FEBS Letters 580 (2006) 1919–1924

Specificity in DNA recognition by a peptide from papillomavirus E2 protein

Joana Faber-Barata[a,b], Ronaldo Mohana-Borges[c], Luís Maurício T.R. Lima[a,*]

[a] *Faculdade de Farmácia, Universidade Federal do Rio de Janeiro, CCS, sala Bs-34, Ilha do Fundão, 21941-590, Rio de Janeiro, RJ, Brazil*
[b] *Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Ilha do Fundão, 21941-590, Rio de Janeiro, RJ, Brazil*
[c] *Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Ilha do Fundão, 21941-590, Rio de Janeiro, RJ, Brazil*

Received 10 January 2006; revised 13 February 2006; accepted 17 February 2006

Available online 28 February 2006

Edited by Ivan Sadowski

Abstract The E2 proteins of papillomavirus specifically bind to double-stranded DNA containing the consensus sequence ACCG-N_4-CGGT, where N is any nucleotide. Here, we show the binding and recognition of dissimilar DNA sequences by an 18 amino-acid peptide (α1E2), which corresponds to the DNA-recognition helix, α-helix-1. Isothermal DNA binding assays performed with the DNA consensus sequence show saturable curves with α1E2 peptide, and the α1E2 peptide is converted to an ordered conformation upon complexation. Measurements performed with non-specific DNA sequence fail to saturate, a behavior characteristic of non-specific binding. Binding of the α1E2 peptide to these DNA sequences display a different counter-ion dependence, indicating a dissimilar, sequence-dependent mechanism of interaction. Quantitative stoichiometric measurements revealed the specificity in α1E2 peptide recognition of the ACCG half-site, demonstrating capacity for discrimination of nucleic acid bases sequences without the need of a whole protein architecture.

Keywords: Peptide; DNA; Molecular recognition; Papillomavirus; E2 protein

1. Introduction

Molecular recognition in biological systems is a very complex phenomenon. In protein–DNA interactions, recognition of DNA targets is directed by chemical and physical properties of both the protein and DNA, such as hydrogen bonding, electrostatic and van der Waals effects, surface complementarity and ability for structural rearrangement of both macromolecules upon complexation. Moreover, it is generally accepted that a global context, i.e., a DNA-binding motif incorporated into a complete folded protein, is necessary for high affinity and proper specificity. Based on the systematic analysis of solved high-resolution structures of DNA–protein complexes, it has been suggested that there is a recognition code for base-preferences [1–3]. Determination of the capacity for specific molecular recognition in a peptide fragment derived from the interaction region of a DNA-binding protein may increase our understanding of the principles that govern protein–nucleic acid interactions. Moreover, it may provide insights for design of potentially selective ligands for biological targets.

*Corresponding author. Fax: +55 21 2564 7380.
E-mail address: Mauricio@pharma.ufrj.br (L.M.T.R. Lima).

Our approach was to test the ability of a model system, in this case an 18 amino-acid peptide (α1E2), corresponding to the DNA-recognition region (α-helix-1) from the regulatory E2 proteins of papillomavirus, to bind to and recognize its consensus double-stranded DNA. E2 proteins of papillomavirus act as origin recognition factors as well as regulators of early viral transcription in infected cells [4–6]. These proteins have three domains: an N-terminal region (the transactivation domain), responsible for interaction with other viral and cellular factors; an unstructured, highly mobile hinge region linking the transactivation and the DNA-binding domains [7] with major phosphorylation sites, leading to targeted degradation of the E2 protein [8]; and a C-terminal domain, responsible for homo- and hetero-dimerization and DNA binding. This domain recognizes and specifically binds to double-stranded (ds) DNA comprising the palindromic consensus sequence ACCG-N_4-CGGT, where N_4 is a spacer composed of any four nucleotides. It presents an unique architecture of a β-strand barrel composed of four antiparallel β-sheets from each monomer, with two α-helices (α-helix-1 and α-helix-2) aligned perpendicular to the β-barrel [9]. Recognition of a cognate sequence by E2 proteins is mediated by both direct and indirect readout mechanisms. Direct contacts are formed between α-helix-1 of E2 proteins and the major grooves of the DNA, with each helix binding an ACCG site, corresponding to half of a palindrome [9–11]. However, binding is sensitive to the DNA bending induced by the four-base spacer, with no contribution from direct contacts between the large flexible loop (β-hairpin) and the spacer sequence [9,10,12–14].

Here we show that the α1E2 peptide can specifically discriminate between specific and non-specific DNA and is fully capable of ACCG site recognition. Our data reveal that DNA binding and discrimination of nucleic acid bases can be performed by a short peptide fragment derived from a protein, indicating that the greater part of the specificity determinants reside in the peptide, and that a minimal sequence without a full-protein context is sufficient to drive specificity. It may be that such a mechanism is a general feature of other nucleic acid-binding proteins.

2. Materials and methods

2.1. Chemicals

All reagents were of analytical grade. Distilled water was deionized to less then 1.0 μS and filtered through a 0.22 μm pore-size membrane in a water purification system prior to use. All solutions were prepared just before use.

0014-5793/$32.00
doi:10.1016/j.febslet.2006.02.047

2.2. Synthetic peptide

The 18 amino-acid peptide corresponding to the DNA recognition helix (α1E2), with sequence H_2N-GDANTLKCLRYRFKKHCT-COOH, and a non-related peptide (NR-peptide) H_2N-DRGWG-NGCGLFGKGG-COOH (both >95% pure), were synthesized and HPLC-purified by Genemed (CA, USA). Molecular weight was confirmed by mass spectrometry.

2.3. Synthetic oligonucleotides

The synthetic fluorescein-labeled and unlabeled single-stranded oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA, USA). The DNA sequences were 5′GTAACCG-AAATCGGTTGA3′ (E2DBS), 5′GCAACCGACATATATATA3′ (Half-E2DBS), and a non-specific DNA 5′ ACTGTATGAGCAT-ACAGTA3′ (NS-DNA), and their complementary strands. Annealing and determination of their concentrations were performed as described elsewhere [15,16]. Annealed oligonucleotides were stored in the annealing buffer at −20 °C and checked for complete hybridization by subjecting the products to native polyacrylamide gel electrophoresis.

2.4. Circular dichroism spectroscopy

All CD measurements were performed in a Jasco-715 spectropolarimeter, coupled with a Peltier thermocontroller. All spectra were an average of 3 accumulations, with 50 nm/min scan rate, measured at 22 °C, in a 1.00 mm pathlength cuvette. Samples were in 5 mM bis-tris buffer containing 1 mM DTT, pH 7.0, and the buffer contribution was subtracted.

2.5. Titrimetric assay of peptide:DNA interaction

Isothermal titrimetric assay of DNA binding by α1E2 was performed by fluorescence anisotropy measurements as described [17], in 5 mM bis-tris, 1 mM DTT, pH 7.0 (DNA binding buffer), at 22 °C. Aliquots of a concentrated solution of α1E2 peptide were added sequentially to a fixed amount of fluorescein-labeled DNA, homogenized and equilibrated for 2 min prior to anisotropy measurements. In all cases maximal dilution was less than 10%. Anisotropy measurements were performed in an ISS-PC1 spectrofluorimeter (ISS, Champaign, IL, USA), assembled in "L" geometry. Excitation was set to 485 nm and emission was recorded through an orange short-wave cut-off filter WG3-69. Anisotropy values were calculated by the ISS program. For each sample, anisotropy was measured until absolute errors were less than 0.002. All experiments were performed at least three times and presented exactly similar profiles. Fluorescence anisotropy is typically a report of hydrodynamic properties of a fluorophore particle in solution, including in this case both global macromolecule (i.e., DNA) and local (i.e., fluorescein moiety) behaviors. However, both physical (e.g., temperature, pressure) and chemical (e.g., salt, pH) variables may change some properties of the labeled molecule such as local interaction between the fluorescent probe and the DNA, which in the end leads to changes in fluorescence intensity, lifetime and thus in absolute anisotropy value [18].

2.6. Analysis of α1E2c:E2DBS interaction

Isothermal titrimetric DNA binding assays were analyzed considering a simple two-state reversible equilibrium between peptide and DNA, employing the model-free Hill formalism [17,19]. For this analysis, the formalism is

$$\text{Fraction bound} = ([L]^n/K_d^n)/(1 + ([L]^n/K_d^n)) \quad (1)$$

where L is the free peptide concentration and K_d is the apparent dissociation constant for interacting sites. Adjusting Eq. (1) to the binding data gives the Hill coefficient "n" and the K_d. The Hill "n" parameter may not reflect the stoichiometry of binding, since it can include an allosteric component due to conformational transition of one or more of the macromolecules involved in the complex. The equation was adjusted to the data by non-linear least-squares regression using SigmaPlot 2002 (version 8.0, Jandel Scientific Co.).

3. Results

In order to investigate the mechanism involved in the association of α1E2 peptide with DNA, we monitored binding isotherms using fluorescence anisotropy spectroscopy. We titrated an 18 base-pairs (b.p.) DNA fragment with a fluorescein moiety attached to one 5′ DNA strand (F-E2DBS). In this assay, an increase in fluorescence anisotropy means association of peptide to labeled DNA, and formation of species displaying hydrodynamic properties different from free DNA [17]. Data representative of several binding isotherms are shown in Fig. 1. We observed an increase in anisotropy values, with well-defined saturation plateaus. Increasing NaCl concentration leads to a shift in the binding curves to higher peptide concentration (Fig. 1 and Table 1), indicative of a progressive decrease in affinity, as expected for systems where ionic contacts are involved in the interaction [20]. A similar large dependence of the specificity of DNA binding on counter-ion concentration is observed for the full E2c protein [21]. At present, we have no explanation for the salt dependence of Hill "n" values in the binding isotherms. For binding isotherms performed in the absence of salt, the high "n" values might be associated with a large stoichiometry of binding due to non-specific, non-saturable interaction.

It is not surprising that interaction occurs between the E2DBS and α1E2 peptide due to electrostatic attraction, since there is a good correlation between DNA interaction and the distribution of positively charged surfaces of proteins and other polypeptides [20,22]. In addition, dimeric E2c protein can bind to the completely unrelated polyanion heparin [23] as well as to an 18-mer polyAT ds DNA [15], a common feature shared with other regulatory nucleic-acid proteins. In or-

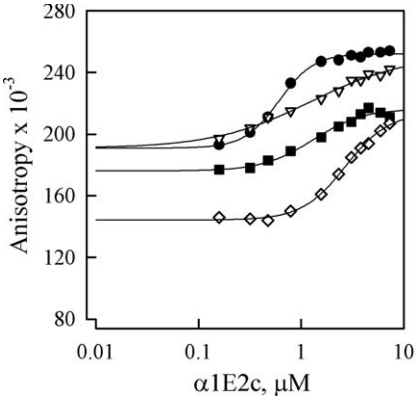

Fig. 1. Binding of α1E2 peptide to a consensus DNA. Representative data for DNA binding isotherms, followed by fluorescence anisotropy of fluorescein-labeled DNA. Assays were carried out in the presence of fluorescein-labeled 5 nM dsDNA (F-E2DBS) and in the absence (circles) or the presence of 20 mM (triangles), 50 mM (squares) and 70 mM (diamonds) NaCl. Solid lines are non-linear regression fits to the raw data using Eq. (1). Details in Section 2.

Table 1
Salt effects on α1E2-peptide: E2DBS interaction

NaCl (mM)	K_d (μM)	Hill "n"
–	0.61 ± 0.028	2.6 ± 0.3
20	1.24 ± 0.28	0.9 ± 0.1
50	1.31 ± 0.20	1.8 ± 0.5
70	2.71 ± 0.16	2.0 ± 0.2

Dissociation constants were calculated from data presented in Fig. 1. Data shown are means ± S.E. Titrimetric analysis were performed as described in Section 2, with 5 nM dsDNA (F-E2DBS), 5 mM bis-tris, 1 mM DTT, pH 7.0, at 22 °C, and varying amounts of α1E2 peptide.