Histone Acetyltransferase GCN5 Enters the Nucleus via Importin-α in Protozoan Parasite Toxoplasma gondii*‡

Received for publication, September 16, 2004, and in revised form, December 7, 2004
Published, JBC Papers in Press, December 9, 2004, DOI 10.1074/jbc.M410656200

Micah M. Bhatti and William J. Sullivan, Jr.‡
From the Department of Pharmacology & Toxicology, Indiana University School of Medicine, Indianapolis, Indiana 46202

The histone acetyltransferase GCN5 acetylates nucleosomal histones to alter gene expression. How GCN5 gains entry into the nucleus of the cell has not been determined. We have mapped a six-amino acid motif (RKRVKR) that serves as a necessary and sufficient nuclear localization signal (NLS) for GCN5 in the protozoan pathogen Toxoplasma gondii (TgGCN5). Virtually nothing is known about nucleocytoplasmic transport in these parasites (phylum Apicomplexa), and this study demonstrates the first demonstrated NLS delineated for members of the phylum. The TgGCN5 NLS has predictive value because it successfully identifies other nuclear proteins in three different apicomplexan genomic databases. Given the basic composition of the T. gondii NLS, we hypothesized that TgGCN5 physically interacts with importin-α, the main transport receptor in the importin/karyopherin nuclear import pathway. We cloned the importin-α gene from T. gondii (TgIMPα), which encodes a protein of 545 amino acids that possesses an importin-β-binding domain and armadillo/e-catenin-like repeats. In vitro co-immunoprecipitation experiments confirm that TgIMPα directly interacts with TgGCN5, but this interaction is abolished if the TgGCN5 NLS is deleted. Taken together, these data argue that TgGCN5 gains access to the parasite nucleus by interacting with TgIMPα. Bioinformatics analysis of the T. gondii genome reveals that other components of the importin pathway are present in the organism. This study demonstrates the utility of T. gondii as a model for the study of nucleocytoplasmic trafficking in early eukaryotic cells.

Eukaryotic DNA is highly compacted in chromatin, which is largely comprised of histone proteins. Once thought to play only a scaffolding role, histones are now recognized as a critical component of gene regulation and epigenetic inheritance. Many chromatin remodeling complexes contain histone-modifying enzymes that alter the nucleosomal structure and modulate the expressed genome (1, 2). GCN5, originally described as a transcriptional adapter in yeast (3), was later discovered to be a histone acetyltransferase (HAT)1 (4). HATs covalently modify chromatin by acetylating specific lysine residues within the N-terminal tails of histones, thereby attenuating the nucleosome-DNA interaction. GCN5 and the related HAT, P/CAF (p300/CBP-associating factor), are critical for growth and development in mice (5, 6) and plants (7). Although much study has been invested in the gene regulatory functions of GCN5s, little attention has been directed at resolving how the HAT gets into the cell nucleus. One study suggests that P/CAF requires importin-α for nuclear localization, but the precise mechanism of the interaction was not determined (8).

Importin-α proteins serve as adaptors that bind both the nuclear destined cargo protein and importin-β. Importin-α recognizes a nuclear localization signal (NLS) on the cargo, which is typically a mono- or bi-partite cluster of basic amino acids (9). Importin-β is capable of binding nucleoporins, thus escorting the importin-α-importin-β-cargo complex to the nuclear pore. RanGTP and nuclear transport factor 2 are involved in translocation of the ternary complex through the nuclear pore, whereupon RCC1 exchanges GDP with GTP to dissociate the cargo (10, 11). RanGTP recycles importin-β, and importin-α is exported back to the cytoplasm by a distinct importin-β homologue, cellular apoptosis susceptibility protein (12). Exceptions exist to this generalized pathway, such as the ability of importin-β to translocate cargo without importin-α (13) and the presence of transportin, a distant relative of importin-β that recognizes nonclassical (M9) NLSs (14).

The GCN5 HAT identified in the protozoan parasite Toxoplasma gondii (phylum Apicomplexa) revealed unexpected features, most notably an N-terminal "extension" that is unusually long (820 amino acids) for an early eukaryote (15). Early eukaryotes including yeast and the fellow alveolate protist Tetrahymena possess a short N-terminal domain of <100 residues (4). The N-terminal extension in plant GCN5s range from 150 to 250 amino acids and have little homology to each other or other HATs (16). Metazoan GCN5 family members contain N-terminal extensions of ~500 residues, and they are fairly similar in composition (17, 18). TgGCN5 contains the second longest N-terminal extension of all GCN5 HATs described to date, the longest (1126 amino acids) having recently been noted in another apicomplexan, Plasmodium falciparum (19), the causative agent of severe malaria. Interestingly, the lengthy N-terminal extensions present in these apicomplexan GCN5s bear no similarity to each other or any known protein and are devoid of known protein motifs. We hypothesized that at least one role for the N-terminal extension may be nuclear localization. Consistent with this hypothesis is the recent observation that maize GCN5 needs its N-terminal extension to access the nucleus; however, the precise NLS was not mapped (16).

1 The abbreviations used are: HAT, histone acetyltransferase; NLS, nuclear localization signal; βgal, β-galactosidase; contig, group of overlapping clones.

* This work was supported by National Institutes of Health Grant GM065051 and a Biomedical Research grant from Indiana University School of Medicine (to W. J. 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.
† To whom correspondence should be addressed: Dept. of Pharmacology & Toxicology, Indiana University School of Medicine, 635 Barnhill Dr., MS A-525, Indianapolis, IN 46202. Tel.: 317-274-1573; Fax: 317-274-7714; E-mail: wjsullivan@iupui.edu.

‡ The on-line version of this article (available at http://www.jbc.org) contains supplemental material.

© 2005 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.
Understanding GCN5 and how it enters the parasite nucleus is significant for a number of reasons. *T. gondii* can cause congenital birth defects and is an important opportunistic pathogen in AIDS and immunosuppressed patients (20–22). Blocking transcription factors from entering the parasite nucleus will subvert parasite differentiation and other processes essential for its survival. Novel elements found in parasite transcription factors and/or the nuclear import pathway may be exploited in the design of more selective therapeutic agents. Additionally, the study of these pathways in protozoa provides a unique perspective on how these systems evolved in early eukaryotic cells.

Here we report the use of *T. gondii* GCN5 (TgGCN5) as a tool to investigate the understudied area of nuclear trafficking in protozoan parasites. We have mapped the first NLS to be identified and validated for any apicomplexan protein and in the process defined a function of the long N-terminal extension of TgGCN5. We demonstrate the predictive value of this NLS by identifying other nuclear proteins from the *T. gondii* genome as well as two additional apicomplexan databases. We have also characterized a *T. gondii* importin-α orthologue (TgIMp0) and demonstrated that this nuclear receptor interacts with TgGCN5 via the NLS that we have elucidated. Knowledge that GCN5 family HATs are translocated to the nucleus by the importin pathway may be useful in developing therapies that prevent this factor from modulating transcription in eukaryotic cells.

**EXPERIMENTAL PROCEDURES**

**Parasite Culture and Methods—** *T. gondii* tachyzoites were cultivated in human foreskin fibroblast cells and transfected by electroporation as previously described (23). The RH1HXGPRT clone and pminiHXGPRT vector were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, National Institutes of Health, from Dr. David Ross (24–26). RH1HXGPRT were grown in culture medium supplemented with 320 μg/ml 6-thioxanthine (Sigma). Mycophenolic acid at 25 μg/ml plus xanthine at 50 μg/ml (Sigma) was used to select for parasites transfected with pminiHXGPRT-derived vectors.

**Plasmids—** All of the constructs for immunolocalization were based on a *T. gondii* expression vector that was built into pminiHXGPRT. The *T. gondii* TUB1 promoter was amplified from putbP30-GFP/sagCAT/17 and used to generate the BamHI and HindIII site. The *T. gondii* dihydrofolate reductase was amplified using these primers 3 and 4. The resulting PCR product was digested with AvrII and NotI sites that were designed for a three-piece ligation with the TUB1 ampiclon above (cut with BamHI and AvrII) and the pminiHXGPRT vector (cut with BamHI and NotI). We have designated this expression construct as putbXFLAG::HX, the X representing where the gene of interest is to be inserted (Fig. 1). It is important to note that although a C-terminal GFP fusion is available in this vector, it was used only for the βgal construct. An N-terminal FLAG tag was fused to TgGCN5 sequences to be tested for localization as outlined below. The primers 5–10 are the sense primers used in conjunction with the antisense primer 11 to create the various FLAG-tagged forms of TgGCN5 or βgal. All of the inserts were ligated into putbXFLAG::HX at the NdeI and AvrII sites except the βgal constructs, which used BglI and AvrII. Prior to electroporation, ~50 μg of plasmid DNA was linearized using NotI.

For the production of *in vitro* translated protein, pGAD77 and pGBK7 vectors were used (Clontech). Full-length TgGCN5 and TgGCN5 lacking the first 99 amino acids (Δ99TgGCN5) were ligated into the pGBK7 vector using the NdeI and XmaI sites. Sense primers 12 and 13 were used with antisense primer 14 to amplify inserts encoding full-length TgGCN5 and Δ99TgGCN5, respectively. ΔNLS-TgGCN5 was generated by a ligation of three DNA fragments. (i) The first piece encoded TgGCN5 amino acids 1–93 and was amplified using the primer 12 and 15, the latter of which contains a NheI site. (ii) The second fragment encoded amino acids 100–1169 and was generated with primer 12 (contains a NheI site) and primer 14 (contains a XmaI site). (iii) The third piece was pGBK7 digested with NdeI and XmaI. The resulting vector encodes a fusion of TgGCN5 in which KRRKRRKKR was replaced by the AS (encoded by the NheI site). TgIMp0 was ligated into pGAD77 in frame with the N-terminal hemagglutinin tag using NdeI and XmaI sites, and the insert was generated by PCR using the primers 17 and 18.

All of the PCRs were conducted with the proofreading DNA polymerase *Pfu* Ultra™ (Stratagene). DNA sequencing was performed at the Indiana University Biochemistry Biotechnology Facility (Indiana University School of Medicine, Indianapolis, IN).

**Molecular Methods—** *T. gondii* mRNA was isolated from ~2 × 10⁸ tachyzoites using Poly(A)Pure™ (Ambion). Tachyzoites were allowed to completely lyse the host cell monolayer and were passed through a 3.0-μm filter to exclude residual host cell debris. mRNA was treated with DNase prior to Northern blotting, which was performed using standard methods (28). A cDNA-derived probe representing the final 2.3 kb of TgGCN5 was labeled with [32P]dATP (PerkinElmer Life Sciences) using random primers and a Prime-A-Genêt® kit (Promega). Reverse transcription-PCRs were carried out using the SuperScript II™ One-Step reverse transcription-PCR system (Invitrogen). The rapid amplification of cDNA ends was performed using an Invitrogen GeneRacer™ kit. Sequencing alignments and phylogeny trees (neighbor joining method) were generated using AlignX, a component of Vector NTI Advance 9.0 (Informax). Protein motif searches were performed using the Pfam data base (version 14.0; pfam.wustl.edu/) and/or PROSITE data base (version 18.35; us.expasy.org/prosite/).

**Immunofluorescence Assays—** Well plates containing confluent human foreskin fibroblast cells grown on glass coverslips were infected with parasites. The cultures were fixed 18–24 h post-infection with ice-cold methanol at −20 °C for 10 min and blocked with phosphate-buffered saline containing 3% Fraction-V bovine serum albumin.

### Table 1

| Number | Primer sequence (5’–GACTACAAGGACGACGACGACAAG) |
|--------|-------------------------------------------------|
| 1      | GGATCCACACACTTGGTGCGAGTGGCC                  |
| 2      | CTAATGATGCAGATATGCTTTAAGAAAGGATTCAGAAAAATGC   |
| 3      | CCATATGAAATGCG/ACTGGGACGACGACGACAAG          |
| 4      | TGGCCGCGCTAGTACAAGTCGGATCG                   |
| 5      | CATATGAAATGCG/ACTGGGACGACGACGACAAG          |
| 6      | CATATGAAATGCG/ACTGGGACGACGACGACAAG          |
| 7      | CATATGAAATGCG/ACTGGGACGACGACGACAAG          |
| 8      | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 9      | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 10     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 11     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 12     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 13     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 14     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 15     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 16     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 17     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |
| 18     | AGATCTAAATGGTGCTTTTTCACAACGTGCTAGTG         |

*The primers in Table 1 were used to amplify the dihydrofolate reductase gene from T. gondii. The primers incorporate a BamHI site on the 5′-end of the primer sequence.*
RESULTS

Full-length TgGCN5 Localizes to Parasite Nucleus—TgGCN5 was initially described as a 474-amino acid protein, as deduced from a 2.3-kb cDNA clone isolated from a library screen (29). Our parallel cloning of TgGCN5 suggested that the open reading frame is much larger, encoding a protein of 1169 amino acids (15). To verify the message length and determine whether alternatively spliced variants of TgGCN5 are possible for nuclear localization was narrowed to be between 94–99 TgGCN5, the resulting protein is not sufficient NLS, RKRVRK was fused to the N terminus of the parasite nucleus, except when fused to the TgGCN5 NLS (Fig. 5). The 5′-untranslated region of 1.3 kb. To resolve the 5′ end of the TgGCN5 message, 5′ rapid amplification of cDNA ends was performed. Two amplicons were produced, indicating transcriptional start sites at −1101 and −1316 nucleotides from our proposed translational start. The 215-nucleotide difference between the two putative transcriptional starts is not resolvable on the Northern.

A recombinant version of the previously reported 474-amino acid version of TgGCN5 (corresponding to amino acids 697–1169 of full-length TgGCN5) fused to a c-Myc epitope accumulated in the parasite cytoplasm, an unexpected localization for a GCN5 family member HAT (29). To test whether the additional N-terminal sequence alters localization to the parasite nucleus, full-length TgGCN5 and a truncated version lacking the N-terminal extension (amino acids 1–697) tagged with a FLAG epitope were expressed in the parasites (FLAG-TgGCN5 and FLAGΔN5904-TgGCN5, respectively). Immunofluorescence assays using anti-FLAG demonstrate that the N-terminal extension is indeed required for nuclear localization of TgGCN5 (Fig. 3). In addition to nuclear DNA, Diamino-2-phenylindole also stains the 35-kb extrachromosomal element housed in the apicoplast. We did not observe any form of TgGCN5 trafficking to the apicoplast organelle.

Mapping the Nuclear Localization Signal (NLS) of TgGCN5—Protein motif searches of the TgGCN5 sequence revealed a putative bipartite NLS at amino acid positions 488–505 (15). However, removal of this 18-amino acid motif from full-length TgGCN5 fused to a C-terminal green fluorescent protein tag did not subvert nuclear localization. Therefore, a series of mapping constructs were designed to identify the region responsible for targeting TgGCN5 to the parasite nucleus. After several truncation studies (data not shown), the region involved in nuclear localization was narrowed to be between amino acids 58 and 120. Visual inspection of this region revealed a hexapeptide enriched with basic residues beginning at residue 94. NLSs often contain basic amino acids and average 6–8 residues in length (30). As shown in Fig. 4, FLAG-tagged TgGCN5 lacking the first 99 residues (FLAG59-TgGCN5) does not enter the parasite nucleus. However, when only the first 93 residues are removed (FLAGΔ93-TgGCN5), the resulting protein is no longer restricted from the parasite nucleus.

This demonstrates that the hexapeptide RKVRKR (amino acids 94–99) is required for the nuclear localization of TgGCN5. To verify whether the hexapeptide is a complete and sufficient NLS, RKRVRK was fused to the N terminus of E. coli βgal for expression in T. gondii. The FLAG epitope tag was moved to the C-terminal end of the βgal fusion proteins to ensure that the FLAG-NLS chimeric sequence was not responsible for nuclear localization. βgalΔFLAG is excluded from the parasite nucleus unless fused to the TgGCN5 NLS (Fig. 5). The

2 W. J. Sullivan, Jr. and B. Striepen, unpublished observations.
The N-terminal extension of TgGCN5 is required for nuclear localization. Recombinant protein in transgenic parasites expressing either full-length \( \text{FLAG}_{\text{TgGCN5}} \) or TgGCN5 lacking the N-terminal extension \( \text{FLAG}_{\text{TgGCN5}} \) or TgGCN5 was detected by immunofluorescence assay using anti-FLAG antibodies followed by staining with anti-rabbit Alexa 488 (green). 4’,6’-Diamino-2-phenylindole (DAPI) was used as a nuclear stain (red). hN, host cell nucleus; TgN, parasite nucleus.

The hexapeptide RKRVKR (amino acids 94–99) is necessary for nuclear localization of TgGCN5. Parasites were transfected with FLAG-tagged forms of TgGCN5 lacking the first 93 \( \text{FLAG}_{\text{Δ99TgGCN5}} \) or 99 \( \text{FLAG}_{\text{Δ99TgGCN5}} \) amino acid residues. Immunofluorescence assay was carried out as described in the legend to Fig. 3. DAPI, 4’,6’-diamino-2-phenylindole.

**Importin-α-mediated localized Localization of TgGCN5**

TgGCN5 NLS (e.g., (R/K)(R/K)(R/K)NLS-βgal) is likely to be found in other nuclear proteins in apicomplexa.

*T. gondii* Possesses an Importin-α Homologue—Nuclear proteins harboring a NLS rich in basic residues typically interact with members of the importin (also known as karyopherin) family of proteins, particularly importin-α (30, 32). The *T. gondii* database (Release 2.1) was screened for sequences encoding possible importin-α homologues using the BLAST algorithm and text queries. The contig Tgg_7395 (corresponding to Tgg_994532 in Release 3.0) exhibited a high degree of similarity to importin-α proteins from other species. Primers designed to the genomic sequence data were employed to obtain the full-length cDNA for T. *gondii* importin-α (TgIMPα, AY267654). Based on the results of the cDNA sequence, we have deduced that the −9.5-kb TgIMPα genomic locus is comprised of 11 exons and 10 introns.

TgIMPα possesses an open reading frame of 1638 nucleotides with a consensus start ATG (33) that is preceded by an in-frame stop codon −156 nucleotides upstream. The deduced 545-amino acid sequence has a calculated molecular mass of ~60 kDa and shows unequivocal similarity to orthologues present in other eukaryotic organisms, especially the fellow apicomplexan parasite *P. falciparum* (34). Apicomplexan IMPα proteins harbor the well conserved importin-α binding domain and armadillo/b-catenin-like repeats (35) (Fig. 6A). Additionally, they contain the leucine-rich nuclear export signal within the final armadillo/b-catenin-like repeat that is required to recycle importin-α back to the cytoplasm (36). An autoinhibitory motif (KRR) is present in the importin-β-binding domain of metazoan importin-α proteins, but the corresponding sequence in TgIMPα is more akin to the one in plants (KKK). Phylogenetic analysis of apicomplexan importin-α proteins (Fig. 6B) also illustrates that they are more similar to those found in plants, a characteristic that has been noted before for other apicomplexan proteins (37).

TgGCN5 Interacts with TgIMPα via the RKRVKR NLS—Co-immunoprecipitation studies were performed to assess whether TgGCN5 physically interacts with TgIMPα and, furthermore, whether such an interaction occurred by virtue of the NLS elucidated by the immunolocalization studies. TgIMPα and various forms of TgGCN5 were transcribed and translated in *vitro* in the presence of radiolabeled methionine. After mixing them together, the proteins were co-immunoprecipitated using antisera generated against the C terminus of TgGCN5 and resolved on SDS-PAGE. As shown in Fig. 7, full-length TgGCN5 interacts with TgIMPα. However, a truncated version of TgGCN5 starting just downstream of the RKRVKR NLS (Δ99TgGCN5) is no longer capable of binding TgIMPα, consist-
ent with the *in vivo* localization data. To ensure that only the NLS hexapeptide is required to mediate the interaction with TgIMPα, a version of full-length TgGCN5 was generated in which RKRVKR was replaced by an alanine and serine (ΔNLS-TgGCN5). ΔNLS-TgGCN5 fails to interact with TgIMPα (Fig. 7), confirming that TgIMPα cannot bind any region of TgGCN5 outside of the NLS RKRVKR.

**DISCUSSION**

Two understudied research topics in the clinically and economically relevant group of pathogens in phylum Apicomplexa include nucleocytoplasmic trafficking and the regulation of gene expression. We have attempted to expand our knowledge base in both of these related areas by studying how the important HAT, TgGCN5, is directed to the parasite nucleus to carry out its gene regulatory functions. **Functions of the Unique N-terminal Extension of TgGCN5**—The strikingly divergent N-terminal extensions present in both TgGCN5 and PGCN5 may be indicative of parasite-specific functions. This study has discovered that at least one function of the N-terminal extension is to localize the HAT to the nucleus via the importin pathway. Additional possible functions of the TgGCN5 N-terminal extension may include the participation in protein-protein interactions or the recognition of substrate. *In vivo*, GCN5 operates as the enzymatic component of large multiprotein complexes such as SAGA and ADA (1). One or more proteins may bind to the N terminus of TgGCN5 to nucleate an analogous histone acetylase complex in the parasites. The divergent nature of the apicomplexan N-terminal sequences implies that they interact with novel, parasite-specific proteins, perhaps different sets of proteins under varying conditions or life cycle stages.

The N-terminal extension of TgGCN5 may also influence enzymatic activity and substrate recognition. Recombinant mammalian GCN5 devoid of the N terminus is capable of acetylating free histones *in vitro* but requires the N-terminal sequence to be able to recognize and acetylate nucleosomes (18). The N-terminal extension is dispensable for GCN5-mediated acetylation of free histone H3 in both *P. falciparum* (19) and *T. gondii*, but further study is required to assess the role of this domain, if any, in recognizing nucleosomal substrates.

The *Predictive Value of the TgGCN5 NLS*—The completion of three apicomplexan genomes in recent years has revealed many predicted open reading frames of no known function. Defining motifs such as an NLS will greatly assist in annotation efforts and functional genomics. The TgGCN5 NLS sequence exhibited no exact matches to predicted genes, but minor variations produced encouraging results when screened against the database. Substituting the only nonbasic residue (valine) with any amino acid identified a DNA polymerase, a probable ribonucleoprotein, and numerous parasite-specific proteins containing zinc fingers, which are likely transcription factors. Table II also displays results from similar searches of the *Plasmodium* database with permutations of the TgGCN5 NLS. For example, searching the *P. falciparum* annotated protein data set with RKRVKR revealed 14 potential hits, including a predicted nuclear protein containing three PHD fingers (PFC0425w; RKKRKK). Searches with the less stringent criteria (RK)(RK)(RK)(RK) revealed 1500 and 1180 potential hits in the annotated protein data sets from *P. falciparum* and *Plasmodium yoelii*, respectively. Cursory analysis of some of these candidate nuclear proteins revealed a putative histone deacetylase (PY07179; KRKNKK), nuclear protein tRNA pseudouridylate synthase (PF10_0175; KKKMKK), and the putative DNA polymerase ζ subunit (PF10_0362; KKKMKK). Searches of the *Cryptosporidium* expressed sequence tag database for instances of (RK)(RK)(RK)(RK) were also successful in identifying nuclear proteins; among the nine hits were histone H2B (CpEST_AA224688; RKKRKK) and a putative nucleolar protein (CpEST_AA253596; KKKRRK). It should be cautioned, however, that some hits in each database contained the NLS but are not likely to be nuclear based on their homology assignment.

Searches for the TgGCN5 NLS on other GCN5 proteins have been less successful. Stringent matches are not evident on higher eukaryotic GCN5s or the highly related PCAF HATs, but PCAF is known to interact with importin-α (8). Similarly, strict matches to the TgGCN5 NLS are also not apparent on fellow apicomplexan GCN5s from *P. falciparum* (19) or *Cryptosporidium parvum* (EAK89017). However, runs of 4–5 basic residues exist in the N termini of these GCN5 HATs that are plausible candidates for importin-α binding. Further experimental evidence must be obtained to conclude whether other GCN5s are translocated to the nucleus via the importin pathway.

**Nucleocytoplasmic Trafficking in Apicomplexan Parasites**—Like other single-celled eukaryotes, it appears that both *T. gondii* and *P. falciparum* possess only one importin-α, in contrast to multicellular organisms, which have expanded their repertoire of importin-α proteins considerably. Bioinformatics searches using keywords or sequence data from other species reveals that *T. gondii* possesses homologues of every importin pathway component examined, including importin-β, Ran, nuclear transport factor 2, RCC1, CAS, RanGAP1, RanBP2, and Crm1/exportin. Several components of the importin pathway have been cloned previously in *P. falciparum*, namely importin-α, importin-β (34), Ran (38), RCC1 (39), and RanBP1 (40).

| DB identifier | Search result | Predicted homology/motifs |
|---------------|---------------|---------------------------|
| Toxoplasma    |               |                           |
| TgTigrScan_3884 | RKRTKR        | DNA polymerase X family   |
| TgTigrScan_2877 | RKRDKR        | Zinc finger               |
| TgTigrScan_5941 | RKRGRK        | Zinc finger               |
| TgTigrScan_6704 | RKRRKR        | Zinc finger               |
| TgTigrScan_7743 | RRKRRK        | RNA recognition motif     |
| TgTigrScan_4179 | RKRRKR        |                           |
| Plasmodium    |               |                           |
| PY07179       | RKKNKK        | Histone deacetylase domain|
| PF10_0175     | RKKRKK        | tRNA pseudouridylate synthase |
| PF10_0362     | KKKMKK        | DNA polymerase B family   |
| PFC0425w      | RKKRKK        | PHD fingers               |
| Cryptosporidium|               |                           |
| CpEST_AA224688| RKRKNK        | Histone 2B                |
| CpEST_AA253596| KKKRRK        | Putative nucleolar protein |
FIG. 6. T. gondii importin-α (TgIMPα). A. Protein sequences of importin-α proteins from various species were aligned using the AlignX module in Vector NTI Advance 9.0. Tg, T. gondii (AY267540); Pf, P. falciparum (AAO85774); At, Arabidopsis thaliana (Q96321); Sc, Saccharomyces cerevisiae (Q02821); Hs, Homo sapiens (P52294). Black inversed letters indicate identical residues in all species, gray inversed letters are functionally conserved residues, and gray highlighted letters are identical or functionally conserved residues in more than 50% of the species examined. \( \downarrow \) denotes residues not conserved in T. gondii; * denotes conservation in apicomplexans only, and ¥ denotes elements conserved between apicomplexans and A. thaliana. IBB, importin-β-binding domain; ARM, armadillo repeats 1–8. The boxed region within the importin-β-binding domain is the autoinhibitory sequence, and the boxed region within armadillo/β-catenin-like repeat 8 is the putative nuclear export signal. B, phylogenetic analysis of importin-α orthologues from representative species with the calculated distance values in parentheses.
These observations are consistent with the idea that the importin-mediated nuclear transport pathway is ancient in origin and well conserved among eukaryotes (32). However, it is notable that the apicomplexan versions are typically more plant-like, which may have relevant functional implications. For example, plant importin-α proteins are able to translocate their cargo into the nucleus without being bound by importin-β (41). Further study is required to test whether this is true in apicomplexa nuclear transport.

Recent studies have also implicated that components of the importin pathway have roles beyond nuclear trafficking. Importins and the Ran GTPase cycle are required for spindle formation and embryonic mitosis in Caenorhabditis elegans (42, 43). Importin-α can associate with Xenopus laevis egg membranes and is involved in the formation of the nuclear envelope (44). There is also evidence that importins function to prevent the aggregation of cytosolic proteins harboring basic rich domains by binding to, and hence shielding, that exposed region (45). It would be of interest to determine whether similar functions exist for the apicomplexan importin pathway members.

Apicomplexans can be considered “minimal” eukaryotes, because their evolutionary trajectory branches before the divergence of fungi, plants, and animals (46). The possession of a streamlined nucleocytoplasmic transport system with only one importin-α makes this early eukaryote attractive to study for dissecting the fundamentals of this pathway. The continued study of both parasite nuclear trafficking and chromatin remodeling in T. gondii promises to reveal insight into how these important biological processes evolved and may reveal differences that can be exploited therapeutically.

Acknowledgments—Preliminary genomic data was accessed via TogoDB.org. The genomic data were provided by the Institute for Genomic Research (supported by National Institutes of Health Grant AI05093) and by the Sanger Center (Wellcome Trust). The plasmid ptpubP30-GFPlacZCAT was a gift from Dr. Boris Striepen (University of Georgia). The authors thank Drs. Tony Sinai and Sherry Queener, as well as members of the Sullivan lab for helpful discussions.

FIG. 7. TgGCN5 interacts with TgIMPα via the NLS RRKVKR. Autoradiogram of immunoprecipitation reactions. Lane 1, full-length TgGCN5 + TgIMPα; lane 2, Δ99TgGCN5 + TgIMPα; lane 3, ΔNLS-TgGCN5 + TgIMPα; lane 4, full-length TgGCN5; lane 5, Δ99TgGCN5; lane 6, ΔNLS-TgGCN5; lane 7, TgIMPα. TgGCN5 proteins were immunoprecipitated with TgGCN5 antisera. TgIMPα was precipitated with anti-hemagglutinin. The large arrow points to the form of TgGCN5 used (~120–130 kDa), and the small arrow points to TgIMPα (~60 kDa). Molecular masses are displayed in kilodaltons.

REFERENCES
1. Stermer, D. E., and Berger, S. L. (2000) Microbiol. Mol. Biol. Rev. 64, 435–459
2. Roos, T. Y., Denz, J. M., and Allis, C. D. (2001) Annu. Rev. Biochem. 70, 81–129
3. Georgakopoulos, T., and Theil, C. G. (1999) EMBO J. 18, 4145–4152
4. Brownell, J. E., Zhou, J., Ranalli, T., Kubayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996) Cell 84, 843–851
5. Yamauchi, T., Yamauchi, J., Kawata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K., and Nakatani, Y. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 11303–11306
6. Xu, W., Edmondson, D. G., Evvard, Y. A., Wakamiya, M., Behringer, R. R., and Roth, S. Y. (2000) Nat. Genet. 22, 229–232
7. Vlachonasios, K. E., Thomasaw, M. F., and Tienzenberg, S. J. (2003) Plant Cell 15, 626–638
8. Kohler, M., Speck, C., Christiansen, M., Bischof, P. R., Prehn, S., Haller, H., Gorlich, D., and Hartmann, E. (1999) Mol. Cell. Biol. 19, 7782–7791
9. Gorlich, D., and Kutay, U. (1999) Annu. Rev. Cell Dev. Biol. 15, 607–660
10. Bischof, P. R., and Ponting, H. (1991) Nature 354, 80–92
11. Gorlich, D., Pante, N., Kutay, U., Aebi, U., and Bischof, P. R. (1996) EMBO J. 15, 5584–5594
12. Kutay, U., Bischof, F. R., Kostka, S., Kraft, R., and Gorlich, D. (1997) Cell 90, 1061–1071
13. Palmeri, D., and Malim, M. H. (1999) Mol. Cell. Biol. 19, 1218–1225
14. Pollard, W. V., Michael, W. M., Nakielny, S., Siomi, M. C., Wang, F., and Dreyfuss, G. (1996) Cell 86, 985–994
15. Sullivan, W. J., Jr., and Smith, C. K., II (2000) Curr. Drug Targets Immune Endocr. Metabol. Disord. 45, 455–469
16. Smith, R. E., Belote, J. M., Stilitz, R. L., Yang, X. J., Moore, P. A., Berger, S. L., Nakatani, Y., and Allis, C. D. (1998) Nucleic Acids Res. 26, 2948–2954
17. Cohen, W., Edmondson, D. G., and Roth, S. Y. (1996) Mol. Cell. Biol. 16, 5659–5669
18. Fan, Q., An, L., and Cui, L. (2004) Curr. Genet. 45, 211–215
19. Seeber, F. (2003) Curr. Drug Targets Immune Endocr. Metabol. Disord. 3, 99–109
20. Wong, S. Y., and Remington, J. S. (1993) Methods Cell Biol. 299–316
21. Seeber, F. (2003) Curr. Genet. 45, 193–200
22. Wong, S. Y., and Remington, J. S. (1993) AIDS 7, 299–316
23. Ross, D. S., Sullivan, W. J., Striepen, B., Bohne, W., and Donald, R. G. (1997) Methods 13, 112–122
24. Pfefferkorn, E. R., and Borotz, S. E. (1994) Exp. Parasitol. 79, 374–382
25. Ross, D. S., Donald, R. G., Morriette, N. S., and Moulton, A. L. (1994) Methods Cell Biol. 45, 27–63
26. Donald, R. G., Carter, D., Ullman, B., and Donald, R. S. (1996) J. Biol. Chem. 271, 14010–14019
27. Striepen, B., He, C. Y., Matrajti, M., Soldati, D., and Ross, D. S. (1998) Mol. Biochem. Parasitol. 92, 325–338
28. Sambrook, J., Russell, D., Irwin, N., and Hanseen, K. (2001) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
29. Hettmann, C., and Soldati, D. (1999) Nucleic Acids Res. 27, 4344–4352
30. Boulikas, A. (1999) J. Cell. Biochem. 76, 32–58
31. Dungwall, C., and Laskey, R. A. (1991) Trends Biochem. Sci 16, 478–481
32. Chook, Y. M., and Blobel, G. (2001) Curr. Opin. Struct. Biol. 11, 703–715
33. Kozak, M. (1991) J. Cell Biol. 115, 887–903
34. Mohamed, A., Kissore, S., Dasaradhi, P. V., Patra, K., Malhotra, P., and Chauhan, V. S. (2003) Mol. Biochem. Parasitol. 127, 199–203
35. Herold, A., Tuurant, R., Wiegand, H., and Cullen, B. R. (1998) J. Cell Biol. 143, 309–318
36. Koe, B. (1999) Nat. Struct. Biol. 6, 388–397
37. Huang, J., Mullapudi, N., Sicheritz-Ponten, T., and Kissinger, J. C. (2004) Int. J. Parasitol. 34, 266–274
38. Doi, T., Okamura, A., Ando, H., Rikihisa, Y., and Kato, H. (1995) J. Bacteriol. 177, 2996–3001
39. Ji, D. D., Sultan, A. A., Chakrabarti, D., Horrocks, P., Doerig, C. (2002) Mol. Biochem. Parasitol. 123, 67–71
40. Bubner, S., Smith, H. M., Hara, K., Chakrabarti, D., Banerjee, S., Gales, M. V., and Kato, H. (1999) J. Bacteriol. 177, 2996–3001
41. Askjaer, P., Galy, V., Hanak, E., and Mattaj, I. W. (1992) Mol. Biol. Cell 3, 839–848
42. Chook, Y. M., and Blobel, G. (2001) Curr. Opin. Cell Biol. 11, 229–232
43. Sogin, M. L. (1991) Curr. Opin. Genet. Dev. 1, 457–463
