Structural Insights into Parasite eIF4E Binding Specificity for m\(^7\)G and m\(^{2,2,7}\)G mRNA Caps

Received for publication, July 29, 2009, and in revised form, August 19, 2009. Published, JBC Papers in Press, August 26, 2009, DOI 10.1074/jbc.M109.049858

Weizhi Liu,† Rui Zhao*, Craig McFarland‡, Jeffrey Kieft‡, Anna Niedzwiecka§, Marzena Jankowska-Anyszka‖, Janusz Stepinski‡, Edward Darzynkiewicz§, David N. M. Jones**, and Richard E. Davis††

From the Departments of †Biochemistry and Molecular Genetics and **Pharmacology, University of Colorado School of Medicine, Aurora, Colorado 80045, ‡Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, 02-089 Warsaw, Poland, §Faculty of Chemistry, University of Warsaw, 02-093 Warsaw, Poland, and ‖Biological Physics Group, Institute of Physics, Polish Academy of Sciences, 32/46 Lotnikow Avenue, 02-668 Warsaw, Poland.

The eukaryotic translation initiation factor eIF4E recognizes the mRNA cap, a key step in translation initiation. Here we have characterized eIF4E from the human parasite Schistosoma mansoni. Schistosome mRNAs have either the typical monomethylguanosine (m\(^7\)G) or a trimethylguanosine (m\(^{2,2,7}\)G) cap derived from spliced leader trans-splicing. Quantitative fluorescence titration analyses demonstrated that schistosome eIF4E has similar binding specificity for both caps. We present the first crystal structure of an eIF4E with similar binding specificity for m\(^7\)G and m\(^{2,2,7}\)G caps. The eIF4E:m\(^7\)GpppG structure demonstrates that the schistosome protein binds monomethyl cap in a manner similar to that of single specificity eIF4E proteins and exhibits a structure similar to other known eIF4Es. The structure suggests an alternate orientation of a conserved, key Glu-90 in the cap-binding pocket that may contribute to dual binding specificity and a position for mRNA bound to eIF4E consistent with biochemical data. Comparison of NMR chemical shift perturbations in schistosome eIF4E on binding m\(^7\)GpppG and m\(^{2,2,7}\)GpppG identified key differences between the two complexes. Isothermal titration calorimetry demonstrated significant thermodynamics differences for the binding process with the two caps (m\(^7\)G versus m\(^{2,2,7}\)G). Overall the NMR and isothermal titration calorimetry data suggest the importance of intrinsic conformational flexibility in the schistosome eIF4E that enables binding to m\(^{2,2,7}\)G cap.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1 and 2.

The atomic coordinates and structure factors (codes 3HXI and 3HXG) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).  

1 To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, 12801 East 17th Ave., Aurora, CO 80045. Fax: 303-724-3215; E-mail: richard.davis@ucdenver.edu.

2 The abbreviations used are: eIF4E, eukaryotic initiation factor 4E; m\(^7\)G, 7-methylguanosine; m\(^{2,2,7}\)G, N\(^2\),N\(^3\),N\(^7\)-methylguanosine; 4E-BP, 4E-binding protein; ITC, isothermal titration calorimetry; HSQC, heteronuclear single quantum correlation; MOPS, 4-morpholino propane sulfonic acid; NOESY, nuclear Overhauser effect spectroscopy; CSP, chemical shift perturbation.
eIF4E. The eIF4E core resembles a “cupped hand” within which the cap-binding pocket residues are located. A key component of eIF4E cap-binding involves sandwiched stacking of the methylated guanine ring between two aromatic tryptophan residues at the top and bottom of the cupped hand that holds the guanine in place through interactions between the ring systems. Several other cap-binding proteins also involve stacking of the cap between aromatic residues as part of their mechanism of cap binding (35–37). The N-7 methyl group on the guanosine cap introduces a positive charge to this moiety that greatly enhances the stacking interaction and stabilizes the complex (25, 30). eIF4E does not effectively recognize unmethylated guanosine phosphate and requires more than one phosphate residue on the N-7 methylated guanosine (1). The mRNA cap is further stabilized inside the eIF4E cap-binding pocket by hydrogen bonds derived from the side chain carboxylate of a conserved glutamic acid, a backbone NH of one of the stacking tryptophans, and additional bonds derived from conserved eIF4E residues with the three phosphate groups and the ribose of the cap (30). van der Waals contacts also contribute to cap stabilization including contact from a third conserved tryptophan with the N-7 methyl group.

Despite this understanding of how eIF4E recognizes the monomethyl cap, the mechanism by which eIF4E proteins from trans-splicing organisms recognize and bind the alternate m2,2,7G cap is not known. Here we describe the cloning, expression, and characterization of the sole form of eIF4E in the human parasitic flatworm Schistosoma mansoni. Schistosomes infect ~200 million people resulting in significant morbidity, and an estimated 800 million people are at risk of infection (16, 17). Schistosome eIF4E exhibits dual cap specificity with similar binding affinity for both m7GpppG and m7GpppA. In addition, we used NMR, fluorescence titration, and isothermal titration calorimetry studies to obtain insights into the ability of schistosome eIF4E to bind the atypical m2,2,7G cap. These studies provide the basis for understanding recognition of the m2,2,7G cap by a diverse group of important parasitic worms and provide general insights into the key cap-binding protein eIF4E.

**EXPERIMENTAL PROCEDURES**

**Cloning of the S. mansoni eIF4E**—Sequences encoding the S. mansoni eIF4E were identified in public genomic and expressed sequence tag databases (38, 39). These sequences were used to design primers corresponding to the full eIF4E open reading frame (5' = ATGACGGCTGTTGAGAGT and 3' = CTA-AATTTCATATTTTCCAGTAC). The primers were used to PCR amplify the open reading frame from cDNA prepared by random priming of adult schistosome RNA (40). The PCR product was initially cloned into a pET-30 Ek/LIC vector (Novagen, Madison, WI), and the sequence was confirmed.

**Protein Expression and Purification**—Initial protein expression and purification were carried out in Rosetta™(DE3) cells induced at 30 °C for 3 h with 0.4 mm isopropyl β-D-thiogalactoside. The protein was purified by nickel-nitrilotriacetic acid-agarose and used for cap analog-Sepharose chromatography as described previously (19). Protein for fluorescent titration studies was prepared from bacterial cultures induced with 0.2 mm isopropyl β-D-thiogalactoside at A600 = 0.7 and incubated overnight at 25 °C. The cells were collected by centrifugation at 4,000 × g, suspended in binding buffer (20 mm Hepes, pH 7.4, 100 mm KCl, 1 mm dithiothreitol, 1 mm EDTA), sonicated, and clarified by centrifugation at 16,000 × g. The supernatant was loaded onto m7GTP-Sepharose resin (GE Healthcare) and rotated at 4 °C for 2 h. The resin was washed with the binding buffer, and eIF4E was eluted with 2 m KCl in binding buffer. The eluted protein was applied to a Superdex 75 (GE Healthcare) column equilibrated in binding buffer, and monomeric protein was collected, analyzed by SDS-PAGE, and concentrated using an Amicon 10,000 molecular weight Ultra Centrifugal Filter Device (Millipore, Billerica, MA).

**Cap Analog-Sepharose Assays**—Sepharose affinity assays were carried out as described previously (19). Briefly ~15 μg of protein in binding buffer (described above) was mixed with 20 μl of m7GTP-Sepharose resin (GE Healthcare) or m2,2,7GTP-Sepharose (22) pre-equilibrated in binding buffer. After incubation on a Nutator at 4 °C for 1 h, the resin was washed three times with 100 μl of binding buffer, and the protein was then eluted by addition of 20 μl of binding buffer with 200 μM cap analog. Eluted protein was analyzed by SDS-PAGE.

**Fluorescence Titration**—The fluorescence titration was performed using a FluoroMax-3 spectrophotometer (Horiba Ltd.) with full-length eIF4E at 20 °C and a protein concentration of 0.4 μM in binding buffer. The excitation wavelength was 280 nm (slit, 1 nm), and the emission wavelength was 335 nm (slit, 2 nm). The temperature was kept at 20 °C with a thermostcule inside a thermostated cuvette, and the sample was stirred magnetically. Assays were performed by addition of 1 μl of increasing concentrations of cap analog to 1400 μl of protein solution using an integration time of 30 s and a gap of 60 s. Dilution during the titration did not exceed 2.5%. Fluorescence intensity was corrected for the inner filter effect, and the emission of the free cap analogs was explicitly included in the numerical analysis. A theoretical curve for the fluorescence intensity as a function of the total ligand concentration was fitted to the experimental data points by means of non-linear, least squares method using GraphPad Software as described previously (9, 30, 41).

**Protein Preparation for Crystallization**—Because of the flexibility of the eIF4E N terminus, most eIF4E crystals have been produced and the structures have been solved using N-terminal truncated proteins (25, 27, 31). From eIF4E sequence alignments and limited trypsin digestions, we chose to truncate the schistosome protein by 22 residues at the N terminus to generate schistosome eIF4E (23–203). This coding region was subcloned into pGEX6P-1 (GE Healthcare) using EcoRI and XhoI restriction sites and then transformed into Escherichia coli strain XA90. Protein expression was induced, the cells were collected and sonicated, and the supernatant was recovered as described above except that phosphate-buffered saline buffer was used. The cell lysate was incubated with Glutathione-Sepharose 4B (GE Healthcare) for 4 h at 4 °C with rotation.
resin was washed with phosphate-buffered saline buffer, and eIF4E was cleaved from the glutathione S-transferase fusion protein by incubation with PreScission Protease at 4 °C overnight. Eluted protein was further purified as described above on a Superdex 75 column.

**Crystallization, Data Collection, and Refinement**—Schistosome eIF4E-(23–203) was concentrated to 10 mg/ml (Amicon 10,000 molecular weight Ultra Centrifugal Filter Device) and then incubated with 1 mM 4E-BP peptide (SGSGRIIYDRK-FLMECRNSPV, corresponding to residues 51–67 of the human 4E-BP1 sequence) and 0.5 mM m7GpppA or m7GpppG cap analogs at 4 °C for 1 h before crystallization setup using the hanging drop vapor diffusion method. Thin platelike crystals developed in 20% polyethylene glycol 4000, 0.2M MgCl2, 100 mM MOPS, pH 6.0–6.5 at 4 °C within 1 week using 0.5 ml of 750 mM NaCl as the well solution. Before the data collection, the crystal was transferred into the equivalent mother solution containing 30% polyethylene glycol 4000 for 15 min and then flash-cooled in liquid nitrogen. Crystallographic data were collected using the mail-in data collection program at the Advanced Light Source beamline 4.2.2 at the Lawrence Berkeley National Laboratory. The structure of schistosome eIF4E-(23–203) was solved by the molecular replacement method using human eIF4E (Protein Data Bank code 2V8W) as a model and the program PHASER in the CCP4 suite (42). Model building and manual refinement were performed in COOT (43) and REFMAC from the CCP4 suite (42), respectively. Figures were generated using Pymol (44).

**Accession Numbers**—Atomic coordinates and structure factors have been deposited in the Protein Data Bank for schistosome eIF4E-m7GpppG+4E-BP complex (Protein Data Bank code 3HXI) and schistosome eIF4Em7GpppA+4E-BP complex (Protein Data Bank code 3HXG).

**NMR Sample Preparation**—Full-length schistosome eIF4E protein was purified by a refolding method. The induction was done as described above, and the bacterial pellet was dissolved in denaturant solution (20 mM Hepes, pH 7.5, containing 6M guanidine HCl, 1 mM dithiothreitol, 1 mM EDTA). The protein solution was then diluted 10-fold into refolding buffer (20 mM Hepes, pH 7.5, containing 1 mM arginine, 300 mM NaCl, 1 mM dithiothreitol) and incubated at 4 °C for 2 h. Following centrifugation at 16,000 g, the supernatant was dialyzed overnight against binding buffer and then mixed with m7GTP-Sepharose. Protein was eluted with 2M KCl and further purified on Superdex 75 using NMR buffer (50 mM pH 7.4 phosphate, 100 mM KCl, 1 mM dithiothreitol, 1 mM EDTA).

Isotopically enriched schistosome eIF4E was prepared by growing cells in minimal M9 medium containing 15NH4Cl and/or [13C6]glucose. To prepare uniformly labeled schistosome eIF4E, host cells were grown in M9 minimum D2O medium supplemented with 1g/liter 2H,15N,13C-labeled ISOGRO® (ISOTEC™, Miamisburg, OH).
NMR Experiments and Data Processing—All NMR spectra were acquired on a 900-MHz Varian spectrometer using 0.6 mm triple labeled sample (2H,15N,13C) at a probe temperature of 25 °C. For the backbone assignment, standard three-dimensional resonance NMR experiments were conducted including HNCACB, HNCOA, HNOCACB, 15N NOESY-HSQC, and 15N-HSQC-NOESY-HSQC. For NMR chemical shift perturbation experiments, two-dimensional 15N HSQC were collected using 0.2 mM 15N-labeled sample supplemented with 7GpppG or 2,2,7GpppG cap analogs (45, 46) in a 1:1 ratio, leading to >90% of the protein in complex with cap according to the fluorescence binding constants. NMR data were processed using NMRPipe (43) and analyzed using CcpNmr (47).

Isothermal Titration Calorimetry—The N-terminal truncated protein (25 μm) used for crystallization studies in 20 mM Hepes, pH 7.4, 1 mM EDTA, 100 mM KCl was used in titration experiments carried out at 20 °C using a VP-ITC calorimeter (MicroCal Inc., Northampton, MA). Each titration experiment consisted of a 5-μl injection followed by 29 injections of 10 μl of 700 μM cap analog. All the data were processed using the single binding site model in Origin (Version 7.0, MicroCal Inc.). Control titration experiments were performed and subtracted (cap titrated into buffer) from the titrations for enthalpy changes due to dilution.

RESULTS

Cap Binding Specificity of Schistosome eIF4E—Sequence searches for eIF4E in comprehensive genomic and expressed sequence tag databases identified only a single isoform of eIF4E in S. mansoni or a second schistosome species, Schistosoma japonicum (38, 39, 48, 49). The presence of only a single eIF4E isoform is atypical for metazoa (50, 51). The schistosome eIF4E has 32% identity and 51% similarity with human eIF4E. Notably schistosome eIF4E is more highly divergent from human eIF4E than are nematode eIF4Es. Schistosome mRNAs exhibit two different types of mRNA caps, the typical m7GpppN cap and the m2,2,7GpppA cap added during spliced leader trans-splicing (Fig. 1A) (12). Although the exact percentage of schistosome mRNAs that are trans-spliced is not known, a rough estimate of ~10% has been proposed (52, 53). Most eIF4E proteins have a lower affinity for trimethylguanosine caps compared with the monomethylguanosine form (30). The presence of only a single eIF4E isoform and the presence of two different mRNA caps in schistosomes prompted us to determine whether the S. mansoni eIF4E was capable of binding both monomethyl- and trimethylguanosine mRNA caps.

We first examined the cap specificity of recombinant schistosome eIF4E using a qualitative assay, cap analog-Sepharose chromatography (19). As shown in Fig. 1, B and C, purified schistosome eIF4E binds both monomethyl- and trimethylguanosine cap. To quantitatively investigate the binding affinity of schistosome eIF4E to cap analogs, a fluorescence titration assay was conducted. The equilibrium dissociation constants calculated for schistosome eIF4E are KD = 0.27 μM for m7GpppG and KD = 1.27 μM for m2,2,7GpppG (Fig. 1D). The affinity of schistosome eIF4E for m7GpppG is similar to that observed for truncated murine and full-length human eIF4E, KD = 0.14 and 0.10 μM, respectively (30). Schistosome eIF4E binds the m2,2,7GpppG cap with a slightly lower affinity (~5-fold). Schistosome eIF4E binds the m7GpppA cap (the native schistosome trans-spliced cap) with a similar affinity (data not shown). Notably, previous studies on murine eIF4E have demonstrated a several hundred-fold lower affinity for the m2,2,7G compared with the m7G mononucleotide cap (30, 54). We conclude that the schistosome eIF4E protein binds both m7G and m2,2,7G cap and discriminates less against the m2,2,7G cap compared with other eukaryotic eIF4E proteins, consistent with its biological requirement of binding both caps.

Overall Crystal Structure of Schistosome eIF4E—To understand the structural basis for the ability of schistosome eIF4E to bind both the m2,2,7G and m7G caps, we solved the crystal structure of a schistosome eIF4E m7GpppG-4E-BP complex to 1.8-Å resolution (Table 1). The electron density is well defined except for residue Asp-196 and the side chains of Lys-39, Arg-70, Glu-71, Arg-145, Ser-174, and Asp-195. The overall structure of schistosome eIF4E is very similar to that of other eIF4Es (Fig. 2, B and C) (5, 25–27, 30, 33, 34, 55), containing eight β-strands and three α-helices that form a glove-shaped structure with the m7GpppG cap analog situated on the concave surface of the β-strands. The 4E-BP1 peptide binds to the opposite surface of the protein and forms interactions with schistosome eIF4E almost identical to those previously observed in the human eIF4E-4E-BP complex structure (56). Therefore, these are not discussed further here. The 4E-BP1 peptide was used to enable crystal formation. In comparison with the human eIF4E

| TABLE 1 Crystal data and refinement statistics for schistosome eIF4E ternary complex (eIF4E-m7GpppG-4E-BP peptide) |
|---------------------------------------------------------------|
| Data collection                                              | m7GpppG | m7GpppA |
| Cell dimensions                                             |         |         |
| a (Å)                                                        | 45.31   | 45.56   |
| b (Å)                                                        | 125.33  | 125.4   |
| c (Å)                                                        | 37.33   | 37.34   |
| α, β, γ (°)                                                  | α = β = γ = 90° | α = β = γ = 90° |
| Measured reflections                                        |         |         |
| R_rms or R_merger                                           | 0.082 (0.472) | 0.113 (0.275) |
| I/σ                                                         | 7.8 (2.0) | 6.6 (3.2) |
| Completeness (%)                                            | 94.8 (96.2) | 98.3 (98.1) |
| Redundancy                                                  | 4.0 (3.0) | 3.7 (3.7) |
| Refinement                                                  | 1.8     | 2.1     |
| Number of reflections                                       | 17,407  | 13,064  |
| Number of atoms                                             | 238.3/28.9 | 23.6/29.1 |
| B factors (Å²)                                              |         |         |
| Protein                                                     | 22.8    | 19.9    |
| Peptide                                                     | 26.4    | 23.3    |
| Ligand                                                      | 22.8    | 17.6    |
| Water                                                        | 29.8    | 25.6    |
| r.m.s. deviations                                           |         |         |
| Bond length (Å)                                             | 0.011   | 0.009   |
| Bond angles (°)                                             | 1.348   | 1.222   |
| Ramachandran plot                                           |         |         |
| Residues in most favorable regions (%)                      | 94.8    | 93.8    |
| Residues in disallowed regions (%)                          | 0       | 0       |

* Values in parentheses are for the highest resolution shell.

b Root mean square.
Parasite eIF4E Structure and m\(^7\)G and m\(^{2,2,7}\)G Cap Binding

(Protein Data Bank code 1WKW), the schistosome eIF4E complex structure has a root mean square deviation of 0.66 Å. Although the overall structure of the schistosome and other eIF4Es is similar, there is a marked difference between the schistosome and human eIF4E (Protein Data Bank code 1WKW) in the position of two loops (S1-S2 loop and S7-S8 loop) (Fig. 2C). The different position for these loops appears to be related to the orientation of the second base of the cap in the schistosome eIF4E structure, which is very different in the two proteins (see below).

**Schistosome eIF4E Cap-binding Site**—Overall the cap-binding pocket and interactions between the cap and residues in the pockets are very similar between schistosome and mammalian eIF4E. The electron density of the dinucleotide cap analog, m\(^7\)GpppG, is well defined in our structure (Fig. 3A), allowing its position to be precisely determined. The residues involved in dinucleotide cap binding include Trp-43, Trp-89, Glu-90, Gln-141, Arg-143, Lys-148, and Arg-192. The major features that contribute to the interaction of the schistosome eIF4E with the cap include the following (Fig. 3 and supplemental Fig. 1): 1) cation–π stacking of the positive charged m\(^7\)G moiety with two highly conserved aromatic Trp residues (Trp-43 and Trp-89) (Fig. 3A and supplemental Fig. 1); 2) two hydrogen bonds between the N-1 hydrogen and N\(^2\) hydrogen of the m\(^7\)G moiety and the carboxyl group of conserved Glu-90 and one more between O\(^6\) of m\(^7\)G and the backbone amide nitrogen of Trp-89 (Fig. 3A and supplemental Fig. 1); and 3) interaction between the three phosphates in the cap dinucleotide and residues (Gln-141, Arg-143, Lys-148, and Arg-192) that constitute the phosphate access slot as illustrated in Fig. 3B and supplemental Fig. 1. The presence of the N-1 H–O bond (2.93 Å) suggests that the cationic form of the m\(^7\)G moiety is preferred. In the complex of m\(^7\)GpppG, N\(^2\) hydrogen of the second base forms a weak hydrogen bond (3.1 Å) with the backbone oxygen of Leu-41. This bond is not present in the complex with m\(^7\)GpppA because of the lack of a hydrogen donor in adenosine at a suitable position. This is consistent with a slightly higher affinity of mammalian eIF4E for m\(^7\)GpppG versus m\(^7\)GpppA and m\(^7\)GpppC (30, 57). All of the cap-binding residues, except Leu-41 that interacts with the cap only through the backbone, are highly conserved in

One observed difference in the geometry of the interatomic contacts within the schistosome eIF4E cap-binding center that might contribute to the similar specificity for the m\(^7\)G and m\(^{2,2,7}\)G forms of the mRNA 5′ cap is the variant conformation of the carboxylate of a conserved Glu-90. This carboxylate is turned by ~80° compared with the orthologous Glu-103 carboxylate in the mammalian eIF4E (Fig. 2C). The result of this rotation is that Glu-90 forms only one hydrogen bond with m\(^7\)G of the cap, whereas in human eIF4E, the equivalent residue Glu-103 forms two hydrogen bonds with both N-1 and N\(^2\) of the cap m\(^7\)G. This almost perpendicular position of the schistosome Glu-90 carboxylate versus mammalian Glu-103 results from the interaction with N\(^7\) of Arg-188 (bond lengths of 2.97 and 3.29 Å in the m\(^7\)GpppG and m\(^7\)GpppA complexes, respectively) fixing the position of the Glu. In human eIF4E, there is no equivalent interaction as this region of the protein is moved out away from Glu-103 due to the positioning of the second base (Fig. 2C) (see below). Mutation of Glu-90 (Ala, Asp, or Arg)
results in protein oligomerization suggesting that this residue contributes structurally to the apo form of the protein. Consequently direct analysis of the contribution of this residue to cap discrimination is difficult. The conformation of this conserved Glu and its contribution to m$^7$G/m$^{2,2,7}$G specificity in other eIF4E isoforms remains an open question.

Position of the Second Base of the Bound Cap Analog—Although the positions of the m$^7$G moiety, the ribose, and the first phosphate group of the cap as well as the conserved eIF4E tryptophans align well between schistosome and mammalian eIF4E (Fig. 2C), the orientation of the second base in the schistosome structure differs from that observed in the mammalian eIF4E structure (Fig. 2C). In previous eIF4E structures with a dinucleotide cap, the second base of the cap was only well resolved in the human eIF4E complex structure (22, 23) apparently due to the flexibility of this nucleotide. In the human structure, the second base (m$^7$GpppA) points to the C-terminal flexible loop and has some hydrogen bonding with the residues located in this loop. However, in the schistosome eIF4E complex structure, the orientation of the second base (m$^7$GpppG) points to the N-terminal S1-S2 loop as shown in Fig. 2C. To test whether the orientation of the second base in the two structures was a consequence of the different purines at the second base (A versus G), the structure of schistosome eIF4E in complex with m$^7$GpppA was also determined (supplemental Fig. 2). In the m$^7$GpppA structure, the location of the second base also points to the N-terminal S1-S2 loop as observed for the m$^7$GpppG complex.

The position of the second base in the schistosome eIF4E structure allows stacking of the base (A or G) onto a Phe (derived from vector sequence) from a symmetry-related molecule in the crystal as shown in Fig. 4A. This Phe may mimic the third base in an RNA and stabilize the position of the second base in our structure. The cap $\gamma$-phosphate in the schistosome eIF4E structure forms three hydrogen bonds with the side chain of Arg-143 and Arg-192 (Fig. 3B). These additional hydrogen bonds to the $\gamma$-phosphate in the schistosome eIF4E structure are consistent with the higher eIF4E binding affinity observed for m$^7$GTP compared with m$^7$GDP in mammalian eIF4E (30). Furthermore mammalian eIF4E binds the dinucleotide m$^7$GpppG cap with lower affinity (~20-fold) compared with m$^7$GTP. The schistosome eIF4E has a similar affinity for the mono- and dinucleotide cap. Overall these data suggest that the second base is not likely to have extensive interactions with eIF4E protein and might suggest that the main interaction for the second base would be stacking with the third base of the RNA.

Consequences of the Position of the Second Base—The orientation of the second cap base is associated with differences in the positions of the two loops (S1-S2 loop and S7-S8 loop) between the schistosome and human eIF4E structures. Specifically in the human structure loop S7-S8 is swung away from the rest of the protein due to the position of the second base of the cap (Fig. 2C). In the schistosome structure, the second base of the cap is positioned in a nearly opposite orientation that allows S7-S8 to move closer to the cap-binding pocket in a “clamp”-like position. Loop S1-S2 in the schistosome structure also shifts slightly (compared with the human structure) to accommodate the position of the second base. The orientation of the second base and its influence on loop S7-S8 likely does not play a role in cap binding specificity but would impact the path of the RNA.

To explore possible pathways that the mRNA may take, we modeled additional bases into the schistosome and human structures (Fig. 4, B and C). Although the structure of this part of the mRNA in different mRNAs may vary and is not known, we modeled the RNA using canonical A-form base stacking, the favored conformation for RNA, even in the single-stranded state (58). When the RNA is modeled onto the human structure so that the third base stacks directly on the second, the predicted RNA has significant steric clash with the protein in loop S7-S8 and the adjacent helix. Manual manipulation of the modeled mRNA revealed a position without steric clash, but this position required severe contortion of the RNA backbone between the second and third bases (Fig. 4C, blue RNA). In contrast, when the mRNA was modeled on the schistosome eIF4E structure so that the third base stacks directly on the second base, the path of the sugar phosphate backbone roughly parallels loop S1-S2. A single steric clash is seen between

Parasite eIF4E Structure and m$^7$G and m$^{2,2,7}$G Cap Binding
Parasite eIF4E Structure and m\textsuperscript{7}G and m\textsubscript{2,2,7}G Cap Binding

![Diagram of eIF4E structure and mRNA paths](image)

**FIGURE 4. Proposed mRNA path based on eIF4E complex crystal structure.** A, a Phe derived from vector sequence in the symmetry-related molecule stacks with the second base, mimicking the third base of the RNA. The tan color denotes the model from a neighboring molecule. B, potential RNA paths based on the schistosome eIF4E complex crystal structure. C, potential RNA paths based on the human eIF4E complex crystal structure (Protein Data Bank code 1WKW). The two RNA orientations (red and blue) are based on a C-3 or C-2 endo ribose conformation with the extended RNAs in an idealized A-form conformation. D, illustration of potential distance change of human Ser-209 from the \( \gamma \)-phosphate of the cap based on the location of the schistosome S7-S8 loop. The red dashed arrow represents the movement of the loop based on the schistosome structure with the new location of the Ser-209 at the arrowhead. The dashed dark blue line represents the 12.6-Å distance of Ser-209 to the phosphate, and the light blue line represents the 6.6-Å distance to the phosphate with the altered location of the S7-S8 loop (yellow, schistosome; blue, human eIF4E).

Phe-17 and the RNA. A second modeled position for the mRNA on the schistosome eIF4E structure in which the third base does not stack directly on the second base also allowed the mRNA to exit the protein without any steric clash (Fig. 4B). Both of these modeled positions on schistosome eIF4E are compatible with our structure and illustrate minimal direct interaction of the RNA with eIF4E. This is consistent with a variety of studies suggesting that eIF4E does not interact specifically with the body of the mRNA (59).

Importantly the position of the second base in the schistosome eIF4E structure is also consistent with data related to the phosphorylation of Ser-209 in mammalian eIF4E. Phosphorylation of Ser-209 can reduce eIF4E cap binding affinity (57, 60–62), suggesting that electrostatic repulsion between the cap phosphate and phosphorylated Ser-209 might reduce cap affinity (62). Ser-209 is located in a very flexible loop (S7–S8) in the eIF4E structure. In the human structure (Protein Data Bank code 1WKW), the distance between the Ca of with the key Glu-90 residue. Our concerted efforts to obtain schistosome eIF4E crystals with a m\textsubscript{2,2,7}G cap were not successful. One possible reason for this is that the N\textsuperscript{2} of m\textsubscript{7}G cap forms crystal packing contacts with Asp-84 in a symmetry-related molecule. Thus, the methylation of N\textsuperscript{2} would likely lead to steric hindrance.

Therefore, to further explore the mechanism of schistosome m\textsubscript{2,2,7}G cap binding specificity, we undertook NMR chemical shift perturbation studies to compare schistosome eIF4E binding to the m\textsuperscript{7}GpppG versus m\textsuperscript{2,2,7}GpppG cap. NMR experiments could not be done for free protein because the free schistosome eIF4E protein was not stable at the requisite high concentration for NMR experiments. Superimposition of the \(^{1}H^{15}N\) HSQC spectrum of m\textsuperscript{7}GpppG-bound (red) versus m\textsuperscript{2,2,7}GpppG-bound (blue) schistosome eIF4E is shown in Fig. 5A, and a summary of the induced chemical shift perturbations with the different caps is provided in Fig. 5B. Surprisingly there are \( \sim 15 \) residues with significant chemical shift perturbations.
(CSPs) that differ on schistosome eIF4E binding to the m^{2,2,7}GpppG versus the m^{7}G cap (Fig. 5B). The only difference between the caps is the addition of two methyl groups at the N^2 position (Fig. 1A), yet around 10% of the schistosome residues show major chemical shift differences between the m^{7}G and m^{2,2,7}G complexes. Residues with major CSPs were mapped onto the schistosome eIF4E crystal structure (Fig. 5, C–E). Most of these residues are distributed around the cap-binding pocket (Fig. 5C; Gln-36, Asp-44, Glu-45, Leu-47, Glu-48, Val-50, Lys-93, Asn-94, Phe-140, Arg-143, Arg-145, Ser-147, and Ser-185). The residues Gln-36, Asp-44, Glu-45, Leu-47, Glu-48, and Val-50, which are around Trp-43 and located in the S2 strands, also undergo major CSPs in the trimethylguanosine compared with the m^{7}G cap-bound complex. Based on the cap binding mechanism recently suggested by Volpon et al. (28), this region is predicted to move like a hinge to lock the capped guanine initially and then hold the m^{7}G moiety in the appropriate position to form the interaction with Trp-56. This suggests that this

**FIGURE 5.** NMR-determined chemical shift perturbations on schistosome eIF4E binding to m^{2,2,7}GpppG compared with m^{7}GpppG. A, superposition of the 1H-15N HSQC spectrum of the schistosome eIF4E bound to m^{7}GpppG (red) and bound to m^{2,2,7}GpppG (blue). B, specific chemical shift perturbations on binding m^{2,2,7}GpppG compared with m^{7}GpppG. The reported chemical shift Δδ represents the weighted chemical shift by applying the Pythagorean theorem: Δδ(1H, 15N) = (Δδ(1H)^2 + Δδ(15N)^2)^1/2, where Δδ(1H) and Δδ(15N) are the chemical shift difference of the amide proton and nitrogen, respectively (28, 69). Amino acid residues with large chemical shifts are labeled. C and D, residues with major chemical shift perturbation on binding m^{2,2,7}G compared with the m^{7}G cap mapped onto the crystal structure of the m^{7}GpppGeIF4E complex. The backbone residues with major chemical shifts are colored red. Residues with major chemical shifts in the vicinity of the second base of the cap are shown in C, and those with major chemical shifts in the vicinity of the third phosphate of the cap are shown in D. E, Clustal sequence alignment of several eIF4E illustrating conserved residues, residues undergoing CSP, and mouse pepsin cleavage sites on cap binding. The black boxes denote residues with major chemical shift perturbation in our NMR experiments as described above. The vertical black arrows denote the specific pepsin cleavage sites based on the mammalian eIF4E (21). Starred residues are those involved in cap binding with the two key Trp residues in red. The red box indicates two key residues for *C. elegans* eIF4E reported to be important for m^{2,2,7}G cap binding specificity (24). Note that *C. elegans* eIF4E-5 and *Ascaris* eIF4E-3 binds both types of caps, whereas mouse eIF4E3 and *C. elegans* eIF4E-3 are specific for m^{7}G cap.
region plays a crucial role for enabling the m7G cap to enter the binding pocket. Significant CSPs for residues in this region were observed on binding m7GpppG compared with m2,2,7GpppG as shown in Fig. 5. One interpretation of these differences could be that the path or mechanisms through which the m7G and m2,2,7G caps enter into the binding pocket have discrete differences. Several residues (Ile-72, Leu-117, Val-151, Phe-181, and Thr-198) with major CSPs are present in a region distal from the cap-binding pocket (Fig. 5, C and D) indicating that long distance conformational changes also occur upon m2,2,7G cap binding. Overall these data suggest that some features of the overall binding mechanism and specific interactions within the binding pocket for each cap are different.

**DISCUSSION**

**Binding Specificity of Schistosome eIF4E and Translation**—Cap analog-Sepharose chromatography demonstrates that schistosome eIF4E can bind both m7G and m2,2,7G caps. The fluorescent titration analyses demonstrate that 1) schistosome eIF4E recognizes both the m7G and m2,2,7G caps with similar affinity and that 2) the schistosome affinity for m7G cap is comparable to that of mammalian eIF4E (Fig. 1). However, the affinity of mammalian eIF4E for the m2,2,7GTP cap is several hundred-fold less than for the m7G (30). Thus, schistosome eIF4E has relatively high affinity for the m2,2,7G cap. Unlike other trans-splicing organisms examined (22, 23, 50, 64), only a single eIF4E isoform is identifiable in schistosomes. This suggests that observed. The binding affinity is likely to be influenced by the different entropic changes observed suggesting that there are potential differences in the overall binding mechanism. The eIF4E cap binding affinity differences observed using ITC and fluorescence titration (Fig. 1D and Table 2) is likely due to methodological differences (63).
Parasite eIF4E Structure and m\textsuperscript{7}G and m\textsuperscript{2,2,7}G Cap Binding

A

B

C

human K159/Schistosome R145

human K119

S4-H2 loop

human Q120

S7-H3 loop
this single eIF4E likely initiates translation for both non-trans-spliced (m7G) and trans-spliced mRNAs (m2,2,7G).

Overall Crystal Structure of Schistosome eIF4Eem1 GpppG/A—Although the schistosome eIF4E has limited sequence identity with other eIF4Es, the overall structure in complex with either m7GpppG or m7GpppA is very similar to that observed for other eukaryotic eIF4Es (Fig. 2, B and C). The mode of binding m7G and the interactions within the cap-binding pocket are essentially identical to that observed for other forms of eIF4E (5, 25, 26, 28, 30, 33) (Fig. 3, A and B). One exception is different orientation of the conserved Glu-90 carboxylate in the cap-binding pocket that is stabilized by Arg-188. It remains to be determined whether this configuration and the orientation of the Glu-90 carboxylate play a role in the binding specificity.

Position of the Second Cap Base and RNA—A major difference in the schistosome eIF4E structure compared with human eIF4E bound to the dinucleotide m7GpppA is the position of two loops, S1-S2 and S7-S8 (Figs. 2C and 6C). The position of these loops appears to be a function of the different location of the second nucleotide in the cap in the schistosome compared with the human eIF4E structure (Fig. 2C). Several aspects of the position of the second nucleotide in the schistosome eIF4E structure are consistent with other data. Modeling of a longer RNA into the schistosome eIF4E structure indicates that the position of the second nucleotide leads to an unobstructed RNA path leaving the schistosome cap-binding pocket (Fig. 4B). In contrast, similar modeling of the human structure leads to steric hindrance with the protein or a highly contorted RNA (Fig. 4C).

The schistosome structure is also consistent with data showing that phosphorylation of Ser-209 in mammalian eIF4E can reduce the affinity of eIF4E for the cap (57, 60 – 62). In addition, mutation of Ser-209 to Glu decreases cap binding affinity of mammalian eIF4E (62), which is consistent with our structure. Although the role of Ser-209 phosphorylation remains unclear in eIF4E function and translation, the schistosome eIF4E crystal structure provides a reasonable model to explain how electrostatic repulsion between the human Ser-209 and cap might affect cap binding. In sum, the orientation of the second cap base in the schistosome structure is consistent with a likely RNA path and some functional data on cap binding affinity.

Comparison of Schistosome Binding to m7G Versus m2,2,7G Cap—Given that the structure and apparent mechanism of schistosome eIF4E binding to m7G cap is strikingly similar to mammalian eIF4E binding to m7G cap, what then enables the schistosome protein to bind to the structurally different m2,2,7G cap? To begin to address this, we used the NMR CSP technique to compare schistosome eIF4E binding to the m7G and m2,2,7G caps. The major CSP differences observed on binding the two different caps can be divided into three categories based on their location in the schistosome eIF4E complex structure with m7G cap: 1) residues in the vicinity of the two key Trp-43 and Trp-89 aromatics involved in stacking the first guanine of the cap (Fig. 5C), 2) residues in the vicinity of the phosphates in the cap-binding region of the protein (Fig. 5D), and 3) residues that are at a significant distance from the cap-binding pocket. Whereas residues in the first two categories likely are involved in direct interactions with the two different caps, residues at a significant distance from the cap-binding pocket may be involved in allosteric conformational and/or stability changes of the protein as a whole on binding the two different caps. These conformational change differences observed by NMR throughout the protein may reflect a difference in conformational dynamics, which contributes to differences in binding affinity as suggested by the thermodynamics studies (Fig. 6).

Recent studies have shown that murine eIF4E interaction with m2,2,7G versus m7G cap leads to differences in pepsin susceptibility within the protein (21). The appearance of new cleavage sites in the vicinity of Trp-56 (murine eIF4E) on interaction with m2,2,7G cap is consistent with our NMR chemical shift data (Fig. 5E). These data may indicate that the m2,2,7G moiety can enter the cap-binding slot even in mammalian eIF4E but that productive binding does not ensue (Fig. 5E). However, we also observed significant chemical shift perturbations for the residues located around Trp-102 and residues in the region that interact with the phosphates (Fig. 5, C and D). These perturbations differ from those observed in the pepsin cleavage studies on mammalian eIF4E. These differences are likely due to the selective binding affinity for the m2,2,7G cap and may suggest that the m2,2,7G cap does not have free access to the whole cap-binding slot in murine eIF4E.

The new schistosome m7G crystal structure described here suggests that the ability of this protein to bind both caps with similar affinity is not a result of major differences in the overall cap binding mechanism. Thus, all residues involved in binding the monomethyl cap are identical between mammalian and schistosome eIF4E. Therefore, based on the available data, we favor a model in which intrinsic protein flexibility or conformational changes likely play a crucial role in m2,2,7G cap binding. This hypothesis is supported by NMR and in part by our ITC data. Using data from our NMR chemical shift perturbation studies, mammalian eIF4E pepsin cleavage data on binding different cap analogs (21), and the mammalian eIF4E free protein structure (28), we suggest the following m2,2,7G cap binding mechanism model. When the schistosome eIF4E associates with the m2,2,7G cap, the stacking interaction between the guanine ring and the Trp (Trp-56 in human) induces a specific protein conformational change. This conformational change specifically promotes formation of a m2,2,7G cap-binding pocket, analogous to an “induced fit” mechanism. Thus, residues far away from the cap-binding pocket are also observed to have significant chemical shift perturbations. In mammalian eIF4E, association and interaction with the m2,2,7G cap may not lead to the requisite conformational changes and formation of the necessary m2,2,7G cap-binding pocket. As a consequence, mammalian eIF4E has a much lower binding affinity for m2,2,7G.
Parasite eIF4E Structure and m7G and m2,2,7G Cap Binding

cap compared with m7G cap. This model suggests that intrinsic and specific conformational flexibility of the schistosome eIF4E plays the crucial role for m2,2,7G cap binding.

Steric hindrance of the cap due to the additional two methyl groups at the N2 position of the m2,2,7G cap has been suggested previously to likely be important in the reduced affinity of mammalian eIF4E in binding m2,2,7G cap (9, 21). Previous mutagenesis studies for C. elegans eIF4E-5, a form of eIF4E that can bind both types of caps, showed that two residue changes (N64Y/V65L) (Fig. 5E) led to a decrease in cap specificity for the m2,2,7G cap (24). It was suggested that these two residues could impact the width and depth of cap-binding slot. However, another nematode eIF4E (Ascaris eIF4E-3) has these Tyr-64/Leu-65 residues and readily binds the m2,2,7G cap (19). Furthermore, our NMR chemical shift perturbation data show that more than 15 residues have major chemical shifts as shown in Fig. 5, A and B, indicating that relatively significant intrinsic conformational changes throughout the molecule are also likely to play a role in cap-binding specificity rather than simply the dimensions of the cap-binding slot and steric considerations due to the additional methyl groups in the m2,2,7G cap.

Other Insights from the Schistosome eIF4E Structure on Cap Binding—The addition of a second base to an N-7 methylated nucleotide in general does not significantly increase mammalian eIF4E affinity for the cap (9). In fact, human eIF4E binding affinity for dinucleotide triphosphate cap (mGpppG) is significantly lower than that observed for a mononucleotide (m7GTP) (9, 45). Interestingly, the schistosome eIF4E binding affinity for mono- and dinucleotide substrates is similar as determined by ITC (Table 2 and data not shown). Comparison of the schistosome and human 4E sequences (Fig. 2A) shows that two residues (Lys-119 and Gln-120) located in the S4-H2 loop were replaced by Ala-86 and Arg-87 in schistosome eIF4E. Mutation of Lys-119 or Gln-120 in the human eIF4E was reported to cause an increase in eIF4E binding affinity to cap analogs (65). In particular, the K119A mutation leads to a 10-fold increase in affinity for dinucleotide cap, whereas only a 3-fold increase is observed for a mononucleotide cap (66). An NMR structural analysis showed that the mutations in this region led to changes in the neighboring S5-S6 loop that directly interacts with the cap phosphates (28). From these data, Volpon et al. (28) suggested that the reduced interaction between these two loops (H2-S4 and S5-S6) might lead to greater conformational flexibility in the H2-S4 leading to the observed increase in cap binding.

The substitution of two residues (K119A/Q120R) and absence of three more residues in the schistosome (RRS) (Figs. 2A and 5E) H2-S4 loop results in a shorter H2 helix compared with human eIF4E. In addition, the schistosome H2-S4 loop is shifted away from the S5-S6 loop, and the S7-H3 loop is shifted inward (Fig. 6C). As a consequence, the interaction between these two loops might be stronger in the schistosome eIF4E enabling the S5-S6 loop to interact more directly with the cap phosphates. This may explain why the schistosome affinity for mono- and dinucleotide caps is similar. These changes may enable dinucleotide caps to more readily enter into the cap-binding slot in schistosome eIF4E. This might also explain why we were able to obtain a relatively high quality schistosome crystal with a dinucleotide cap.

Conclusion—We have identified and characterized the cap binding characteristics of the sole isoform of schistosome eIF4E. Several lines of data indicate that schistosome eIF4E is able to bind both m7G and m2,2,7G caps with similar affinity. We have determined the co-crystal structure of eIF4E with mGpppG and mGpppA. The schistosome eIF4E cap binding mechanism for m7G cap is very similar to that described for mammalian eIF4E. Our data demonstrate one potential difference in the cap-binding pocket and suggest a likely position for the RNA when the cap is bound by eIF4E consistent with biochemical data. We have also identified that significant conformational changes occur in eIF4E on binding m2,2,7G cap that provide insight into schistosome eIF4E binding to m2,2,7G compared with m7G cap. Interestingly, analysis of schistosome eIF4E binding to mG and m2,2,7G caps demonstrates significant thermodynamic differences. Although a detailed molecular mechanism for binding m2,2,7G cap is not yet available, our data indicate that major conformational differences occur on schistosome binding the two types of cap, and aspects of the binding mechanisms for each cap are likely to be different.

Mammalian eIF4E affinity for the m2,2,7G cap is several hundred-fold lower than observed for schistosome eIF4E (30, 54). The large difference in affinity for the m2,2,7G cap between schistosome and mammalian eIF4E suggests that this protein may represent a potential target for rational drug design against schistosomes (16, 17). Nematodes also have trans-splicing and eIF4E proteins with high affinity for the m2,2,7G cap. Nematodes infect upward of 2 billion people and are important agricultural pests (15, 17). A better understanding of the mechanisms of parasite eIF4E binding of the m2,2,7G cap might enable development of new compounds that are efficacious against a broad spectrum of important parasites.

Acknowledgments—We thank all members of the structural biology community at the University of Colorado Denver School of Medicine, University of Colorado Denver x-ray and NMR facility (supported in part by the University of Colorado Cancer Center), and Shaun Bevers at the University of Colorado Denver biophysical core without which this work would not have been possible. We thank Jay Nix at the Advanced Light Source at the Lawrence Berkeley National Laboratory for eIF4E x-ray data collection and Ryszard Stolarski, Elan Eismesser, and members of the Davis laboratory for comments on the manuscript.

REFERENCES
1. Gingras, A. C., Raught, B., and Sonenberg, N. (1999) Annu. Rev. Biochem. 68, 913–963
2. Pestova, T. V., Lorsch, J. R., and Hellen, C. U. (2007) in Translational Control in Biology and Medicine (Mathews, M. B., Sonenberg, N., and Hershey, J. W. B., eds) pp. 87–128, Cold Spring Harbor Laboratory Press, Woodbury, NY
3. von der Haar, T., Gross, J. D., Wagner, G., and McCarthy, J. E. (2004) Nat. Struct. Mol. Biol. 11, 503–511
4. Schneider, R. J., and Sonenberg, N. (2007) in Translational Control in Biology and Medicine (Mathews, M. B., Sonenberg, N., and Hershey, J. W. B., eds) pp. 401–432, Cold Spring Harbor Laboratory Press, Woodbury, NY
5. Brown, C. J., McNae, I., Fischer, P. M., and Walkinshaw, M. D. (2007) J.

NOVEMBER 6, 2009•VOLUME 284•NUMBER 45 JOURNAL OF BIOLOGICAL CHEMISTRY 31347
