FEBS Letters 355 (1994) 57–60

FEBS 14784

The DNA recognition subunit of a DNA methyltransferase is predominantly a molten globule in the absence of DNA

David P. Hornby[a], Alan Whitmarsh[a], Hatice Pinarbasi[a], Sharon M. Kelly[b], Nicholas C. Price[b,*], Paul Shore[a,**], Geoffrey S. Baldwin[a,**], Jonathan Waltho[a]

[a]*Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, UK*
[b]*Department of Biological and Molecular Sciences, University of Stirling, Stirling, FK9 4LA, Scotland, UK*

Received 15 September 1994

Abstract Enzyme-catalysed DNA methylation provides an opportunity for the modulation of protein-DNA recognition in biological systems. Recently we have demonstrated that the smaller of the two subunits of the heterodimeric, cytosine-specific DNA methyltransferase, M.*Aqu*I, is largely responsible for sequence-specific DNA recognition. Here we present evidence from a series of NMR, fluorescence and circular dichroism spectroscopy experiments that the DNA binding subunit of M.*Aqu*I has the characteristics of a molten globule in the absence of the catalytic machinery. In this metastable state this subunit retains its ability to bind DNA in a sequence-specific manner. We believe this finding offers an insight into the structural flexibility which underpins the mechansim of action of these enzymes, and may provide a possible biological role for molten globules in protein function.

Key words: DNA methylase; Molten globule; Binding specificity; Nuclear magnetic resonance; Circular dichroism

1. Introduction

The specific interaction of proteins with DNA is of fundamental importance in several biological processes, including replication, transcription and recombination. Enzymes that interact with DNA must combine both sequence-specific recognition and turnover during the catalytic cycle. The molecular details of the interaction of cytosine (C5)-specific DNA methyltransferases (C5 MTases) have recently begun to emerge from the elegant X-ray studies of M.*Hha*I [1,2]. In view of the pattern of conserved sequence motifs, especially in the catalytic domain, the latter enzyme is likely to serve as an excellent model for both the prokaryotic and eukaryotic C5 MTases [1].

We have recently demonstrated that the related C5 MTase, M.*Hga*I-2, undergoes a wide spectrum of conformational transitions [3]. In order to dissect further the functional components of these enzymes, we have chosen the heterodimeric C5 MTase from *Agmanellum quadruplicatum* as a model system, since it contains the characteristic set of conserved sequence motifs found in all C5 MTases but differs from all the other enzymes in comprising two different subunits [4]. We have recently demonstrated [5] that the β subunit (comprising sequence motifs IX and X and the target recognition domain, TRD) of M.*Aqu*I is capable of DNA recognition in the absence of the α subunit (sequence motifs I–VIII [6]).

In this paper we present the results of a series of biophysical experiments aimed at examining the structural features of the isolated β subunit of M.*Aqu*I in solution. The conclusion we have drawn is that in the absence of DNA the β subunit exists predominantly as a molten globule [7,8]. This fulfils a prediction [9,10] that nucleic acid binding proteins may adopt a dynamic, highly flexible form in the absence of ligand in order to achieve specificity of binding without the need for very high affinity interaction.

*Corresponding author. Fax: (44) (786) 464 994.

***Present addresses:* P. Shore, Department of Biochemistry and Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne, NE2 4HH, UK: G.S. Baldwin, Department of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.

Abbreviations: C5 MTase, cytosine (C5)-specific DNA methyltransferase; NMR, nuclear magnetic resonance; CD, circular dichroism; AdoMet, *S*-adenosyl methionine; ANS 1-anilinonaphthalene-8-sulphonate; TRD, target recognition domain.

2. Materials and methods

2.1. Protein preparation

The β subunit of M.*Aqu*I was prepared as described elsewhere [5], as was M.*Hga*I-2 [3]. For CD and fluorescence experiments the final concentration of M.*Aqu*I was between 3 and 30 μM, whilst for NMR experiments the protein was at a final concentration of 300 μM. All samples of M.*Aqu*I were dissolved in either distilled water or 10 mM Tris-HCl, pH 7.6, where indicated, and were supplemented, where appropriate, with Aristar grade urea (BDH, UK). M.*Hga*I-2 was used at a concentration of 300 μM dissolved in 100 mM potassium phosphate buffer, pH 7.4.

2.2. Spectroscopy

All NMR spectra were recorded on a Bruker AMX500 spectrometer and CD experiments were performed on a Jasco J600 spectropolarimeter as described previously [3]. Fluorescence spectra were recorded on a Perkin Elmer LS 50 spectrofluorimeter. The fluorescence of protein was excited at 290 nm. Enhancement of ANS (20 μM) fluorescence in the presence of the β subunit was measured with excitation and emission wavelengths of 380 nm and 470 nm, respectively.

3. Results

3.1. Spectroscopic studies on the β subunit of M.AquI

We have recently used CD spectroscopy to identify conformational transitions that accompany ligand binding to the DNA hemimethyltransferase M.*Hga*I-2 [3]. These experiments demonstrated that substantial conformational changes accompany both cofactor and DNA binding. These observations can be rationalised in part in the light of the determination of the structures of liganded and unliganded M.*Hha*I. However, the origin of the predominantly α-helical signal in the far UV spectrum, its reduction by *S*-adenosyl-L-methionine (AdoMet), and

SSDI 0014-5793(94)01171-0

its enhancement by a cognate DNA duplex is unclear since both domains contain α-helical regions. The ability to separate the contribution of both domains by using the individual polypeptides from M.*Aqu*I enables us to dissect the contribution of both domains to the CD spectrum. That the isolated β subunit is capable of sequence-specific DNA recognition is demonstrated in Fig. 1. By titration of a complex between the β subunit and a specfic, end-labelled oligonucleotide duplex with both an unlabelled specific (a 32mer centred on the sequence CCCGGG) and a non-specific (a 32mer centred on the sequence ACCGGT) oligonucleotide duplex, a discrimination factor between specific and non-specific sequences of around 100-fold is observed: a factor which has been observed in all other C5 MTases so far investigated [11].

As can be seen in Fig. 2a, the far UV CD spectrum of the β subunit indicates a significant α-helical content estimated as 23% by the CONTIN procedure [12] and is very similar to that of M.*Hga*I-2 [3]. In contrast, no significant signal in the near UV CD spectrum was obtained with the β subunit of M.*Aqu*I (Fig. 2b). This apparent lack of a coherent near UV CD spectrum with the retention of a distinct far UV spectrum is a characteristic feature of an expanded form of tertiary structure often called a molten globule [7], and we therefore carried out a strategic set of experiments in order to establish whether the β subunit was indeed a molten globule.

Initial attempts to obtain high resolution proton NMR spectra for the β subunit of M.*Aqu*I proved unsuccessful owing to the tendency of the protein to aggregate at concentrations above 50 μM. In order to overcome these problems we tried two strategies. The first involved constructing a deletion derivative of the polypeptide which eliminated the last 25 amino acids. This region of the C-terminus is particularly hydrophobic and the corresponding sequence in M.*Hha*I forms a major interaction with the catalytic domain in M.*Hha*I. Unfortunately, although the truncated polypeptide was significantly more soluble in water than the wild-type subunit, it has little or no affinity for DNA (A.W. and D.P.H., unpublished results). Next, we examined the influence of urea upon the structure and function of the β subunit. Titration of the β subunit with urea revealed that the single tryptophan present in the

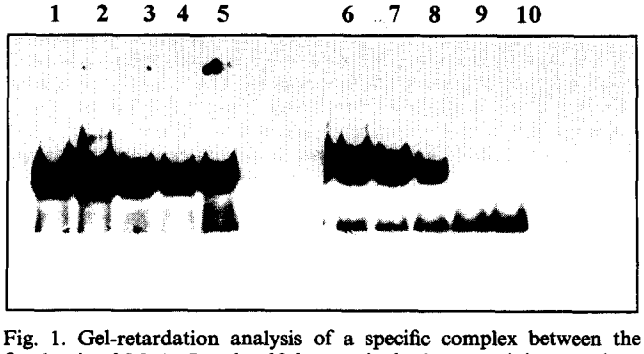
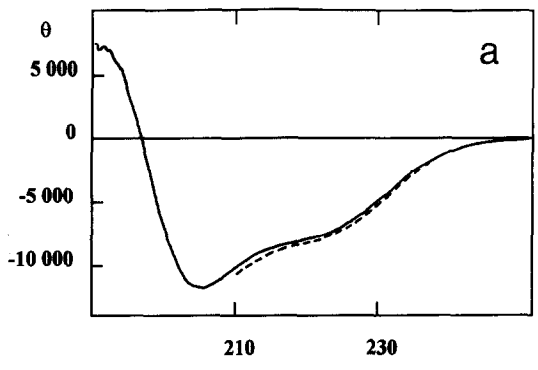


Fig. 1. Gel-retardation analysis of a specific complex between the β subunit of M.*Aqu*I and a 32 base pair duplex containing a unique *Aqu*I site (GTATCTAGCATCGGCCCGGGCCGATGCTAGAT). Lanes 1 and 6 contain the protein and the labelled duplex alone whilst lanes 2–5 contain increasing amounts of a duplex containing the sequence ACCGGT in place of the *Aqu*I site. In lanes 7–10 the complex has been challenged with unlabelled specific duplex. In all lanes 100 ng of protein in 50 mM Tris-HCl, pH 7.6, including 1 mM EDTA have been added to 1 ng of a ^{32}P-labelled specific duplex. In lanes 2–5 and lanes 7–10 the competing duplex is present in the following amounts: 0.1, 1, 10 and 100 ng, respectively.



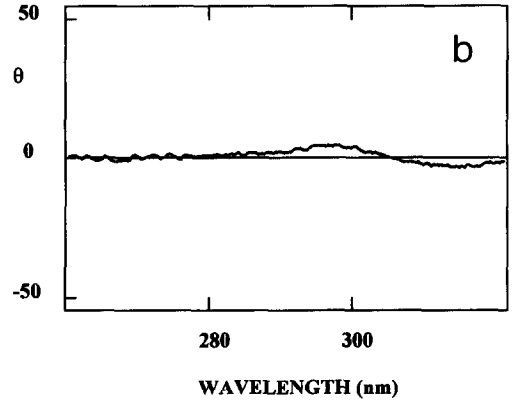


Fig. 2. CD spectra of the M.*Aqu*I β subunit. (a) Far UV CD spectrum. The protein concentration is 3 μM in 10 mM Tris-HCl, pH 7.4, containing 1 mM EDTA in a cell of path length 0.1 cm at 20°C. The solid line represents the protein in the absence of urea whilst the broken line corresponds to the sample in the presence of 2 M urea. In the presence of urea the spectrum is only shown down to a wavelength of 210 nm owing to interference. (b) Near UV CD spectrum. The protein concentration is 30 μM in 10 mM Tris-HCl, pH 7.4, containing 1 mM EDTA and 2 M urea, in a cell of path length 0.5 cm at 20°C. The addition of urea has a negligible effect on this spectrum. On the ordiante, the units of θ are deg·cm^2·dmol^{-1}.

subunit became more exposed at concentrations of urea greater than 3 M (as confirmed by a 4 nm red shift in the emission maximum of the protein fluorescence). Moreover, using a filter binding assay under the conditions described in Fig. 1, there was no decrease in the level of binding of the specific oligonucleotide to the β subunit when up to 2 M urea was included in the assay. This suggested to us that by a judicious choice of urea concentrations we might be able to prevent subunit aggregation whilst retaining the structural and functional integrity of the indiviual subunits. However, the NMR spectrum of the β subunit in 2 M urea shows little chemical shift dispersion; this is typical of a polypeptide in which there is a high degree of rotational freedom of the side chains. This suggests that the β subunit is largely unstructured in a classical sense in the absence of DNA. The comparison in Fig. 3 between the NMR spectra of M.*Aqu*I and M.*Hga*I-2 illustrates the relatively low chemical shift dispersion in the upfield region of the β subunit

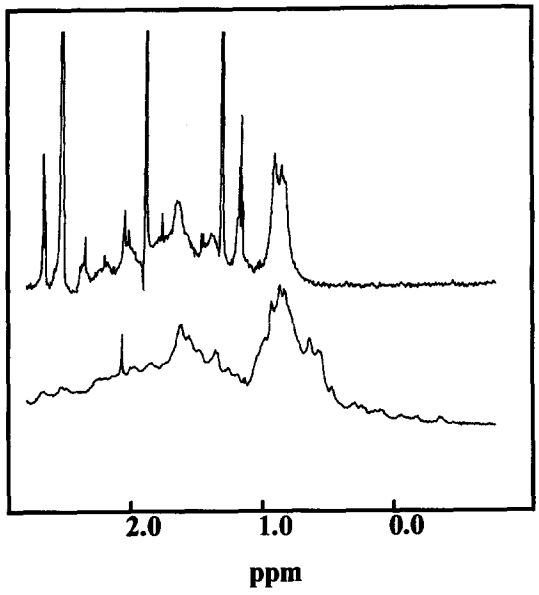


Fig. 3. 500 MHz ^{1}H NMR spectra of DNA methyltransferases. The lower spectrum is that of M.*Hga*I-2 (at a concentration of 300 μM in 100 mM sodium phosphate, pH 7.4): the upper spectrum corresponds to that of the β subunit of M.*Aqu*I (at a concentration of 350 μM in 2 M urea).

spectrum compared with that in M.*Hga*I-2. The difference in the NMR spectra between these species resembles, for example, the loss of chemical shift dispersion observed in the acid denaturation of α-lactalbumin to its molten globule state [13] and is also similar to that observed in studies of recombinant bacterial luciferase and T4 lysozyme [14].

The CD spectrum in the far UV region (Fig. 2a) suggests that the M.*Aqu*I β subunit possesses a substantial degree of secondary structure. The addition of urea to a final concentration of 2 M has no observable effect on the CD spectra in the far and near UV. These observations are consistent [8] with the subunit adopting a molten globular state in 2 M urea. As an additional test to demonstrate that this is the case, the influence of the subunit upon the fluorescence of 1-anilinonaphthalene-8-sulfonate (ANS) was determined. This reagent has been widely used as a spectral probe of the molten globule state since preferential binding (as indicated by enhanced fluorescence) occurs to this state compared with the native or fully unfolded states [7,15]. ANS, when added to the β subunit in 20 mM Tris-HCl buffer at pH 7.4 in the presence or absence of 2 M urea, gives rise to a 10-fold enhancement of ANS fluorescence which is a typical value for a molten globule such as that formed as an intermediate during the folding of carbonic anhydrase [15].

3.2. *The influence of the* β *subunit of M.*AquI *upon DNA*

Owing to the high level of noise observed at wavelengths below 240 nm upon addition of the cognate duplex to the β subunit of M.*Aqu*I, the CD spectra of the complex formed between the β polypeptide and an equivalent concentration of DNA are only shown at wavelengths above 260 nm. However, it is clear from the near UV spectra shown in Fig. 4a and b, that the addition of the β subunit only gives rise to an alteration in the CD spectrum of the DNA if the duplex contains the enzyme's recognition site. Indeed from six independent experiments the $\Delta\theta_{280}$ values determined for the addition of specific and non-specific duplexes are 7.96 ± 1.29% and 0.48 ± 1.11%, respectively. These data are consistent with the finding that the recognition of DNA by the β subunit is sequence-specific and also suggest that a small conformational change in the DNA (presumably representing a localised unwinding of the duplex) accompanies sequence-specific DNA recognition. Similar results have been obtained with the intact, related DNA methyltransferase M.*Hga*I-2 (G.S.B., S.M.K., N.C.P. and D.P.H., unpublised results).

4. Discussion

The results we have obtained from these experiments suggest that the β subunit of M.*Aqu*I adopts a highly dynamic structure with the characteristics of a molten globule [7,8]. It is likely that in vivo the interaction with the α subunit of M.*Aqu*I serves to stabilise the conformation of the β subunit, in a manner somewhat similar to that observed in the case of the heterodimeric luciferase from *Vibrio harveyi* where the individual subunits exist as molten globules but associate to form a complex with stable tertiary structure [14]. Reference to the crystal structure of M.*Hha*I in the absence of DNA [1] shows that there is a network of interactions across the domain interface which

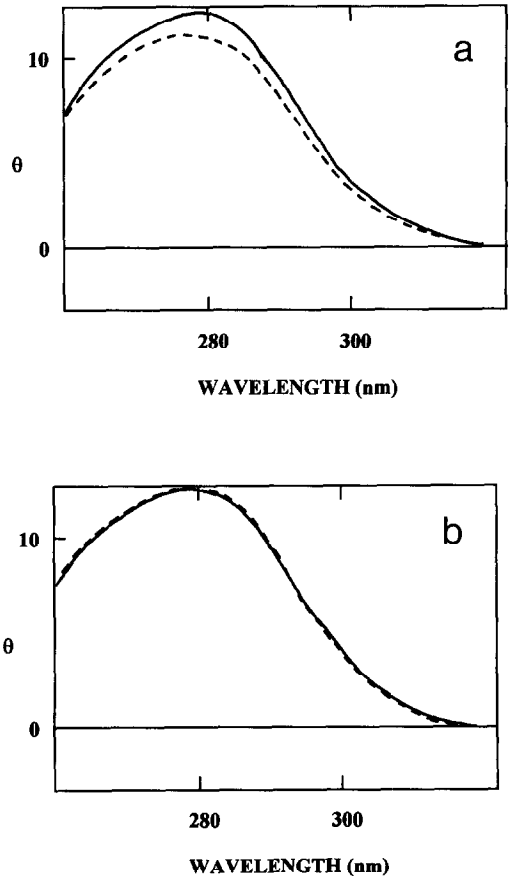


Fig. 4. The influence of the β subunit of M.*Aqu*I upon the near UV CD spectra of specific (a) and non-specific (b) DNA duplexes. Both DNA and protein samples are at a concentration of 3 μM in 10 mM Tris-HCl, pH 7.4, containing 1 mM EDTA and 2 M urea in a cell of path length 0.5 cm at 20°C. The solid and broken lines represent the DNA in the absence and presence of the β subunit respectively. On the ordinate, the units of θ are mdeg.

could be involved in such a stabilisation in the M.*Aqu*I heterodimer. In addition, the C-terminal helix (comprising conserved element X) of M.*Hha*I (and by analogy the related sequence in the β subunit of M.*Aqu*I) is inserted into a hydrophobic channel formed by residues in the large, catalytic domain.

The above types of intermolecular interactions may not be possible in the group of bacteriophage-encoded DNA methyltransferases which contain more than one DNA binding motif, but only a single catalytic site [16]. It is clear that the multi-specific C5 MTases like M.SPR, which can methylate up to three different DNA sequences and contain three corresponding TRDs, possess a great deal of structural plasticity since domain exchange experiments have shown that the various TRDs in this group of enzymes can be shuffled around without loss of function [17]. Since, at any one time, only one of the three TRDs of, for example, M.SPR can be accommodated by the equivalent of the large domain of M.*Hha*I, it is quite possible that these domains adopt a molten globular structure in order to facilitate the conformational transitions which are thought to accompany their activity.

The results in this paper demonstrate for the first time that a polypeptide which exists predominantly as a molten globule is capable of sequence-specific DNA binding. As pointed out previously, this type of flexible, dynamic structure could be important in achieving specificity of binding without requiring undesirably high affinity [9,10]. The conversion of the molten globule to a more rigid tertiary structure on binding to DNA, may in some cases be an alternative, or preferable option to, for example, the helix–coil transition seen in the case of the DNA binding domain of the transcription factor GCN4 [18], or the perturbation of an equilibrium between two rigid structures. We are currently investigating the structural changes which appear to occur on formation of the complex between the M.*Aqu*I β subunit and the cognate DNA duplex.

Acknowledgements: We thank Drs. Xiadong Cheng and Rich Roberts for their kind provision of the coordinates of M.*Hha*I in advance of publication, Dr. Christian Karreman for providing the original plasmids encoding M.*Aqu*I, and Dr. Dan Herschlag for many helpful discussions. We thank the Medical Research Council for their financial support. The Krebs Institute within the Department of Molecular Biology and Biotechnology at Sheffield and the CD laboratory within the Department of Biological and Molecular Sciences at Stirling are supported by the Science and Engineering Research Council.

References

[1] Cheng, X., Kumar, S., Posfai, J., Pflugrath, J.W. and Roberts, R.J. (1993) Cell 74, 299–307.
[2] Klimasauskas, S., Kumar, S., Roberts, R.J. and Cheng, X. (1994) Cell 76, 357–369.
[3] Baldwin, G.S., Kelly, S.M., Price, N.C., Wilson, G., Connolly, B.A. and Hornby, D.P. (1994) J. Mol. Biol. 235, 545–553.
[4] Karrreman, C. and de Waard, A. (1990) J. Bacteriol. 172, 266–272.
[5] Pinarbasi, H., Whitmarsh, A. and Hornby, D.P. (1994) Gene (submitted).
[6] Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R.J. and Wilson G.G. (1994) Nucleic Acids Res. 22, 1–10..
[7] Christensen, H. and Pain, R.H. (1991) Eur. Biophys. J. 19, 221–229.
[8] Ptitsyn, O.B. (1992) in: Protein Folding (Creighton, T.E. ed.) pp. 243–300, Freeman, New York.
[9] Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyanimov, S.Yu. and Ptitsyn, O.B. (1981) FEBS Lett. 136, 311–315.
[10] Schulz, G.E. (1979) in: Molecular Mechansims of Biological Recognition (Balaban, M. ed.) pp. 79–94, Elsevier, Amsterdam.
[11] Dubey, A.K. and Roberts, R.J. (1991) Nucleic Acids Res. 20, 3167–3173.
[12] Provencher, S.W. and Glöckner, J. (1981) Biochemistry 20, 33–37.
[13] Baum, J., Dobson, C.M., Evans, P.A. and Hanly, C. (1989) Biochemistry 28, 7–13.
[14] Flynn, G.C., Beckers, C.J., Baase, F.W. and Dahlquist, F.W. (1993) Proc. Natl. Acad. Sci. USA 90, 10826–10830.
[15] Semisotnov, G.V., Rodionova, N.A., Kutyshenko, V.P., Ebert, B., Blank, J. and Ptitsyn, O.B. (1987) FEBS Lett. 224, 9–13.
[16] Trautner, T.A., Balganesh, T.S. and Pawlek, B. (1988) Nucleic Acids Res. 16, 1649–1658.
[17] Lange, C., Jugel, A., Walter, J., Noyer-Weidner, M. and Trautner, T.A. (1991) Nature 352, 645–648.
[18] Weiss, M.A., Ellenberger, T.E., Wobbe, C.R., Lee, J.P., Harrison, S.C. and Struhl, K. (1990) Nature 347, 575–578.