Significance of the first transcribed nucleoside of capped RNA for ligand-induced folding of the cap-binding complex

2005 J. Phys.: Condens. Matter 17 S1495

(http://iopscience.iop.org/0953-8984/17/18/007)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **17** (2005) S1495–S1502 doi:10.1088/0953-8984/17/18/007

Significance of the first transcribed nucleoside of capped RNA for ligand-induced folding of the cap-binding complex

Remigiusz Worch[1], Anna Niedzwiecka[1,2], Janusz Stepinski[1], Marzena Jankowska-Anyszka[3], Catherine Mazza[4], Edward Darzynkiewicz[1], Stephen Cusack[4] and Ryszard Stolarski[1,5]

[1] Department of Biophysics, Institute of Experimental Physics, Warsaw University, 93 Zwirki & Wigury St., 02-089 Warsaw, Poland
[2] Biological Physics Group, Institute of Physics, 32/46 Lotnikow Ave., Polish Academy of Sciences, 02-668 Warsaw, Poland
[3] Faculty of Chemistry, Warsaw University, 1 Pasteura St., 02-093 Warsaw, Poland
[4] EMBL, 6 rue Jules Horowitz, BP181 38042 Grenoble Cedex 9, France

E-mail: stolarsk@biogeo.uw.edu.pl

Received 26 October 2004, in final form 10 March 2005
Published 22 April 2005
Online at stacks.iop.org/JPhysCM/17/S1495

Abstract

Many proteins, including those that bind RNA, change conformation upon binding a ligand, a phenomenon known as induced fit. CBP20, the small subunit of the nuclear cap-binding complex (CBC), recognizes specifically the 5′ cap of eukaryotic mRNA and snRNA. The N- and C-terminal regions of the CBP20 subunit of the human nuclear cap-binding complex only acquire a proper fold in complex with capped RNA. The cap is composed of 7-methylguanosine linked by a 5′-to-5′ triphosphate bridge to the first transcribed nucleoside of the RNA. The significance of the latter for the capped RNA–CBC association and local folding of CBC has been characterized by emission spectroscopy. Fluorescence titration of CBC has been performed for three selected, mono- and dinucleotide mRNA 5′ cap analogues. The measured values of the equilibrium association constant and the corresponding Gibbs free energy depend on the type of the first transcribed nucleoside (purine or pyrimidine), and decrease ∼10-fold in the case of a mononucleotide analogue, 7-methylguanosine triphosphate. However, the total quenching of the intrinsic protein fluorescence is similar for each analogue. Changes of the solvent-accessible CBC hydrophobic surface of CBC on binding of the structurally different cap analogues have been followed using bis-ANS fluorescent probe.

[5] Author to whom any correspondence should be addressed.

1. Introduction

It is well known that the activities of biomolecules are related to their structure. However, more and more examples of proteins intrinsically disordered under physiological conditions have been reported recently [1–3]. Disordered fragments of 30 amino acids and longer were found in 30% of eukaryotic proteins [3], which are often involved in regulation of some metabolic processes and cell signalling [3, 4]. Flexibility of proteins in solution can result in substantial conformational changes upon binding to target molecules, and thus binding and folding are thought of as coupled processes [5]. Intriguingly, large conformational changes seem to be characteristic for RNA-binding proteins, which possess a wide spectrum of specificity and binding affinity [6].

The heterodimeric, human nuclear cap-binding complex (CBC) consists of two subunits: CBP80 (780 residues) and CBP20 (156 residues) [7] (figure 1). Interaction between the two subunits is necessary to confer the RNA-binding ability of CBP20. The latter interacts directly with a unique RNA 5′ terminus (the so-called 'cap', scheme 1) [8], which is composed of 7-methylguanosine (m^7G) connected by a 5′-to-5′ triphosphate bridge with the second nucleoside i.e. the first transcribed nucleoside, guanosine (G), adenosine (A), cytosine (C), uridine (U) or their methylated derivatives. The cap structure, which is attached co-transcriptionally to Pol II transcripts, ensures optimal pre-mRNA splicing and poly-adenylation [11, 12], snRNA nuclear export, mRNA stability [13], and mRNA translation [9, 10].

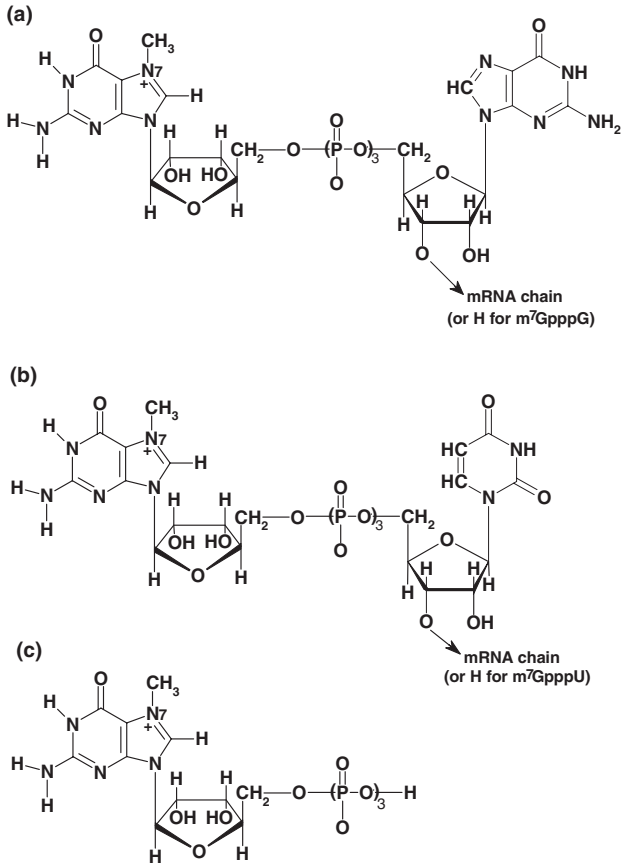
The central part of CBP20 forms a ribonucleoprotein (RNP) domain, found in many other RNA-binding proteins [14]. The 3D structures of *apo*-CBC [7, 15] and CBC in complex with m^7GpppG [15, 16] revealed a large conformational transition of the some 50 residues of the disordered N- and C-termini of *apo*-CBP20 to a folded form upon binding of a cap analogue (figure 1). The observed stabilization of the ligand complex by sandwich stacking of the 7-methyl-guanine with two aromatic amino acid side-chains (Tyr20 and Tyr43), enhanced by numerous hydrogen bonds, is characteristic of other cap-binding proteins, such as eukaryotic translation initiation factor eIF4E [17] and vaccinia virus 2′-O-methyltransferase VP39 [18]. However, stacking of the first transcribed nucleoside of the cap on Tyr138 of CBP20 is a novel element in the recognition of the capped RNA 5′ terminus by cap-binding proteins [15, 16].

To investigate the role of the first transcribed nucleoside in binding of capped RNA to CBC, especially regarding the induced partial folding of the CBP20 disordered regions, fluorescence spectroscopy has been applied. Measurements of quenching of the intrinsic protein fluorescence during titration experiments [19] with selected cap analogues (scheme 1), yielded values of equilibrium association constants (K_{as}) and of the corresponding Gibbs free energies (ΔG°) for both the wild type (wt) CBC and the mutated protein, with tyrosine at position 138 of CBP20 replaced by alanine (Y138A CBC). The conformational changes of CBC upon cap binding have been followed by fluorescence measurements of the environment-sensitive bis-ANS probe as cap analogues are titrated into *apo*-CBC.

2. Materials and methods

Protein and cap analogue preparation. Individual subunit expression, purification and complex assembly of the wild type and Y138A CBC were performed according to Mazza *et al* [7, 16]. Cap analogues were synthesised as described previously [20, 21]. All chemicals used were of analytical grade, purchased from Sigma-Aldrich, Merck, Carl Roth (Germany) or Fluka and Molecular Probes (USA).

Prediction of CBP20 native disorder. The disorder profile of CBP20 was obtained as an output of the DISOPRED2 server with 5% false/positive rate threshold [3].



Scheme 1. Dinucleotide mRNA 5′ cap analogues with (a) guanosine (G) or (b) uridine (U) as the first transcribed nucleoside; (c) mononucleotide cap analogue, m^7GTP.

Spectroscopic measurements and data analysis. Fluorescence time-synchronized titration was performed on LS-50B and LS-55 spectrofluorometers (Perkin Elmer Co., Norwalk, CT, USA) in 50 mM HEPES/NaOH, pH 7.5, 200 mM NaCl, 0.2 mM EDTA and 10 mM DTT, at $20 \pm 0.2\,^{\circ}C$, as described previously [19]. Aliquots (1 μl) of the cap analogue solutions of increasing concentrations were added to 1000 μl of the 0.05 μM CBC solution. The fluorescence signal (excitation at 275 nm, detection at 336 nm) was integrated by 30 s, with 30 s breaks for addition of the ligand. The data have been corrected for the sample dilution (<4.5%) and for the *inner filter* effect [22]. Equilibrium association constants (K_{as}) were determined from each titration by fitting of the theoretical curve describing the fluorescence intensity (F) as a function of the total concentration of the cap analogue ($[L]$) to the experimental data points according to the equation:

$$F = F(0) - [cx](\Delta\phi + \phi_{\text{lig-free}}) + [L]\phi_{\text{lig-free}} \tag{1}$$

where $F(0)$ is the initial fluorescence intensity; $\Delta\phi$, difference between the fluorescence efficiencies of the *apo*-protein and the complex; $\phi_{\text{lig-free}}$, the fluorescence efficiency of the free cap analogue; and the equilibrium concentration of the cap-CBC complex is given by

$$[cx] = \frac{[L] + [P_{\text{act}}]}{2} + \frac{1 - \sqrt{(K_{\text{as}}([L] - [P_{\text{act}}] + 1)^2 + 4K_{\text{as}}[P_{\text{act}}])}}{2K_{\text{as}}} \tag{2}$$

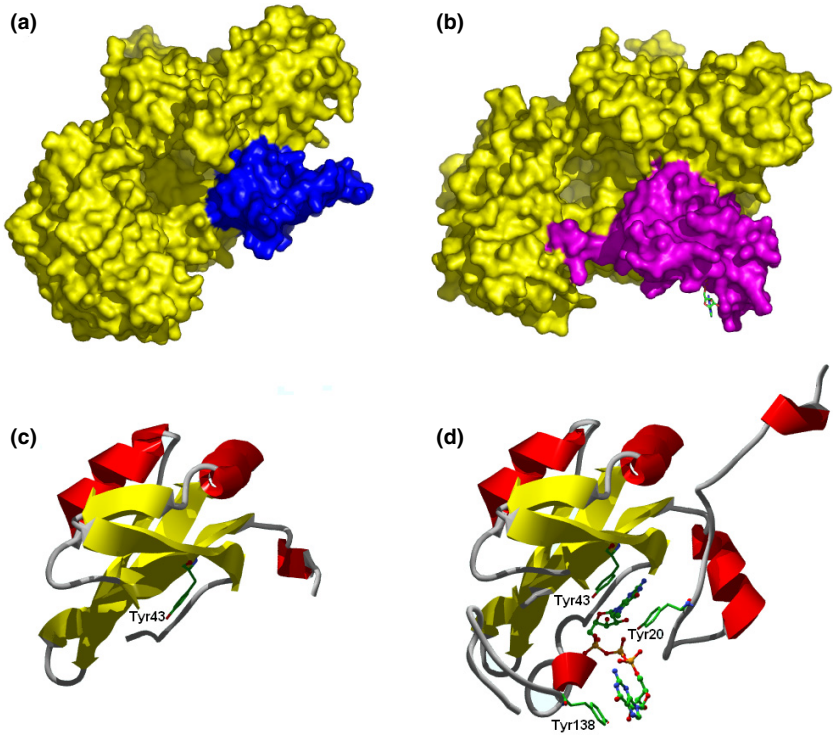


Figure 1. Molecular surfaces of CBC generated by PyMOL (DeLano Scientific LLC, USA); (a) *apo*-form of CBC, (b) CBC in the complex with m^7GpppG. Ribbon diagrams of the CBP20 subunit generated by Swiss-Pdb Viewer [27] (http://www.expasy.org/spdbv) and POV-Ray (http://www.povray.org); (c) *apo*-form of CBP20, (d) CBP20 in the complex with m^7GpppG. Tyr20 and Tyr138 are localized in the disordered regions of the *apo*-protein.

(This figure is in colour only in the electronic version)

where $[P_{act}]$ is the concentration of the active protein. The numerical least-squares nonlinear regression was performed using ORGIN 6.0 (Microcal Software Inc., USA). The final K_{as} values were calculated as weighed averages of $\ln(K_{as}^i)$, where K_{as}^i was the association constant from a single titration. The error of the averaged K_{as} value was calculated from the Student t distribution [23]:

$$\Delta(\ln K_{as}) = t_{\alpha,N} \cdot \sqrt{\frac{\sum_{i=1}^{N}(\ln K_{as}^i - \langle \ln K_{as} \rangle)^2}{N(N-1)}} \tag{3}$$

where N is the number of independent K_{as} values and $t_{\alpha,N}$ is the factor of the Student t distributions for confidence interval $\alpha = 0.6826$, i.e. one standard deviation [23]. The procedure of averaging was chosen to avoid biasing of the results toward lower values, which would occur for the weighed averages of the K_{as} values. The statistically estimated errors of $\ln K_{as}$ allowed for the reliable comparison of the individual results.

The Gibbs free energy of binding ΔG° was calculated as

$$\Delta G^\circ = -RT \cdot \ln K_{as}. \tag{4}$$

Total quenching (Q) was calculated as

$$Q = \frac{[P_{act}] \cdot \Delta\phi}{F(0)} \tag{5}$$

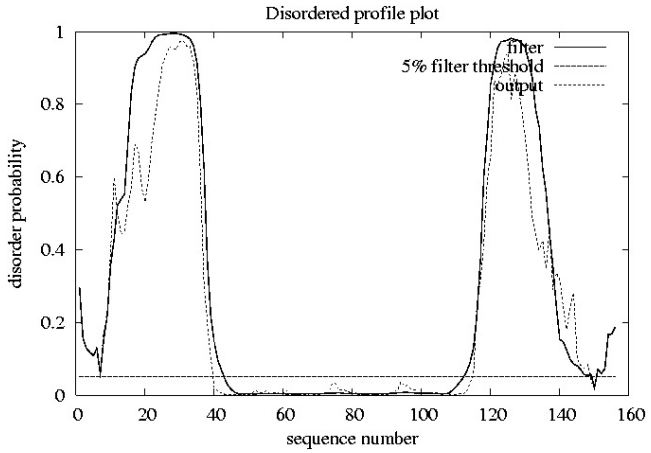


Figure 2. Graphical representation of CBP20 disorder probability derived from DISOPRED2 [3]. The output (short dashed curve) was also filtered using 5% false/positive rate threshold (long dashed line and solid curve) to recover at least 57% of the disordered residues (higher sensitivity).

and the error was calculated due to propagation rules using numerical errors from the fitting procedure.

Experiments with bis-ANS (4, 4′-dianilino-1,1′-binaphthyl-5,5′-di-sulfonic acid, di-potassium salt; Molecular Probes, USA, catalogue no B153) were performed in the same buffer as for the fluorescence titrations. The buffer background was subtracted and the spectra of bis-ANS at 0.40 μM were corrected for the dilution caused by addition of the CBC to the final concentration of 0.05 μM, m^7GTP to 3.57 μM or m^7GpppG to 0.99 μM. The CBC and bis-ANS were incubated for 30 min before addition of the cap analogues. The excitation wavelength was 385 nm, and the excitation and emission slits were 2.5 and 8 nm, respectively.

3. Results and discussion

The crystallographic data showed high flexibility of the N- and C- termini of *apo*-CBC [15, 16]. Prediction of the natively unstructured protein fragments using an algorithm implemented at the DISOPRED2 server [3] confirmed a high disorder probability (nearly 100%) of the regions that contain both Tyr20 and Tyr138 (figure 2). Tyr20 is crucial for the cap recognition in the CBP20 binding centre by stacking with the 7-methyl-guanine moiety. Partial folding induced in CBP20 is related to binding and stabilization of the cap structure, as shown from the comparison of the crystal structures of *apo*-CBC and the CBC–cap complexes [7, 15, 16]. Emission spectroscopy is a successful method to analyse the affinity and energy of CBC–cap interaction by following changes of intrinsic protein fluorescence (figure 3(a)), and to study the changes of the molecular surface by means of bis-ANS, a probe which binds to nonpolar segments in proteins, especially in proximity to positive charges (figure 3(b)). The binding leads to a large increase in the bis-ANS fluorescence quantum yield [24]. The latter feature is commonly applied to monitor protein conformational changes [25].

The equilibrium association constants (K_{as}), the corresponding Gibbs free energies (ΔG°), and the total fluorescence quenching (Q) for three cap analogues, m^7GpppG, m^7GpppU, and m^7GTP, have been derived from the titration experiments (figure 4) and were collected in table 1. The analogues were selected to differ in the structure or the presence of the first transcribed nucleoside. The highest affinity is observed for m^7GpppG, while the

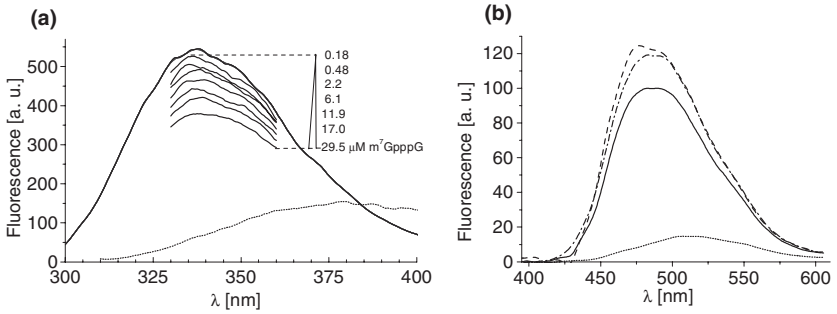


Figure 3. (a) Fluorescence spectra of *apo*-CBC at 1 μM (full length solid curve), CBC in the presence of concentrations of m^7GpppG, from 0.18 to 29.5 μM (solid line fragments), and free m^7GpppG at 13.4 μM (dotted curve), at excitation wavelength of 275 nm. The spectra shown here were not corrected for the inner filter effect due to clarity (for further details see section 2). (b) Fluorescence spectra of free bis-ANS at 0.36 μM (dotted curve), bis-ANS at 0.36 μM in the presence of *apo*-CBC at 0.05 μM (solid curve), and after 99.9% saturation of CBC with m^7GTP (dashed–dotted curve) or with m^7GpppG (dashed curve), at excitation wavelength of 385 nm; addition of a cap analogue to the bis-ANS solution has no effect on the bis-ANS emission.

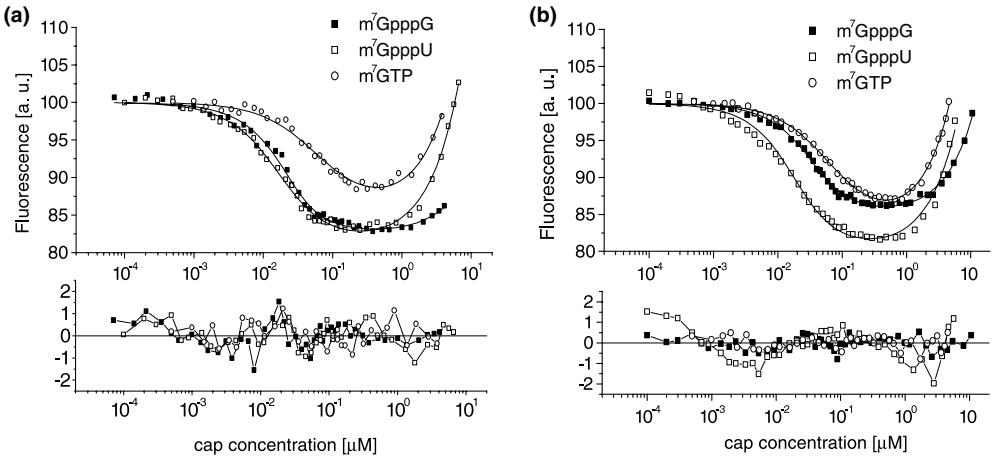


Figure 4. Titration curves and fitting residuals describing interactions of (a) wild type CBC and (b) the Y138A CBC mutant with m^7GpppG, m^7GpppU, and m^7GTP. The experiments were performed at 20 ± 0.2 °C in 50 mM HEPES/NaOH, pH 7.5, 200 mM NaCl, 0.2 mM EDTA and 10 mM DTT. An increasing fluorescence signal at higher concentrations of cap analogues originates from emission of the free cap analogues in solution.

association constants of m^7GTP and of m^7GpppU are ∼8-fold lower (ΔG° less negative by ∼1.2 kcal mol^{-1}) and ∼2.5-fold lower (ΔG° less negative by ∼0.7 kcal mol^{-1}), respectively. Replacement of Tyr138 by alanine results in a ∼4-fold decrease of K_{as} (∼0.8 kcal mol^{-1}) decrease of (ΔG°) for m^7GpppG but does not affect the K_{as} and ΔG° values of both m^7GpppU and m^7GTP. The observed differences in affinity can be ascribed to weaker stacking between Tyr138 and a pyrimidine base (U) compared to a purine base (G), or to the lack of such stacking interaction in the case of m^7GTP.

The total quenching of the CBC fluorescence (Q) due to binding of different cap analogues does not vary significantly. In general, the total quenching is weak, up to 18%. Since fluorescence of the single tryptophan in CBP20 (Trp115) is likely eliminated by the cation-π

Table 1. Association constants (K_{as}), Gibbs energies of binding (ΔG°) and total fluorescence quenching (Q) for interactions of the wild type CBC and its Y138A mutant with three mRNA 5′ cap analogues.

	wt CBC			Y138A CBC		
Cap analogue	K_{as} (10^{-6} M^{-1})	ΔG° (kcal mol^{-1})	Q (%)	K_{as} (10^{-6} M^{-1})	ΔG° (kcal mol^{-1})	Q (%)
m^7GpppG	231 ± 28	-11.205 ± 0.071	18 ± 11	61.4 ± 5.5	-10.434 ± 0.052	17 ± 13
m^7GpppU	97.1 ± 6.9	-10.701 ± 0.041	17 ± 10	89.0 ± 4.2	-10.650 ± 0.027	19 ± 27
m^7GTP	29.7 ± 2.3	-10.012 ± 0.045	14.3 ± 9.7	19.6 ± 0.9	-9.770 ± 0.027	15 ± 18

interaction with the proximal, positively charged Arg91, the observed fluorescence quenching most probably arises from changes of emission of those tyrosines that are engaged in the binding, and/or of other tryptophans due to global conformational rearrangement of CBC. Thus, small changes of the emission of the single Tyr138 upon binding of different cap analogues are probably 'lost' in the high signal background coming from 49 fluorescent residues of CBC. Additionally, large uncertainties of Q are caused by a numerical correlation between the P_{act} and $\Delta\phi$ parameters in the fitting procedure (see materials and methods).

The fluorescence intensity of bis-ANS increases in the presence of CBC compared to the fluorescence of free bis-ANS in the aqueous buffer. This is accompanied by a blue shift of the emission maximum (figure 3(b)) due to interaction of the fluorescent probe with the hydrophobic protein residues. Addition of a mono- or a dinucleotide cap analogue at a concentration high enough to saturate CBC at the level of 99.9% causes further increase of the bis-ANS fluorescence intensity, 24% for m^7GpppG and 19% for m^7GTP. The result is in accordance with the tighter binding of m^7GpppG than m^7GTP to CBC.

Control experiments showed no increase of bis-ANS fluorescence due to the presence of a cap analogue in solution without the protein. These observations can be ascribed to an increase of the solvent-accessible hydrophobic surface area of the whole protein complex upon structuring of the flexible regions of CBP20 due to its association with the cap. Similar changes of bis-ANS fluorescence caused by the binding of either mono- or dinucleotide cap analogues to CBC reflect similar exposure of the CBC hydrophobic surface to the solvent in each case, though the corresponding association constants are substantially different. It was shown by crystallography [15] that the transition from the open form of CBP20 to the closed, cap-bound state buries 667 $Å^2$ of the CBP20 surface area. Our spectroscopic data indicate that, as the net result, the hydrophobic surface area of the entire CBC becomes greater after cap binding. In this regard, fluorescence spectroscopy provides qualitative but unique information on the conformational changes of the entire CBC. The information is hard to obtain from crystallography, since there are no two protein structures for suitable comparison, i.e. *apo* and cap-bound structures, which would possess well defined electron density for all amino acid residues (see e.g. figure 1). The observed increase of the CBC hydrophobic surface can reflect a stimulating effect of the cap on possible interactions of CBC with other proteins, e.g. eukaryotic translation initiation factor eIF4G [26].

4. Conclusions

The type of the first transcribed nucleoside in the mRNA 5′ terminal structure affects the strength of binding to the nuclear cap-binding complex, CBC. Association of a mononucleotide cap analogue (m^7GTP) that lacks the next nucleoside is characterized by a marked, eightfold

decrease of K_{as} (1.2 kcal mol^{-1} decrease of ΔG°) compared with the dinucleotide caps. Conformational rearrangement of the CBC on binding of various cap analogues leads to an increase of the solvent-accessible hydrophobic protein surface. The effect corresponds to the differences in the affinity of the cap analogues for CBC.

Acknowledgments

The project was supported by the state committee for scientific research (KBN, Poland) 3 P04A 021 25, PBZ-KBN 059/T09/10, BST 833/BF.

References

[1] Wright P E and Dyson H J 1999 *J. Mol. Biol.* **293** 321–31
[2] Dunker A K *et al* 2001 *J. Mol. Graph. Modelling* **19** 26–59
[3] Ward J J, Sodhi J S, McGuffin L J, Buxton B F and Jones D T 2004 *J. Mol. Biol.* **337** 635–45
[4] Iakoucheva L M, Brown C J, Lawson J D, Obradovic Z and Dunker A K 2002 *J. Mol. Biol.* **323** 573–84
[5] Dyson H J and Wright P E 2002 *Curr. Opin. Struct. Biol.* **12** 54–60
[6] Leulliot N and Varani G 2001 *Biochemistry* **40** 7947–56
[7] Mazza C, Ohno M, Segref A, Mattaj I W and Cusack S 2001 *Mol. Cell* **8** 383–96
[8] Shatkin A J 1976 *Cell* **9** 645–53
[9] Muthukrishnan S, Both G W, Furuichi Y and Shatkin A J 1975 *Nature* **255** 33–7
[10] Sonenberg N 1988 *Prog. Nucleic Acid Res. Mol. Biol.* **35** 173–207
[11] Konarska M M, Padgett R A and Sharp P A 1984 *Cell* **38** 731–6
[12] Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E and Mattaj I W 1994 *Cell* **78** 657–68
[13] Lewis J D and Izaurralde E 1997 *Eur. J. Biochem.* **247** 461–9
[14] Nagai K, Oubridge C, Ito N, Avis J and Evans P 1995 *Trends Biochem. Sci.* **20** 235–40
[15] Calero G, Wilson K F, Ly T, Rios-Steiner J L, Clardy J C and Cerione R A 2002 *Nat. Struct. Biol.* **9** 912–7
[16] Mazza C, Segref A, Mattaj I W and Cusack S 2002 *EMBO J.* **21** 5548–57
[17] Marcotrigiano J, Gingras A C, Sonenberg N and Burley S K 1997 *Cell* **89** 951–61
[18] Quiocho F A, Hu G and Gershon P D 2000 *Curr. Opin. Struct. Biol.* **10** 78–86
[19] Niedzwiecka A *et al* 2002 *J. Mol. Biol.* **319** 615–35
[20] Darzynkiewicz E, Ekiel I, Tahara S M, Seliger L S and Shatkin A J 1985 *Biochemistry* **24** 1701–7
[21] Stepinski J *et al* 1995 *Nucleosides Nucleotides* **14** 717–21
[22] Parker C A 1968 *Photoluminescence of Solutions* (Amsterdam: Elsevier)
[23] Brandt S 1999 *Data Analysis. Statistical and Computational Methods for Scientists and Engineers* (New York: Springer)
[24] Rosen C G and Weber G 1969 *Biochemistry* **8** 3915–20
[25] Chatterjee S, Ghosh K, Dhar A and Roy S 2002 *Proteins Struct. Funct. Genet.* **49** 554–9
[26] Fortes P, Inada T, Preiss T, Hentze M W, Mattaj I W and Sachs A B 2000 *Mol. Cell* **6** 191–6
[27] Guex N and Peitsch M C 1997 *Electrophoresis* **18** 2714–23