Insertion of an N7-methylguanine mRNA Cap between Two Coplanar Aromatic Residues of a Cap-binding Protein Is Fast and Selective for a Positively Charged Cap*

Guanghui Hu‡§, Ah-Lim Tsai¶, and Florante A. Quiocho‡¶**

From the †Graduate Program in Structural and Computational Biology and Molecular Biophysics, ‡Howard Hughes Medical Institute and Mars and Verna McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine Houston, Texas 77030 and the ¶Hematology Division, Internal Medicine, University of Texas Houston Medical School, Houston, Texas 77030

The N7-methylated guanosine (m7G) cap structure, which is found at the 5' ends of mature eukaryotic mRNAs, is critical to a myriad of biological processes. The twenty structures of complexes of cap nucleosides and nucleotides and methylated bases with the vaccinia virus VP39, a cap-specific RNA 2'-O-methyltransferase, which we have determined previously, have revealed the atomic basis of cap binding. The precise insertion and tight fitting of the m7Gua moiety of the cap between two parallel aromatic residues that are spaced only 6.8 Å apart governs the high specificity of binding. Here we report the investigation of the reaction mechanism of VP39 with three capped ligands (m7G, m7GppG, and m7GpppGAP) by fluorescence stopped-flow technique. Cap binding is a simple one-step mechanism with very fast association rate constant (∼107 M−1 s−1). Moreover, the pH dependence on the association rate constant of m7G binding indicates that only the positively charged keto tautomer of the cap is recognized and bound. The association and dissociation rate constants and affinity constants of the three ligands do not vary greatly, demonstrating that binding is achieved almost entirely by the interactions of m7Gua with two aromatic residues in a cation-π sandwich.

The ability of proteins to discriminate alkylated from nonalkylated nucleic acids is of paramount importance in numerous biological processes, including DNA repair, pre-mRNA splicing, nucleocytoplasmic transport, mRNA translation, cap-dependent ribose methylation, and influenza virus transcriptional priming (briefly summarized in (1)). VP39, a vaccinia virus protein, provides an excellent and well defined system to study mRNA recognition at the molecular level. The protein acts in the processing of both ends of nascent mRNA transcripts. At the 3' end, the protein serves as a processivity factor for the vaccinia virus poly(A) polymerase (1). Moreover, the N7-methylguanosine (m7Gua) moiety of the cap. This binding mode is mediated almost entirely by the interactions with the m7Gua nucleobase as the proximal groups (ribose, phosphates, and the second G) show little or no electron density and thus appear sequence-nonspecific. The failure of VP39 to catalyze

Received for publication, July 16, 2003, and in revised form, October 20, 2003
Published, JBC Papers in Press, October 22, 2003, DOI 10.1074/jbc.M307654200

* This work was supported by grants from NIH (GM-44911 and GM56818) to A.-L. T. and from the Welch Foundation (Q-0581) and Howard Hughes Medical Institute (to F. A. Q.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1794 solely to indicate this fact.

This paper is available on line at http://www.jbc.org

This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1794 solely to indicate this fact.

1 The abbreviations used are: m7Gua, N7-methylguanine; m7G, N7-methylguanosine.
reactions are fundamentally analogous to those of VP39, some
cation-
activities of the mutant protein (named F180W VP39) have
crystallographic, fluorescence measurements, isothermal calo-
tions concerning which tautomer (keto or enolate) of the cap is
Fluorescence titration measurements with F180W VP39 were
cap analog cap, are fundamentally analogous to those of VP39, some
cation-
Equilibrium Ligand Binding to F180W VP39 as a Function of pH—
the alkylated nucleobases about an axis perpendicular to that of
the plane of the nucleotide ring.
Kinetic studies of ligand binding are necessary for a deeper
understanding of the cap mRNA recognition process and reac-
tion mechanism. In the present study, we investigate the ki-
etics of cap binding to VP39 by rapid mixing, fluorescence
stopped-flow technique. This study has been facilitated by the
replacement of Phe-180 by a tryptophan to provide a fluores-
cence (Fig. 1B) (6, 8). In addition we also address a key ques-
tion concerning which tautomer (keto or enolate) of the cap is
recognized and bound by investigating the effect of pH on the
kinetics.

EXPERIMENTAL PROCEDURES

Materials—Site-directed F180W VP39 was purified as previously
described (6). Nucleosides mG was purchased from Sigma-Aldrich,
mGpppG from New England Biolabs. The compound mGpppGAAA
was kindly provided by the late Dr. Alec E. Hodel.

Equilibrium Ligand Binding to F180W VP39 as a Function of pH—
Fluorescence titration measurements with F180W VP39 were
cap analog cap, are fundamentally analogous to those of VP39, some

cation-

One-step Binding Mechanism—Cap analog binding studies
in solution have been greatly facilitated by replacing one of
the two aromatic residues (Phe-180) that sandwich the mG
moiety of capped nucleotides and oligonucleotides with a Trp
residue to provide a fluorescent reporter group (Fig. 1) (6, 8).
The fluorescence emission maximum is at 330 nm with 55%
quenching, with no shift in the emission maximum, occurring
upon the addition of excess mG. Ligand binding and functional
activities of the mutant protein (named F180W VP39) have
been characterized by a variety of techniques, including x-ray
crystallographic, fluorescence measurements, isothermal calo-
rimetry, and surface plasmon resonance. The F180W VP39
showed no apparent defects in methyltransferase catalytic ac-
tivity or VP39-capped RNA interactions and no perturbation in
the cap-binding pocket structure (6, 7). Moreover, the struc-
tures of the wild-type and F180W VP39s complexed with mG

Fig. 1. Stacking interaction or cation–π sandwich of the
mGua moiety of capped nucleotides by aromatic residues of
VP39. A, mGua moiety of mG sandwiched between Tyr-22 and Phe-
180 of the wild-type VP39 projected parallel to the base ring (adapted
from Ref. 6). The atomic surface is shown in mesh representation. B,
mGua group of mGpppG sandwiched between Tyr-22 and Trp-180 of
the F180W VP39 projected perpendicular to the base ring (adapted
from Ref. 8).

e14FE was elucidated by both x-ray crystallography and NMR
techniques (9, 10). Although the features of cap recognition by
e1F4E, based on its crystal structure with bound mGppp cap
analog, are fundamentally analogous to those of VP39, some
details of the recognition differ between the two proteins in the
following manner (see also Ref. 1 for details). i) The mGua
moiety is also sandwiched by two aromatic side chains, but they
come from Trp residues. ii) The N7-methyl group is engaged in
a nonpolar van der Waals contact with a carbon atom of a Trp
residue. iii) The polar groups of mGua moiety are involved in
only three hydrogen bonds. The C6 position carbonyl oxygen
accepts a hydrogen bond from a backbone NH group. Only one
carboxylate group from a Glu residue makes hydrogen bonds
with the N(1)H and N(2)H2 groups. Asp-182, one of the two
acidic residues in VP39 that makes hydrogen bond with the
mGua moiety, has no counterpart in e1F4E, and it appears to
be essentially dispensable to cap binding and VP39 function
(6, 7).

The prevailing view emanating from a variety of studies,
including extensive crystallographic and functional analysis of
wild-type and mutant VP39, is that the enhanced double stack-
ing or cation–π (or more appropriately “cation–π sandwich”) interactions between the two aromatic residues and methyl-
ated nucleobases play a dominant role in cap recognition and
binding (1, 6, 7). This mechanism requires a cap in the posi-
tively charged form. Although the two acidic residues in the
cap-binding slot are apparently essentially dispensable to cap
binding and VP39 function (5, 6, 11), they play primarily a role
in molecular recognition by fixing the rotational orientation of
The kinetics of m7G cap analog binding to the F180W VP39 measured in a stopped-flow apparatus, was followed as a decrease in fluorescence intensity. As the binding was too fast to be measured at 25 °C, the kinetic measurements were performed at 4–5 °C. Typical kinetic data are depicted in Fig. 2A. At faster rates (i.e., higher mG concentration), a significant proportion of the binding reaction occurred during the dead time (1.5 ms) of the apparatus. The observed fluorescence changes clearly follow a standard single exponential kinetics (Fig. 2A). This kinetic behavior is consistent with a simple one-step reaction mechanism (Reaction 1), with only one event shown in citrate buffer. The observed rates, 200–240 s⁻¹, however, did not show any significant changes in all four buffer combinations. These control experiments eliminate the possibility that the different ionic strength or ion species introduced by the buffer affect the observed rates. As long as the pH is maintained at the intended value the observed rate remains the same regardless of the buffer strength and composition. We have also shown previously that ionic strengths do not affect significantly the equilibrium binding of mG (6).

**Table 1**

| pH | \(k_1\times10^{-7}(\text{M}^{-1}\text{s}^{-1})\) | \(k_{-1}\times10^{-7}(\text{M}^{-1})\) | \(K_a\times10^{-5}(\text{M}^{-1})\) |
|----|-----------------|-----------------|-----------------|
| 5.5 | 3.00 ± 0.05     | 69 ± 5          | 4.4             |
| 6.0 | 3.00 ± 0.17     | 45 ± 17         | 6.6             |
| 6.5 | 2.32 ± 0.07     | 51 ± 13         | 4.6             |
| 7.0 | 1.30 ± 0.03     | 41 ± 11         | 3.2             |
| 7.5 | 0.557 ± 0.005   | 60 ± 6          | 0.93            |
| 8.0 | 0.180 ± 0.008   | 70 ± 18         | 0.26            |
| 8.5 | 0.078 ± 0.003   | 56 ± 9          | 0.14            |

*Standard error for \(k_1\) and \(k_{-1}\) values were obtained from linear regression to the observed rates (\(k_{obs}\)) measured at various ligand concentrations at a given pH. The various buffers have no significant effect on the observed rates (see “Results”).

To check whether changing buffers at different pH affects the kinetic data by the changes in cation concentrations from the buffer, we performed the following stopped flow control experiment. mG binding kinetics to 1 μM VP39 was measured in 0.05 M or 0.1 M buffers at either pH 5.5 (citrate or cacodylate) or pH 7.0 (HEPES or Tris). The mG concentrations after mixing were 10 and 15 μM for pH 5.5 and 7.0, respectively. The four different buffer combinations at pH 7.0 did not make any significant difference in measured kinetics. The observed rates at 6 °C were in the range of 180–220 s⁻¹, and the observed fluorescence change was very similar for all four conditions. Thus changing buffers between Tris and HEPES and varying the buffer concentrations from 0.05 to 0.1 M produced no effect on the outcome of the kinetic measurements. Similar results were obtained at pH 5.5 using citrate and cacodylate buffers at concentrations of 0.05 and 0.1 M, but the observed fluorescence change in the cacodylate buffer was about 30% lower than that found in citrate buffer. The observed rates, 200–240 s⁻¹, however, did not show any significant changes in all four buffer combinations. These control experiments eliminate the possibility that the different ionic strength or ion species introduced by the buffer affect the observed rates. As long as the pH is maintained at the intended value the observed rate remains the same regardless of the buffer strength and composition. We have also shown previously that ionic strengths do not affect significantly the equilibrium binding of mG (6).
The apparent $K_a (=k_1/k_{-1})$ values derived from the kinetic data at 5°C are not greatly different from those obtained by equilibrium fluorescence titration at 25°C (see “Experimental Procedures”) (Fig. 3B). More importantly, the pH dependence of $K_a$ derived from both methods shows a similar pK of about 7 (Fig. 3, A and B).

**Binding Kinetics of Longer Cap Analogues**—We have extended our kinetic study to include capped analogs longer than m7G. Restricted by reagent availability, binding of only two analogs (the dinucleotide m7GpppG and the substrate m7GpppG(A)3) was investigated at pH 6.0 (Table II). This study was further prompted by the different functional and structural features that each ligand exhibits as a potential substrate for the methylation activity of VP39. Although the dinucleotide has the methylation target (2′-OH) from the second G, it is not a substrate for VP39 (11). This finding is in full agreement with the crystallographic study showing that the second G is not visibly anchored in the catalytic site in the bound structures of both wild-type and F180W proteins (7, 8). In fact, like m7G, only the entire m7Gua moiety of the dinucleotide showed a well resolved electron density in these structures. In sharp contrast, m7GpppG(A)3 is an active methylation substrate (5). Moreover, the crystal structure of a ternary complex with the m7GpppG(A)3, a substrate that behaves like m7GpppG(A)2, and S-adenosyl homocysteine revealed a well defined electron density for the entire substrate and several polar interactions, including those with the β and γ phosphates of the triphosphate linkage and the entire region of the first three transcribed bases (5). The anchoring provided by the interactions with the three transcribed bases is apparently a prerequisite for an active substrate.

Our kinetic analysis demonstrates that both the $k_1$ and $k_{-1}$ rate constants decrease in the order m7G > m7GpppG > m7GpppG(A)3 (Table II). Both rate constants decrease in such a manner that the apparent association constant ($K_a = k_1/k_{-1}$) values do not vary greatly.

**DISCUSSION**

The one step mechanism of ligand binding to VP 39 is consistent with many crystallographic data showing the absence of structural rearrangement following binding of m7G and several other cap analogs to both wild-type and F180W VP39s (1, 5–8). The uppermost association rate constant ($k_1$ of $-3 \times 10^7$ M$^{-1}$ s$^{-1}$) observed at pH 5.5 and 6.0 is about 2 orders of magnitude slower than the diffusion rate limit, which appears fast for a binding process that depends almost entirely on the precise and tight insertion of the m7Gua nucleobase ring between two aromatic side chains (Trp-180 and Tyr-22) that are spaced only 6.8 Å apart (Fig. 1). The electron-rich π clouds of both aromatic side chains, which possess a permanent electrostatic quadrupole (12), must therefore impart a strong attractive force for the electron-deficient m7Gua moiety. It is this force that in fact dominates cap binding by cap-specific proteins that rely on stacking interactions (5, 6). Although hydrogen bonds contribute little to the strength of cap binding, they play an important role in dictating rotational orientation in the plane of the stack that apparently depends solely upon the arrangement of polar moieties of the nucleoside with respect to the hydrogen bondable side chains arrayed around the cap-binding slot of VP39 (6).

The stopped flow reaction kinetics of the binding of m7GpppG to eIF4E, although it could not be ascribed strictly to either a one-step or two-step mechanism, indicates also a rapid association rate constant (13), about 10-fold faster than that to VP39 (Table II). The difference in rate constants may in reality be insignificant because the experiment with eIF4E was conducted at higher temperature (20°C). The slightly faster bind-
ing could also be attributed to the involvement of two tryptophan residues of eIF4E in the stacking interactions (9). Tryptophan is believed to be the most efficacious in stacking with the m'G residues of eIF4E in the stacking interactions (9).

In light of the effect of pH (from 5.5 to 8.5) on \( k_1 \) (Fig. 3A), which indicates binding of only the protonated m'G species, Reaction 1 can be modified as follows,

\[
P + \text{LH} \rightleftharpoons \text{PLH}
\]

\[
K_{1} = \frac{[\text{PLH}]}{[\text{P}][\text{LH}]} \quad \text{(Equation 2)}
\]

\[
k_{\text{obs}} = [\text{LH}] \times k_1 + k_{-1} = L_0/(1 + K[H^+]^2) \times k_1 + k_{-1} \quad \text{(Equation 3)}
\]

where LH and L represent the N1 protonated (keto) and deprotonated (enolate) forms, respectively, of m'G, \( L_0 \) corresponds to the sum of both forms (Equation 2), and P refers to the protein. A nonlinear regression was conducted for the pH dependence of \( k_{\text{obs}} \) according to Equation 3 to obtain the optimal values for the three floating parameters \( k_1, k_{-1}, \) and \( K \) as defined in Reaction 3. The \( k_1 (3 \times 10^7 \text{ M}^{-1} \text{s}^{-1}) \) at pH 5.5 resulting from the nonlinear regression is very close to that determined experimentally (Table I). The \( k_{-1} \) of 52 s \(^{-1} \) also agrees very well with that of 56 s \(^{-1} \), which is the average of the experimental values obtained at different pHs (Table I). Moreover, plots of \( k_{\text{obs}} \) versus pH derived from the nonlinear regression and the experimental measurement are very similar (Fig. 3C). This finding strongly supports Reaction 2 or one step binding of only the protonated m'G to VP39. That these plots closely resemble the one shown in Fig. 3A for the effect of pH on \( k_1 \) indicates that the changes in \( k_{\text{obs}} \) are correlated with changes in \( k_1 \).

To fortify our conclusion on the binding mechanism represented by Reaction 2, numerical integration of Equations 2 and

---

**TABLE II**

| Cap analog | \( k_1 \times 10^{-2} \text{ (M}^{-1} \text{s}^{-1}) \) | \( k_{-1} \text{ (s}^{-1}) \) | \( K_{1} \times 10^{-5} \text{ (M}^{-1} \text{s}^{-1}) \) | \( K_{1} \times 10^{-5} \text{ (M}^{-1} \text{s}^{-1}) \) |
|------------|------------------------------------------|-------------------------------|---------------------------------------------|---------------------------------------------|
| m'G        | 3.00 ± 0.17                              | 45 ± 17                       | 6.6                                         | 1.2                                         |
| m'GGpppG   | 0.38 ± 0.01                              | 11 ± 2                        | 3.4                                         | 0.7                                         |
| m'GGpppG(A) | 0.26 ± 0.01                              | 2.4 ± 0.7                     | 10.4                                        | 8.3                                         |
| m'GGpppG(A) | 0.015                                    | 1.7 × 10^{-4}                 | 8.6 × 10^{3}                                |                                             |

*For the stopped flow measurements of the binding of the first three ligands, standard error for \( k_1 \) and \( k_{-1} \) values were obtained from linear regression to the observed \( (k_{\text{obs}}) \) rates measured at various ligand concentrations.

*From equilibrium fluorescence titration \( (K_{1} = 1/K_{p}) \) at pH 5.9 (8).

*From surface plasmon resonance (BIAcore) assay at pH 7.0 assuming a bimolecular reaction (8).

---

**TABLE III**

| pH | m'G | F180W VP39 | Protein-m'G (\( \mu \)) | \( h_{\text{obs}} \text{ (s}^{-1}) \) (simulated) | \( h_{\text{obs}} \text{ (s}^{-1}) \) (measured) |
|----|-----|------------|-------------------------|---------------------------------|---------------------------------|
| 5.5| 5   | 1          | 0.72                    | 207                             | 216 ± 3                        |
| 7.0| 7.5 | 1          | 0.80                    | 277                             | 299 ± 4                        |
| 7.5| 10  | 1          | 0.85                    | 355                             | 368 ± 6                        |
| 12.5| 12.5 | 1   | 0.88                    | 425                             | 444 ± 10                       |
| 15  | 15  | 1          | 0.90                    | 495                             | 519 ± 12                       |
| 7.0| 7.5 | 1          | 0.63                    | 160                             | 128 ± 1                        |
| 15  | 15  | 1          | 0.78                    | 251                             | 246 ± 3                        |
| 30  | 30  | 1          | 0.89                    | 439                             | 443 ± 7                        |
| 45  | 45  | 1          | 0.92                    | 587                             | 614 ± 14                       |
| 60  | 60  | 1          | 0.94                    | 866                             | 826 ± 26                       |
| 8.5| 100 | 2.5        | 1.11                    | 161                             | 130 ± 2                        |
| 200 | 200 | 2.5        | 1.86                    | 210                             | 218 ± 3                        |
| 300 | 300 | 2.5        | 1.92                    | 265                             | 299 ± 4                        |
| 400 | 400 | 2.5        | 2.06                    | 325                             | 360 ± 6                        |
| 500 | 500 | 2.5        | 2.15                    | 396                             | 450 ± 9                        |
Reaction Kinetics of Capped mRNA Binding to a Protein

3 was performed to simulate the kinetic data obtained at different pH values. Initial total concentrations of VP39 and mG are known (Table III), as are the amounts of protonated keto (LH) and deprotonated enolate (L) species at a given pH from the known K values (Reaction 3). Table III lists the kobs values obtained from the simulation at three pHs (5.5, 7.0, and 8.5) and different total ligand concentrations and, for comparison, those obtained experimentally. All kobs values from the simulations, which used a fixed set of kL and kH values, are very similar to those obtained experimentally. The amount of protonated mG predicted by the simulation indicates that, even at pH as high as 8.5, there is more than sufficient amount of the active ligand to saturate all binding sites. This additional cross-check by computer simulation provides further compelling evidence that the reaction between VP39 and mG is represented by Reactions 2 and 3. It is also abundantly clear that the pH effect is simply a perturbation that modulates the concentration of the protonated keto or positively charged form of mG.

How does the deprotonation of the N(1)H of the mGua in the basic pH range and resulting formation of the enolate tautomer diminish ligand binding? We attribute this to the formation of a mesoionic or charge neutral mGua (14), bearing the positive charge that is locked into the five membered ring of the enolate tautomer, whereas the decrease in affinity at pH >7.6 was ascribed to the deprotonation of the positively charged side chain, believed to be from a histidine residue. However, the cap-binding slot of eIF4E contains no histidine residue (9). A more serious flaw of the model of cap recognition and binding whose major features require the enolate tautomer form of the cap and electrostatic interaction between the negatively charged O6 and a positively charged form of the mGua group of the capped ligands. These results further demonstrate for the first time that incorporation of an additional negatively charged center following deprotonation of the N(1)H group at high pH and concomitant formation of the enolate tautomer disallow binding.

Acknowledgments—We thank Dr. Carmen Mouré for technical assistance.

REFERENCES

1. Quiocho, F. A., Hu, G., and Gershon, P. D. (2000) Curt. Opin. Struct. Biol. 10, 78–88
2. Gershon, P. D., Ahn, B.-Y., Garfield, M., and Moss, B. (1991) Cell 66, 1269–1278
3. Schnierle, B. S., Gershon, P. D., and Moss, B. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 2987–2991
4. Barbas, E., and Moss, B. (1978) J. Biol. Chem. 253, 7693–7697
5. Hodel, A. E., Gershon, P. D., and Quiocho, F. A. (1998) Mol. Cell 1, 443–447
6. Hu, G., Gershon, P. D., Hodel, A. E., and Quiocho, F. A. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 7149–7154
7. Hodel, A. E., Gershon, P. D., Shi, X., Wang, S. M., and Quiocho, F. A. (1997) Nat. Struct. Biol. 4, 350–354
8. Hu, G., Oguro, A., Li, C., Gershon, P. D., and Quiocho, F. A. (2002) Biochemistry 41, 7677–7685
9. Marcotrigiano, J., Jingras, E.-C., Sonenberg, N., and Burley, S. K. (1997) Cell 90, 951–961
10. Matsuo, H., Li, H., McGuire, A. M., Pletcher, C. M., Gingras, E.-C., Sonenberg, N., and Wagner, G. (1997) Nat. Struct. Biol. 89, 717–724
11. Lockless, S., Cheng, H.-T., Hodel, A. E., Quiocho, F. A., and Gershon, P. D. (1998) Biochemistry 37, 8564–8574
12. Dougherty, D. A. (1996) Science 271, 163–168
13. Blachut-Okrasinska, E., Bojarks, E., Niedzwiecka, A., Chlebicka, L., Darzynekiewicz, E., Stolarski, R., Stepinski, J., and Antosiewicz, J. M. (2000) Eur. J. Biochem. 267, 487–498
14. Glennon, R. A., and Tejani-Butt, S. (1993) in Chemistry of nucleosides and nucleotides (Townsend, L. B., ed) Vol 2, pp. 1–23, Plenum Press, New York, London
15. Rhoads, R. E., Hellmann, G. M., Remy, P., and Ebel, J.-P. (1983) Biochemistry 22, 6084–6088
16. Cresswell, S. E., Rhoads, R. E., and Guss, D. J. (1989) Biochemistry 28, 8078–8083
17. Jencks, W. P. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 4046–4050