Galactose Recognition by the Apicomplexan Parasite
Toxoplasma gondii

Jan Marchant†1, Ben Cowper†1, Yan Liu§, Livia Lai‡, Camila Pinzan*, Jean Baptiste Marq¶, Nikolas Friedrich,** Lloyd Liew†, Wengang Chai†, Robert A. Childs§, Savvas Saouros†, Peter Simpson‡, Maria Cristina Roque Barreira†, Wengang Chai§, Ten Feizi‡, Dominique Soldati-Favre‡, and Stephen Matthews‡

From the †Division of Molecular Biosciences, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom, the §Glycosciences Laboratory, Department of Medicine, Imperial College London, London W12 0NN, United Kingdom, the ¶Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos, Universidade de São Paulo, Ribeirão Preto, São Paulo CEP 14.049-900, Brasil, the **Department of Microbiology and Genetics, Faculty of Medicine, University of Geneva CMU, 1 rue Michel-Servet, 1211 Geneva 4, Switzerland, and the National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom

Background: TgMIC4 is an important microneme effector protein from Toxoplasma gondii.

Results: The structure of TgMIC4 together with carbohydrate microarray analyses reveal a broad specificity for galactose-terminating sequences.

Conclusion: Lectin activity within the fifth apple domain of TgMIC4 is reminiscent of the mammalian galectin family.

Significance: TgMIC4 may contribute to parasite dissemination within the host or down-regulation of the immune response.

Toxoplasma gondii is the model parasite of the phylum Apicomplexa, which contains numerous obligate intracellular parasites of medical and veterinary importance, including Eimeria, Sarcocystis, Cryptosporidium, Cyclospora, and Plasmodium species. Members of this phylum actively enter host cells by a multistep process with the help of microneme protein (MIC) complexes that play important roles in motility, host cell attachment, movement junction formation, and invasion. T. gondii (Tg)MIC1-4-6 complex is the most extensively investigated microneme complex, which contributes to host cell recognition and attachment via the action of TgMIC1, a sialic acid-binding adhesin. Here, we report the structure of TgMIC4 and reveal its carbohydrate-binding specificity to a variety of galactose-containing carbohydrate ligands. The lectin is composed of six apple domains in which the fifth domain displays a potent galactose-binding activity, and which is cleaved from the complex during parasite invasion. We propose that galactose recognition by TgMIC4 may compromise host protection from galectin-mediated activation of the host immune system.

Toxoplasma gondii is an obligate intracellular parasite of the phylum Apicomplexa, and is prevalent among human populations, with a worldwide infection rate of up to 50%. Infection in humans primarily occurs following consumption of undercooked infected meat or contact with feces from infected domestic cats. In healthy adults the infection is generally either asymptomatic or results in a mild, flu-like illness, which marks the beginning of a lifelong chronic infection (1). In immunocompromized individuals infection can lead to acute disease, in which T. gondii awakens from a semidormant state causing blindness (2) or potentially fatal encephalitis (3, 4). Infection in pregnant women can result in a range of fetal birth defects or death (5). Due to its remarkably high infection rate, T. gondii constitutes the third most common cause of food-related death in both the United States (6) and France (7) after Salmonella and Listeria.

T. gondii and other apicomplexan parasites, including Plasmodium species, rely on an active, phylum-specific host cell invasion process to establish infection. Host cell entry consists of several sequential steps initiated by the release of proteins from secretory organelles named micronemes and rhoptries (8). Some of the microneme protein complexes contribute to host cell attachment and provide a link between host cell receptors and the parasite actomyosin motor and hence provides the motive force necessary for host cell invasion (9, 10). Among the microneme protein complexes operating in invasion, the TgMIC1-4-6 complex is important for the efficiency of this process and has been shown to contribute to the virulence of the parasite in mice (11, 12). The structure of the N-terminal region of TgMIC1 revealed a pair of novel domains termed the microneme adhesive repeat region (MARR) (13), in complex...
with a wide range of sialylated glycans (14). TgMIC1 not only interacts with a range of sialylated glycans on host cell receptors, but also recruits TgMIC4, which is anticipated to exert adhesive function during invasion (15). Previous studies suggested that the first two apple (A) domains of TgMIC4 bind to the N terminus of TgMIC1 (16, 17), whereas the C-terminal fragment (including the sixth apple domain) exhibited cell binding activity of unknown specificity (supplemental Fig. S1) (16). Although lectin activity has not yet been reported for TgMIC4, speculation has recently arisen due to an earlier report showing that a TgMIC1-4 subcomplex could be recovered from a lactose-affinity column (18) and our previous studies revealing TgMIC1 specificity for sialylated oligosaccharides only (13).

Here, we combine atomic resolution studies with data from carbohydrate microarrays to reveal the basis of the interaction between TgMIC4 and a variety of galactose (Gal)-terminating oligosaccharides, and further define the interaction between TgMIC4 and TgMIC1. This reveals new features regarding both parasite-receptor interactions and the stoichiometry of the TgMIC1-4-6 complex.

**EXPERIMENTAL PROCEDURES**

**Cloning, Expression, and Purification from Escherichia coli—**

15N,13C-Labeled samples of TgMIC4-A12 (spanning residues 58 to 231 of TgMIC4), TgMIC4-A5 (residues 410–491), and TgMIC4-A3 (residues 225–304) were expressed using the pET32Xa/LIC plasmid (Novagen) in E. coli Origami (DE3) (Strategene), at 30°C in minimal media containing 0.07% 15NH4Cl and 0.2% [13C6]glucose. Protein expression was induced with 500 μM isopropyl β-D-thiogalactopyranoside. The hexahistidine-thioredoxin fusion protein was purified using nickel-nitritolotriacetic acid His Bind resin (Novagen), and cleaved with Factor Xa (Invitrogen). The cleaved protein was re-applied to the same column, and pure target protein was recovered in the flow-through. The protein was buffer exchanged into 20 mM K2HPO4/KH2PO4, 100 mM NaCl, pH 6.5, and concentrated to ~10 mg/ml. TgMIC1-MARR was expressed and purified as previously described (13).

**Solution Structure Determination of TgMIC4-A12 and TgMIC4-A5—**Backbone and side chain assignments were completed using our in-house, semiautomated assignment algorithms and standard triple resonance assignment methodology (19). Hα and Hβ assignments were obtained using HBBHa(CB-CACOaNH). The side chain assignments were completed using HCCCH-total correlation (TOCSY) spectroscopy and HCCCHNH TOCSY. 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. The ARIA protocol (20) was used for completion of the NOE assignment and structure calculation, including water refinement of final structures using a thin layer of explicit solvent (21). Dihedral angle restraints derived from TAALOS were also implemented (22). The structural statistics are presented in the supplemental data.

**Calculation of TgMIC4-A5/Lacto-N-biose Structural Model—**TgMIC4-A5 was incrementally titrated with galactose (Melford Laboratories), lactose (Melford Laboratories) (Dextra Laboratories), 2,3-sialyl-N-acetyllactosamine (Sigma), lacto-N-biose (Dextra Laboratories), and GM1-penta (Enzo Life Sciences). Chemical shift perturbations were monitored via 1H-15N-HSQC.

A structural model for the TgMIC4-A5 complex was calculated using the molecular docking program HADDOCK (23) Selected residues that exhibited a significant shift perturbation, namely Lys-428, Asn-460, Tyr-467, Lys-469, Tyr-476 and Tyr-478, were defined as “active” for docking. Additionally intermolecular NOEs between TgMIC4-A5 and lacto-N-biose were measured via a three-dimensional 13C-filtered 1H-13C-HSQC-NOESY experiment. TgMIC4-A5 aromatic nuclei were re-assigned in the presence of lacto-N-biose via titration of the protein with the ligand and observance of chemical shift perturbations in the 1H-13C-HSQC spectrum (selective for aromatic nuclei). The presence of a 5-fold excess of lacto-N-biose enabled NOE-bearing nuclei to be assigned using a sample of free disaccharide. Lacto-N-biose 1H chemical shift assignment was carried out using data from 1H-1H TOCSY, 1H-1H COSY, 1H-13C NOESY, and 13C-HSQC NMR spectra. Intermolecular NOEs were implemented as distance restraints alongside chemical shift perturbation-derived ambiguous interaction restraints to drive structure calculation using HADDOCK version 2.1. The docking process utilized an ensemble of 10 low-energy conformers from the TgMIC4-A5 structure determination, and a lacto-N-biose structure (and parameter and topology files) created using the GlyCaNS server.

**Isothermal Titration Calorimetry—**Isothermal titration calorimetry data for the apple-5/lacto-N-biose interaction was collected using a VP-Isothermal Titration Calorimetry Microcalorimeter (MicroCal™). Apple-5 (650 μM) was incrementally titrated with 5-μl volumes of lacto-N-biose (9.75 mM) at 5-min intervals for 10 injections, followed by 10-μl volumes for a further 24 injections. All samples were constituted in 20 mM K2HPO4/KH2PO4, 100 mM NaCl, pH 6.5, and the experiment was performed at 303 K. Data were analyzed using Origin® version 7.0 software.

**T. gondii Culture—**T. gondii tachyzoites were grown in confluent human foreskin fibroblasts or Vero cells maintained in Dulbecco’s modified Eagle’s medium (Invitrogen) supplemented with 10% fetal calf serum, 2 mM glutamine, and 25 μg/ml of gentamicin.

**Immunofluorescence Assay and Confocal Microscopy—**Parasite-infected human foreskin fibroblasts were fixed with 4% paraformaldehyde or 4% parafomaldehyde, 0.05% glutaraldehyde in PBS, depending of the antigen to be labeled and processed as previously described (24). Confocal images were collected with a Leica laser scanning confocal microscope (TCS-NT DM/IRB and SP2) using a 1003 Plan-Apo objective with NA 1.4.

**Plaque Assay—**Host cells were infected with parasites for 7 days before fixation with paraformaldehyde/glutaraldehyde followed by Giemsa staining.

**Antibodies—**The antibodies used in this study were previously described. In short, α-GAP45 (25), α-MIC4 (15), α-MIC1, and α-MIC2 were kindly provided by J. F. Dubremetz. α-MIC6CT (16), the c-Myc tag, was detected with α-Myc (mAb 9E10).
**Galactose Recognition by T. gondii**

**Invasion Assay**—Invasion assays were performed as described previously using the RH-2YFP strain as an internal standard (26). Briefly, confluent human foreskin fibroblasts have been heavily infected with a mixture of the strain of interest and RH-2YFP parasites and washed out after 1 h. Parasites were incubated for 36 h. Extracellular parasites were then collected and the ratio of non-YFP to YFP parasites was determined. At the same time these parasites were transferred on new host cells and allowed to invade for 1 h at 37 °C before washing. Then, incubation continued for 24 h and cells were fixed. Parasites were stained with α-GAP45 and the ratio between non-YFP and YFP parasite vacuoles was calculated. The efficiency of invasion was determined by counting vacuoles in 20 fields (e.g. around 400 vacuoles) for each condition and for two independent experiments.

**Cloning of pSAG1MIC1myc δβ-Linker**—Deletion in the pSAG1MIC1myc was performed using the QuikChange site-directed mutagenesis kit (Stratagene catalog number 200518) and the primer pair 2940, CCCGAATTGCTGGTGTCGACGGATCCG and the primer pair 2941, CCGCAATTGAGAACATGGCGTCGACGGATCCG. The full-length PCR product was digested by the restriction enzyme MfeI and ligated, and transformed into bacteria.

**Carbohydrate Microarray Analyses**—The microarray analyses were performed using the neoglycolipid-based microarray system that contains sequence-defined lipid-linked oligosaccharide probes: glycolipids and neoglycolipids (27, 28). The repertoire of 400 probes is described in supplemental Table S2. Among these are 248 neutral sequences, 55 acidic but nonsialylated, and 97 sialylated probes with differing sialic acid linkage glycan backbone, chain length, and sequence. The lipid-linked glycan probes were printed onto 16-pd nitrocellulose-coated glass slides, in duplicate at two levels, 2 and 5 fmol/spot, as described (28). The screening microarray binding analyses were performed at ambient temperature. His-tagged TgMIC4-A12, TgMIC4-A5, and TgMIC1-MARR were assayed essentially as described (12, 13). In brief, the arrayed slides were blocked for 1 h with 1% (w/v) bovine serum albumin (Sigma) in Pierce Casein Blocker solution (casein/BSA). The His-tagged proteins were precomplexed with mouse monoclonal anti-polyhistidine (Sigma) and biotinylated goat anti-mouse IgG antibodies (Sigma) in a ratio of 1:2.5:2.5 (by weight) and overlaid onto the arrays. TgMIC4-A12 and TgMIC1-MARR were tested at 40 μg/ml and TgMIC4-A5 was tested at 20 μg/ml. Binding was detected using Alexa Fluor 647-conjugated streptavidin (Molecular Probes). Microarray data analysis and presentation were carried out using dedicated software. The binding to oligosaccharide probes was dose-related, and results at 5 fmol/spot are shown.

To closely compare the binding preferences of TgMIC4-A5 a focused microarray "dose-response" format was used, which included glycolipids asialo-GM1, GM1, SM1a, and SB1a (29). For this, the probes were quantified at the same time and printed in duplicate at four levels: 0.3, 1, 2, and 5 fmol/spot.

**Blue Native PAGE and SDS-PAGE of Native TgMIC1-4 Complex**—Native TgMIC1-4 complex from tachyzoites of the virulent RH strain of T. gondii was purified as described previously (18). BN-PAGE experiments were performed with the Native PAGE Novex BisTris Gel System (Invitrogen) according to the instructions, using a 4–16% gradient gel. In addition to natively purified complex (Lane”), protein samples were denatured either in 8 M urea or 6 M guanidinium chloride for 4 h at room temperature. Samples were mixed with Native PAGE loading buffer (Invitrogen), and 5 μg of protein was loaded into the gel. Electrophoresis was performed in an ice bath, at 150 V until completion of dye migration. The gel was stained using the Colloidal Blue Staining Kit (Invitrogen). A high molecular weight calibration kit (GE Healthcare) was used to indicate the protein size.

Natively purified complex was analyzed via SDS-PAGE using a 4–12% BisTris polyacrylamide gel (Invitrogen) in NuPAGE MOPS-SDS buffer, pH 7.7, according to the manufacturer’s instructions. In addition, a 220 kDa band from blue native PAGE was excised from the native gel and the proteins were eluted by incubating the gel slice with SDS-sample buffer overnight. This was incubated at 90 °C for 5 min and the supernatants were loaded into the gel. SDS-PAGE was performed at 170 V until completion of sample migration. Protein bands were visualized via silver staining using a Silver Staining Plus kit (Bio-Rad).

**RESULTS AND DISCUSSION**

**The Overall Structure of Apple Domain Pair from TgMIC4**—Most previous structural studies of apple domains of various proteins have been focused on individual apple domains (30–32). Sequence analysis reveals that TgMIC4 comprises six apple domains that occur in intimate pairs, with only 3 residues separating the first and second (A12), third and fourth (A34), and fifth and sixth (A56) domains (15). A phylogenetic analysis of the individual apple domains from TgMIC4 showed a clear divergence between the odd and even domains, indicating that they may naturally form pairs (33) (supplemental Fig. S2).

To investigate the arrangement of apple domains in TgMIC4, we determined the solution structure of the first two apple domains, namely A1 and A2 comprising residues 58–231 (TgMIC4-A12), using heteronuclear multidimensional nuclear magnetic resonance (NMR) spectroscopy. The structure reveals an intimately associated pair of apple domains (Fig. 1 and supplemental Fig. S3), each consisting of a sheet formed of 4–5 antiparallel β-strands that cradles an α-helix. On the other face of this β-sheet lies an additional smaller sheet that is formed from two short β-strands, and in the second domain is extended with an additional 4-residue helix. Two disulfide bridges connect the helix to the central strands of the sheet (C2:C5 and C3:C4), with a further disulfide bridge connecting the N terminus to the C terminus (C1:C6). These structural features correspond to previously determined apple domain structures (supplemental Fig. S4). The ensemble of the 10 lowest energy structures has been deposited in the Protein Data Bank under accession number 4A5V (Fig. 1A and Supplemental Table S1). The interface between the domains is mediated in the main by a collection of hydrophobic residues, most notably a patch of four alanine residues found on the outer face of the helix, Phe-213 and Met-158 in A2 and Leu-124 and Pro-122 from A1. A large number of NOEs can be assigned between the two domains across the interface (supplemental Fig. S3).
Overall Arrangement of TgMIC4 Apple Domains—In TgMIC4, residues 232–380 (TgMIC4-A34) and 419–565 (TgMIC4-A56) show sequence identities of 41 and 43%, respectively, with TgMIC4-A12. These high sequence similarities suggest that the structures of these domains can be generated by homology modeling. Models of MIC4-A34 and MIC4-A56 were created using MODELLER (34). Both models have a GA341 score of 1.0, corresponding to “native-like” (Fig. 2).

Analysis of all apple domain structures determined to date reveals only one in which the arrangement of tandem pairs has been established: the crystal structure of the four apple domains found in coagulation factor XI (35). Although the apple domains found in coagulation factor XI have a relatively low sequence similarity to TgMIC4-A12, a comparison of the arrangement of apple domains in this protein with TgMIC4-A12 reveals a very similar structural interface (Fig. 2). Interestingly, the relative orientation of the domains is reversed. In coagulation factor XI, helix 1 of the odd numbered domains is buried at the interface, whereas in TgMIC4-A12, helix 1 of the even numbered domains is found at the interface.

The similarity in the arrangement of the apple domain pairs to that found in coagulation factor XI suggests that the full-length TgMIC4 protein may adopt a similar disc-like structure (Fig. 2). The significantly longer linker between the fourth and fifth apple domains in TgMIC4, and the fact that this region is cleaved by a parasite-encoded protease, TgSUB1, at the surface of the parasite releasing TgMIC4-A56 (36), suggests a model in which coagulation factor XI arrangement is maintained between the first two pairs, with the third pair (A56) more loosely associated (Fig. 2).

TgMIC4 Interacts with the β-Finger of TgMIC1 via Second Apple Domain—Previous studies on mutant parasites indicate that TgMIC4-A12 interacts with the N-terminal region of TgMIC1 within the TgMIC1-4-6 complex (16). To investigate this interaction we performed an NMR titration using recombinantly expressed TgMIC1-MARR and TgMIC4-A12. No interaction could be observed by NMR, which was also confirmed by isothermal titration calorimetry and analytical gel filtration. The disparity between in vivo and in vitro results lead us to reanalyze the crystal structure of TgMIC1-MARR (13). There are two additional cysteine residues present in the β-finger motif at the C-terminal end of MAR2, which introduce a rearrangement of the disulfide bond pattern in MAR2 compared with that seen in MAR1. The expected pairing between C4 and C6 is broken and two new disulfide bonds are made with the cysteine residues from the β-finger. This results in the β-finger motif being pinned against the surface of the MAR2 domain and we hypothesize that this may block a potential interaction between recombinant TgMIC1 and TgMIC4 (supplemental Fig. S5). The normal protein folding environment within the ER of the parasite and subsequent quality control checks would enforce an alternative bonding pattern in MAR2 similar to the one observed for MAR1 happening in vivo and allow correct assembly of the complex.

To investigate the possibility that this disulfide rearrangement in recombinant TgMIC1-MARR is responsible for the
lack of binding in vitro, a peptide corresponding to the TgMIC1 β-finger (residues 237–256) was synthesized and an intramolecular bond was formed between the two β-finger cysteine side chains. An NMR titration experiment was performed by recording the 1H-15N-HSQC spectra of 15N-labeled TgMIC4-A12 in the presence of increasing amounts of peptide. A number of chemical shift perturbations are seen in this spectrum (Fig. 3A). The residues undergoing chemical shift perturbation are localized exclusively to A2 (Fig. 3B), suggesting that an interaction exists between the second apple domain of TgMIC4 and the β-finger region of TgMIC1. Further studies will be required to elucidate the precise binding mode of the peptide.

The “β-Finger” Region of TgMIC1 Is Required for Correct Trafficking of TgMIC4 in T. gondii—It has been shown previously that an interaction exists between TgMIC1 and TgMIC4 that facilitates the correct targeting of TgMIC4 to the micronemes (16, 37). In these studies complementation of the mic4ko strain with either TgMIC4 or TgMIC4-A12 results in their correct sorting to the micronemes, suggesting that the first pair of apple domains is sufficient to mediate interaction with TgMIC1 within the parasite. To determine the significance of the potential interaction between TgMIC4-A2 and the β-finger motif from TgMIC1, the transport of the components of the complex was analyzed on mic1 knockout strain (mic1ko) complemented with a construct expressing TgMIC1-Δβfinger (lacking residues 216 to 237). The expression of TgMIC1-Δβfinger was assessed by Western blot (Fig. 4A). When mic1ko is complemented with full-length TgMIC1, TgMIC6 and TgMIC4 are successfully targeted to the micronemes (16) (Fig. 4B). In contrast, when mic1ko parasites are complemented with
TgMIC1-Δβfinger, the TgMIC1 mutant protein and TgMIC6 are correctly transported to the micronemes, whereas TgMIC4 remains mislocalized (Fig. 4B). As observed in mic1ko, TgMIC4 is blocked in earlier compartments of the secretory pathway, suggesting that the β-finger forms a necessary part of the TgMIC4-MIC1 interface and is crucial for the sorting of TgMIC4. These in vivo results are in excellent agreement with the structural data.

We next compared these strains in cell invasion assays (Fig. 4C). In the mic1ko strain, invasion is significantly reduced compared with wild-type confirming the role for MAR domains in invasion (12, 13). Efficient invasion can be restored when this strain is complemented with TgMIC1wt, whereas complementation with TgMIC1-Δβfinger also significantly improves invasion efficiency. This suggests that the folding and activity of MAR domains is retained in the absence of the β-finger, ruling out the possibility that the loss of TgMIC4 recruitment is merely due to a loss of TgMIC1 structural integrity. Additionally, the observed deficiency of invasion recovery by TgMIC1-Δβfinger compared with TgMIC1wt may be understandable by the loss of cell adhesion via TgMIC4.

**Figure 4.** Functional characterization of mic1ko parasites complemented with either full-length TgMIC1 or TgMIC1 with the β-finger region excised (mic1ko + TgMIC1Δβ-finger). A, Western blot analysis of parasite lines expressing TgMIC1wt or TgMIC1Δβ-finger on the mic1ko background. The recipient strain for knockouts RHΔHX and mic1ko and mic4ko are included as controls. The transgenic expression of TgMIC1wt or TgMIC1Δβ-finger is detected by anti-Myc and compared with the endogenous level of MIC1 with anti-MIC1 antibodies. Detection of catalase is used as loading control. B, upper panel, immunofluorescence assays showing that TgMIC1Δβ-finger is stably expressed and correctly targeted to the micronemes (anti-MIC1; green). Localization is shown relative to the myosin motor complex under the parasite plasma membrane using anti-TgGAP45 (red). Middle panel, trafficking of TgMIC4 (anti-MIC4; red) to the micronemes is disrupted (not rescued) in the mutant parasite carrying TgMIC1Δβ-finger and appears distributed through the early secretory pathways as observed in mic1ko. Lower panel, in contrast to TgMIC4, the trafficking of TgMIC6 is restored to the micronemes by the TgMIC1Δβ-finger. C, cell invasion assays comparing the T. gondii mutant strains demonstrate that the expression of TgMIC1Δβ-finger restores invasion in mic1ko to the wild-type level. The considerable overexpression of TgMIC1wt lead to a significant increased efficiency of invasion. Error bars indicate S.D. D, the plaque assay recapitulates several lytic cycles of the parasites. The mic1ko parasites expressing either TgMIC1wt or TgMIC1Δβ-finger form plaques of comparable size, whereas mic1ko form smaller plaques.
TgMIC6 and the C-terminal galactin-like domain of TgMIC1. TgMIC6 possesses three EGF domains in which the TgMIC1-binding interface is conserved and therefore could recruit up to three molecules of TgMIC1 (supplemental Fig. S6) (26, 37). This would suggest that although the first TgMIC6-EGF domain is eventually cleaved during trafficking of the complex through the secretory pathway, the TgMIC1 trimer would remain intact until secreted onto the parasite surface.

A disulfide-bonded 440-kDa TgMIC4/TgMIC1 species corresponds in molecular mass to approximately six molecules, with a possible arrangement being three of TgMIC4 joining a trimeric platform of TgMIC1. Although the arrangement of intermolecular disulfide bonds remains undefined, it is worth noting the proximity of a seventh, unpaired cysteine (Cys-263; known hereafter as C3) from TgMIC4-A3 to the conserved disulfide linkages within the internal α-helix/β-hairpin loop and to A2 in our TgMIC4 model (Fig. 2). Recombinant production of TgMIC4-A3 yields a polydisperse sample and NMR analysis identifies three folded species in approximately equal quantities (supplemental Fig. S7). Peptide fingerprinting via MALDI (matrix-assisted laser deabsorption/ionization) mass spectrometry (MS) under nonreducing conditions indicates that one species contains the expected pattern of disulfide linkages (C1:C6, C2:C5, and C3:C4 with C3 free; supplemental Fig. S7C). The two additional species contain mismatched and free C3, C3', and C4. Due to the high homology with A1 and A5 (in which alanine replaces the free cysteine residue), it was reasoned that mutation of C3 should prevent disulfide scrambling, resulting in a monodisperse sample. However, combined NMR and MALDI-MS data reveals that recombinant TgMIC4-A3C263A adopts two stable conformations, in which C3/C4 are disulfide-linked and free, respectively (supplemental Fig. S7C). The dynamic nature of this region together with the labile nature of the disulfide pattern would provide a mechanism for covalent association with TgMIC1. Surface-associated disulfide isomerases may also contribute the necessary shuffling of sulfide bonds. Interestingly, the homologue of MIC4 in the closely related species Neospora caninum does not contain this additional cysteine residue, and NcMIC1 (the homologue of TgMIC1) does not co-purify from parasite lysates (38, 39). In summary, we propose a model (Fig. 5C) in which the TgMIC1-4 complex forms an array on the parasite surface composed of trimeric TgMIC1 and heterohexameric TgMIC1-4 subcomplexes, anchored via TgMIC6.

FIGURE 5. Analysis of TgMIC1-4 native complex by SDS-PAGE and blue native PAGE. Native TgMIC1-4 complex was purified via lactose-affinity chromatography. A, the lactose-binding (Lac+) fraction was subjected to SDS-PAGE. Molecular mass standards were loaded in lane 1. TgMICs migration positions are indicated (lane 2). B, blue native PAGE. The gel was loaded with NativeMark™ standards (lane 1), alongside 5 μg of native proteins (Lac+) (lane 2) and 5 μg of Lac− solubilized in 8 M urea or 6 M guanidinium chloride (lanes 3 and 4, respectively). C, a proposed model of TgMIC1-4 assembly. The molecular masses of the native subcomplexes are consistent with the existence of TgMIC1-4 hetero-hexamer and TgMIC1 trimer species.
The Fifth Apple Domain of TgMIC4 Is a Lectin with Specificity for Galactose-terminating Oligosaccharides—To assess the carbohydrate-binding properties of TgMIC4, we carried out carbohydrate microarray analyses using the recombinant proteins TgMIC4-A5, TgMIC4-A56, and TgMIC4-A12. The microarrays encompassed a panel of 400 lipid-linked oligosaccharide probes representing diverse mammalian glycan sequences and their analogs, as well as sequences derived from fungal and bacterial polysaccharides. These are arranged based on negative charge (neutral and acidic), sialyl linkages, and backbone sequences (Fig. 6A and supplemental Table S2). TgMIC4-A12 showed no significant binding to any of the probes in the microarray (data not shown); whereas, TgMIC4-A56 and TgMIC4-A5 showed good binding to a diverse range of oligosaccharide probes terminating in β-galactose (Gal) with a similar binding profile (results for TgMIC4-A5 shown in Fig. 6 and supplemental Table S2). The probes bound include a wide range of neutral sequences and several acidic sequences that are sialylated and sulfated at inner residues. This is clearly distinct from the binding specificity of TgMIC1-MARR, which bound exclusively to sialic acid-terminating sequences (Fig. 6B and supplemental Fig. S8) (13).

Among the neutral sequences bound by TgMIC4-A5 are several short lactose and N-acetyllactosamine (LacNAc)-based probes (numbers 15–26), linear or branched mammalian milk oligosaccharide-related sequences (numbers 39–41, 43, 61–63, 66–70, 75, and 77–78), branched polyLacNAc sequences (numbers 81–84), complex-type N-glycan sequences with at least one Gal-terminating arm (numbers 125–131 and 133–136), asialo-GM1-related probes (numbers 138–139), and short Galβ1–3/6GalNAc sequences (numbers 144–146). These sequences are widely distributed in mammalian cells and tissues. There was no detectable binding to sequences terminating in α-linked Gal, e.g. probes 32–34 and blood group B/“B-like” sequences (numbers 42, 47, and 85–90). β-Gal terminating, the Lewis^a/x sequences, LNFPII (number 50), and LNFPIII (number 55), which have fucose (Fuc) on the adjoining N-acetylgalactosamine (GlcNAc) sequences, were not bound. The exceptions were Lewis^a/x trisaccharides (numbers 23–26) and B-trisaccharide (numbers 30–31), which elicited binding signals. These are likely to be a reflection of the flexible presentation of the core Gal arising from the fully or partially ring-opened state in these probes (40). Of note, the Lewis^a/x trisaccharide probe, which has fully ring-closed monosaccharide core (number 27) (40), was not recognized.

TgMIC4-A5 gave strong binding signals with several acidic oligosaccharide sequences with terminal Gal, and sialic acid or sulfate on inner residues. Among these are SM1a (number 286), which is the sulfoglycolipid analogs of GM1, several GM1-related probes with N-acetyl or N-glycolyl (NeuGc) forms of sialic acid (numbers 344–347), and GD1b (number 380) (Fig. 6B). Also bound are two sialyl tetrasaccharides having NeuAcα2–6–internally linked to penultimate GlcNAc or N-acetylgalactosamine (GlcNAc), probes 360 and 369, respectively (supplemental Table S2).

The reciprocity in the binding signals of TgMIC4-A5 and TgMIC1-MARR to N-glycans and gangliosides is clearly shown in the matrix presentation (Fig. 6B). Whereas the Gal-terminating N-glycan NA2F was bound by TgMIC4-A5, the disialylated analog A2F(2–3) was bound only by TgMIC1-MARR. Unlike TgMIC4-A5, TgMIC1-MARR did not bind to GM1-related probes, but it bound strongly to the closely related members of the ganglioside family, e.g. GM2 and GT1b, which were not recognized by TgMIC4-A5.

A closer comparison of TgMIC4-A5 binding to GM1-related sequences was performed by microarray analyses in dose-response format using asialo-GM1, SM1a, SB1a, and GM1 glycolipids (Fig. 6C). Here also, TgMIC4-A5 elicited no binding signals with SB1a, indicating the importance of unmodified terminal Gal for binding, but that a negative charge at position 3 of the internal Gal residue contributes positively the binding strength.

Atomic Resolution Insight into TgMIC4-A5 Oligosaccharide Ligand Interactions—The solution structure of the fifth apple domain of TgMIC4, comprising residues 410–491 (TgMIC4-A5), was determined using NMR spectroscopy, revealing the expected canonical apple domain fold (supplemental Fig. S9 and Table S3). The ensemble of the 10 lowest-energy structures has been deposited in the Protein Data Bank under accession number 2LL3. To investigate the binding mode of TgMIC4-A5 in more detail, NMR titration experiments were performed with Gal, lactose (Galβ1–4Glc), LacNAc (Galβ1–4GlcNAc), lacto-N-biose (Galβ1–3GlcNAc), and GM1-penta (Galβ1–3GlcNAcβ1–4(Neu5Aca2–3Galβ1–4Glc)). Each ligand induced a significant number of chemical shift perturbations in the TgMIC4-A5 H^1,3H-N-HSQC spectrum, indicative of an interaction with each of the ligands (Fig. 7, A and E). The pattern of shift perturbations was similar for each ligand, indicating a conserved binding pocket, the core of which lies at the junction between the two-stranded and four-stranded β-sheets. Where possible (i.e. fast-exchange), shift perturbations were used to estimate the dissociation constant (Kd) for the interaction (Fig. 7C). Galactose binds with a Kd of ~2.6 × 10^−4 m, whereas lactose and LacNAc each bind with Kd values of ~1.6 × 10^−4 m. The β1,3-linked analog (lacto-N-biose) binds more tightly, in the intermediate-exchange regime, with a Kd of ~1.1 × 10^−4 m determined using isothermal titration calorimetry (supplemental Fig. S10). Of all the ligands tested, the ganglioside oligosaccharide GM1-penta (lacking the ceramide tail) was found to bind most tightly, in the slow-exchange regime (i.e. Kd ~10^−5 m) (supplemental Table S4). This is in overall agreement with the results observed in the microarray analyses.

To further characterize the mechanism of galactose recognition by TgMIC4-A5, a structural model of a TgMIC4-A5 lacto-N-biose complex was calculated using HADDOCK; a computer program for data-driven molecular docking (23) (Fig. 7, structure statistics in supplemental Table S6). An ensemble of the 10 low-energy structures has been deposited in the Protein Data Bank under accession number 2LL4. Based on the NMR titration data, ambiguous interaction restraints were implemented for residues Lys-425, Asn-460, Tyr-467, Lys-469, Tyr-476, and Tyr-478. These data were complemented by the detection and assignment of seven intermolecular NOEs (nuclear Overhauser enhancements) between TgMIC4-A5 and lacto-N-biose, measured in a 13C-filtered 13C-HSQC-NOESY (nuclear Overhauser
**Galactose Recognition by T. gondii**

A

Fluorescence

Galβ-4Glc- and Galβ-3/4GlcNAc-terminated
asiagloM1-related (#138, #139)
Galβ-3/6GalNAc (#144-146)
GM1-related (#344-#347)
GD1b (#380)
LSTb (#360)
SM1a (#286)

B

| Pos. | Probe | Structure |
|------|-------|-----------|
| 39   | LNT   | Galβ-3GlcNAcβ-3Galβ-4Glc |
| 40   | LNN7  | Galβ-4GlcNAcβ-3Galβ-4Glc |
| 50   | LNFP-II | Galβ-3GlcNAcβ-3Galβ-4Glc | Fucα-4 |
| 55   | LNFP-III | Galβ-4GlcNAcβ-3Galβ-4Glc | Fucα-3 |
| 360  | LSTb  | Galβ-4GlcNAcβ-3Galβ-4Glc | NeuAcα-6 |
| 133  | NA2F  | Galβ-4GlcNAcβ-2Manβ-6 | Fucα-4 | Manβ-4GlcNAcβ-4Glcα4c |
| 338  | A2F(2-3) | NeuAcα-3Galβ-4GlcNAcβ-2Manβ-6 | Fucα-4 | Manβ-4GlcNAcβ-4Glcα4c |
| 138  | Asialo-GM1 | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer |
| 286  | SM1a  | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer |
| 287  | SB1a  | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer | Sβ-3 |
| 343  | GM2   | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer | NeuAcα-3 |
| 344  | GM1   | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer | NeuAcα-3 |
| 346  | GM1(Gc) | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer | NeuAcα-3 |
| 380  | GD1b  | Galβ-3GalNAcβ-4Galβ-4Glcβ-Cer | NeuAcα-3 |
| 391  | GT1b  | NeuAcα-3Galβ-4Galβ-4Glcβ-Cer | NeuAcα-3 |

C

Relative binding intensity

0-10%
>10-30%
>30-70%
>70-100%

Fluorescence

fmol/spot

Asialo-GM1
GM1
SM1a
SB1a
Galactose Recognition by T. gondii

In terms of TgMIC4-A5, intermolecular NOEs were restricted entirely to aromatic nuclei, of which chemical shifts in the bound state were assigned via an NMR titration using a 1H,13C-HSQC (aromatic-selective) experiment (supplemental Fig. S11A). In terms of lacto-N-biose, the presence of a 5-fold excess resulted in detection of effectively free-state chemical shifts, as verified via comparison of conventional and 13C-edited 1H-13C TOCSY spectra of, respectively, free and bound disaccharide (supplemental Fig. S11B). 1H-Chemical shift assignment of the free-state molecule was therefore carried out (supplemental Fig. S11, C and D), enabling the intermolecular NOE assignment to be completed. Assignments were implemented as distance restraints for molecular docking of TgMIC4-A5 and lacto-N-biose.

Due to the low molecular weight of the protein and intermediate-exchange regime of the interaction, it was not possible to obtain data regarding the bound-state conformation of lacto-N-biose using established transferred NOE methods. Coupled with the availability of only a small volume of unambiguous distance restraint data from intermolecular NOEs, this precluded the calculation of a full experimental structure using a program such as ARIA, and instead HADDOCK was utilized in a similar manner to previous studies (41–43).

The structural model of TgMIC4-A5/lacto-N-biose suggests that the protein adopts a similar mechanism of galactose recognition to the Galectin family proteins (44), despite TgMIC4-A5 being structurally distinct (supplemental Fig. S12A). The galactose ring stacks against the aromatic ring of Tyr-467 (in an equivalent position to the conserved tryptophan residue of Galectins) and forms hydrogen bonds with the side chains of Asn-460, Lys-469, and Tyr-476. The side chains of Lys-428 and Tyr-478 form steep walls at either end of the pocket and may provide additional ligand contacts. To confirm our structural description for galactose recognition we created binding site mutants (namely K428A, K469M, and Y478L), pocket and may provide additional ligand contacts. To confirm our structural description for galactose recognition we created binding site mutants (namely K428A, K469M, and Y478L), checked their foldedness, and reassessed their carbohydrate binding by NMR (supplemental Fig. S13). Affinity for galactose remained unaffected by the K428A mutation, which is consistent with the interaction with the backbone atoms in this region. Diminished binding of Gal was observed for TgMIC4-A5K478L (a Kd of ~1.7 mM was determined via NMR titration data) suggesting that this residue provides important contacts, consistent with the observance of intermolecular NOEs to its aromatic ring. The interaction was completely abolished in TgMIC4-A5K469M, suggesting that this residue forms a key hydrogen bond.

The only other structural insight that is available for an apple domain-carbohydrate complex is from the crystal structure of a hepatoocyte growth factor-NK1-heparin complex (45). Whereas general features of the binding site are shared, such as hydrogen bonding to lysine residues, heparin binds to a face of the apple domain that is completely different from that of TgMIC4-A5 bound by Gal-terminating sequences (supplemental Fig. S12). This study therefore identifies a new mode of oligosaccharide recognition by apple domains.

These data prompt us to revisit two other microneme proteins for which lactose binding has been suggested. The Sarcozystis mursis lectin, SML-2 (46), displays the same arrangement of Gal-binding residues and would be predicted to use the same mode of recognition as TgMIC4 (supplemental Fig. S2A). The microneme proteins EtMIC4 and EtMIC5 of Eimeria tenella form a high molecular weight complex that is also pulled down by lactose chromatography and binds host cells (47). EtMIC5 contains 11 apple domains and by comparison with TgMIC4 we deduce that the seventh apple domain may be a lectin similar to TgMIC4-A5 (supplemental Fig. S2A). Gal/GalNAc-specific lectins have also been identified in other protozoan parasites; examples include surface proteins from Cryptosporidium spp. (48), Entameba histolytica (49), and Trypanosoma cruzi (50). It is possible that the mode of galactose recognition characterized here is conserved in these more distantly related organisms.

TgMIC1 and TgMIC4 Cannot Simultaneously Bind GM1—As previously reported, TgMIC1 binds to a range of sialylated glycans with a preference for α2–3 sialic acid linkage. It has been suggested that recognition of sialylated sequences, such as those found on gangliosides, may be important for the tropism of the parasite to the brain with formation of cysts in the intermediate hosts (13). The binding studies carried out in this work have revealed that TgMIC4-A5 has galactose-binding activity and binds strongly to the oligosaccharide moiety of ganglioside GM1. GM1 possesses both terminal galactose and side chain sialic acid moieties, and is often targeted by microbial pathogens; for example, it is recognized by the cholera toxin from Vibrio cholerae (51) and the major capsid VP1 from simian virus 40 (52, 53). Given that GM1 contains α2–3-linked sialic acid on the inner Gal residue, which is bound by TgMIC1, it was reasoned that TgMIC1 should also be capable of binding to GM1-penta, although the affinity is likely to be weak as there was no significant binding to GM1 by TgMIC1-MARR in the solid phase microarray analyses (Fig. 6B). The capability of TgMIC1-MARR to interact with GM1-penta in solution was indeed demonstrated via the NMR chemical shift perturbation analysis (supplemental Fig. S14). To test the ability of TgMIC1-MARR and TgMIC4-A5 to bind to GM1 simultaneously, we performed a sequential NMR titration experiment. GM1 was first titrated into 15N-labeled TgMIC1-MARR and binding was monitored by specific peak perturbations in 1H–15N-HSQC spectra. After saturation, 13C–15N-labeled TgMIC4-A5 was then titrated to the complex and the interaction of both microneme proteins was monitored by 1H–15N-HSQC (for TgMIC1-MARR) and 1H–15N two-dimensional HNCO (for TgMIC4-A5) spectra (supplemental Fig. S14). The data show...
that TgMIC4 efficiently displaces TgMIC1 at an equimolar ratio of TgMIC4 to glycan, suggesting that the affinity of TgMIC4–A5 for GM1 is higher than TgMIC1-MARR. Furthermore, we can deduce that the interaction of TgMIC4–A5 with GM1 occludes the sialic acid branch thereby preventing binding by TgMIC1-MARR. Interestingly, microarray data reveal stronger binding of TgMIC4–A5 to GM1 and SM1a (sulfated analog of GM1) than to asialo-GM1 (Fig. 6C) suggesting a contribution from the acidic moiety. Examination of the surface electrostatics reveals several regions of significant positive charge adjacent to the galactose-binding pocket that would likely stabilize an interaction with the negative charge of sialic acid (supplemental Fig. S15).

**Glycan Recognition by TgMIC1-4-6 Complex and Its Biological Relevance**—Our observation that TgMIC4 is capable of displacing TgMIC1 from GM1 suggests that even though these
adhesins are present within the same adhesive complex they are likely to exploit different carbohydrate ligands on the host cell surface. If the distinct sialyl and galactose-binding preferences of TgMIC1 and TgMIC4 are purely adhesive then this dual recognition may allow the parasite to exploit both sialic acid-dependent and -independent invasion mechanisms. Modulation of the sialic acid dependence has been observed in Plasmodium falciparum (54) and the proteolytic trimming of TgMIC4 on the parasite surface and subsequent loss of galactose binding could provide the necessary switch (36). Alternatively, the different binding specificities of TgMIC1 and TgMIC4 may have a special implication in the preferential tissue/cell tropisms of the parasite in the brain, where different ganglioside molecules are abundant and their oligosaccharide moieties exposed on the cell surface. It is possible that on those cells that express high affinity ganglioside ligands for both TgMIC1 and TgMIC4 there would be an amplification of the binding strength.

The Gal-specific lectin activity may also fulfill a role that is independent of cell adhesion and TgMIC1. Interestingly, a novel role has been suggested for TgMIC4 or a TgMIC4-like protein in the oocyst stage (55), where it is released into the parasitophorous vacuole, the space in which the parasite replicates inside the host cell (56). As revealed by our microarray and structural studies, the carbohydrate binding profile of TgMIC4 resembles that of the galectins, a family of eukaryotic lectins with roles in regulating cell adhesion, receptor activation, intracellular signaling, apoptosis, and immune system function. Furthermore, a growing body of evidence suggests that parasites can actively usurp galectin activity to help propagate an infection as well as keep the immune system in check (57). Galectin function can also contribute to any stage of an infection by altering the magnitude and quality of the immune response. Specifically, galectin activity controls the balance between anti-apoptotic and pro-apoptotic signals, activation of immune cells, and cytokine secretion. It is conceivable that proteolytic maturation of TgMIC4 provides a mechanism to liberate a soluble galectin-like lectin, which could subsequently contribute independently to parasite dissemination or down-regulation of the host immune response. This would be reminiscent of the Gal/GalNac-binding surface protein from the intestinal parasite E. histolytica, which is essential for adhesion to target cells, cytotoxicity, and the inhibition of human complement (49). Recently, it has been shown that engagement of glycosylphosphatidylinositol-anchored proteins present on the surface of the T. gondii tachyzoite by galectins may serve to activate immunity (58). It is also worthwhile noting that mice vaccinated with the NcMIC4 antigen were more susceptible to neosporosis (59). Exhaustive analysis of apicomplexan genomes reveals several other MIC4-like proteins secreted by organelles involved in invasion and one secreted into the parasitophorous vacuole postinvasion, suggesting that galactose recognition might be a ubiquitous strategy by which the parasites control the host response (60).

Concluding Remarks—This work complements previous models of the TgMIC1-4-6 complex, providing new insight into the location of the interaction between the second apple domain of TgMIC4 and the β-finger of TgMIC1, and a possible stoichiometry of the macromolecular complex. Glycan binding has been localized to the fifth apple domain of TgMIC4, and its specificity for galactose-terminating oligosaccharides has been discovered. These findings are summarized in a schematic model of the TgMIC1-4-6 subcomplex (supplemental Fig. S16). The similarity of the carbohydrate recognition of TgMIC4-A5 lectin activity to those of galectins and the biological significance is a subject for future functional studies.

Acknowledgments—We thank Mark Stoll for design of software and collaboration in microarray data analysis. The Glycosciences Laboratory acknowledges with gratitude colleagues and collaborators over the years with whom our microarray probes were studied. Peptide fingerprinting via MALDI mass spectrometry was carried out by Dr. Jeff Keen at the Proteomics facility, University of Leeds.

REFERENCES

1. Hill, D., and Dubey, J. P. (2002) Toxoplasma gondii, transmission, diagnosis, and prevention. Clin. Microbiol. Infect. 8, 634–640
2. Rothova, A. (2003) Ocular manifestations of toxoplasmosis. Curr. Opin. Ophthalmol. 14, 384–388
3. Richards, F. O., Jr., Kovacs, J. A., and Luft, B. J. (1995) Preventing toxoplasmic encephalitis in persons infected with human immunodeficiency virus. Clin. Infect. Dis. 21, 549–556
4. Sell, M., Klingebiel, R., Di Iorio, G., and Sampalo, S. (2005) Primary cerebral toxoplasmosis. A rare case of ventriculitis and hydrocephalus in AIDS. Clin. Neuropathol. 24, 106–111
5. Carruthers, V. B. (2002) Host cell invasion by the opportunistic pathogen Toxoplasma gondii. Acta Trop. 81, 111–122
6. Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M., and Tauxe, R. V. (1999) Food-related illness and death in the United States. Emerg. Infect. Dis. 5, 607–625
7. Vaillant, V., de Valk, H., Baron, E., Ancelle, T., Colin, P., Delmas, M. C., Dufour, B., Pouillot, R., Le Strat, Y., Weinbrech, P., Jougla, E., and Desenclos, J. C. (2005) Foodborne infections in France. Foodborne Pathog. Dis. 2, 221–232
8. Carruthers, V. B., and Boothroyd, J. C. (2007) Pulling together, an integrated model of Toxoplasma cell invasion. Curr. Opin. Microbiol. 10, 82–89
9. Sibley, L. D. (2011) Invasion and intracellular survival by protozoan parasites. Immunol. Rev. 240, 72–91
10. Soldati-Favre, D. (2008) Molecular dissection of host cell invasion by the apicomplexans, the glideosome. Parasite 15, 197–205
11. Cérède, O., Dubremetz, J. F., Soête, M., Deslée, D., Vial, H., Bout, D., and Soldati-Favre, D. (2009) Atomic resolution insight into host cell recognition by Toxoplasma gondii virulence. J. Exp. Med. 201, 453–463
12. Friedrich, N., Santos, J. M., Liu, Y., Palma, A. S., Leon, E., Saouros, S., Kiso, M., Blackman, M. J., Matthews, S., Feizi, T., and Soldati-Favre, D. (2010) Members of a novel protein family containing microneme adhesive repeat domains act as sialic acid-binding lectins during host cell invasion by apicomplexan parasites. J. Biol. Chem. 285, 2064–2076
13. Blumenschein, T. M., Friedrich, N., Childs, R. A., Saouros, S., Carpenter, E. P., Campanero-Rhodes, M. A., Simpson, P., Chai, W., Kourtoukides, T., Blackman, M. J., Feizi, T., Soldati-Favre, D., and Matthews, S. (2007) Detailed insights from microarray and crystallographic studies into carbohydrate recognition by microneme protein 1 (MIC1) of Toxoplasma gondii. Protein Sci. 16, 1935–1947
14. Brecht, S., Carruthers, V. B., Ferguson, D. J., Giddings, O. K., Wang, G., Jakle, U., Harper, J. M., Sibley, L. D., and Soldati, D. (2001) The Toxoplasma micronemal protein MIC4 is an adherin composed of six conserved apple domains. J. Biol. Chem. 276, 4119–4127
15. Reiss, M., Viebig, N., Brecht, S., Fournoux, M. N., Soete, M., and Cristina,
Galactose Recognition by *T. gondii*

M. Dubremetz, J. F., and Soldati, D. (2001) Identification and characterization of an escortor for two secretory adhesins in *Toxoplasma gondii*. *J. Cell Biol.* **152**, 563–578

17. Saouros, S., Blumenschein, T. M., Sawynok, K., Marchant, J., Koutroukides, T., Liu, B., Simpson, P., Carpenter, E. P., and Matthews, S. J. (2007) High-level bacterial expression and purification of apical complexen micronen proteins for structural studies. *Protein Pept. Lett.* **14**, 411–415

18. Lourenço, E. V., Pereira, S. R., Faça, V. M., Coelho-Castelo, A. A., Mineo, J. R., Roque-Barreira, M. C., Greene, L. J., and Panunto-Castelo, A. (2001) *Toxoplasma gondii* micronemal protein MIC1 is a lactose-binding lectin. *Glycobiology* **11**, 541–547

19. Marchant, J., Sawynok, K., Saouros, S., Simpson, P., and Matthews, S. (2008) Complete resonance assignment of the first and second alpha domains of MIC4 from *Toxoplasma gondii*, using a new NMRView-based assignment aid. *Biomol. NMR Assign.* **2**, 119–121

20. Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T. E., and Nilges, M. (2007) ARIA2, automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382

21. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Refinement of protein structures in explicit solvents. *Proteins Struct. Funct. Bioinform.* **50**, 496–506

22. Shen, Y., Delaglio, F., Cornilascu, G., and Bax, A. (2009) TALOS+; a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223

23. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK, a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737

24. Hettmann, C., Herrn, A., Geiter, A., Frank, B., Schwarz, E., Soldati, T., and Soldati, D. (2000) A dibasic motif in the tail of a class XIV apicomplexan myosin is an essential determinant of plasma membrane localization. *Mol. Biol. Cell* **11**, 1385–1400

25. Plattner, F., Varovinsky, F., Romero, S., Didry, D., Carlier, M. F., Sher, A., and Soldati-Favre, D. (2008) Toxoplasma profilin is essential for host cell invasion and TLR11-dependent induction of an interleukin-12 response. *Cell Host Microbe* **3**, 77–87

26. Sawynok, K., Saouros, S., Friedrich, N., Marchant, J., Simpson, P., Blejielevs, F., Blackman, M. J., Soldati-Favre, D., and Matthews, S. (2008) Structural insights into micronemate protein assembly reveal a new mode of EGF domain recognition. *EMBO Rep.* **9**, 1149–1155

27. Feizi, T., and Chai, W. (2004) Oligosaccharide microarrays to decipher the galactose-binding protein. *Nat. Rev. Mol. Cell Biol.* **5**, 119–121

28. Palma, A. S., Liu, Y., Childs, R. A., Herbert, C., Wang, D., Chai, W., and Matthews, S. J., and Soldati-Favre, D. (2005) A novel galectin-like domain from *Toxoplasma gondii* micronemal protein 1 assists the folding, assembly, and transport of a cell adhesion complex. *J. Biol. Chem.* **280**, 38583–38591

29. Keller, N., Naguleswaran, A., Cannas, A., Vonlaufen, N., Bienz, M., Björkman, F., Bohne, W., and Hemphill, A. (2002) Identification of a *Neospora caninum* microneme protein (NeoMIC1), which interacts with sulfated host cell surface glycosaminoglycans. *Infect. Immun.* **70**, 3187–3198

30. Liu, Y., Feizi, T., Campanero-Rhodes, M. A., Childs, R. A., Zhang, Y., Mulloy, B., Evans, P. G., Osborn, H. M., Otto, D., Crocker, P. R., and Chai, W. (2007) Neoglycolipid probes prepared via oxime ligation for microarray analysis of oligosaccharide-protein interactions. *Chem. Biol.* **14**, 847–859

31. Matta-Camacho, E., Kozlow, G., Trempe, J. F., and Gehring, K. (2009) Atypical binding of the Swa2p UBA domain to ubiquitin. *J. Mol. Biol.* **386**, 569–577

32. Long, J., Garner, T. P., Pandya, M. J., Craven, C. J., Chen, P., Shaw, B., Williamson, M. P., Layfield, R., and Searle, M. S. (2010) Dimerization of the *Para Slice* complex reveals a pathway for transacylation. *J. Biol. Chem.* **285**, 36969–36976

33. Lietha, D., Chirgadze, D. Y., Mulloy, B., Blundell, T. L., and Gherrardi, E. (2001) Crystal structures of NK1-heparin complexes reveal the basis for NK1 activity and enable engineering of potent agonists of the MET receptor. *EMBO J.* **20**, 5543–5555

34. Muller, J. J., Muller, E. C., Montag, T., Zyto, N., Loschner, B., Klein, H., Heimann, U., and Otto, A. (2001) Characterization and crystallization of a novel *Sarcocystis murris* lecin. *S. Acta Crystalog* **6**, 1042–1045

35. Periz, J., Gill, A. C., Hunt, L., Brown, P., and Tomley, F. M. (2007) The micronemate proteins ETMIC4 and ETMIC5 of *Eimeria tenella* form a novel, ultra-high molecular mass protein complex that binds target host cells. *J. Biol. Chem.* **282**, 16891–16898

36. Bhat, N., Joe, A., PereiraPerrin, M., and Ward, H. D. (2007) Cryptosporidium p30, a *gallactose/N-acetylgalactosamine*-specific lectin, mediates infection in vitro. *J. Biol. Chem.* **282**, 34877–34887

37. Frederick, J. F., and Petri, W. A. (2005) Roles for the lactose- and *N-acetylgalactosamine*-binding lectin of *Entamoeba* in parasite virulence and differentiation. *Glycobiology* **15**, 53R–59R

38. Berlan, S. M., Marcipar, I. S., Roodveldt, C., Cabeza Meckert, P., Laguens, L., and Marcipar, A. J. (2002) *Trypanosoma cruzi*, identification of a lactose-binding protein that binds to cell surface of human erythrocytes and is involved in cell invasion by the parasite. *Exp. Parasitol.* **100**, 217–225
51. Holmner, A., Mackenzie, A., Okvist, M., Jansson, L., Lebens, M., Teneberg, S., and Krengel, U. (2011) Crystal structures exploring the origins of the broader specificity of Escherichia coli heat-labile enterotoxin compared to cholera toxin. J. Mol. Biol. 406, 387–402

52. Campanero-Rhodes, M. A., Smith, A., Chai, W., Sonnino, S., Mauri, L., Childs, R. A., Zhang, Y., Ewers, H., Helenius, A., Imberty, A., and Feizi, T. (2007) N-Glycolyl GM1 ganglioside as a receptor for simian virus 40. J. Virol. 81, 12846–12858

53. Neu, U., Woellner, K., Gauglitz, G., and Stehle, T. (2008) Structural basis of GM1 ganglioside recognition by simian virus 40. Proc. Natl. Acad. Sci. U.S.A. 105, 5219–5224

54. Dolan, S. A., Miller, L. H., and Wellems, T. E. (1990) Evidence for a switching mechanism in the invasion of erythrocytes by Plasmodium falciparum. J. Clin. Investig. 86, 618–624

55. Ferguson, D. J., Brecht, S., and Soldati, D. (2000) The microneme protein MIC4, or an MIC4-like protein, is expressed within the macrogamete and associated with oocyst wall formation in Toxoplasma gondii. Int. J. Parasitol. 30, 1203–1209

56. Suss-Toby, E., Zimmerberg, J., and Ward, G. E. (1996) Toxoplasma invasion, the parasitophorous vacuole is formed from host cell plasma membrane and pinches off via a fission pore. Proc. Natl. Acad. Sci. U.S.A. 93, 8413–8418

57. Young, A. R., and Meeusen, E. N. (2004) Galectins in parasite infection and allergic inflammation. Glycoconj. J. 19, 601–606

58. Debierre-Grockiego, F., Niehus, S., Coddeville, B., Elass, E., Poirier, F., Weingart, R., Schmidt, R. R., Mazurier, J., Guérardel, Y., and Schwarz, R. T. (2010) Binding of Toxoplasma gondii glycosyolphosphatidylinositols to galectin-3 is required for their recognition by macrophages. J. Biol. Chem. 285, 32744–32750

59. Srinivasan, S., Mueller, J., Suana, A., and Hemphill, A. (2007) Vaccination with microneme protein NcMIC4 increases mortality in mice inoculated with Neospora caninum. J. Parasitol. 93, 1046–1055

60. Chen, Z., Harb, O. S., and Roos, D. S. (2008) In silico identification of specialized secretory-organelle proteins in apicomplexan parasites and in vivo validation in Toxoplasma gondii. PLoS One 3, e3611

61. Lee, F. (2007) NMR methods for the determination of protein-ligand dissociation constants. Prog. Nucl. Magn. Reson. Spectrosc. 51, 219–242