The "Cap-Binding Slot" of an mRNA Cap-Binding Protein: Quantitative Effects of Aromatic Side Chain Choice in the Double-Stacking Sandwich with Cap[†,‡]

Guanghui Hu,[§,||] Akihiro Oguro,[⊥,#] Changzheng Li,[⊥] Paul D. Gershon,*[,⊥] and Florante A. Quiocho*[,§,@]

Graduate Program in Structural and Computational Biology and Molecular Biophysics, Howard Hughes Medical Institute, and Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas 77030, and Institute of Biosciences and Technology, Texas A&M University System Health Science Center, Houston, Texas 77030

Received March 11, 2002

ABSTRACT: The N7-methylguanine portion of the mRNA cap structure interacts with cap-binding proteins via an unusual double-stacking arrangement in which the positively charged cap is sandwiched between two parallel-oriented aromatic protein side chains. Three-dimensional costructures of cap with two mRNA cap-binding proteins, namely, translational initiation factor eIF4E and VP39 (the vaccinia virus-encoded mRNA cap-specific 2′-*O*-methyltransferase), have heretofore been reported. Despite striking similarities between the two proteins in the double stack with the cap, the stack differs most notably in the species of stacked side chain donated by the protein. Whereas eIF4E employs two tryptophans, VP39 uses a tyrosine and a phenylalanine. Here, we have generated tryptophan substitutions in VP39. Tryptophan substitution was shown, crystallographically, not to disrupt the maintenance of a bona fide parallel stack. However, the single-tryptophan and double-tryptophan substitutions were associated with increased affinity for cap nucleoside by factors of 10 and 50, respectively. VP39 interacted more strongly with a true substrate (containing portions of RNA downstream of the cap in addition to the cap itself) than with isolated cap nucleoside, by several orders of magnitude. VP39 mutants with tryptophan substitution at position 180 exhibited apparent defects in substrate catalytic rate during the first turnover cycle, indicating the possibility of an exquisite sensitivity of the catalytic center to subtle changes in substrate position brought about by alterations in the cap-binding slot. The X-ray structure of VP39 with a genuine nucleobase analogue of N7-methylguanosine, namely, N7,9-dimethylguanine, indicated that the N7-methylguanosine rotational orientation within the stack is a property of the cap nucleobase itself.

An N7-methylated guanosine (m^7G)[1] cap structure is found at the 5′ ends of mature eukaryotic mRNAs, and is central to a number of cellular processes, including translation, mRNA degradation, splicing, and mRNA export. A number of cap-binding proteins have been identified, including eukaryotic translation initiation factor eIF4E (*1*), VP39 [the vaccinia virus-encoded cap-specific mRNA 2′-O-methyltransferase (*2*)], CBP20/CBP80 ["cap-binding complex" (*3*)], and mRNA decapping enzyme (*4*). For two of these, namely, VP39 and eIF4E, cocrystal structures have been reported for the proteins with either cap dinucleotide, capped RNA, or both (*5*−*9*). The two proteins show striking similarity in their respective modes of binding to the m^7Gua moiety of the cap, with the latter being sandwiched between two aromatic side chains of the protein a near-perfect van der Waals distance from each, in a parallel stacked fashion (*9*). In addition, for each of the two proteins, side chains are observed to form hydrogen bonding interactions with ring nitrogens and exocyclic functional groups of the cap heterocycle.

For VP39, interactions have been examined with bona fide capped RNA substrates as well as essential yet insufficient portions thereof such as cap dinucleotide or analogues of the cap-terminal m^7Gua nucleobase (*2*, *5*, *6*, *9*−*12*). A model has been suggested (*10*) whereby VP39 interacts with its capped RNA substrate via four sites, namely, (i) an m^7G-binding "pocket" or "slot", (ii) the methyltransferase catalytic center (which also includes a binding site for the *S*-adenosylmethionine cofactor), (iii) a site for the binding of up to three nucleotides of RNA immediately downstream of the triphosphate bridge of the cap [which has been shown, crystallographically, to form a shallow groove in the protein surface (*6*), to bind RNA entirely via backbone phosphoribose moieties (*6*), and to be sensitive to pH values above neutrality (*10*)], and (iv) a site that interacts with the capped mRNA somewhere beyond its sixth transcribed nucleotide

† This work was supported by National Science Foundation Grant MCB-9604188 to P.D.G. and the Howard Hughes Medical Institute with which F.A.Q. is an Investigator. G.H. was supported by a Predoctoral Fellowship of the Welch Foundation through a grant to F.A.Q.

‡ Atomic coordinates have been deposited with the Protein Data Bank, as entries 1JTE, 1JTF, and 1JSZ.

* To whom correspondence should be addressed. E-mails: pgershon@ibt.tamu.edu and faq@bcm.tmc.edu.

§ Graduate Program in Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine.

|| Current address: Tanox Inc., 10301 Stella Link, Houston, TX 77025.

⊥ Texas A&M University System Health Science Center.

Current address: Department of Tumor Biology, The Institute of Medical Science, University of Tokyo, Tokyo, Japan.

@ Howard Hughes Medical Institute and Department of Biochemistry and Molecular Biology, Baylor College of Medicine.

[1] Abbreviations: AdoHcy, *S*-adenosylhomocysteine; m^7G, N7-methylguanosine; m^7Gua, N7-methylguanine; $m^{7,9}$Gua, N7,9-dimethylguanine; ITC, isothermal titration calorimetry.

10.1021/bi0201926 CCC: $22.00
Published on Web 05/23/2002

[the existence of which was inferred via the downward shift in K_m^{RNA} observed when examining mRNA substrates greater than six nucleotides in length (*10*)]. Weakening or eliminating interaction at individual sites leads to an increased K_m for VP39's RNA substrate (*10*).

VP39's interactions with the nucleobase moiety of the m^7G cap, in isolation, have been examined in detail (*5*, *9*, *11*). Thus, crystallographic and biochemical studies employing a variety of positively charged nucleobases have demonstrated that possession of a positive charge by the nucleobase is essential for maintaining affinity within the stack (*11*). It was shown that the positively charged nucleobases adopt a rotational orientation in the plane of the stack that apparently depends solely upon the interaction of polar moieties of the nucleobase with the hydrogen bondable and charged side chains arrayed around VP39's cap-binding slot (*11*). By substitution of one or both of the stacked aromatics of the cap-binding slot with alanine, it was shown that either of these two side chains alone can maintain specificity for m^7G-capped RNA with respect to G-capped or A-capped RNA (albeit VP39 retains only low levels of enzymatic activity), whereas the combination of two aromatics contributes enormously to affinity (*10*). A "cation$-\pi$" (*13*) interpretation of the cap−protein stacking interaction has been suggested (*9*, *11*) in which the positively charged nucleobase cation interacts electrostatically with the π-clouds or quadrupole moment of the stacked aromatics. The most significant evidence for this would be the requirement for the nucleobase to possess a positive charge for VP39 binding.

Despite striking similarities, a notable difference in character between the cap-binding slots of eIF4E and VP39 is the species of aromatic side chain donated by the protein to the stack. Whereas eIF4E donates a pair of tryptophans (*7*, *8*), VP39 provides a phenylalanine and a tyrosine (*5*). Explanations for this difference might include the following. (i) Some functional group of the aromatic (such as VP39's tyrosine hydroxyl) plays a role in affinity or in contributing to an electrostatic field that "steers" the cap as it enters the slot. (ii) VP39 employs aromatics with a smaller surface area and less intense regions of electrostatic potential than eIF4E's because VP39, being an enzyme, has an inherent turnover requirement. A lower-affinity cap−protein interaction would enable the protein to release the capped RNA at some point during its catalytic cycle. (iii) Evolution of the overall protein fold (which is entirely different for eIF4E and VP39) has mandated a different choice of aromatic to fit the respective folding pathways or permit the aromatic side chains to fit into the respective slots. (iv) The choice of aromatic is functionally neutral, and/or the interconversion of the respective codons presents a serious mutational barrier.

To address the above issues, VP39's cap-binding slot has been converted to one that more closely resembles eIF4E's by substituting VP39's cap-stacking phenylalanine and tyrosine either individually or together with tryptophan. We show that the substituted tryptophans fit into the VP39 slot and maintain a perfectly parallel planar orientation without affecting the overall protein fold. Moreover, the mode of cap binding is unchanged (e.g., the rotational orientation of the bound m^7Gua moiety is unaltered), and the VP39 mutants are catalytically active. Although the VP39 Trp mutants exhibit a 10−50-fold greater affinity than wild-type VP39 for the m^7Gua moiety of the capped RNA substrate, affinity for the overall substrate is altered only slightly. Mutants possessing the Phe180 to Trp substitution appear to be compromised in methyltransferase chemistry and possibly also in RNA substrate turnover.

MATERIALS AND METHODS

Proteins, RNA, Nucleosides, and Nucleotides. Wild-type VP39 comprised VP39-ΔC26 (a fully active truncation mutant). A GST-tagged version of VP39-ΔC26 was used for all experiments except ITC, for which a six-His-tagged version of the protein was used. Point mutations were made as described previously (*11*, *14*). GST-tagged VP39-ΔC26 and mutant fusion proteins were expressed, purified, and detagged as described previously (*11*, *14*). $m^7GpppG(A)_3$ was kindly provided by A. Hodel (Department of Biochemistry, Emory University School of Medicine, Atlanta, GA), and $m^{7,9}Gua$ was kindly provided by B. Lippert (Fachbereich Chemie, Universitaet, Dortmund, Germany) (*11*, *14*). m^7G, m^7Gp, and m^7Gpp were purchased from Sigma-Aldrich, and $m^7G(5')pppG$ was bought from New England BioLabs Inc.

X-ray Structure Analysis. Crystallization and X-ray structure determination for wild-type and mutant VP39 proteins with bound AdoHcy and cap analogues were performed as previously described (*11*, *14*). Since the structures described in this paper are isomorphous with previous structures (*11*, *14*), they were determined by direct phasing using only the protein model with bound AdoHcy.

For structural analysis of F180W VP39 with bound AdoHcy and m^7GpppG, one or two F180W VP39−AdoHcy complex crystals were soaked in a solution containing m^7GpppG. The soaking solution was made by mixing 5 μL of 100 mM m^7GpppG (dissolved in 0.25 M HCl) with 1 μL of 100 mM AdoHcy in 5 mM HCl and 44 μL of 15% PEG 8000 in 0.125 M ammonium sulfate and 0.1 M sodium cacodylate (pH 6.5). The pH of the resulting mixture was around 5.8. Soaking was allowed to continue for 2 days at room temperature. For X-ray data collection, crystals were flash-frozen in a cryoprotectant comprising 15% PEG 8000, 0.125 M ammonium sulfate, 0.1 M sodium cacodylate (pH 5.8), and 25% glycerol at −160 °C using a cryogenic cooler. Data were collected using a Siemens SMART 2K CCD detector with a Göbel mirror. The detector system was mounted on a Rigaku RU200 rotating anode (Cu Kα) operated at 100 mA and 50 kV. The minimum resolution for the data collection was set at 9999 Å. The software package Siemens SAINT was used to reduce and scale the data. Resolution and data merging statistics are shown in Table 1. Data used in the refinement extended from 8 to 2.6 Å. Water molecules were added to the model, and structure was refined through multiple rounds of positional and temperature factor refinement using the program XPLOR (*15*). The bound ligand was fitted to density in the final stages of the refinement. A similar procedure was used in the structure determination for wild-type VP39 soaked in $m^{7,9}Gua$. The final values of R_{factor}, associated R_{free}, and rmsd for the refined structures are given in Table 1.

Calorimetric Measurement. Protein samples for ITC assays were dialyzed extensively against 10% glycerol, 1 mM 2-mercapthoethanol, 200 mM NaCl, and 20 mM sodium cacodylate (pH 5.9). Calorimetric experiments were conducted using the ITC module of a Microcal MCS VP-ITC

Table 1: Statistics of Crystallographic Analysis for Mutant F180W VP39 in the Absence and Presence of m^7GpppG, and for the Wild-Type Protein Complexed with $m^{7,9}$Gua

	F180W	F180W	wild type
ligand	none	m^7GpppG	$m^{7,9}$Gua
space group	*C*2	*C*2	*C*2
resolution (Å)	2.0	2.6	1.93
cell parameters	$a = 84.6$ Å, $b = 67.9$ Å, $c = 80.0$ Å, $\beta = 117.9°$	$a = 84.2$ Å, $b = 67.7$ Å, $c = 79.6$ Å, $\beta = 118.1°$	$a = 84.3$ Å, $b = 67.9$ Å, $c = 79.9$ Å, $\beta = 118.1°$
R_{merge}[a]	0.070	0.107	0.081
completeness (%)	87.5	94.1	73.5
R_{factor}[b]	0.232	0.258	0.21
R_{free}[c]	0.285	0.279	0.252
rmsd for bonds (Å)	0.011	0.018	0.01
rmsd for angles (deg)	1.63	1.96	1.58
rmsd for dihedrals (deg)	24.43	24.46	24.38
rmsd for impropers (deg)	1.45	1.81	1.36

[a] $R_{merge} = \sum|I_i - \langle I\rangle|/\sum|I_i|$, where I_i is the intensity of a reflection and $\langle I\rangle$ is the average intensity of that reflection. [b] $R_{factor} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{calc}|$, where F_{calc} and F_{obs} are the calculated and observed structure factors, respectively. [c] R_{free} is equivalent to R_{factor}, but calculated for a randomly chosen 10% of reflections omitted from the refinement process.

calorimeter (Microcal Inc., Amherst, MA) and the Origin software supplied with the instrument for data acquisition and analysis. To eliminate possible spurious heat effects from buffer mixing, the ligand (m^7G) used in the titration was dissolved in the buffer against which the protein had been dialyzed. Each titration was carried out at 25 °C with precision control being achieved via a computer-controlled thermostat within the calorimeter. Using the Origin software, binding constants were obtained by nonlinear regression of the ITC data using a model that assumes possession of a single binding site by the protein (*16*). The fitting equation, assuming one-site binding, was $(1/V_0)(dQ/dL_0) = {}^1/_2\Delta H\{1 + (1 - r - x)/[x^2 - 2x(1 - r) + (1 + r)^2]^{1/2}\}$, where V_0, Q, L_0, ΔH, r, M_0, and x represent the volume of the reaction cell, the heat absorbed or evolved, the total ligand concentration, the molar heat of binding (enthalpy), $(M_0K)^{-1}$, the total protein concentration, and L_0/M_0, respectively.

Fluorescence Assay for Cap Analogue Binding. Fluorescence measurements with VP39 Trp mutants were conducted using an SLM/Aminco model 4800 spectrofluorimeter (*11*, *17*) with the excitation wavelength set at 282 nm. Ligand titrations were carried out at 20 °C by following the decrease in fluorescence at 330 nm after adding microliter aliquots of a concentrated ligand solution to a 2.5 μM solution of appropriately buffered protein [2 mL; protein buffer typically comprising 0.2 M NaCl and 0.1 M sodium cacodylate (pH 5.9)]. To correct for possible photo-oxidation, the fluorescence decrease was measured, in parallel, for an equivalent protein solution to which aliquots of buffer only were added. For each addition of ligand or buffer, an average of 10 individual measurements was recorded. K_d values were determined by least-squares nonlinear curve fitting to the quadratic expression for a generalized equilibrium equation $Y = [-b + (b^2 - 4L_0/P_0)^{1/2}]/2$, where $b = 1 + K_d/P_0 + L_0/P_0$ (P_0 and L_0 refer to the starting concentrations of the protein and ligand, respectively) and Y refers to the fractional degree of saturation computed from $\Delta F_i/\Delta F_{max}$. ΔF_i is the fluorescence change after the ith addition ($F_i - F_0$), and ΔF_{max} is the maximum total change in going from the completely unsaturated to the completely saturated protein (*17*). As confirmation of the method, K_d values were also obtained by linear curve fitting to Scatchard plots.

BIAcore Assay. This was performed as described previously (*5*, *10*, *18*, *19*). Running/injection buffer comprised 10 mM HEPES-NaOH (pH 7.0) supplemented with NaCl at a concentration of 90, 150, or 200 mM. After injection of ~700−800 RU of $m^7GpppG(A)_{20}U_{biotin}$ or $pppG(A)_{20}U_{biotin}$ to streptavidin-coated flow-cell surfaces, proteins were injected serially in buffer containing NaCl at different concentrations. Kinetic constants did not vary significantly in the range of 90−200 mM NaCl (although, at NaCl concentrations below 90 mM, the level of nonspecific binding was too great, and at NaCl concentrations above 200 mM, all binding was lost). Each series of automated mutant protein injections was preceded and followed by cycles of wild-type VP39 injection (to ensure that the capped RNA substrate had not been degraded and/or lost from the surface during the experiment). Therefore, data for wild-type (wt) VP39 are derived from double the number of individual data points obtained for the mutants. Double-exponential curves were fit to portions of sensorgrams resulting from each binding−unbinding cycle, from which two K_d values were obtained. The major one, which always represented at least 95% of the protein bound, is shown here. A number of parameters have been modified from previous BIAcore measurements (*10*) to improve data collection. In the current experiments, the flow rate has been increased during the dissociation phase from 40 (*10*) to 100 μL/min to eliminate rebinding during dissociation. The NaCl concentration was increased from 60 to 90−200 mM to avoid any possibility of nonspecific binding or rebinding during dissociation. The amount of data from the sensorgram used for curve fitting was also increased.

Methyltransferase Assays. These were performed as described in ref *10*.

Methyltransferase Turnover Assay. The substrate for the assay was generated in 100 μL in vitro transcription reaction mixtures, as described previously (*10*, *20*). Prior to transcription, the following two oligonucleotides were annealed to one another in TE buffer: 5′-CCTAATACGACTCACTATAGA-3′ and 5′-TTTCTATAGTGAGTCGTATTAGG-3′. The nucleoside triphosphates/analogues thereof present during transcription reactions comprised rATP (4 mM), [α-^{33}P]ATP (167 nM), and N^7mG(5′)pppG (1.2 mM). After 3 h at 37 °C, pyrophosphate was pelleted. The supernatant was diluted 23-fold with water, and 0.5 mL of the resulting material was applied to a 4.6 mm × 250 mm reverse-phase C18 HPLC column, eluting with a two-part gradient (1 mL/min): from 0 to 15% B (10 min) and from 15 to 60% B (120 min). A is 0.1 M triethylamine acetate (pH 5.8), and B is 0.1 M triethylamine acetate (pH 5.8) and 25% CH_3CN. Material from the portion of the "forward-capped" peak furthest removed from the "reverse-capped" peak was lyophilized and then dissolved in 25 μL of water. RNA concentrations were quantitated from OD_{260} values.

A 2 μL aliquot of a solution comprising the purified labeled RNA (43 μM), HEPES-NaOH (pH 7.5, 50 mM), DTT (1 mM), and fresh *S*-adenosylmethionine (10 mM) was transferred to an equal volume of deionized formamide (substrate control), and a second aliquot was transferred to a tube containing 2 μL of RNAse T1 (14 units/μL in 3.2 M ammonium sulfate, Roche) and incubated at 55 °C for 5 min,

before this was also transferred to an equal volume of deionized formamide (time zero sample). To initiate the methylation reaction, the remainder of the assay mixture was added to VP39 or mutants thereof (6−14-fold molar excess of substrate over protein) at 30 °C. At various times thereafter, 2 μL aliquots of the reaction mixture were incubated with additional aliquots of RNase T1 and then mixed with formamide, as described above. At the end of the assay, all formamide-containing samples were loaded onto a pre-electrophoresed (90 min) 25% polyacrylamide, TBE−urea gel. After electrophoresis at 35 V/cm for 90−120 min, a storage phosphorgram was made of the wet gel. Characteristic T1-cleaved and uncleaved radio-emitting bands on the gel (representing VP39 substrate and product, respectively) were quantitated; the resulting values were normalized for the numbers of radioactive phosphates, and "turnover proportion" was calculated as [P]/[S] + [P], where [P] represents the counts for the T1-resistant band and [S] represents normalized counts for a T1-cleavage product in the same lane. Turnover proportion values were factored by the molar ratio of substrate to VP39 at the outset of the methylation reaction, and the resulting values were plotted against time. The plotted values represent the numbers of enzymatic turnovers at each given time. These values assume an enzyme specific activity of 100%. Since specific activities were unknown, the plotted numbers can be regarded as minimum values.

RESULTS AND DISCUSSION

Converted Cap-Binding Pocket: Maintenance of the Parallel-Stacking Arrangement with the Cap. The two stacking aromatics of VP39's cap-binding pocket were converted either individually or together from tyrosine or phenylalanine to tryptophan, yielding three mutants (Y22W, F180W, and Y22W/F180W). After protein expression in *Escherichia coli*, only low levels of Y22W and Y22W/F180W were obtained, but levels of F180W were much higher, comparable to those obtained with wild-type VP39. Since relatively large amounts of protein were required for crystallography and microcalorimetry, those studies centered mainly upon the F180W mutant of VP39 and the wild-type enzyme.

The structures of the F180W−AdoHcy binary complex in the absence and presence of bound m^7GpppG dinucleotide were refined against 2.0 and 2.6 Å resolution data, respectively, to good statistics as shown in Table 1. With the exception of the Trp substitution (e.g., Figure 1A), these structures were very similar to those for the wild-type protein crystallized under identical conditions (*5*). Moreover, wild-type and F180W structures showed no distinguishable difference in binding mode for the dinucleotide, with the m^7Gua ring's density being resolved (the N7-methyl group being clearly visible) and the ribose−pppG moiety of the dinucleotide being weaker (Figure 1A). In both wild-type and F180W proteins, the m^7Gua ring of the bound ligand was stacked in an equivalent position, both translationally and rotationally. The stacking position was equivalent to that of the m^7Gua rings of other m^7G-based ligands bound in all VP39 structures determined thus far (Figure 1A) (*5*, *6*).

Our prior crystallographic and biochemical studies showing VP39 binding to methylated nucleobases, nucleosides and nucleotides, demonstrated that a positively charged nucleobase is essential for cap recognition and binding (*9*, *11*). For the m^7Gua moiety of the standard m^7G cap, positive charge arises from the substitution of both nitrogens (N7 and N9) in the five-membered portion of m^7Gua (N7 due to methylation and N9 due to the attached ribose). Since an N=C double bond must also be accommodated within the N7−C8−N9 portion of m^7Gua, the resulting resonance structure is consistent with a partial quaternization of N7 and N9 and a delocalization of the resulting net positive charge within m^7Gua's N7−C8−N9 region (i.e., within the five-membered ring). Although positive charge trapped within this region presumably has some limited electron-withdrawing effect on m^7Gua's six-membered ring (since 7-methylated Gua has a lower p*K* at N1 than unmethylated Gua), maintenance of the resonance structure dictates that the charge not be delocalized significantly outside of the N7−C8−N9 region.

Views perpendicular to the stack for wild-type and mutant VP39, and for eIF4E, are shown in panels B−E of Figure 1. Despite the striking similarities between wild-type VP39 (Figure 1B) and the F180W mutant (Figure 1C), there is a clear shift in the position of the aromatic at position 180 with respect to the m^7Gua heterocycle. While the overall area of overlap remained similar for the two structures, a shift in overlap position for the F180W mutant left the stacked aromatic side chain covering significantly less of the six-membered portion of m^7Gua and significantly more of the five-membered portion. As a result, with respect to the F180 side chain in wild-type VP39, the W180 side chain in the mutant lies closer to m^7Gua's positively charged N7−C8−N9 region and particularly close to methylated N7.

Although mutants containing the Y22W substitution could not be expressed for crystallization (above), the Y22W substitution was modeled upon the F180W mutant crystal structure (Figure 1D). If it is assumed that the model is accurate, the effect of Y22W substitution appears to be the converse of that for F180W. While the region of overlap (i.e., with the six-membered ring of m^7Gua) is unchanged, the area of overlap is increased. Comparison of the VP39 Y22W/F180W model with the eIF4E structure (panels D and E of Figure 1, respectively) indicates no conservation in the portions of the trio of stacked components that mutually overlap. Thus, despite the two different cap-binding proteins donating identical aromatic side chains to their double stacks with m^7Gua, there is no preferred or default stacking arrangement.

Converted Cap-Binding Pocket: Changes in Solution Binding Constants for m^7G. The differences between wild-type VP39 and the F180W mutant in their geometry of interaction with m^7Gua (above) led us to investigate the possibility of associated quantitative differences, using ITC and fluorescence methods to assess the binding of m^7G. Raw data from an ITC titration of wild-type VP39 with m^7G at 25 °C, along with the corresponding binding isotherm (integration with respect to time with appropriate molar correction), are shown in panels A and B of Figure 2, respectively. Data from an equivalent titration against the F180W mutant are shown in panels C and D of Figure 2. Best fits to the isotherms of panels B and D of Figure 2 provided K_d ($1/K_a$) values for the interaction of m^7G with wild-type VP39 and the F180W protein of 86.2 and 10.9

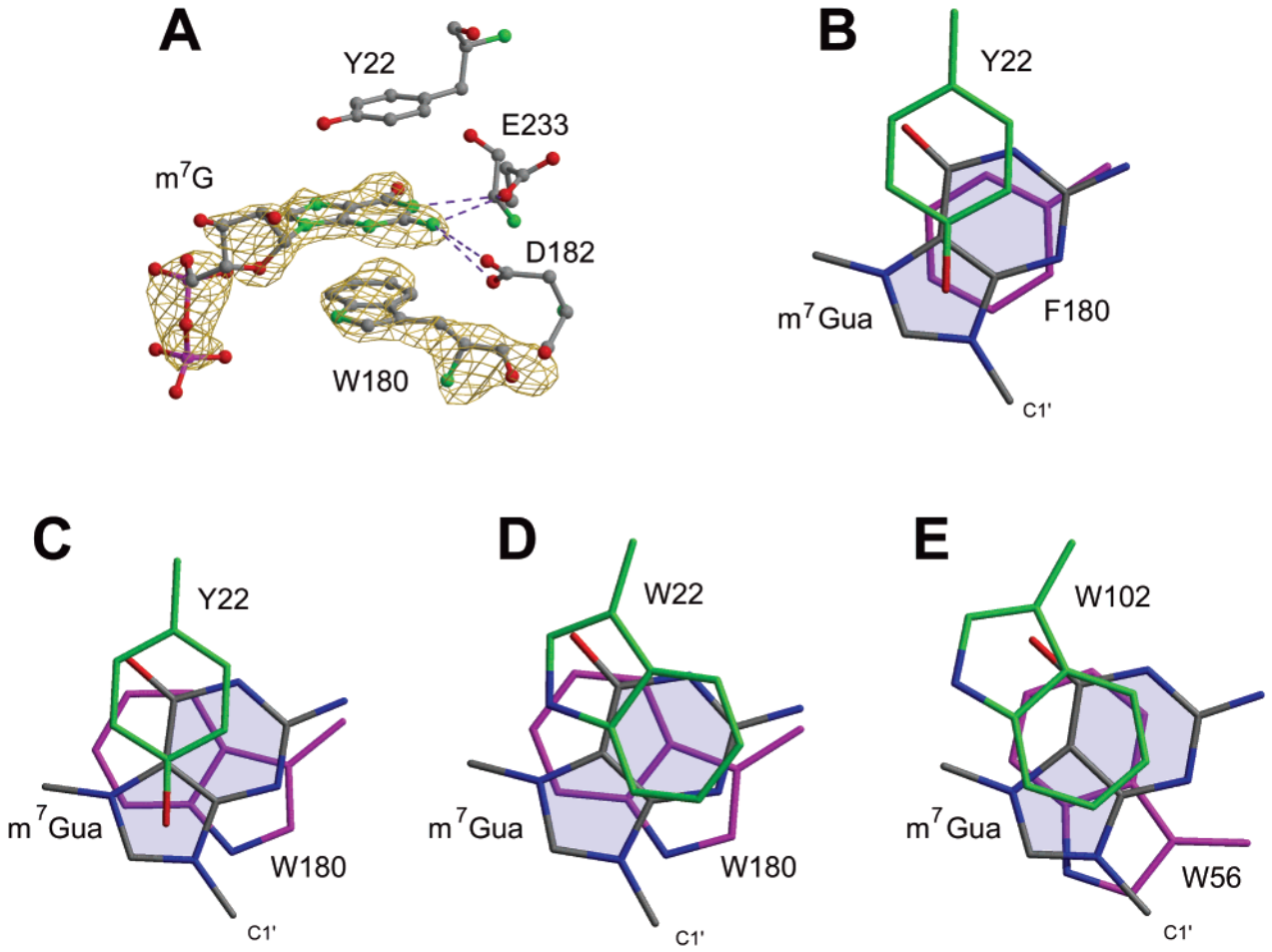

FIGURE 1: Cap-binding site of VP39 and eIF4E. (A) $F_o - F_c$ omit map for the cap-binding slot of mutant F180W after soaking in 10 mM m^{7}GpppG, contoured at the 3σ level. The map was calculated using a model that excluded the ligand and Trp180. The difference density is well-resolved for the m^{7}Gua ring but is incomplete for the attached ribose−phosphate−phosphate moiety. No electron density was visible for the terminal phosphate−guanosine moiety. Dashed lines represent hydrogen bonds. Panel A was generated using BOBSCRIPT (*33*) and RASTER3D (*34*). (B−E) Views down the m^{7}Gua-bound double-stacking sandwiches of cap-binding proteins. The atoms of the model for the m^{7}Gua portion (shaded) of the bound ligand are colored as follows: gray for carbon, red for oxygen, and blue for nitrogen. The carbon atoms of aromatic sandwiching residues located above and below the sandwiched m^{7}Gua are colored green and magenta, respectively: (B) wild-type VP39 (from ref *5*), (C) mutant F180W (from panel A), (D) Y22W/F180W double mutant [Y22 in the F180W structure (C) was replaced with Trp in silico], and (E) eIF4E (data from ref *7*; PDB entry 1EJ1). Panels B−E were generated using MIDAS (*35*).

μM, respectively. In a second experiment in which entirely different batches of reagents were employed, wild-type VP39 exhibited a K_d of 95 μM (only an ~10% difference, attesting to the reproducibility of the determination). The calculated free energy (ΔG) values of binding to wild-type VP39 (−5.58 kcal/mol) and the mutant (−6.81 kcal/mol) indicate that the replacement of Phe180 with Trp enhanced the stability of the stacking interaction with the m^{7}Gua moiety of m^{7}G by ~1.2 kcal/mol, or almost 1 order of magnitude. Apparently, shifting the position of overlap in the stack toward the region of m^{7}Gua containing the highest positive charge density (comparing panels B and C of Figure 1, above) is associated with a sizable increase in affinity.

We next turned to the fluorescence quenching assay. Although this method cannot be applied to wild-type VP39 due to the absence of a fluorescence change upon addition of ligand (data not shown), the assay afforded a rapid and accurate approach to the confirmation of binding constants for our three tryptophan mutant proteins. For all three proteins, fluorescence excitation and emission maxima were observed at 282 and 330 nm, respectively. The emission spectra (e.g., Figure 3A) are attributed largely to Trp residues, including those introduced into the cap-binding slot by site-directed substitution. Quenching of the emission maxima with no corresponding wavelength shift was observed upon addition of excess m^{7}G to each of the proteins. This quenching is attributed solely to the Trp residue in the cap-binding slot since (i) fluorescence quenching has previously been attributed to $\pi-\pi$ stacking interactions between an aromatic side chain and a nucleobase (*21*, *22*), (ii) VP39's three naturally occurring Trp residues are considered to be too remote from the substrate-binding site to be responsive to ligand binding (crystallographic structures of m^{7}G-liganded and -unliganded VP39 indicate the primary m^{7}G-binding site and indicate no conformational changes upon m^{7}G binding), (iii) pH titration of the binding of m^{7}G led to a decrease in quenching above pH 7.5 corresponding to the expected rise in K_d due to the loss of m^{7}Gua's N1 proton above its pK (*11*, *23*; data not shown), and (iv) no quenching was observed when using guanosine (which is N7-unmethylated and therefore refractory to VP39 binding, data not shown).

A 55% quenching of the emission maximum was observed upon addition of excess m^{7}G to the VP39 F180W mutant

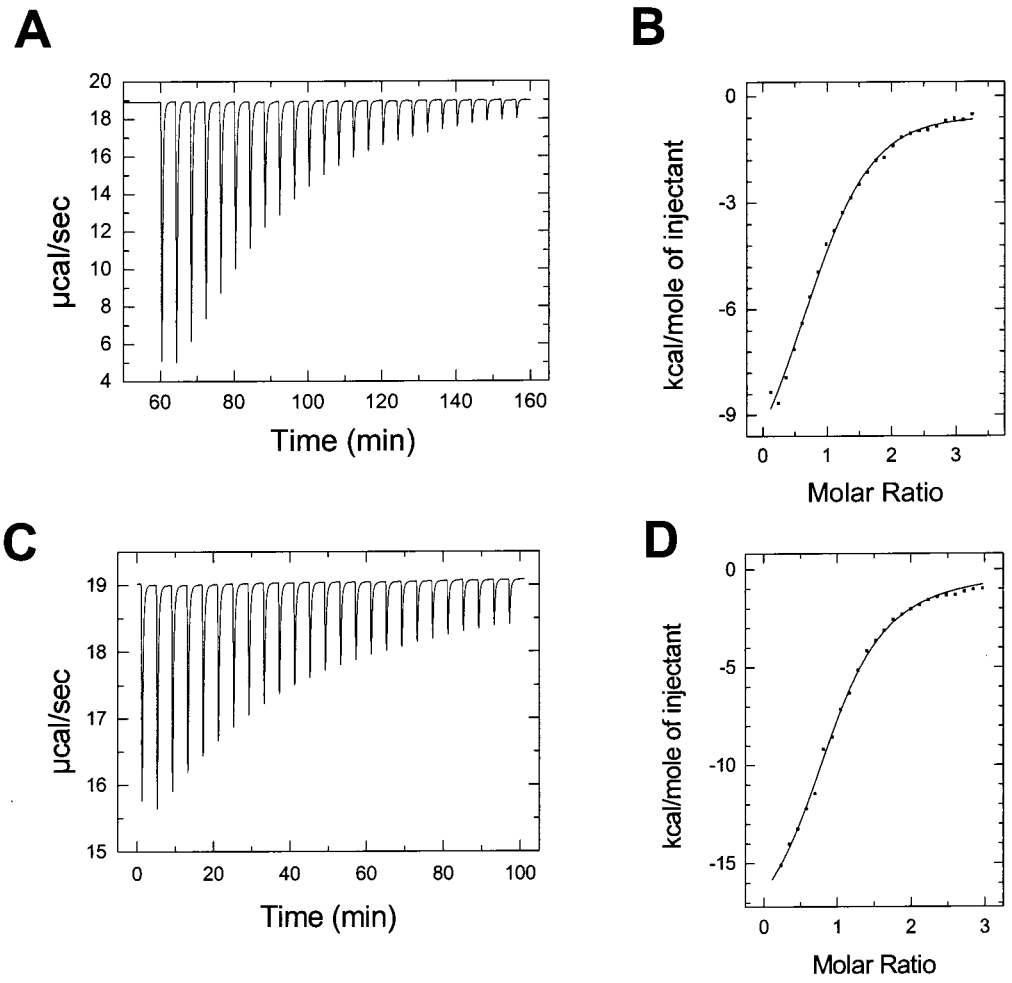

FIGURE 2: ITC titrations of wild-type VP39 (A and B) and mutant F180W (C and D) with m^7G at 25 °C. (A) Raw data from the titration of the wild-type protein. Each peak shows the heat produced by serial injections of m^7G (10 μL of a 5 mM solution per injection) into 2 mL of wild-type VP39 (0.3 mM in the same buffer) at 240 s intervals. In total, 25 sequential injections were made. (B) Binding isotherms derived by integration with respect to time, with appropriate molar correction. The points show the experimental data, and the solid line shows the calculated best fit, which yielded an equilibrium association constant (K_a) of $(1.16 \pm 0.12) \times 10^4$ M^{-1}. (C) Raw data from the titration of the F180W mutant. It is identical to panel A except that 0.05 mM mutant protein was used and 25 sequential injections (4 μL of a 2 mM solution per injection) were made. (D) Binding isotherms are analogous to those of panel B. The observed K_a value is $(9.18 \pm 0.59) \times 10^4$ M^{-1}.

(Figure 3A). Equilibrium titration experiments, in which the $m^7G-F180W$ interaction was monitored via quenching at 330 nm followed by fitting to a simple bimolecular binding process (Figure 3B), yielded a K_d value of 7.2 μM. This is highly comparable to the value of 10.9 μM obtained by ITC (Table 2). The slightly lower K_d obtained by fluorescence quenching might be attributable to the slightly lower temperature that was employed (20 °C for fluorescence titration vs 25 °C for ITC).

Taken together, the data of Table 2 indicate that a dramatic rise in affinity results from either moving the stacked area toward the region of positive charge in the stacked nucleobase (F180W) or increasing the surface area within the stack (Y22W), and that combining the two effects (Y22W/F180W) leads to a further ~5-fold increase in affinity.

So far, the quantitative effects of Trp substitution have been interpreted entirely via effects on the geometry of the stack. However, the data may also be interpretable in terms of "intrinsic" effects of the Phe → Trp and Tyr → Trp substitutions. Thus, ab initio calculations have indicated that the maximum intensity of negative electrostatic potential over the center of the ring system for Trp's indole side chain is greater than for Phe or Tyr side chains (*13*, *24*, *25*). Other calculations have indicated that Trp possesses a greater predisposition to π-electron donation (*21*). Moreover, studies of free amino acid stacking with free cap analogue have concluded the following ranking of stacking efficacy: Trp > Tyr > Phe [in which the ranking of Tyr was simply inferred (*26*)]. Coincidentally or not, the data with VP39's Trp mutants reflect the above ranking in that the VP39 mutant with a double-Trp substitution (Y22W/F180W) possessed the highest affinity (Table 2); next in affinity was mutant F180W, in which the natural Tyr22 was combined with a Trp substitution at Phe180, and third was Y22W, in which the native Phe180 was combined with a Trp substitution at Tyr22. As expected, wild-type VP39 (no Trp in the cap-binding slot) exhibited the lowest affinity for m^7G. This is the first demonstration of such a ranking within an intact protein (which differs from small molecule systems in that the geometry of the stacking "sandwich" is preformed).

Converted Cap Pocket: Binding and Methylation of Bona Fide VP39 Substrates. A BIAcore-based assay (described in detail in refs *5*, *10*, *18*, and *19*) was conducted to show, directly, the effect of Trp substitutions on substrate−enzyme

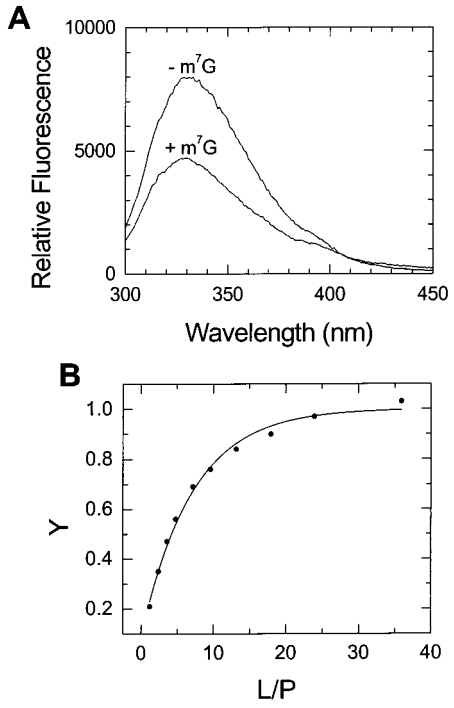

FIGURE 3: Fluorescence spectra and titration of the mutant F180W with m^7G. (A) Fluorescence emission spectra of F180W [2.5 μM, in 2 mL of NaCl (0.2 M), with sodium cacodylate (0.1 M, pH 5.9)] in the absence and presence of saturating m^7G (15 μM in water), generated at 20 °C. (B) Fluorescence titration of F180W (P) with m^7G (L) at 20°C (see Materials and Methods). Y refers to the fractional change in fluorescence, $\Delta F_i/\Delta F_{max}$ (Materials and Methods). The best fit of the data to a simple binding process, which is shown as a smooth line, yielded a K_d value of 7.2 μM. A Scatchard analysis of the binding data (not shown) yielded a similar K_d value, namely, 7.8 μM.

Table 2: Affinities of VP39−Ligand Complexes[a]

VP39	ligand	K_d (μM)
Phe180Trp	$m^{7,9}$Gua	2.1
Phe180Trp	m^7G	7.2
Phe180Trp	m^7Gp	15.9
Phe180Trp	m^7Gpp	19.6
Phe180Trp	m^7GpppG	16.7
Phe180Trp	$m^7GpppG(A)_3$	0.93
Tyr22Trp	m^7G	10.7
Tyr22Trp/Phe180Trp	m^7G	2.2
wild type	m^7G	86.2
Phe180Trp	m^7G	10.9

[a] The first eight entries were determined by a fluorescence quenching technique (e.g., see Figure 3 and Materials and Methods), and the last two were obtained by ITC (see the text and panels A and B of Figure 2). Duplicate fluorimetric titrations deviated by less than 5%. The affinity of mutant F180W for either guanosine, GTP, or GpppG was too low to be measurable by fluorescence quenching, though K_d values were estimated to be >200 μM. RNAs longer than $m^7GpppG(A)_3$ could not be generated in sufficient amounts for direct binding measurements.

affinity. In this assay, capped and uncapped poly(A) [$m^7GpppG(A)_{20}$-biotin (henceforth identified as "long RNA") and $pppG(A)_{20}$-biotin, respectively] were anchored in separate streptavidin-coated flow cells of a BIAcore biosensor chip, and VP39 or its mutants were injected. For each protein, a clear difference in affinity for capped and uncapped RNA was observed (data not shown) as demonstrated previously for wild-type VP39 (*18*). In replicate experiments, apparent dissociation and association rate constants (k_{diss} and k_{ass}, respectively) for protein interaction with the capped poly(A) were derived by fitting double-exponential functions to the dissociation and association portions, respectively, of sensorgrams. Figure 4A shows the major k_{diss} and k_{ass} values derived from multiple, replicate experiments at several ionic strengths in the range of 90−200 mM NaCl. Apparent dissociation constant ($K_d^{long\ RNA}$) values were obtained by dividing the apparent k_{diss} by k_{ass} values for each of the proteins. These each fell in the range of 8.84×10^{-10} to 1.26×10^{-9} M (i.e., on the order of 1 nM). This is comparable with the K_m value of 5 nM deduced for brome mosaic virus RNA, a standard mRNA substrate (*2*).

Previously, a four-site model for VP39−substrate interaction was suggested (*10*; see the introductory section) in which VP39 interacts with its RNA substrate via (i) the cap-binding site, (ii) the catalytic center, (iii) a downstream RNA-binding site for up to three nucleotides of cap-adjacent RNA, and (iv) a far-downstream RNA binding site interacting with the mRNA substrate beyond its sixth transcribed nucleotide. Whereas the m^7G nucleoside (as examined by ITC and fluorescence titration, above) interacts with only the cap-binding pocket (no ribose contacts were observed crystallographically), the 20-nucleotide capped poly(A) (long RNA substrate) anchored to the BIAcore chip would interact with all four sites. The difference in affinity (3−4 orders of magnitude) would reflect the cumulative affinities of the different binding sites, and/or the "connection Gibbs free energy" arising from multisite substrate anchoring (*27*). k_{diss} and k_{ass} values for a long RNA substrate [the 20-nucleotide capped poly(A) used in BIAcore experiments] did not differ substantially between wild-type VP39 and the Trp mutants (Figure 4A), despite 10−50-fold differences in affinity for the m^7Gua portion of it (Tables 1 and 2). This is consistent with a negligible proportion of the overall RNA substrate binding energy arising from the m^7Gua portion alone.

To assess whether the quantitative effects of Trp substitutions could be detected using a shorter RNA substrate [one that is too short to reach RNA binding site (iv) of the four-site model], we examined the interaction of wild-type VP39 and the Trp mutants with capped RNA of the form $m^7GpppG(A)_{3-4}$ (identified as "short RNA"). Since RNA of this length was considered to be too short for anchoring within the BIAcore instrument without the possibility of steric hindrance effects, and insufficient RNA was available for ITC studies, we instead compared methyltransferase kinetic constants for wild-type VP39 and the Trp mutants as described (*10*), and assayed solution binding by fluorescence quenching (Table 2).

With regard to methyltransferase kinetic constants, Michaelis−Menten kinetics can be an accurate model for an enzyme, provided the rate constants for the chemical step (k_{cat}) and steps relating to product release are negligible with respect to the rate constant for substrate unbinding. Since early reports of VP39 methyltransferase activity have described rates of product formation in units of detectable radioactive "counts" as opposed to rates of product formation and substrate depletion in molar units, substrate turnover rates were not being measured directly. Michaelis−Menten kinetics were simply assumed, insofar as these studies showed Lineweaver−Burk plots (relating initial observed rates of product formation to substrate concentration) to be both linear and intersecting (*2*). More recently, apparent k_{cat} values were derived indirectly (*10*) by applying the Michaelis−Menten

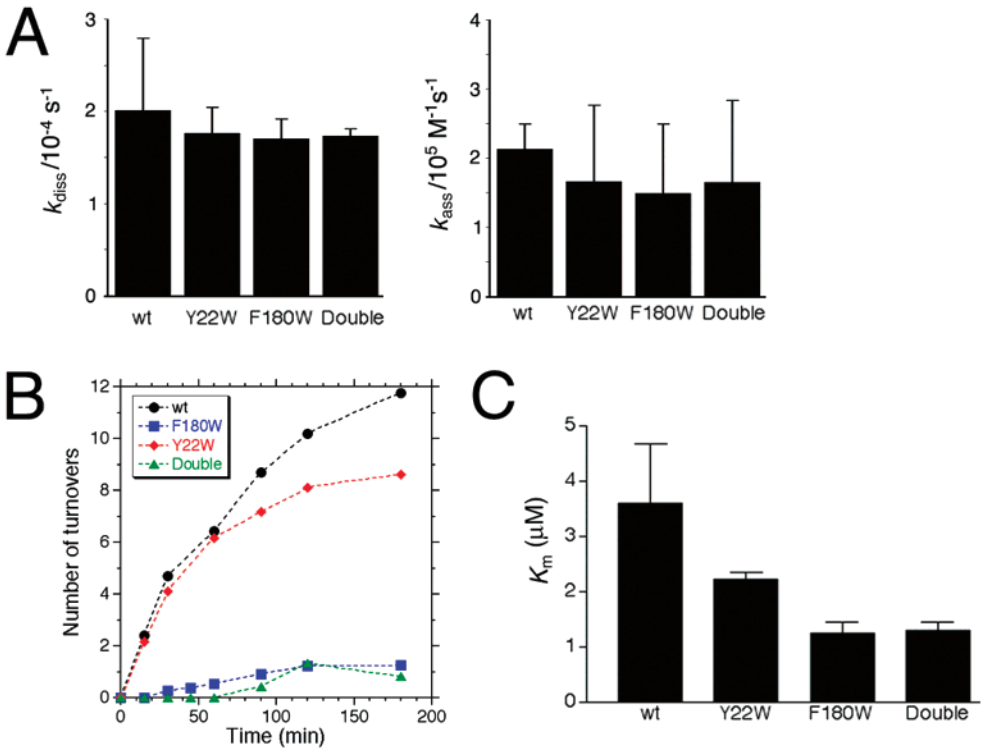

FIGURE 4: VP39 interaction with the intact substrate. (A) BIAcore mean k_{diss} (dissociation rate constant) and k_{ass} (association rate constant) values from replicate experiments: wt (wild-type VP39), 10 or 11 replicates at 90 and 150 mM NaCl; Y22W, 5 replicates at 90 and 150 mM NaCl; F180W, 5 replicates at 90 and 200 mM NaCl; and double (Y22W/F180W double mutant), 5 replicates at 90 and 200 mM NaCl (see Materials and Methods). Nonlinear fits to the exponential phases of sensorgrams were excellent, with residuals ranging from (at worst) one part in 500 to as low as 200 ppm. Error bars denote the standard deviation. The anchored ligand was $m^7GpppG(A)_{20}U_{biotin}$. (B) Short substrate [$m^7GpppG(A)_3$]-to-product conversion by wild-type (wt) VP39 and the Y22W, F180W, and Y22W/F180W (double) mutants. Assays contained substrate at a concentration of 4.36×10^{-5} M. Protein concentrations were as follows: 3.1×10^{-6} M for wt, 4.29×10^{-6} M for Y22W, 7.4×10^{-6} M for F180W, and 6×10^{-6} M for double. See the text for further details. (C) K_m values for wild-type VP39 and Trp mutants using the short, capped VP39 substrate $m^7GpppG(A)_4$. For each protein, several replicate experiments were performed. Mean values are shown, and error bars denote the standard error. The apparent K_m for wild-type VP39 (3.5×10^{-6} M) was ~10-fold greater than that determined in ref *10*, probably due to the use of an entirely different set of reagents. Despite this shift in absolute values after a 2 year hiatus, emphasis is placed on the relative values in the comparison of the wild type and mutants in parallel, replicate assays.

equation to K_m values for short capped RNA substrates (derived from Lineweaver−Burk "counts" plots) along with values for initial rates of product formation (v) that depended upon the quantitation of absolute molar amounts of tritiated product in bands accurately excised from dried polyacrylamide gels, where the following factors were influential: accurate alignment of gel and X-ray film for band excision, correction for the high degree and variability of quenching during scintillation counting, and background counts in the gel. Despite difficulties in obtaining accurate values for k_{cat} in this way, it was clear that VP39's rate of substrate turnover was very low with respect to the turnover rates for other enzymes. We now introduce a new assay, employing ^{33}P-labeled short RNA substrate [$m^7GpppG(A)_3$], in which rates of VP39 substrate−product conversion (and hence k_{cat}) can be measured directly from known starting amounts of substrate. Under the conditions of the new assay (substrate concentrations significantly above the apparent $K_m^{\text{short RNA}}$ value, below, assuming Michaelis−Menten kinetics), wild-type VP39-catalyzed conversion was monitored through the 12th turnover (Figure 4B). Through at least the eighth turnover, substrate−product conversion proceeded more-or-less linearly, indicating no apparent obstacle to multiple turnovers. The observed rate during the linear phase was ~1 turnover per 7.5 min, indicating an apparent k_{cat} value of 2.2×10^{-3} s^{-1}. Interestingly, since enzyme-bound as well as free RNA products would have been subject to our analysis, this modest value for k_{cat} apparently reflects a slow chemical step as opposed to a rate-limiting release of RNA product or cofactor/product. Since the observed reaction rate for the wild-type protein remained constant through at least the first approximately eight turnovers, the progressive accumulation of ~8 molar equiv of AdoHcy was not inhibitory with respect to the ongoing reaction under the assay conditions that were employed. The drop in rate at ~10 turnovers could be attributable to substrate depletion, product accumulation, or loss of enzyme activity after 2 h of assay time. The observed substrate−product conversion rate (k_{cat}) for mutant Y22W was comparable to that for wild-type VP39. Interestingly, mutants F180W and Y22W/F180W ("double") required ~108 min to effect the first turnover (Figure 4B), corresponding to an apparent k_{cat} value of ~1.5 $\times 10^{-4}$ s^{-1}. This was 14-fold lower than the k_{cat} values for wild-type VP39 and Y22W. Although equivalent specific activities are assumed for all protein preparations, it is notable that mutant F180W, for which the lower k_{cat} value was observed, nonetheless could be expressed at high levels and crystallized readily, indicating an absence of heterogeneity in the protein preparation. The turnover data therefore indicated that the chemical steps may be slowed in the context of an F80W mutation. Perhaps this mutation leads to a very slight reorientation of VP39's 2′-OH with respect

to the catalytic center, leading to a critical deviation from optimal hydrogen bonding angles and/or lengths. In addition to the change in the apparent k_{cat}, mutants F180W and Y22W/F180W exhibited no more than a single complete turnover during the assay (Figure 4B). This would not have been due to product inhibition. Despite protein concentrations in the assay being ~2-fold higher for F180W and Y22W/F180W than for wild-type VP39 and Y22W, the amount of product generated in a single turnover by the former two mutants would be equivalent to that produced in two turnovers by the latter two, which was not an inhibitory concentration (see above). Moreover, we do not believe that the inability of F180W and Y22W/F180W to exhibit more than a single turnover resulted from substrate depletion, because substrate and cofactor concentrations at the outset of the assay were well above the apparent $K_m^{\text{short RNA}}$ and K_m^{AdoMet} (below, assuming Michaelis−Menten kinetics). Therefore, assuming proteins retained activity during the latter part of the assay, we are left with the possibility that mutants F180W and Y22W/F180W may exhibit some blockade of product release which begins to predominate over the slow catalytic rate upon approaching the completion of the first turnover. This would presumably result from the higher affinity of the two mutants for the m^{7}G portion of the cap (above).

Due to the low turnover rates (above), we consider that "steady-state" kinetic conditions are approximated over significant periods of time in the in vitro assay, a prerequisite for Michaelis−Menten analysis. The Michaelis−Menten mechanism also requires that k_{cat} be negligible with respect to the rate of substrate unbinding from the enzyme. Since the latter rate constant has only been quantitated for a long RNA substrate (Figure 4A), our only indication of the appropriateness of Michaelis−Menten kinetics for the short RNA substrate is the linearity of Lineweaver−Burk plots (which proved to be more-or-less linear in our hands, data not shown), so we proceeded with the quantitation of a Michaelis constant for wild-type VP39 and the Trp mutants. $K_m^{\text{short RNA}}$ values for m^{7}GpppG(A)$_4$ were quantitated over 6−18 replicate assays for each mutant, and mean values were plotted (Figure 4C). Apparent $K_m^{\text{short RNA}}$ values were slightly lower for the mutant proteins than for wild-type VP39 (up to 2.5-fold). If Michaelis−Menten kinetics are assumed, $K_m^{\text{short RNA}}$ should be equivalent to the dissociation constant for the enzyme−substrate complex. The values that were obtained, therefore, indicated that Trp mutations in the cap-binding pocket enhance VP39's affinity for the overall short RNA substrate by only a relatively modest amount.

The data depicted in Figures 2−4 and Table 2 indicate that $K_m^{\text{short RNA}}$ is lower than the K_d^{m7G} for the F180W mutant and wild-type VP39 by factors of only 5−9 and ~25, respectively. In contrast, $K_d^{\text{long RNA}}$ (BIAcore) and $K_m^{\text{long RNA}}$ are 3−3.5 orders of magnitude lower than $K_m^{\text{short RNA}}$ [where short and long RNAs are represented by m^{7}GpppG(A)$_{3-4}$ and m^{7}GpppG(A)$_{20}$, respectively]. Although care must be exercised in comparing affinities arising from connected binding sites (*27*) and those quantitated using different methods, these data nonetheless suggest that site (iv) (the far-downstream RNA binding site) makes a sizable contribution to VP39−substrate interaction with respect to site (iii) (the site for RNA immediately downstream of the cap).

Combination of the fluorescence quenching assay with VP39 Trp mutants (in this case, F180W) afforded a convenient opportunity to derive binding constants for a series of step-by-step truncations of VP39's ligand from short RNA length [in this case, m^{7}G(5′)pppG(A)$_3$] down to only the m^{7}Gua portion of the cap alone (Table 2). $K_d^{\text{short RNA}}$ (0.93 μM, Table 2) was found to be very close to $K_m^{\text{short RNA}}$ (1.5 μM, Figure 4C) for mutant F180W. There was a significant (~17-fold) loss of affinity upon eliminating site (iii) anchoring by removing the three- or four-nucleotide oligo(A) chain of the RNA (leaving only the m^{7}GpppG dinucleotide, Table 2). The VP39−short RNA complex is crystallographically well defined (*6*), whereas in the VP39−m^{7}G(5′)pppG complex, the ribose−pppG moiety is only poorly defined (*5*; Figure 1A). Thus, elimination of anchoring of cap-proximal RNA at site (iii) is clearly accompanied by a loss of contact at site (ii) [site (ii) being the catalytic center, which interacts with the ribose of the G nucleotide of m^{7}G(5′)-pppG]. This is consistent with site (ii) contacts being of only modest affinity with respect to those at site (iii). Due to its apparently negligible affinity, site (ii) will be ignored in the following discussion of overall RNA substrate affinity. The ~17-fold loss of affinity upon removal of site (iii) anchoring does not necessarily correspond directly with the affinity of site (iii) itself, due to additional effects which come into play upon conversion of two-site [site (i) plus site (iii)] to single-site [site (i) alone] anchoring. These effects include a presumed gain of entropy arising from the gain of disorder by the ribose−pppG moiety of m^{7}G(5′)pppG (above) and also the "connection free energy" (*27*) arising from the linkage of two "ligands" {in this case, the m^{7}Gua [site (i)] and oligo(A) [site (iii)] portions of the short capped RNA}.

There was only a very slight (~1−1.2-fold) loss of affinity upon further truncation of the ligand series by removal of the pG moiety of m^{7}GpppG (Table 2), consistent with the disordered (unanchored) terminal guanosine contributing only minimally to affinity. There was no significant change in affinity upon truncation by one additional phosphate leaving only m^{7}Gp. Unexpectedly, however, there was an apparent 2.2-fold gain in affinity upon removal of the final phosphate, and a further ~3.4-fold gain upon removal of the final ribose (leaving only an analogue of the m^{7}Gua portion of the cap, namely, the dimethylated nucleobase m^{7,9}Gua). Apparently, the structurally ill defined (*5*) phosphoribose portion of m^{7}Gp has a modestly deleterious effect on affinity.

When one examines the series overall, m^{7,9}Gua (the smallest member of the ligand series examined in Table 2) appears to have only a 2-fold lower affinity than m^{7}G(5′)-pppG(A)$_3$ (the largest member of the series). It could therefore be supposed that the contribution of the ribose−pppG(A)$_3$ portion of the short, capped RNA to overall binding energy is vanishingly small, especially in light of the fact that m^{7}G(5′)pppG(A)$_3$ contributes a connection free energy (above) to its binding constant whereas m^{7,9}Gua does not, and the pppGpApAp portion of the ligand is known to be anchored by a total of at least 12 hydrogen bonds and/or salt links (Figure 3E of ref *6*). We can only presume, therefore, that the unfavorable "freezing-out" of large numbers of internal rotations and the concomitant loss of entropy upon binding of the ribose−pppG(A)$_3$ portion of the short capped RNA, with respect to the m^{7,9}Gua ligand for which there would be no internal rotations, are thermodynamically highly unfavorable. The true affinity of the pppG(A)$_3$ portion of the ligand could only be assessed by

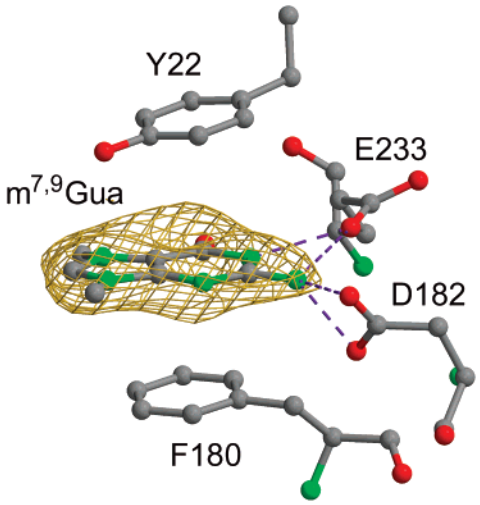

FIGURE 5: Difference Fourier electron density for $m^{7,9}$Gua bound to a wild-type VP39 crystal contoured at the 3σ level. The crystal was soaked in 10 mM $m^{7,9}$Gua. The density was calculated using a model that included only the protein. Water molecules are not shown. This figure was generated using BOBSCRIPT (*33*) and RASTER3D (*34*).

directly measuring the affinity of this portion alone. However, such measurements would present a formidable challenge. First, the K_m for uncapped RNA cannot be determined because this RNA is not a substrate for VP39 methyltransferase (*2*). Second, VP39 possesses multiple RNA binding sites for uncapped RNA (*12*, *28*, *29*). Since the RNA chain portion of short capped RNA is bound only via its phosphoribose backbone (*6*), specificity for VP39's methyltransferase substrate RNA binding site is conferred only by the cap (*12*).

For the above reasons, perhaps the most reliable comparisons of binding constants that can currently be made are those in which connectivity and entropy remain unchanged by, for example, modulating affinity at specific subsites without a complete loss of anchoring. In this regard, 10−50-fold increases in affinity at the cap-binding pocket, by substitution with Trp, led to only a modest change in $K_m^{\text{short RNA}}$ (Figure 4C), indicating that the RNA chain portion of the short capped RNA contributes significantly to the overall binding energy with respect to the cap portion.

Role of Positive Charge in Stacking of a Bona Fide m^7Gua Analogue. Recently, VP39−AdoHcy cocrystals were soaked with methylated nucleobases (m^1cytosine, m^3cytosine, m^1adenine, and m^3adenine) at pH values below their p*K* (i.e., pH values at which the rings would possess positive charge), and binding was shown to be positive charge-dependent (*11*). In each of these nucleobases, the charge was localized to a six-membered aromatic ring, and none of them was based on guanine. We have now soaked VP39−AdoHcy cocrystals with $m^{7,9}$Gua which (i) is based on m^7Gua, (ii) like m^7G possesses a positive charge in a five-membered ring, and (iii) like m^7G possesses a positive charge that is locked into the ring in a pH-independent manner. As shown in Figure 5, we observed clear and complete electron density for $m^{7,9}$Gua. Its geometry of VP39 binding was virtually identical to that of m^7G (including the N9-methyl which corresponds to the C1′ of m^7G) and to the m^7Gua portion of m^7G-containing ligands bound in all VP39 structures determined thus far (Figure 1A) (*5*, *6*).

Although the rotational orientations, within the stack, of positively charged nucleobase analogues can differ dramatically with respect to that of m^7G (*11*), that of bound $m^{7,9}$Gua (Figure 5) did not. Since $m^{7,9}$Gua represents m^7G without an attached ribose, the ribose moiety of m^7G apparently has no effect on the rotational orientation of the attached m^7Gua moiety. It is notable that VP39 complexes with capped ligands show the positively charged five-membered ring of m^7Gua sitting only partially within the stack (e.g., see Figure 1B). The adoption of this position by $m^{7,9}$Gua indicates that the phenomenon does not result from constraints imposed by the ribose moiety or additional downstream substituents, but rather, it is an inherent property of the cap nucleobase−binding slot interaction.

Implications of VP39 and eIF4E's Choice of Stacked Aromatic Side Chain. Despite the similarity in their modes of cap binding, eIF4E and VP39 differ functionally in that the former operates via stable cap binding whereas the latter must turn over its RNA substrate. At first glance, this would correlate with the choice of aromatic side chains, in that Phe and Tyr (VP39) present smaller and less intrinsically stackable aromatic rings than does the indole ring of tryptophan (eIF4E) (*13*, *21*, *24*−*26*) (see also above). However, multiple protein sequence alignments show that some eIF4E analogues (which have been identified and predicted from large-scale genome and cDNA sequencing projects) possess a tyrosine in place of one of the two cap-stacking tryptophans. Examples of such analogues include a novel human eIF4E (*30*) (GenBank entry AAC39871), mouse eIF4E-like protein 4E-LP (GenBank entry AAC19373, unpublished), the cap-binding protein of *Arabidopsis thaliana* (*31*) (GenBank entry BAB09469), and putative initiation factor E4 of *Plasmodium falciparum* (*32*) (GenBank entry CAB38977). If the proteins have a role in translational initiation, then tryptophan would not be essential for eIF4E function. In contrast, the Tyr/Phe combination is absolutely conserved in the VP39 sequences of a variety of poxviruses from diverse hosts (albeit the overall VP39 sequences are quite highly conserved among the various poxviruses anyway). In the current study, we have completed the "quadrant" (eIF4E vs VP39 or Trp vs Phe/Tyr) by making Trp substitutions at VP39's conserved Phe and Tyr, yet we observed no apparent ill effects in protein folding or cap stacking, indicating that VP39's choice of Tyr and Phe does not relate to the maintenance of the integrity of the cap-binding pocket. However, the apparently negative effect of the F180W substitution upon catalysis may point to a significant sensitivity of the catalytic center toward miniscule or subtle changes in substrate orientation. Moreover, the possible inhibition of product release by the F180W substitution may point to an importance of maintaining a lower-affinity Tyr/Phe pocket to facilitate substrate turnover.

In addition, it is possible that VP39's Tyr and Phe have been preserved, to date, simply because the conversion of Phe or Tyr codons to the monocodonic Trp has presented an excessively great an evolutionary jump (requiring two contiguous nucleotide substitutions). The question of why Phe and Tyr have not interconverted in VP39 (only a U ↔ C change would be required) could perhaps be addressed by determining catalytic constants for mutants in which these two residues had been swapped.

ACKNOWLEDGMENT

We thank Professor Bernhard Lippert (Fachbereich Chemie, Universitaet) for the kind gift of $m^{7,9}$Gua, Dr. Alec Hodel (Department of Biochemistry, Emory University School of Medicine) for the kind gift of $m^7GpppG(A)_3$, Ms. Shi-Mei Wang for generating the DNA mutations for VP39 mutant proteins Y22W, F180W, and Y22W/F180W, Mr. Benjamin Lasseter and Dr. Greg Reinhardt (Department of Biochemistry & Biophysics, Texas A&M University, College Station, TX) for use of their microcalorimeter and assistance in the use thereof, and Dr. Yuxin Mao and Mr. William E. Meador for technical assistance.

REFERENCES

1. Sonenberg, N., Morgan, M. A., Merrick, W. C., and Shatkin, A. J. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4843−4847.
2. Barbosa, E., and Moss, B. (1978) *J. Biol. Chem. 253*, 7698−7702.
3. Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994) *Cell 78*, 657−668.
4. LaGrandeur, T. E., and Parker, R. (1998) *EMBO J. 17*, 1487−1496.
5. Hodel, A. E., Gershon, P. D., Shi, X., Wang, S.-M., and Quiocho, F. A. (1997) *Nat. Struct. Biol. 4*, 350−354.
6. Hodel, A. E., Gershon, P. D., and Quiocho, F. A. (1998) *Mol. Cell 1*, 443−447.
7. Marcotrigiano, J., Gingras, A.-C., Sonenberg, N., and Burley, S. K. (1997) *Cell 89*, 951−961.
8. Matsuo, H., Li, H., McGuire, A. M., Fletcher, C. M., Gingras, A. C., Sonenberg, N., and Wagner, G. (1977) *Nat. Struct. Biol. 4*, 717−724.
9. Quiocho, F. A., Hu, G., and Gershon, P. D. (2000) *Curr. Opin. Struct. Biol. 10*, 78−86.
10. Lockless, S. W., Cheng, H.-T., Hodel, A. E., Quiocho, F. A., and Gershon, P. D. (1998) *Biochemistry 37*, 8564−8574.
11. Hu, G., Gershon, P. D., Hodel, A. E., and Quiocho, F A. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7149−7154.
12. Gershon, P. D., Shi, X., and Hodel, A. E. (1998) *Virology 246*, 253−265.
13. Dougherty, D. A. (1996) *Science 271*, 163−168.
14. Hodel, A. E., Gershon, P. D., Shi, X., and Quiocho, F. A. (1996) *Cell 85*, 247−256.
15. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
16. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) *Anal. Biochem. 179*, 131−137.
17. Miller, D. M., Olson, J. S., Pflugrath, J. W., and Quiocho, F. A. (1983) *J. Biol. Chem. 258*, 13665−13672.
18. Shi, X., Yau, P., Jose, T., and Gerson, P. D. (1996) *RNA 2*, 88−101.
19. Gershon, P. D. (1997) Analysis of mRNA formation and function, in *Methods in Molecular Genetics* (Richter, J., Ed.) pp 127−148, Academic Press, San Diego.
20. Milligan, J. F., Groegbe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8798.
21. Pullman, B., and Pullman, A. (1958) *Proc. Natl. Acad. Sci. U.S.A. 44*, 1197−1202.
22. Ishida, T., Katsuta, M., Inoue, M., Yamagata, Y., and Tomita, K. (1983) *Biochem. Biophys. Res. Commun. 115*, 849−854.
23. Hendler, S., Furer, E., and Srinivasan, P. R. (1970) *Biochemistry 9*, 4141−4153.
24. Ma, J. C., and Dougherty, D. A. (1997) *Chem. Rev. 97*, 1303−1324.
25. Mecozzi, S., West, A. P. J., and Dougherty, D. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10566−10571.
26. Ishida, T., Doi, M., and Inoue, M. (1988) *Nucleic Acids Res. 16*, 6175−6190.
27. Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4046−4050.
28. Deng, L., Johnson, L., Neveu, J. M., Hardin, S., Wang, S.-M., Lane, W. S., and Gershon, P. D. (1999) *J. Mol. Biol. 285*, 1417−1427.
29. Johnson, L., and Gershon, P. D. (1999) *Nucleic Acids Res. 27*, 2708−2721.
30. Zhang, Q. H., Ye, M., Wu, X. Y., Ren, S. X., Zhao, M., Zhao, C. J., Fu, G., Shen, Y., Fan, H. Y., Lu, G., Zhong, M., Xu, X. R., Han, Z. G., Zhang, J. W., Tao, J., Huang, Q. H., Zhou, J., Hu, G. X., Gu, J., Chen, S. J., and Chen, Z. (2000) *Genome Res. 10*, 1546−1560.
31. Kotani, H., Nakamura, Y., Sato, S., Asamizu, E., Kaneko, T., Miyajima, N., and Tabata, S. (1998) *DNA Res. 5*, 203−216.
32. Bowman, S., Lawson, D., Basham, D., Brown, D., Chilingworth, T., et al. (1999) *Nature 400*, 532−538.
33. Esnouf, R. M. (1997) *J. Mol. Graphics Modell. 15*, 132−134.
34. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
35. Huang, C. C., Pettersen, E. F., Klein, T. E., Ferrin, T. E., and Langridge, R. (1991) *J. Mol. Graphics 9*, 230−236.

BI0201926