is important for stability and function. No effect was seen...
on the NMR spectrum of TgMIC6-EGF3 acid recorded in the presence of a large excess of \( \text{Ca}^{2+} \) (100 mM), which confirms predictions based on the absence of any \( \text{Ca}^{2+} \) recognition motifs. In contrast, the stepwise addition of TgMIC1-CT had a dramatic effect on the two-dimensional \( ^1\text{H}-^1\text{N} \) NMR spectrum of TgMIC6-EGF3 acid. A number of amide resonances appear at chemical shifts that are characteristic of the folding of a small, structured domain within TgMIC6-EGF3 acid (Fig. 5D, right panel). The structures of classic EGF domains reveal two double-stranded \( \beta \)-sheets and six cysteine residues involved in three disulfide bonds. As residues on both sides of the Galectin fold of TgMIC1-CT are implicated in binding TgMIC6, it is likely that a portion of the acidic tail is also involved in the interaction.

The Galectin-like Domain of TgMIC1 Rescues the Transport of TgMIC6—This, together with those describing the retention of TgMIC4 and TgMIC6 within the ER/Golgi of \( \text{mic} \text{i} \text{k} \text{o} \) parasites (5, 11), implies that correct folding or stabilization of TgMIC6-EGF3 acid assisted by the galectin-like domain of TgMIC1, could provide the necessary quality control mechanism for successful exit from the early compartments of the secretory pathway. To assess this, \( \text{mic} \text{i} \text{k} \text{o} \) parasites were complemented with constructs expressing full-length
A Novel Galectin-like Domain from T. Gondii MIC1

TgMIC1, TgMIC1-TSR1, or TgMIC1-CT (together with a myc epitope for detection). The complementation was analyzed by an indirect immunofluorescence assay and monitored by confocal microscopy. The expression of full-length TgMIC1 was able to rescue fully TgMIC4 and TgMIC6 targeting to the micronemes (Fig. 6B). In contrast, TgMIC1-TSR1 alone did not affect TgMIC6, which was retained in the early secretory compartment (Fig. 6C, bottom panels). Interestingly, however, the presence of the TgMIC1-TSR1 enabled TgMIC4 to exit the ER into the dense granules (Fig. 6C, top panels), which has been previously identified as the default pathway for secretion of soluble proteins in T. gondii (20). This result is consistent with a direct interaction between TgMIC4 and TgMIC1-TSR1 as neither the TSR1-like fragment nor TgMIC4 carry targeting information for the micronemes; the two proteins follow the default pathway and accumulate in the dense granules (Fig. 6C). The presence of TgMIC1-CT does not have any obvious impact on the localization of TgMIC4 (Fig. 6D, top panels). Strikingly, complementation with TgMIC1-CT restores the exit of TgMIC6 from the early secretory compartments and subsequent transport to the micronemes (Fig. 6D, bottom panels).

DISCUSSION

It is well accepted that the binding of micronemal proteins to host cells provides a "molecular bridge" to the parasite, thereby facilitating further steps of invasion. TgMIC1 supplies the platform for assembly of the TgMIC4-MIC1-MIC6 complex by making independent interactions to TgMIC4, TgMIC6, and the host cell (Fig. 1B). Our data reveal that the N-terminal TSR1-like domains from TgMIC1 are able to bind host cells independently of any other parasite factor. Furthermore, we highlight the multifunctionality of the TgMIC1-TSR1, in that it is able to recruit TgMIC4 to the complex and simultaneously anchor TgMIC1 to the host cell surface.

A role for the C-terminal region from TgMIC1 (TgMIC1-CT, Fig. 1) has yet to be revealed, and its annotation been has been neglected in literature studies. Specifically, no functional or sequence annotation for TgMIC1-CT has been reported, which can presumably be attributed to the lack of detectable sequence homology with any other protein. The solution structure of the C-terminal domain from TgMIC1 represents the first atomic resolution insight into a T. gondii micronemal protein and reveals a novel member of Galectin-fold family. Galectins form part of a unique family of soluble, calcium-independent, carbohydrate-binding animal lectins (21, 22). Structures of Galectin-carbohydrate complexes (Fig. 4, C and D) have revealed a consistent picture of carbohydrate recognition, in which critical side chains are located in a central region of the six-stranded β-sheet and comprise an array of hydrophilic residues (in the case of Galectin-3 these include His158, Asn160, Arg162, Glu165, Asn174, Glu184, and Arg186, Fig. 4, C and D) together with the key aromatic side chain of Trp181 (18, 19). These positions are not conserved in the equivalent locations within TgMIC1-CT suggesting that carbohydrate recognition is not retained by TgMIC1-CT. Furthermore, the concave nature of the carbohydrate-binding pocket, formed by sur...
A Novel Galectin-like Domain from T. Gondii MIC1

FIGURE 6. Confocal microscopy of mic1ko mutant parasite complemented with either TgMIC1myc, TgMIC1-5SR1myc, or TgMIC1-CTmyc. A, schematic representation of selected compartments within the secretory pathway of a tachyzoite; ER is shown in yellow, Golgi in blue, dense granules in green, and the micronemes in red. B, the transient expression of TgMIC1myc leads to the correct targeting of TgMIC4 and TgMIC6 in the mic1ko parasites. TgMIC4 and TgMIC6 are retained in the early secretory compartments in non-transfected parasites (indicated by an asterisk). Arrows indicate the presence of co-localization of TgMIC1 with TgMIC4 or TgMIC6 in the micronemes (see A). C, the transient expression of TgMIC-5SR1myc leads to complex formation with TgMIC4. In the absence of TgMIC6, which possesses the microneme targeting information, the complex follows the default secretory pathway to the dense granules (top panels). The arrow indicates co-localization of TgMIC4 and TgMIC1-5SR1myc in the dense granules. However, in the presence of absence of TgMIC1-5SR1myc, TgMIC6 is retained in the early secretory compartments (bottom panels). D, the transient expression of TgMIC1-CTmyc leads to the retention of TgMIC4 in the early secretory compartments (top panels). However, TgMIC-CTmyc induces folding and subsequent correct trafficking of TgMIC6 to the micronemes (bottom panels). The arrow indicates co-localization of TgMIC1-CTmyc and TgMIC6 in the micronemes (see A).

REFERENCES
1. Carruthers, V. B., and Sibley, L. D. (1999) Mol. Microbiol. 31, 421–428
2. Carruthers, V. B., Giddings, O. K., and Sibley, L. D. (1999) Cell. Microbiol. 1, 225–235
3. J. Exp. Med. 1, 201, 201–210
4. J. W., and Soldati, D. (2002) J. Cell Sci. 115, 563–574
5. Cerede, O., Dubremetz, J. F., Soete, M., Di Cristina, M., Dubremetz, J. F., and Soldati, D. (2000) J. Cell. Biol. 152, 563–578
6. Meissner, M., Carruthers, V. B., Tournel, C., Tombor, S., Ajioka, J. W., and Soldati, D. (2002) J. Cell Sci. 115, 563–574
7. Cerde, O., Dubremetz, J. F., Soete, M., Deslee, D., Vial, H., Bout, D., and Lebrun, M. (2005) J. Exp. Med. 201, 453–463
8. Tomley, F. M., and Soldati, D. (2003) Trends Parasitol. 17, 81–88
9. Brecht, S., Carruthers, V. B., Ferguson, D. J. P., Giddings, O. K., Wang, G., Jakle, U., Harper, J. M., Sibley, L. D., and Soldati, D. (2001) J. Biol. Chem. 276, 4119–4127
10. Tomley, F. M., and Soldati, D. (2003) Trends Parasitol. 17, 81–88
11. Huynh, M. H., Opitz, C., Kwok, L. Y., Tomley, F. M., Carruthers, V. B., and Soldati, D. (2004) *Cell. Microbiol.* 6, 771–782
12. Saouros, S., Chen, H. A., Simpson, P., Cota, E., Edwards-Jones, B., Soldati-Favre, D., and Matthews, S. (2005) *J. Biomol. NMR* 31, 177–178
13. Sattler, M., Schleucher, J., and Griesinger, C. (1999) *Prog. NMR Spectrosc.* 34, 93–158
14. Soldati, D., and Boothroyd, J. C. (1993) *Science* 260, 349–352
15. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Watson, T., Formanek, J. D., and Kay, L. E. (1994) *Biochemistry* 33, 5984–6003
16. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2003) *Bioinformatics* 19, 315–316
17. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* 13, 289–302
18. Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S. H., and Rini, J. M. (1998) *J. Biol. Chem.* 273, 13047–13052
19. Umemoto, K., Leffler, H., Venot, A., Valafar, H., and Prestegard, J. H. (2003) *Biochemistry* 42, 3688–3695
20. Karsten, V., Qi, H. L., Beckers, C. I. M., Reddy, A., Dubremetz, J. F., Webster, P., and Joiner, K. A. (1998) *J. Cell Biol.* 141, 1323–1333
21. Barondes, S. H., Cooper, D. N. W., Gitt, M. A., and Leffler, H. (1994) *J. Biol. Chem.* 269, 20807–20810
22. Barondes, S. H., Castronovo, V., Cooper, D. N. W., Cummings, R. D., Drickamer, K., Feizi, T., Gitt, M. A., Hirabayashi, J., Hughes, C., Kasai, K., Leffler, H., Liu, F. T., Lotas, R., Mercurio, A. M., Monsigny, M., Pillai, S., Poirier, F., Raz, A., Rigby, P. W. J., Rini, J. M., and Wang, J. L. (1994) *Cell* 76, 597–598
23. Lourenco, E. V., Pereira, S. R., Faca, V. M., Coelho-Castelo, A. A. M., Mineo, J. R., Roque-Barreira, M. C., Greene, L. J., and Panunto-Castelo, A. (2001) *Glycobiology* 11, 541–547
24. Stebbins, C. E., and Galan, J. E. (2003) *Nat. Rev. Mol. Cell Biol.* 4, 738–743
25. Di Cristina, M., Spaccapelo, R., Soldati, D., Bistoni, F., and Crisanti, A. (2000) *Mol. Cell. Biol.* 20, 7332–7341