Structural Requirements for the Specific Recognition of an m^{7}G mRNA Cap†

Pi-Chen Hsu,‡ Mary R. Hodel,‡ Jordan W. Thomas,§ Laura J. Taylor,§ Curt H. Hagedorn,§ and Alec E. Hodel*,‡

Department of Biochemistry, Emory University School of Medicine, 1510 Clifton Road, Atlanta, Georgia 30322, and Department of Medicine and Genetics Program-Winship Cancer Center, Emory University School of Medicine, Atlanta, Georgia 30322

Received March 17, 2000; Revised Manuscript Received July 12, 2000

ABSTRACT: 7-Methylguanosine (m^{7}G), also known as the mRNA "cap", is used as a molecular tag in eukaryotic cells to mark the 5′ end of messenger RNAs. The mRNA cap is required for several key events in gene expression in which the m^{7}G moiety is specifically recognized by cellular proteins. The configurations of the m^{7}G-binding pockets of a cellular (eIF4E) and a viral (VP39) cap-binding protein have been determined by X-ray crystallography. The binding energy has been hypothesized to result from a $\pi-\pi$ stacking interaction between aromatic residues sandwiching the m^{7}G base in addition to hydrogen bonds between the base and acidic protein side chains. To further understand the structural requirements for the specific recognition of an m^{7}G mRNA cap, we determined the effects of amino acid substitutions in eIF4E and VP39 cap-binding sites on their affinity for m^{7}GDP. The requirements for residues suggested to $\pi-\pi$ stack and hydrogen bond with the m^{7}G base were examined in each protein by measuring their affinities for m^{7}GDP by fluorimetry. The results suggest that both eIF4E and VP39 require a complicated pattern of both orientation and identity of the stacking aromatic residues to permit the selective binding of m^{7}GDP.

With the compartmentalization of genetic material inside a nuclear envelope, eukaryotic cells have by necessity evolved unique mechanisms for chaperoning the transfer of genetic information from the nucleus to the cytoplasm. To facilitate the transfer of genetic information from the nucleus to ribosomes, the messenger RNA of eukaryotes is modified at both the 5′ and 3′ ends (*1*). These terminal modifications provide molecular tags for the cellular machinery that performs splicing, nuclear export, and translation functions (*2*, *3*). Eukaryotic cells utilize the methylation of a guanosine base at the N^7 position as a tag or cap for the 5′ end of messenger RNA (*4*, *5*). The m^{7}G(5′)ppp(5′)N mRNA cap is specifically recognized in the splicing of the first intron in nascent transcripts, transport of mRNA through the nuclear envelope (*6*), and translation of the message by ribosomes (*4*, *7*). Thus, the N^7 methylation of guanosine must present a ligand that is distinct from the large pools of unmethylated guanine nucleotides in cells.

Insight into the mechanism of m^{7}G[1] cap recognition has been advanced by the determination of the atomic structures of two specific m^{7}G−protein complexes. These structures include the eIF4E−m^{7}GDP complex (*8*, *9*) and a complex between an m^{7}G-capped RNA oligonucleotide and the vaccinia methyltransferase VP39 (*10*, *11*). The cocrystal structure of the eIF4E−m^{7}GDP complex revealed that the m^{7}G base is stacked between two tryptophan residues (Trp56 and Trp102) and a glutamate side chain (Glu103) that is hydrogen bonded to N^1 and N^2 protons of the purine ring. Spectroscopic and crystallographic studies of small molecule models suggested that the mechanism for m^{7}G selectivity involved a "molecular complex" between the electron-rich indole group of tryptophan and the positively charged π-ring system of the m^{7}G base in a face-to-face stacking interaction (*12*, *13*). Tryptophan was hypothesized to be essential for the specific recognition of m^{7}G (*12*). However, the structure of vaccinia VP39 bound to m^{7}GDP demonstrated that tyrosine and phenylalanine (Tyr22 and Phe180) could substitute for the stacked tryptophan bases observed in eIF4E (*10*). These structural studies suggested that the three aromatic residues (Trp, Tyr, and Phe) could all serve to specifically recognize the m^{7}G base through stacking interactions.

To date, only limited studies of the structural requirements for m^{7}G recognition in eIF4E have been reported. Mutating Trp102 to Phe ablated all m^{7}G binding activity, while mutating Trp56 to Phe reduced the level of m^{7}G-Sepharose binding by 50−80% (*14*). Morino et al. showed that human eIF4E was unable to bind to an m^{7}G affinity column when either Trp56 or Trp102 was substituted with leucine (*15*). They also showed that cap binding was ablated when Glu103, whose side chain hydrogen bonds to the N^1 and N^2 protons of the m^{7}G base, was replaced with an alanine. These studies suggested that phenylalanine, but not other hydrophobic amino acids such as leucine, could substitute for the tryptophan in stacking with the m^{7}G base. However, the methods used in these assays did not yield quantitative measurements of the effects of these substitutions.

† This work was supported through awards from the National Institutes of Health (Grant GM58692 to A.E.H. and Grant CA63640 to C.H.) and the National Science Foundation (Grant MCB-9874548 to A.E.H.).

* To whom correspondence should be addressed. E-mail: ahodel@emory.edu. Phone: (404) 727-8764. Fax: (404) 727-3746.

‡ Department of Biochemistry.

§ Department of Medicine.

[1] Abbreviations: ant-m^{7}GDP, anthraniloyl 7-methylguanosine 5′-diphosphate; CD, circular dichroism; eIF4E, eukaryotic initiation factor 4E; GDP, guanosine 5′-diphosphate; K_d, dissociation constant; m^{7}G, 7-methylguanosine; m^{7}GDP, 7-methylguanosine 5′-diphosphate.

10.1021/bi000623p CCC: $19.00
Published on Web 10/14/2000

Here, we report a quantitative, structure-based analysis of the protein requirements for specific protein−mRNA cap recognition in both eIF4E and VP39. We have performed an extensive site-directed mutagenesis analysis of the m^7G specific binding sites in these two nonhomologous proteins. The purpose of this study was to determine the effects of amino acid substitutions in the m^7G cap-binding site of both VP39 and eIF4E on their affinity for m^7GDP. The requirements for aromatic stacking residues and acidic hydrogen bonding residues were explored by preparing site-directed variants at each position in both proteins. The ability of each variant protein to bind to m^7G was then assessed by fluorimetry-based binding assays to define the structural requirements for m^7G cap recognition.

EXPERIMENTAL PROCEDURES

Materials. Vectors pGEX-2T and pGEX-4T-3 were obtained from Amersham Pharmacia Biotech, Inc. (Piscataway, NJ). m^7GTP-Sepharose and the Hi-Trap heparin column were purchased from Amersham Pharmacia Biotech, Inc. Glutathione−agarose beads, glutathione, and m^7GDP were purchased from Sigma Chemical Co. (St. Louis, MO). The ISS PC1 photon-counting spectrofluorometer was purchased from ISS Inc. (Champaign, IL).

Construction of the GST-4E Mutants. All the eIF4Es studied here were fused to GST to facilitate the protein purification. The construction of the GST−4E fusion plasmid will be described elsewhere (manuscript in preparation). A Stratagene QuickChange mutagenesis kit was used to prepare the mutants primed with two degenerate oligonucleotides for each position. W56 and W102 were substituted with the degenerate codon HWK potentially encoding Phe, Tyr, and His as well as the nonaromatic amino acids Leu, Ile, and Lys. Glu103 was substituted with the codon VHM which encodes negatively charged Asp, positively charged Lys, and uncharged but polar amino acids Asn and Gln. The sequences of all constructs were verified by sequencing both strands of plasmid DNA.

Expression and Purification of GST−4E Fusion Proteins in Escherichia coli. Expression and purification of the eIF4E mutant proteins will be described in detail elsewhere (manuscript in preparation). Briefly, the protein purified by glutathione−agarose affinity chromatography was further purified by FPLC using a 5 mL heparin-Sepharose column. The protein was applied to a heparin-Sepharose column and washed with a solution of 10 mM Tris-HCl (pH 8), 1 mM EDTA, 1 mM β-mercaptoethanol, and 10% glycerol. Following the wash step, proteins were eluted with a 200 mL linear gradient of 0 to 500 mM NaCl at a rate of 2 mL/min in the same buffer.

Construction and Expression of VP39 Variants. VP39 variants were expressed with a C-terminal six-His tag using a modified version of the vector pET28-a (Novagen). The 31 C-terminal amino acids of VP39 which correspond to a flexible "tail" of the protein have been removed and replaced with the sequence SRGSCGLEHHHHHH. Site-directed mutagenesis was performed using the Stratagene QuickChange kit, and all variant genes were verified by sequencing. Variants were expressed in *E. coli* strain BL21(DE3) by inducing a saturated culture (~16 h after inoculation with 1% of a saturated culture) with 0.2 mM IPTG. Cell pellets were lysed with a French press in a buffer containing 10 mM Tris (pH 8.0), 100 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.1% IGEPAL detergent, and 10% glycerol. Lysed cells were cleared by centrifugation and purified by Ni affinity columns followed by heparin affinity columns. The average yield was 25 mg of pure protein per liter of culture.

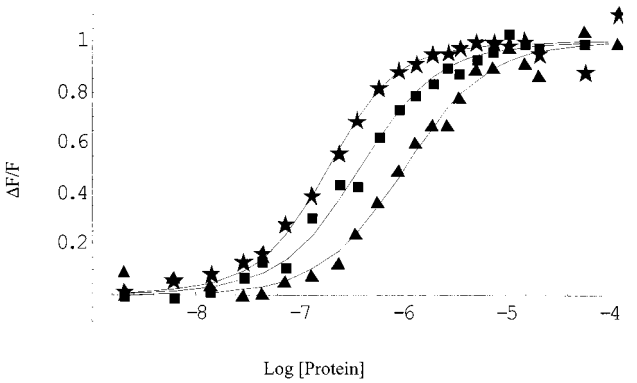


FIGURE 1: Plot of the ratio of the fluorescence change ($\Delta F/F$) of eIF4E proteins to the log value of protein concentration (log C_{ligand}) during titration. Representative binding data of the variant eIF4Es, including wild type (■), a mutant (★) with a higher m7G binding affinity (W56Y), and a mutant (▲) with a lower m^7G binding affinity (W56F).

Circular Dichroism Measurements. CD spectra for all protein variants were collected at 25 °C on an Aviv CD spectrometer. Each protein was dialyzed into a buffer consisting of 10 mM sodium phosphate (pH 7.4) and 100 mM NaCl. Spectra were taken with protein concentrations near 1 μM in a 1 mm cuvette.

Fluorescence Measurements. Fluorescence measurements were taken at 25 °C on the ISS PC1 photon-counting spectrofluorometer equipped with polarizers. eIF4E binding assays were performed in a buffer of 20 mM HEPES (pH 7.5), 150 mM KCl, and 1 mM DTT. VP39 assays were performed in a buffer of 20 mM HEPES (pH 7.0), 20 mM NaCl, and 1 mM DTT. Binding to m^7G was assessed in two different assays depending on the experiment. With one method, we observed changes in the intrinsic fluorescence of the proteins with the addition of m^7GDP. With an excitation wavelength of 290 nm, the fluorescence of a fixed amount of protein was measured at several m^7GDP concentrations. The fluorescence of m^7GDP was measured in a separate experiment with no protein, and the results were subtracted from the protein−m^7GDP measurements to obtain a binding curve (see Figure 1). In a second assay, the changes in the fluorescence intensity and polarization of anthraniloyl-labeled m^7GDP (ant-m^7GDP) were monitored by excitation at 332 nm while varying the protein concentration. Ant-m^7-GDP synthesis was performed as described previously (*16*). The fluorescence intensity and anisotropy of the ant-m^7GDP yielded two independent measurements of protein binding. These values were then both utilized to calculate a binding curve.

Data Analyses. Binding curves from both assays were fit by least-squares nonlinear regression to a simple binding curve quadratic equation using the program Mathematica (Wolfram Research). Mathematica scripts and details of the calculations are available on the Internet (*17*).

In a number of cases, the measurement of very weak binding constants was limited by the experimental conditions. The assay based on the intrinsic fluorescence of the protein was limited by the overlapping fluorescence of the m^7GDP. The assay based on the ant-m^7GDP probe was limited by the solubility of the protein. In the cases where these effects became limiting, an estimated lower boundary on the dissociation constant is reported.

RESULTS AND DISCUSSION

A set of three amino acid side chains that are hypothesized to have similar roles in binding mRNA caps in eIF4E and VP39 was the focus of this study. The cap-binding pocket in eIF4E contains two tryptophans stacked above and beneath the m^7G base and a glutamic acid that forms hydrogen bonds with one edge of the base (Figure 2A). VP39 contains a tyrosine and a phenylalanine that are stacked with the m^7G base and two acidic amino acids, Asp and Glu, that hydrogen bond with the edge of m^7G. Although the m^7GDP binding sites in eIF4E and VP39 are similar (Figure 2A), there remain unique differences in the orientation of the aromatic and acidic side chains. The shared features of the eIF4E and VP39 m^7G-binding site configurations include two aromatic amino acids stacking with different sides of the m^7G ring (face A or face B) and acidic residues interacting with the edge of the base by hydrogen bonding (Figure 2B). Attention was focused on Glu233 of VP39, instead of Asp182, in these studies because it is most similar to Glu103 of eIF4E in the m^7G-binding site geometry.

Requirements for the Acidic Residue. The series of amino acids substitutions that were introduced in eIF4E and VP39 are illustrated schematically in Figure 2C. The first set of mutations tested the requirements for a hydrogen-bonding glutamate residue in each protein. Glu103 in eIF4E and Glu233 in VP39 were replaced with aspartate or glutamine. The aspartate substitution retained the negative charge of glutamate. However, the geometry of the cap-binding site in each protein suggested that an aspartate might not be able to adopt the hydrogen bond with m^7G as glutamate does in the native protein. The complementary substitution of glutamate with glutamine was performed because the hydrogen bonding characteristics of the native glutamate are maintained while removing the negative charge. Analysis of these mutants of eIF4E and VP39 showed that either substitution of the native glutamate ablated m^7GDP binding (Figure 2C). CD analyses of the variant proteins indicate that the observed defects are not a result of misfolding or a large conformational change, but are due to the subtle changes in the local environment of the m^7G binding sites (Figure 3).

These results provide evidence that there is an absolute requirement for a glutamate residue hydrogen bonding to N^1 of the guanine base in the cap-binding site of both eIF4E and VP39. The loss of cap binding with the substitution of this glutamate by the isostructural glutamine indicates that the negative charge of the glutamate provides a significant energetic contribution to m^7G binding. The observation that the aspartate substitution at these positions ablated cap binding suggests that the hydrogen bonds are as important as the negative charge of the acidic side chain. The relative effect of these mutations on eIF4E was more severe than in VP39. The effect of removing the glutamate residue in VP39 is most likely mitigated by the additional hydrogen bonding provided by Asp182.

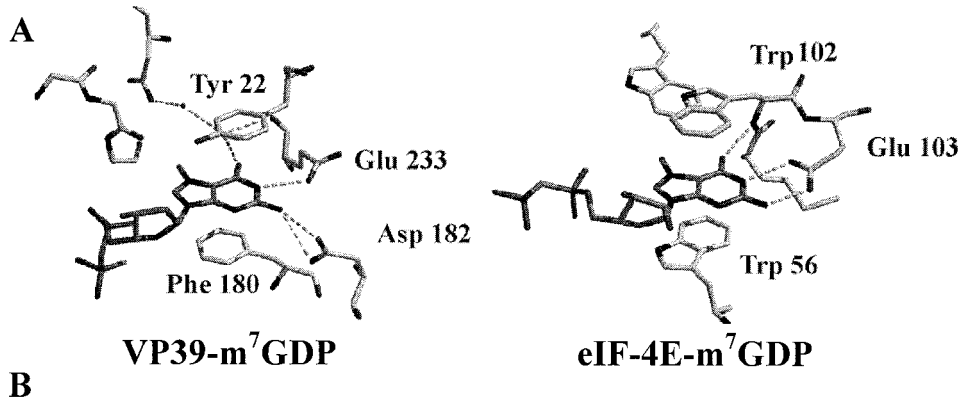
Requirements for Aromatic Stacking Residues. The next set of experiments was performed to test whether the three aromatic amino acids (Trp, Tyr, and Phe) were equivalent in their ability to promote high-affinity stacking interactions with m^7GDP. Aromatic residues of eIF4E and VP39 that stack with the m^7G base were systematically substituted with all three aromatic amino acids, and the effects of these substitutions on m^7GDP affinity were assessed. When assayed through CD analysis, all variant proteins have a signature identical to that of the wild-type proteins, suggesting that all variants fold into the correct native structure (Figure 3). The cap-binding affinity of eIF4E was very sensitive to substitutions at position 102. Substitution of the native Trp102 of eIF4E with Tyr or Phe reduced the m^7GDP binding affinity to an unmeasurable level (data not shown). In contrast, when Trp56 of eIF4E, which stacks with the "B" face of the m^7G base, was changed to Tyr, the m^7G-binding affinity increased slightly (Figures 1 and 2C). However, when Trp56 was replaced with Phe, the m^7G-binding affinity decreased to $^1/_3$ of that of wild-type eIF4E. The effect of substituting Tyr22 of VP39, which stacks on the "A" face of m^7G, with Trp and Phe was also determined. The Trp22-substituted VP39 retained m^7GDP binding affinity. However, substituting the native Tyr22 with the nearly isostructural Phe dramatically reduced the m^7G binding affinity (Figure 2C). Although the effect of Phe substitution in eIF4E was less dramatic than that observed in VP39, these results indicate a pattern where both Trp and Tyr form nearly equivalent interactions with the m^7GDP base and Phe is not capable of mimicking this interaction. Thus, m^7G binding is much weaker in both Phe-substituted proteins.

To test this pattern more thoroughly, the same series of substitutions was performed in a variant of VP39 where Phe180, which stacks on the "B" face of m^7G, was replaced with Trp creating a "4E-like" m^7G binding site (Figure 2A,C). Tyr22 on the A face of the m^7G base was then substituted with Trp or Phe. Like that for eIF4E, a pattern of m^7GDP binding affinities was observed where Trp and Tyr retained m^7G binding, whereas Phe dramatically reduced the affinity for m^7GDP (Figure 2C). Unfortunately, we were unable to generate the corollary "VP39-like" eIF4E variant due to the sensitivity of eIF4E residue 102 to amino acid substitutions.

Besides the mutations introducing alternative aromatic residues at these stacking positions, other mutant proteins with nonaromatic substitutions of Tyr22 in VP39 and Trp56 and Trp102 in eIF4E were prepared and analyzed. These mutations included Tyr22 to Leu, Arg, or Gln for VP39 and either Trp56 to His and Ala or Trp102 to His, Leu, and Ala for eIF4E. All of these substitutions ablated the ability of either protein to bind m^7GDP while maintaining the correct native fold of the protein (see the legend of Figure 3). The m^7GDP dissociation constants for these mutant proteins were at least 50 times higher than those of their respective wild types (data not shown).

In our examination of the requirements for the residues that stack on the faces of the m^7G base, a striking pattern emerged from the mutant proteins that retained the native function. The most noteworthy observation was that both

A



B

Schematic representation of m^7G-binding site configuration		
[Tyr 22]	Aromatic stacked on 'A' face	[Trp 102]
m^7G [Glu 233]	Hydrogen bonding acidic residue	m^7G [Glu 103]
[Phe 180]	Aromatic stacked on 'B' face	[Trp 56]
VP39		**IF-4E**

C

Protein Model	Active Site Configuration	Amino Acid Substitution at position [?] / m^7G affinity (μM)			Assay Method
IF-4E (Glu103)	**[Trp]** m^7G [?] [Trp]	Glu (WT)	Gln	Asp	
		0.3 ± 0.1	>20 ± 10	16 ± 7	Trp
		0.7 ± 0.3	NM	NM	Ant-m^7GDP
VP39 (Glu233)	**[Tyr]** m^7G [?] [Phe]	Glu (WT)	Gln	Asp	
		7 ± 2	18 ± 8	30 ± 13	Ant-m^7GDP
IF-4E (Trp56)	**[Trp]** m^7G [Glu] [?]	Trp (WT)	Tyr	Phe	
		0.3 ± 0.1	0.1 ± 0.1	1.0 ± 0.2	Trp
		0.7 ± 0.3	0.4 ± 0.3	1.0 ± 0.4	Ant-m^7GDP
VP39 (Tyr22)	[?] m^7G [Glu] [Phe]	Trp	Tyr (WT)	Phe	
		4 ± 2	7 ± 2	30 ± 12	Ant-m^7GDP
VP39 '4E like' (Tyr22)	[?] m^7G [Glu] [Trp]	Trp	Tyr	Phe	
		1 ± 0.3	3 ± 0.5	>20 ± 10	Trp

FIGURE 2: (A) Crystallographic structures of the m^7G-binding pocket in the VP39−m^7GDP and eIF4E−m^7GDP complexes. In the VP39−m^7GDP complex, m^7G was bound through stacking interactions with two aromatic residues, Tyr22 and Phe180, and hydrogen bonding with acidic residues, Glu233 and Asp182. In the eIF4E−m^7GDP complex, m^7G was bound by stacking interactions with two tryptophans, Trp102 and Trp56, and hydrogen bonding with acidic residue Glu103. (B) Schematic model of the m^7G-binding pockets of VP39 and eIF4E. Viewing from the edge of the m^7G-base ring, illustrated as the shadowed box, two aromatic residues stacked above and beneath the ring facing different sides of the base. In addition, the acidic residue that interacts with the m^7G by hydrogen bonding was placed to the right side of m^7G in the model to present its relative position in the binding pocket. (C) m^7GDP binding affinity of the protein variants measured by fluorescence spectroscopy, either by the tryptophan titration assay (Trp) or by the anthraniloyl-probe inhibition assay (ant-m^7GDP). Affinities are reported in micromolar with approximate error. NM indicates that no measurable binding was observed in the assay. The residue that was mutated is shown as a question mark in the binding site configuration. Wild-type residues are indicated by WT. Three sets of data are presented in this figure to show the ability of key mutants to bind m^7GDP. The data are not shown for mutations that resulted in a very low m^7G binding affinity.

tryptophan and tyrosine were nearly equivalent in their ability to support m^7GDP binding, whereas phenylalanine was unable to fulfill this role. These data support the hypothesis that there is a specific interaction that forms between the m^7G base and the aromatic side chains of either tryptophan or tyrosine. This interaction either is not supported by a phenylalanine side chain or is much weaker between this amino acid and m^7G.

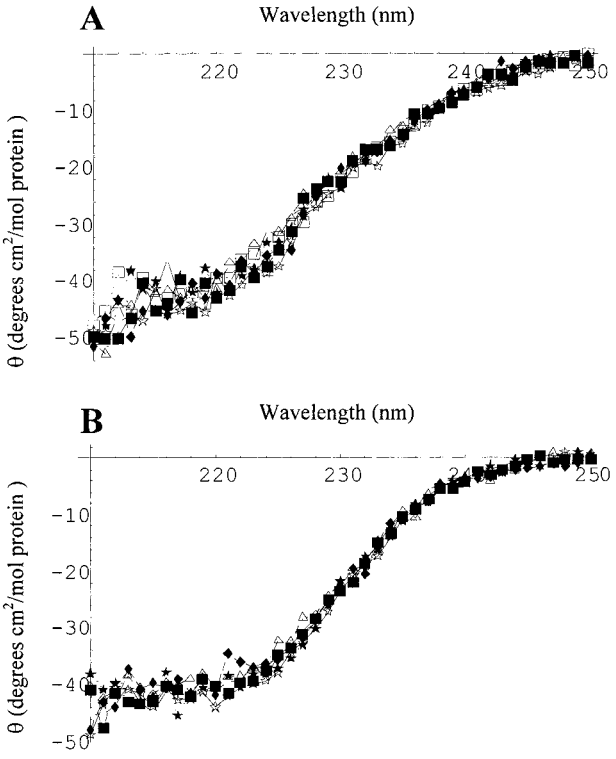


FIGURE 3: CD spectra for variants of (A) eIF4E and (B) VP39. Spectra were obtained for all variant proteins mentioned in the text. All spectra were essentially identical to that of the wild-type proteins. For clarity, only the spectra for key variants are shown above. In panel A, the spectra for wild-type eIF4E (▲) and the E103D (★), E103Q (■), W102Y (◆), W102F (△), W56Y (☆), and W56F (□) variants are shown. In panel B, the spectra for wild-type VP39 (▲) and the E233D (★), E233Q (■), Y22F (◆), F180W (△), and F180W/Y22F (☆) variants are shown.

Charge-Transfer Complexes and the Analogy to Flavins. The pattern of relative interactions between aromatic amino acids and m^7GDP is strikingly similar to that observed for amino acids in complex with flavin molecules (*18*). The face-to-face stacking of aromatic amino acids with flavins has been extensively examined both in small molecule model systems and in flavin-bound proteins (*19*−*21*). A variety of spectroscopic, chemical, and crystallographic data suggest that flavins interact preferentially with tryptophan and tyrosine with little or no interaction with the other aromatic amino acids phenylalanine and histidine (*21*−*23*).

One of the more remarkable phenomena that coincides with such stacking complexes is the formation of a charge-transfer complex. This term was coined to describe a new absorption band in complexes of tryptophan and riboflavin that does not exist in either compound in isolation (*24*). The absorption band was hypothesized to originate from the donation of an electron from the highest-energy occupied molecular orbital (HOMO) of the tryptophan indole ring to the lowest-energy unoccupied molecular orbital (LUMO) of the riboflavin. Although the spectroscopic observation of a charge-transfer band often coincides with the formation of a tight face-to-face stacking complex, there is little evidence that this electronic phenomenon contributes significantly to the binding energy of the complex. The binding energy of the stacking complex is most likely due to a combination of hydrophobic, electrostatic, and van der Waals interactions between the flat ring systems (*18*, *25*). These collective interactions are significantly weaker for phenylalanine and histidine side chains than for tryptophan and tyrosine, which is consistent with the observations reported here with mutants of eIF4E and VP39.

The interaction between 7-methylguanosine and aromatic amino acids has also been studied in small molecule systems through both spectroscopic and X-ray crystallographic techniques (*12*, *26*). So far, these studies indicate that tryptophan is uniquely suited to complex with the m^7G base. Although there is a small molecule crystal structure of a phenylalanine−m^7GMP complex (*27*, *28*), a structural study of the interaction of tyrosine with the m^7G base has not been reported. A tryptophanyl peptide−m^7GMP charge-transfer complex has been reported (*12*) which resembles the model of flavin molecules (*21*). This charge transfer has been often cited as the primary mechanism by which a charged nucleic base, such as m^7G, is specifically recognized by proteins (*8*, *12*). The LUMO energy for an m^7G base is significantly lower than that of the unmethylated guanine as a result of the positive charge conferred to the base by the methylation at N^7 (*26*). Thus, it has been suggested that the LUMO energy of the base facilitates a specific donation of an electron from the electron-rich tryptophan indole to the π-ring system of the charged base (*12*). This is an attractive hypothesis, and there is a large body of evidence for a strong interaction between tryptophan and the m^7G base (*12*, *29*−*31*). However, there is no direct evidence at this time for the formation of a charge-transfer complex in these proteins.

On the basis of the knowledge of interactions between flavins and aromatic amino acids, the tight stacking interactions of Trp or Tyr with an m^7G base are most likely due to a combination of hydrophobic, electrostatic, and van der Waals forces between the flat aromatic rings. The energy in these interactions is not necessarily due to the actual electron donation from the aromatic side chain to the m^7G base. The delocalized positive charge of the m^7G base does provide a unique molecular "handle" for such an interaction. The electron-rich aromatic amino acid side chains may specifically interact with the delocalized positive charge through simple Coulombic interactions similar to the aromatic−cation interactions that have recently gained some attention (*32*). The π-ring system of tyrosine is more electron-rich than that of phenylalanine due to the electron-donating character of the phenolic oxygen. The relative difference in the m^7G interaction between tyrosine and phenylalanine could be due to this difference in electron localization coupled with a short-range Coulombic interaction between the π-electron cloud and the positive charge of the m^7G base. Thus, it is striking that Trp and Tyr residues appear to engage in favorable stacking interactions with both flavins and the m^7G base in a similar manner.

Complex Orientation and Position Requirements for High-Affinity Stacking. Although the pattern of Trp and Tyr interactions with the m^7G base was observed in both eIF4E and VP39, each model system still maintains some unidentified terms that define the general requirements of an m^7G binding site. The requirement for Trp and Tyr was observed in at least one of the two residues stacking with the m^7G base in each protein. By comparing the similar sites in the two proteins, position 56 in eIF4E and position 22 in VP39, we found that these two amino acid positions are on opposite faces of the m^7G base (the A face in VP39 and the B face

in eIF4E). In addition, the relative orientation of the amino acid side chains in the two proteins is very different and the residue stacked on the opposing face in each protein does not show the same amino acid requirements (i.e., only Trp or Tyr) (Figure 2A). Position 102 in eIF4E had the most stringent requirement for tryptophan, whereas position 180 in VP39 apparently had less stringent requirements and accepts phenylalanine as a stacking residue in the wild-type protein. The stringent requirement for a Trp residue at position 102 in eIF4E is probably due to the distance between the α-carbon of Trp102 and the m^7G base. The indole of the side chain only interacts with the m^7G base with its distal edge. Thus, neither Tyr nor Phe residues are large enough to maintain a significant interaction area with the m^7G base when they replace Trp102 (Figure 2A).

The lax requirement at position 180 in VP39 is an unanswered riddle. Phenylalanine exists at this position in wild-type VP39, and m^7G binding is maintained, although with an affinity that is about 10-fold weaker than that of eIF4E. This observation may suggest that of the two residues stacking with the m^7G base, at least one must provide some van der Waals contact with the base, and the other must be either a Trp or a Tyr residue engaging in a specific stacking interaction with the charged base. With this hypothesis, one might view the off-center stacking of Trp102 in eIF4E as being unable to promote a specific stacking interaction and thus establishing a requirement for either Trp or Tyr in the stacking partner at position 56. The primary fault with this hypothesis is found in our attempts to engineer an eIF4E-like m^7G binding site into VP39. Substituting Phe180 with a Trp residue should position the indole of the Trp with a large area of overlap with the m^7G base. If the stated hypothesis is correct, substituting Trp at position 180 in VP39 should relieve the requirement for Trp and Tyr at partner position 22. However, when this was tested, m^7G binding was only maintained when residue 22 was either a Trp or a Tyr and binding was ablated when phenylalanine was substituted at position 22.

There are two possible explanations for these somewhat contradictory observations. The first possibility is that Trp, Tyr, and Phe residues all support a specific interaction with m^7G. Subtle conformational changes or misfolding of mutant proteins might be the cause of the differential binding between tyrosine and phenylalanine substitutions. However, this seems unlikely because phenylalanine and tyrosine are nearly isostructural, the atomic structures of the proteins show no specific interactions where the tyrosine hydroxyl is required, and the same observation was made in two nonhomologous proteins.

The more likely explanation is that there are specific position and orientation requirements in the establishment of a high-affinity stacking interaction between m^7G and Trp or Tyr. With this assumption, we can return to the hypothesis that one of the stacking partners in an m^7G binding pocket must be a Trp or a Tyr residue and form an orientation-dependent high-affinity stacking interaction with the m^7G base. The second stacking residue in the site may simply provide a flat, complementary surface for lower-affinity van der Waals interactions. On the basis of this hypothesis, we can then explain our observations. In eIF4E, Trp102 is not in the proper orientation for a high-affinity stacking interaction and only a tryptophan at this position can provide the necessary van der Waals interactions to act as the low-affinity stacking partner. Position 56 of eIF4E is situated such that either a tyrosine or a tryptophan can engage in a high-affinity stacking interaction with the m^7G base. In VP39, residue 180 is not positioned in a manner that allows a Trp residue at this location to form a high-affinity stacking interaction with the m^7G base, but Trp and Phe at this position can provide the low-affinity van der Waals interactions necessary to support m^7G binding. Then residue 22 in VP39 must be properly positioned to allow either Trp or Tyr to form a high-affinity stacking interaction with the m^7G base. The more difficult aspect of this hypothesis to understand is that position 56 in eIF4E and position 22 in VP39 exist in very different orientations relative to m^7G, yet both support the proposed "high-affinity" stacking interaction for Tyr and Trp residues. The best approach to address this question would be a comparison of the three-dimensional structures of m^7G-specific binding sites in other nonhomologous cap-binding proteins yet to be discovered.

Evolutionary Correlation. An evolutionary correlation to our interpretation exists in a recently reported homologue of eIF4E. All the known orthologues of eIF4E in species as diverse as yeast and humans absolutely conserve all eight tryptophans in the protein sequence (*8*). However, a recently isolated mammalian eIF4E-related cap-binding protein 4EHP retains a cap binding affinity similar to that of human eIF4E (*33*). Remarkably, 4EHP has a single Trp to Tyr substitution at residue 78 that corresponds to Trp56 in human eIF4E. In our experiments, replacing Trp56 in eIF4E with tyrosine resulted in slightly stronger binding to m^7G. Thus, the equivalency of Trp and Tyr at this position with regard to cap binding affinity is supported by this natural orthologue.

In conclusion, we provide experimental evidence for a specific high-affinity stacking interaction between an m^7G base and either Trp or Tyr in both eIF4E and VP39. This interaction is absolutely necessary and must be present in at least one of the aromatic stacking residues in the cap-binding site of both proteins. The interaction appears to have specific requirements for the relative orientation of the m^7G base and the aromatic side chains; however, more than one orientation can support the interaction. As our results indicate for VP39 and eIF4E, both Trp and Tyr appear to be nearly equivalent in their affinity for m^7G at these unique positions. A better understanding of the specific m^7G recognition motifs may provide a foundation for the design of small molecules that either specifically bind to mRNA caps to modulate translation or block the binding of mRNA to specifically targeted viral or cellular cap binding proteins (*34*).

ACKNOWLEDGMENT

We thank Dr. J. Wylie Nichols for advice and the use of his fluorimeter in the initial stages of this work. We also thank Drs. Kim Richie and Anita Corbett for reading the manuscript.

REFERENCES

1. Wells, S. E., Hillner, P. E., Vale, R. D., and Sachs, A. B. (1998) *Mol. Cell 2*, 135−140.
2. Drummond, D. R., Armstrong, J., and Colman, A. (1985) *Nucleic Acids Res. 13*, 7375−7394.
3. Bechler, K. (1997) *Biochem. Biophys. Res. Commun. 241*, 193−199.

4. Varani, G. (1997) *Structure 5*, 855−858.
5. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997) *Genes Dev. 11*, 3306−3318.
6. Lewis, J. D., and Izaurralde, E. (1997) *Eur. J. Biochem. 247*, 461−469.
7. Izaurralde, E., and Mattaj, I. W. (1995) *Cell 81*, 153−159.
8. Marcotrigiano, J., Gingras, A. C., Sonenberg, N., and Burley, S. K. (1997) *Cell 89*, 951−961.
9. Matsuo, H., Li, H., McGuire, A. M., Fletcher, C. M., Gingras, A. C., Sonenberg, N., and Wagner, G. (1997) *Nat. Struct. Biol. 4*, 717−724.
10. Hodel, A. E., Gershon, P. D., Shi, X., Wang, S. M., and Quiocho, F. A. (1997) *Nat. Struct. Biol. 4*, 350−354.
11. Hodel, A. E., Gershon, P. D., and Quiocho, F. A. (1998) *Mol. Cell 1*, 443−447.
12. Ishida, T., Doi, M., Ueda, H., Inoue, M., and Scheldrick, G. M. (1988) *J. Am. Chem. Soc. 110*, 2286−2296.
13. Stolarski, R., Sitek, A., Stepinski, J., Jankowska, M., Oksman, P., Temeriusz, A., Darzynkiewicz, E., Lonnberg, H., and Shugar, D. (1996) *Biochim. Biophys. Acta 1293*, 97−105.
14. Altmann, M., Edery, I., Trachsel, H., and Sonenberg, N. (1988) *J. Biol. Chem. 263*, 17229−17232.
15. Morino, S., Hazama, H., Ozaki, M., Teraoka, Y., Shibata, S., Doi, M., Ueda, H., Ishida, T., and Uesugi, S. (1996) *Eur. J. Biochem. 239*, 597−601.
16. Ren, J., and Goss, D. J. (1996) *Nucleic Acids Res. 24*, 3629−3634.
17. http://www.biochem.emory.edu/hodel/research/capbinding99/hsuetal.html.
18. McCormick, D. B. (1977) *Photochem. Photobiol. 26*, 169−182.
19. MacKenzie, R. E., Fory, W., and McCormick, D. B. (1969) *Biochemistry 8*, 1839−1844.
20. Fory, W., MacKenzie, R. E., Wu, F. Y., and McCormick, D. B. (1970) *Biochemistry 9*, 515−525.
21. Inoue, M., Shibata, M., Kondo, Y., and Ishida, T. (1981) *Biochemistry 20*, 2936−2945.
22. Wu, F. Y.-H., and McCormick, D. B. (1971) *Biochim. Biophys. Acta 229*, 440−443.
23. Johnson, P. G., and McCormick, D. B. (1973) *Biochemistry 12*, 3359−3364.
24. Wright, L. D., and McCormick, D. B. (1964) *Experientia 20*, 501−502.
25. Helene, C., and Lancelot, G. (1982) *Prog. Biophys. Mol. Biol. 39*, 1−68.
26. Nishimura, Y., Takahashi, S., Yamamoto, T., Tsuboi, M., Hattori, M., Miura, K., Yamaguchi, K., Ohtani, S., and Hata, T. (1980) *Nucleic Acids Res. 8*, 1107−1119.
27. Ishida, T., Kamiichi, K., Kuwahara, A., Doi, M., and Inoue, M. (1986) *Biochem. Biophys. Res. Commun. 136*, 294−299.
28. Ishida, T., Doi, M., and Inoue, M. (1988) *Nucleic Acids Res. 16*, 6175−6190.
29. Ueda, H., Iyo, H., Doi, M., Inoue, M., Ishida, T., Morioka, H., Tanaka, T., Nishikawa, S., and Uesugi, S. (1991) *FEBS Lett. 280*, 207−210.
30. Ueda, H., Iyo, H., Doi, M., Inoue, M., and Ishida, T. (1991) *Biochim. Biophys. Acta 1075*, 181−186.
31. Ishida, T., Katsuta, M., Inoue, M., Yamagata, Y., and Tomita, K. (1983) *Biochem. Biophys. Res. Commun. 115*, 849−854.
32. Dougherty, D. A. (1996) *Science 271*, 163−168.
33. Rom, E., Kim, H. C., Gingras, A. C., Marcotrigiano, J., Favre, D., Olsen, H., Burley, S. K., and Sonenberg, N. (1998) *J. Biol. Chem. 273*, 13104−13109.
34. Baker, B. F., Miraglia, L., and Hagedorn, C. H. (1992) *J. Biol. Chem. 267*, 11495−11499.

BI000623P