

# ***N*<sup>ω</sup>-arginine dimethylation modulates the interaction between a Gly/Arg-rich peptide from human nucleolin and nucleic acids**

**Baktisaran Raman, Corrado Guarnaccia, Katalin Nadassy<sup>1</sup>, Sotir Zakhariyev, Alessandro Pintar, Francesco Zanuttin, Dávid Frigyes, Cristina Acatrinei, Alessandro Vindigni, Gábor Pongor<sup>2</sup> and Sándor Pongor\***

International Centre for Genetic Engineering and Biotechnology, Padriciano 99, 34012 Trieste, Italy, <sup>1</sup>EMBL Outstation, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK and Department of Biological and Molecular Sciences, University of Stirling, Stirling FK9 4LA, UK and <sup>2</sup>Department of Theoretical Chemistry, Eötvös Loránd University, PO Box 32, H-1518 Budapest 112, Hungary

Received April 30, 2001; Accepted June 27, 2001

## **ABSTRACT**

**We studied the interaction between a synthetic peptide (sequence Ac-GXGGFGGXGGFXGGXGG-NH<sub>2</sub>, where X = arginine, *N*<sup>ω</sup>,*N*<sup>ω</sup>-dimethylarginine, DMA, or lysine) corresponding to residues 676–692 of human nucleolin and several DNA and RNA substrates using double filter binding, melting curve analysis and circular dichroism spectroscopy. We found that despite the reduced capability of DMA in forming hydrogen bonds, *N*<sup>ω</sup>,*N*<sup>ω</sup>-dimethylation does not affect the strength of the binding to nucleic acids nor does it have any effect on stabilization of a double-stranded DNA substrate. However, circular dichroism studies show that unmethylated peptide can perturb the helical structure, especially in RNA, to a much larger extent than the DMA peptide.**

## **INTRODUCTION**

*N*<sup>ω</sup>-methylation of arginine residues is a widespread post-translational protein modification in eukaryotes (1). Over 50 different proteins have been found to be enzymatically methylated in extracts of human P12 cells and ~90% of methylation occurs at the guanidine side chain of arginine residues (2). The predominant (>90%) product of arginine methylation is *N*<sup>ω</sup>,*N*<sup>ω</sup>-dimethylarginine (asymmetrical dimethylarginine or DMA), produced by type I *N*-methyltransferases (PRMT type I) through an *N*<sup>ω</sup>-monomethylarginine (MMA) intermediate (3). Methylation occurs in conserved glycine + arginine-rich (GAR) sequences (3–5). Most GAR sequences are found in nuclear proteins involved in interactions with RNA (Fig. 1), such as the nucleolar proteins hnRNP A1, fibrillarin and nucleolin.

However, *N*<sup>ω</sup>-methylation of arginine residues is not restricted to this class of proteins and *N*<sup>ω</sup>-methylated arginines have also been identified in myelin basic protein, in the high molecular weight form of fibroblast growth factor-2, in fragile X mental retardation protein and in viral proteins.

Asymmetrical dimethylation of the arginine guanidinium group has several physicochemical effects: (i) it renders the arginine molecule less basic (6), although earlier studies reported pI values for DMA of 10.77, as opposed to 10.02 for arginine; (ii) it increases the hydrophobicity and the solvent-accessible surface of the arginine side chain (6); (iii) it is expected to decrease the hydrogen bonding ability, through replacement of the available H atoms by methyl groups.

At present there is no clear understanding of the exact role that arginine methylation may play in nucleic acid–protein interactions. Valentini *et al.* (7) reported that methylation does not affect specific RNA binding of Hrp1 protein, a member of the hnRNPs. In contrast, Rajpurohit *et al.* (8) reported that binding of recombinant hnRNP A1 protein to single-stranded nucleic acid is reduced upon enzymatic methylation. GAR domains themselves seem to non-specifically bind to nucleic acids, however, the presence of a C-terminal GAR domain is essential for sequence-specific RNA binding to occur in such diverse proteins as nucleolin (7,9), hnRNP A1 (10) and hnRNP U (11). Arginine methylation is also known to facilitate nuclear export of hnRNP proteins (12).

Expression of GAR domains at high levels in prokaryotes is difficult because of the toxicity of these peptides and their susceptibility to protease cleavage. Moreover, prokaryotes lack the protein-arginine methyltransferases involved in post-translational modification of arginines to DMA and only unmethylated GAR domains are accessible in this way. On the other hand, enzymatic methylation *in vitro* of arginine-containing peptides is not quantitative and produces a mixture of products (13,14). These problems have often hampered a direct comparison between methylated and unmethylated peptides in physicochemical and biological studies. We thus opted for a solid phase peptide synthesis approach using either arginine or DMA as building blocks (15). Through this approach we can precisely control the degree of arginine methylation, introduce DMA at discrete sites in the sequence and produce material that is free from any other cellular protein or enzyme that may be difficult to separate due to the very high positive charge of GAR domains.

\*To whom correspondence should be addressed. Tel: +39 040 375 7300; Fax: +39 040 226 555; Email: pongor@icgeb.trieste.it



release of the Gaussian suite of programs (16) for density functional theory at the B3LYP/6-31G\* level, as implemented in Gaussian 98. To model solvation Onsager's reaction field model was used (17 and references therein). The dielectric constant of the solvent (water) was assumed to be 80. Geometry optimizations were performed for all compounds in the reaction field. Natural population analysis was used for generation of natural charges (15,18).

### Statistical analysis of protein–nucleic acid structures

Of the protein–nucleic acid complex structures in the Protein Data Bank (PDB) (19), a non-redundant dataset including 95 structures of protein–dsDNA complexes was selected following the earlier study of Nádassy *et al.* (20). HBPLUS (21) was used to find the hydrogen bonding interactions in these complexes. The program calculates the locus for each donor atom and uses the procedure of Momany and McGuire (22) to position the hydrogen atoms. The default geometric criteria of the program for maximum distances and angles were used in determining potential hydrogen bonds. From the results of HBPLUS the H bonds occurring between the arginine side chain in the protein and backbone phosphate oxygens in DNA were selected and analyzed.

### Synthesis of peptides RGG, DMA-GG and KGG

Peptides were synthesized by a solid phase method using Fmoc chemistry. Peptide DMA-GG was synthesized using Fmoc-*N*<sup>ω</sup>,*N*<sup>ω</sup>-dimethylarginine(Mts)-OH as described (23). Peptides were purified by reverse phase HPLC on a C-18 column (Waters RCM, 25 × 100 mm) using a linear gradient of acetonitrile in water (0–35% over 70 min) containing 0.1% trifluoroacetic acid, their identity confirmed by ESI mass spectrometry (Perkin-Elmer SCIEX API-150EX) and stored as lyophilized powder. Peptide concentrations in double filter binding, DNA melting temperature and circular dichroism experiments were evaluated by amino acid analysis of the corresponding stock solutions.

### Double filter binding assays

The synthetic DNA (5'-GATCTCGCATCACGTGACGAA-GATC and its complement 5'-GATCTTCGTACGTGATGCGAGATC) and RNA [(UG)<sub>12</sub>] oligonucleotides (MWG Biotech, Germany) were labeled using [ $\gamma$ -<sup>32</sup>P]ATP and T4 polynucleotide kinase (New England Biolabs). Samples for filter binding assays were prepared by incubating the nucleic acid substrates (ssDNA, dsDNA and ssRNA) at constant concentration (1 nM) with increasing amounts of peptide (final concentrations 0.2–100  $\mu$ M for RGG and DMA-GG, 9–4500  $\mu$ M for KGG). All binding experiments were carried out in 20 mM sodium phosphate buffer (pH 7.2) at 25°C. Double filter binding experiments were carried out as described by Wong and Lohman (24). Nitrocellulose, DEAE membranes and the 96-well dot-blot apparatus were from Schleicher & Schuell. The amount of free oligonucleotide bound to DEAE membrane and that of peptide–oligonucleotide complex bound to nitrocellulose membrane were quantified with a Canberra Packard Cyclone Phosphorimager system. The estimated error in the  $[DNA_{\text{bound}}]/[DNA_{\text{total}}]$  ratio is  $\pm 0.03$ . Non-specific adsorption on the membranes was not detectable.

### DNA melting curves

Melting temperatures were measured on a Pharmacia Biotech Ultraspec 3000 spectrophotometer equipped with a Peltier heated cell holder and a temperature control unit under computer control. Quartz microcuvettes of 1 cm path length and 200  $\mu$ l working volume were used. Absorbance was monitored at 260 nm between 30 and 60°C, with heating and cooling rates of 0.5°C/min.  $T_m$  values were calculated using SWIFT software (Pharmacia Biotech) with an estimated error of  $\pm 0.5^\circ\text{C}$ . The double-stranded A<sub>25</sub>T<sub>25</sub> DNA was prepared from equimolar amounts of the corresponding synthetic oligonucleotides by annealing at 65°C for 5 min. The concentration of dsDNA in samples was calculated from the absorbance at 260 nm at 60°C, assuming a molar extinction coefficient of 590 000 calculated from the base composition. All samples were in 20 mM sodium phosphate buffer (pH 7). The DNA concentration was 1.7  $\mu$ M and measurements were repeated with peptide:DNA molar ratios of 150, 50 and 10.

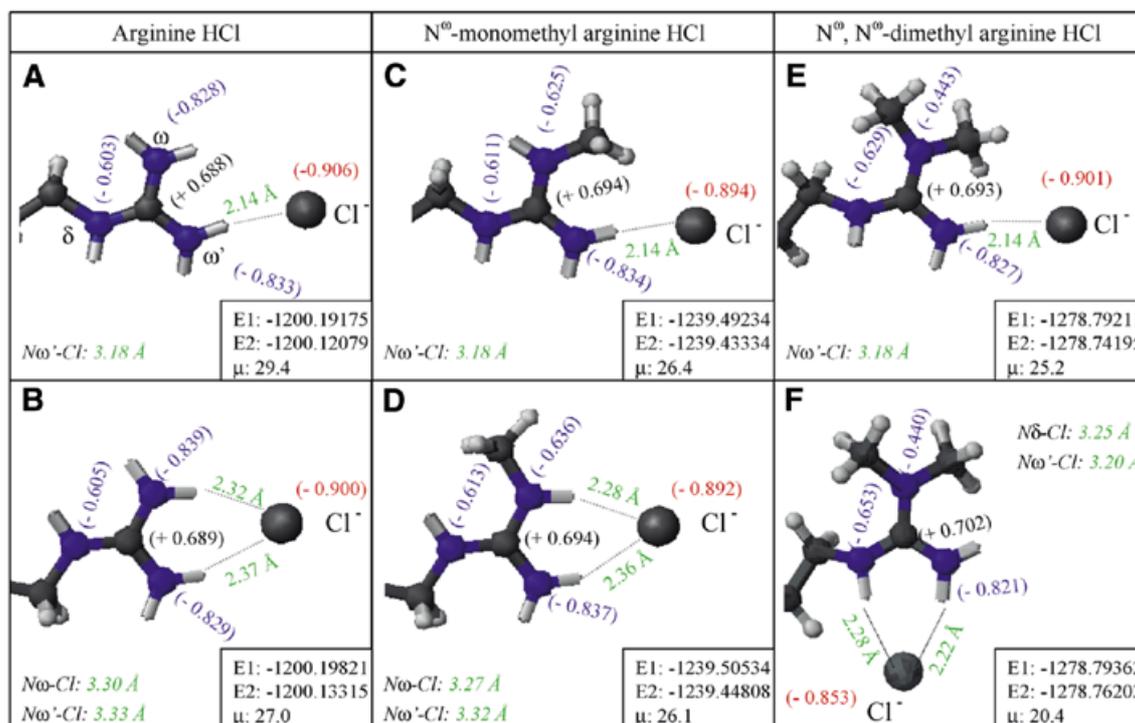
### Circular dichroism spectroscopy

Circular dichroism spectra were recorded between 210 and 310 nm on a Jasco J-650 spectropolarimeter using 10 mm pathlength quartz cuvettes. All experiments were carried out in 20 mM sodium potassium phosphate buffer (pH 7.2) at 25°C. Single-stranded MS2 phage RNA (Fluka) (34  $\mu\text{g}\cdot\text{ml}^{-1}$ ), single- and double-stranded calf thymus DNA (Fluka) (50  $\mu\text{g}\cdot\text{ml}^{-1}$ ) and TAR RNA (GGCCAGAUCUGAGCCUGGGAGCUCUCU-GGCC; MWG Biotech) (30  $\mu\text{g}\cdot\text{ml}^{-1}$ ) were used as nucleic acid substrates. Nucleic acid concentrations were estimated from the OD at 260 nm. Nucleic acid samples were titrated with increasing amounts of each peptide to obtain a final peptide concentration of 0–35  $\mu$ M. Final dilution of the samples upon titration was ~5%. CD spectra were recorded after each addition of peptide and corrected for solvent contributions and dilution. To take into account the different lengths of the substrates, the CD signal was expressed as mean residue mass ellipticity ( $\Delta[\theta]_{\text{MRM}}$ ) at 268 nm.

## RESULTS

### Molecular modeling

The effects of methylation on the partial charges and the H bonding potential of the guanidinium group were first estimated from quantum chemical calculations on the protonated forms of arginine and its methylated derivatives using *N*<sup>ω</sup>-acetyl-L-arginine carboxamide hydrochloride in aqueous solution as the model compound (Fig. 2). The chloride counterion was selected as a simple model of the negatively charged nucleic acid. Although the partial charge of the methylated *N*<sup>ω</sup> atom is slightly altered, the rest of the guanidinium group seems to be little affected by methylation, as is both the length and direction of the electric dipole. Arg, MMA and DMA can all form the single H bonded pattern (Fig. 2A, C and E) and the double H bonded pattern involving hydrogens from  $\delta$  and  $\omega'$  nitrogens (Fig. 2F). In contrast, the double H bonded pattern involving two hydrogens from the  $\omega$  and  $\omega'$  nitrogens (Fig. 2B and D) can be formed by Arg and MMA but not by DMA, as expected. One more effect of *N*<sup>ω</sup>-methylation, as shown by calculation, is loss of planarity of the nitrogen atoms in the guanidinium group. This might result in a decreased delocalization energy



**Figure 2.** Ball and stick models of the side chains of *N*-acetyl-L-arginine amide hydrochloride and its *N*<sup>ω</sup>-methylated derivatives based on quantum chemical calculations in aqueous solution (see Materials and Methods for details). The calculations were performed for the imino-protonated forms and only the head portions of the side chains are shown in the figure. The natural charges on each atom are shown within parentheses in atomic units (1 au = 1.602177 × 10<sup>-19</sup> C). The distance between the NH hydrogen and Cl<sup>-</sup> ion, indicated as a dotted line, is in Å. The distances between the H donor nitrogen atom(s) and Cl<sup>-</sup> ion are also indicated (in italic). (Insets) E1 and E2 are the total energies in hartree units (1 hartree = 4.359748 aJ), including solvent energy and without reaction field, respectively. μ is the dipole moment in Debye units (1 Debye = 3.33564 × 10<sup>-30</sup> Cm).

of the positive charge and, eventually, a decrease in the basicity of the guanidinium group (14).

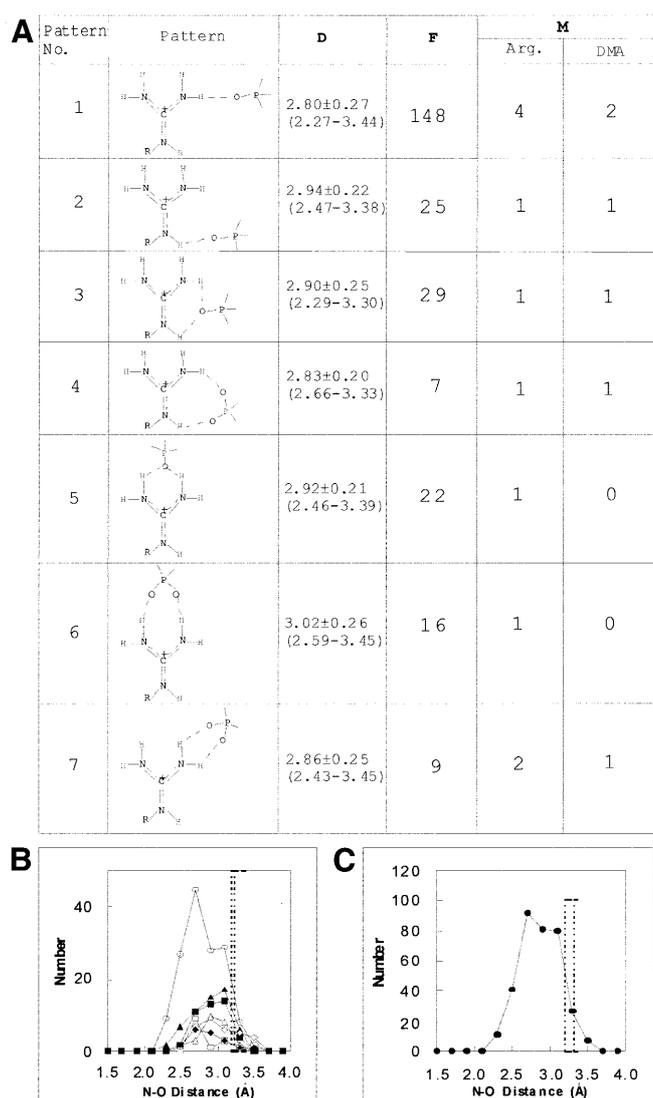
### H bonding patterns of arginine with backbone phosphates of DNA

To verify that the H bonding patterns predicted by calculation are actually observed in protein–DNA complexes, the interactions between the arginine side chain and DNA were examined. The atomic contacts between arginine residues and nucleic acids were collected from high resolution protein–nucleic acid structures from the PDB (19). In the 95 non-homologous protein–DNA complexes selected following the earlier study of Nádassy *et al.* (20) there was a total of 702 contacts between arginine residues and DNA, 658 of which were H bonds formed by the guanidinium group N atoms. Of these 658 H bonds, the majority (395, ~60%) were non-specific interactions with oxygen atoms of the sugar–phosphate backbone and only 10 bonds were found to ribose oxygens. The H bonds were grouped according to patterns (Fig. 3), following the approach of Shimoni and Glusker (25). All H bond patterns that are theoretically possible between a phosphate and a guanidinium group in fact occur in the database. One known structure, the pattern observed between the guanidinium side chain and TAR RNA (26,27), was not found in the selected protein–DNA complexes. This pattern, however, corresponds to a combination of two different patterns (patterns 5 and 7 in Fig. 3A). The arginine side chain usually has more than one possibility to form a particular pattern. We term this number

the ‘multiplicity’ of the pattern; it depends on the number of H atoms available to form the particular H bond pattern. Arg and DMA show notable differences in this multiplicity parameter (Fig. 3A). The multiplicity is significantly lower for DMA than for Arg. Most importantly, some of the patterns (patterns 5 and 6 in Fig. 3A) with double H bonds are excluded for DMA. Figure 3B shows the distribution of the distances between the H donor nitrogen and the phosphate oxygen (N–O distance) in the observed patterns. There is no marked difference in the distribution between the patterns, except that the N–O distances are somewhat higher in patterns having two H bonds. The distances between the H donor nitrogen and the chloride counterion (N–Cl<sup>-</sup> distance) in the structures derived from quantum chemical calculations (Fig. 2) appear to fall within the range but to the higher side of the distribution of N–O distances observed in the database (Fig. 3B and C). This difference could, in part, arise from the difference in size between chloride and oxygen.

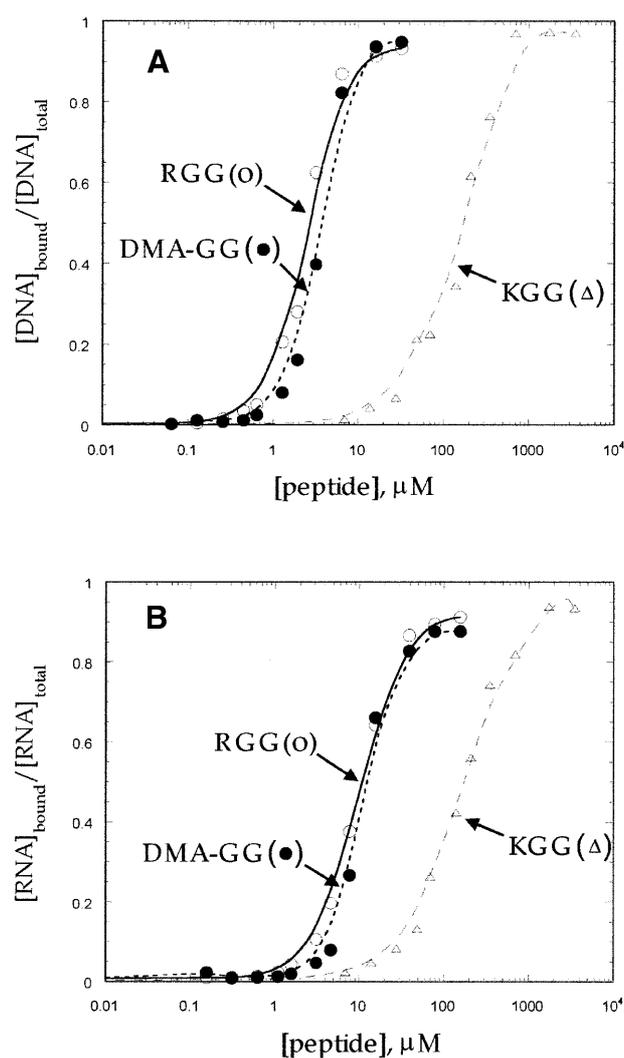
### Nucleic acid binding

Double filter binding assays (24) carried out on ssDNA and ssRNA oligonucleotides (Fig. 4) show that the *K*<sub>d</sub> of the interaction (i.e. the inflection point of the binding curves, assuming a 1:1 stoichiometry) is in the range 5–10 μM for the DMA-GG and RGG peptides and is not substantially altered upon methylation. The RGG peptide seems to bind slightly more strongly in each case, however, the differences are within the estimated experimental error. On the other hand, KGG, a peptide analog



**Figure 3.** Hydrogen bonding interactions between the arginine side chain and a DNA backbone phosphate exhibited in protein-DNA complexes. Protein-DNA complexes (95 structures) in PDB were analyzed (see Materials and Methods for details). (A) Various types of H bonding patterns found. The double and single bonds of phosphate are not distinguished in the structures shown. D is the average H bond distance in Å ± SD. The range of low and high values of the distance is indicated within parentheses. F is the number of the particular pattern found. M is defined as the 'multiplicity parameter' of the pattern, which depends on the number of H atoms available to form the particular H bonding pattern. (B) Distribution of distances between the guanidinium group H donor nitrogens and phosphate oxygens (N-O distances) in the patterns: 1 (open circles), 2 (open triangles), 3 (black triangles), 4 (open squares), 5 (black squares), 6 (open diamonds) and 7 (black diamonds). (C) Distribution of the sum of the N-O distances in all the patterns. Each data point in (B) and (C) represents the indicated distance ± 0.1 Å. The three vertical dotted lines from left to right indicate the average distance between the H donor nitrogen and Cl<sup>-</sup> counterion in the structures obtained by quantum chemistry calculations shown in Figure 6A, B and F, respectively, for comparison.

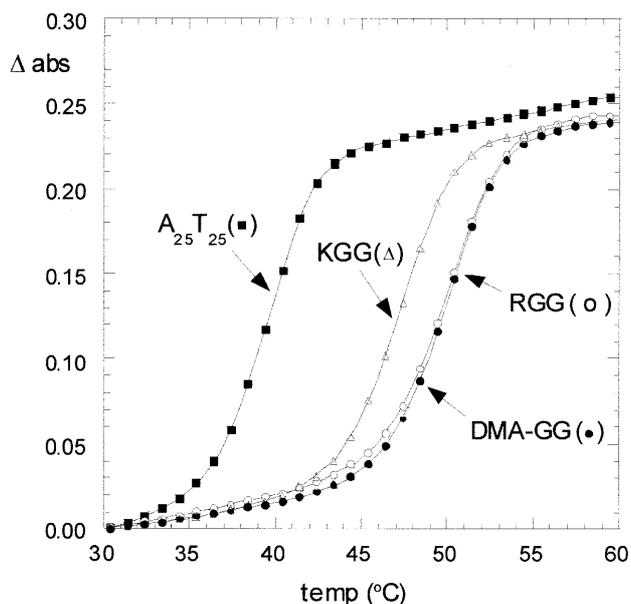
that contained lysine residues in place of all arginine residues, showed an ~50- to 20-fold weaker binding affinity for the RNA and DNA molecules, respectively (Fig. 4). Similar results were obtained with dsDNA (data not shown).



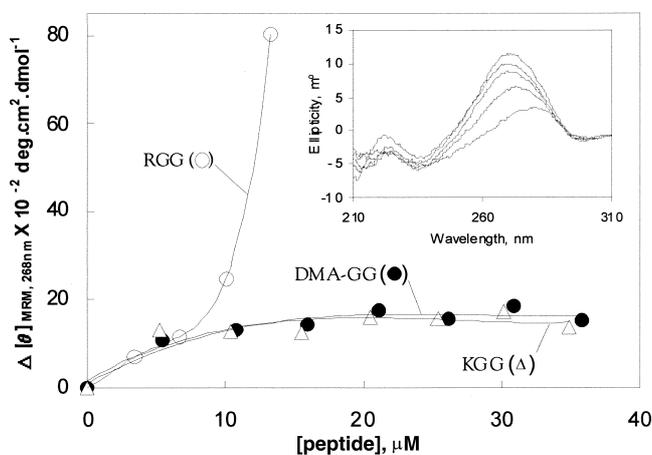
**Figure 4.** Double (nitrocellulose and DEAE) filter binding experiment titrating a nucleic acid fragment (1 nM) with the KGG (open triangles), RGG (open circles) and DMA-GG (black circles) peptides. (A) Experiment in the presence of DNA (5'-GATCTCGCATCACGTGACGAAGATC). (B) Experiment in the presence of RNA [(UG)<sub>12</sub>].

### DNA melting curves

As the filter binding assays suggested that RGG and DMA-GG peptides behave similarly upon nucleic acid binding while binding of KGG peptide is weaker, we conducted an additional test, measurement of DNA melting temperature in the presence of various peptides. These experiments were carried out with a dsDNA oligomer, A<sub>25</sub>T<sub>25</sub>. In the buffer used A<sub>25</sub>T<sub>25</sub> has a reversible melting curve with a unique midpoint at 39.5°C (Fig. 5). The presence of the peptides increased the  $T_m$ , but to different levels. For a 50-fold peptide excess of the RGG and DMA-GG peptides  $T_m$  increased to 50.0°C, whereas the KGG peptide increased the  $T_m$  to 47.5°C. The  $T_m$  showed some dependence on the peptide concentration. In the case of a 150-fold excess of RGG some turbidity appeared and the  $T_m$  could not be reliably measured. Neither DMA-GG nor KGG produced turbidity at similar concentrations. Otherwise, all the transitions appeared to be reversible and unique. The actual magnitude of  $\Delta T_m$  varied with peptide concentration, but the relative



**Figure 5.** Melting curves at 260 nm of the  $A_{25}T_{25}$  dsDNA oligonucleotide alone (black squares) and in the presence of a 50-fold molar excess of KGG (open triangles), RGG (open circles) or DMA-GG (black circles) peptide. The concentration of DNA was  $1.7 \mu\text{M}$ . For each sample the absorbance at  $30^\circ\text{C}$  was subtracted from each data point to account for the peptide contribution and the absorbance difference reported. The cooling curves only are displayed for clarity and every 10th point is marked.



**Figure 6.** Interaction of RGG (open circles), DMA-GG (black circles) and KGG (open triangles) peptides with MS2 phage RNA.  $\Delta[\theta]_{\text{MRM},268 \text{ nm}}$  is the difference in mean residue mass ellipticity at 268 nm of RNA upon interaction with the peptides. (Inset) Circular dichroism spectra of RNA in the presence of increasing concentrations (top to bottom: 0, 3.4, 6.8, 10.0 and  $13.0 \mu\text{M}$ ) of RGG peptide. A  $34 \mu\text{g}\cdot\text{ml}^{-1}$  solution of RNA in 20 mM sodium potassium phosphate buffer (pH 7.2) was used.

differences between the various peptides remained the same (not shown).

### Circular dichroism

The inset in Figure 6 shows the CD spectra of MS2 phage RNA in the absence and presence of increasing concentrations of the RGG peptide described above. Above 250 nm only the RNA and not the peptides exhibited a CD signal, a positive peak

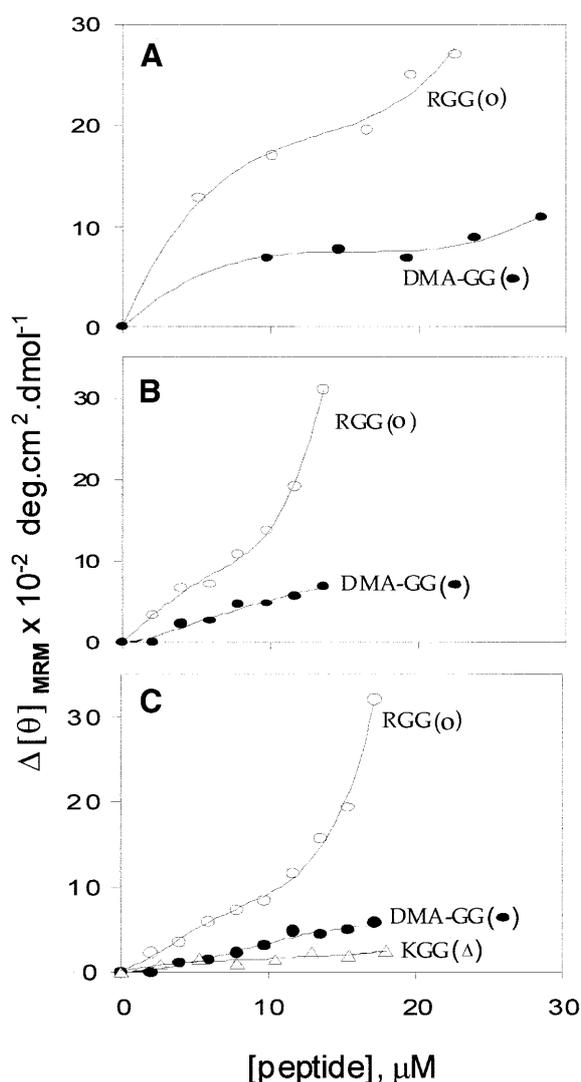
centered at 268 nm. With an increase in peptide concentration the peak shifted towards higher wavelengths, with a reduction in ellipticity. This result indicates that RGG peptide alters the conformation of RNA (28). We used this decrease in RNA  $\Delta[\theta]_{\text{MRM}}$  at 268 nm to characterize the interactions between the peptides and various RNA samples. Figure 6 compares the interactions of RGG and DMA-GG peptides with MS2 phage RNA as a function of peptide concentration. The profile shows a sharp increase in  $\Delta[\theta]_{\text{MRM}}$  above  $10 \mu\text{M}$  RGG peptide. The solution became turbid above the indicated concentrations of peptide, suggesting that the RNA-peptide complex is insoluble or is prone to aggregation. On the other hand, DMA-GG produced a much lower level of  $\Delta[\theta]_{268}$  and neither a sharp increase in  $\Delta[\theta]_{268}$  nor turbidity was observed, even if the peptide was added in high concentrations. Qualitatively similar results were obtained with a short, synthetic TAR RNA molecule (Fig. 7A), with the exception that no turbidity was observed in this case. Finally, we performed CD experiments with single-stranded and double-stranded calf thymus DNA (Fig. 7B and C, respectively): in all cases the DMA-GG and the KGG peptides produced a smaller effect than RGG peptide (Figs 6 and 7C).

### DISCUSSION

As a first step in understanding the physicochemical effects of  $N^{\omega}$ -arginine dimethylation on protein-nucleic acids interactions, calculations were performed on model compounds to derive the partial charges and H bonding capabilities. Dimethylation increases the hydrophobicity, the molecular and accessible surface and the molecular volume of the guanidine moiety (14) and subtly changes the  $\text{p}K_{\text{a}}$  of the arginine side chain (14), but does not alter the total charge and only partially affects the partial charges on the nitrogen atoms and the electric dipole of the guanidinium group. As expected, the most evident effect of  $N^{\omega},N^{\omega}$ -dimethylation is a reduction in the number of possible H bond patterns that can be formed by the guanidinium moiety.

A survey of the available three-dimensional structures of protein-DNA complexes shows that H bonds, especially double H bonded structures, between the arginine side chain and oxygens of the nucleic acid backbone phosphates account for the majority of arginine-DNA contacts. Dimethylation drastically reduces the number of H bonds that arginine can form (Fig. 3) and some of the patterns with double H bonds are excluded. As each H bond can contribute  $1.5 \text{ kcal}\cdot\text{mol}^{-1}$  to binding, we speculated that  $N^{\omega},N^{\omega}$ -dimethylation could eventually reduce non-specific binding of arginine-rich sequences to nucleic acids. However, filter binding assays (Fig. 4) on different substrates using a model peptide revealed no appreciable differences between nucleic acid binding of the  $N^{\omega},N^{\omega}$ -dimethylated and unmethylated peptides. These findings confirm the results of Serin *et al.* (29) and Valentini *et al.* (7), who did not find appreciable differences between binding of dimethylated and unmethylated proteins to nucleic acids. On the other hand, filter binding assays showed that binding of the KGG peptide is 20–50 times weaker. This suggests that non-specific binding of RGG or DMA-GG peptides to nucleic acids is not entirely dictated by the total charge of the side chain.

As the C-terminal region of nucleolin has been reported to be able to partially unwind dsDNA substrates, we also investigated the possible stabilizing/destabilizing effects of the RGG and



**Figure 7.** Interaction of RGG (open circles), DMA-GG (black circles) and KGG (open triangles) peptides with TAR RNA (A) and single-stranded (B) and double-stranded (C) calf thymus DNA. Solutions of 30  $\mu\text{g}\cdot\text{ml}^{-1}$  TAR RNA and 50  $\mu\text{g}\cdot\text{ml}^{-1}$  ssDNA or dsDNA in 20 mM sodium potassium phosphate buffer (pH 7.2) were used.  $\Delta[\theta]_{\text{MRM}}$  is the difference in mean residue mass ellipticity at 265 nm (TAR RNA) or 282 nm (ssDNA or dsDNA) of the nucleotides upon interaction with the peptides.

DMA-GG peptides on a model dsDNA. As measured from melting temperatures (Fig. 5), both the RGG and DMA-GG peptides tended to stabilize dsDNA to the same extent. The reduced stabilization induced by KGG peptide could be interpreted as a consequence of the weaker H bonding capabilities of lysine, but also as a consequence of weaker binding of the KGG peptide. Similar stabilizing effects were observed in the interaction of protamines with DNA (30). Protamines are small, highly positively charged peptides containing rows of arginines and thus bear some similarity to Gly/Phe/Arg-rich peptides from nucleolin. It has been shown that protamines are capable of raising the  $T_m$  of dsDNA and that arginine-containing peptides are more effective than their lysine equivalents. It has also been observed that at proper concentrations

protamines are able to precipitate DNA (30). Interestingly, we also observed some turbidity in the presence of a 150-fold excess of RGG peptide, but not in the presence of DMA-GG or KGG peptide.

Despite the similarities shown in filter binding and melting temperature experiments, circular dichroism spectroscopy (Figs 6 and 7) revealed a clear difference in the behavior of the RGG and DMA-GG peptides.  $N^{\omega},N^{\omega}$ -dimethylation substantially decreases the ability of RGG peptide to modify the conformation of the nucleic acid substrate and this difference is remarkable in the case of the interaction with MS2 phage RNA (Fig. 6). The decrease in intensity of the CD band at 260 nm and its shift to higher wavelengths is usually interpreted as arising from base unstacking. However, DNA melting studies on a simple dsDNA model showed that the presence of the RGG or DMA-GG peptide increases, rather than decreases, the stability of the DNA helix and thus suggest that the changes in the CD spectra are due to conformational changes connected to local perturbations in base stacking, but not necessarily to unwinding and destabilization of the helix structure.

We thus think that binding of RGG peptide, but not that of DMA-GG or KGG peptide, brings about a local deformation of the nucleic acid structure.

RGG motifs in different proteins are found in quite different contexts (Fig. 1). However, there are constant features that suggest that these motifs may form precise contacts with the nucleic acid backbone. For example, a nearly conserved spacing of arginine groups and a phenylalanine residue, which in protein-DNA complexes is frequently found to perturb the helical structure, is invariably present. Phenylalanine residues are usually buried in the hydrophobic core of proteins and rarely occupy positions that are exposed to the solvent. In protein-DNA complexes aromatic residues are frequently found in intercalating positions between two base pairs. It is tempting to speculate that upon DNA binding the aromatic residue comes into contact with the DNA bases. While the positive charges necessary for binding the negatively charged nucleic acid backbone would be provided by arginines, the conserved aromatic residues of the RGG motif might be essential for intercalation into the RNA or DNA structure, but the H bonds necessary to precisely position them could be modulated by  $N^{\omega},N^{\omega}$ -dimethylation of the arginine residues, through the selection of some H bond patterns and not others.

## ACKNOWLEDGEMENTS

We thank Prof. Arturo Falaschi for his advice in the formulation of this project. The work of G.P. was partly supported by the OTKA Research Foundation of Hungary, grant no. T-025830. B.R. is the recipient of an ICGEB post-doctoral fellowship. K.N. is the recipient of a studentship from the University of Stirling, UK.

## REFERENCES

- Aletta, J.M., Cimato, T.R. and Ettinger, M.J. (1998) Protein methylation: a signal event in post-translational modification. *Trends Biochem. Sci.*, **23**, 89-91.
- Najbauer, J. and Aswad, D.W. (1990) Diversity of methyl acceptor proteins in rat pheochromocytoma (PC12) cells revealed after treatment with adenosine dialdehyde. *J. Biol. Chem.*, **265**, 12717-12721.

3. Gary, J.D. and Clarke, S. (1998) RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.*, **61**, 65–131.
4. Klein, S., Carroll, J.A., Chen, Y., Henry, M.F., Henry, P.A., Ortonowski, I.E., Pintucci, G., Beavis, R.C., Burgess, W.H. and Rifkin, D.B. (2000) Biochemical analysis of the arginine methylation of high molecular weight fibroblast growth factor-2. *J. Biol. Chem.*, **275**, 3150–3157.
5. Smith, J.J., Rucknagel, K.P., Schierhorn, A., Tang, J., Nemeth, A., Linder, M., Herschman, H.R. and Wahle, E. (1999) Unusual sites of arginine methylation in poly(A)-binding protein II and *in vitro* methylation by protein arginine methyltransferases PRMT1 and PRMT3. *J. Biol. Chem.*, **274**, 13229–13234.
6. Kennedy, K.J., Lundquist, J.T.T., Simandan, T.L., Kokko, K.P., Beeson, C.C. and Dix, T.A. (2000) Design rationale, synthesis and characterization of non-natural analogs of the cationic amino acids arginine and lysine. *J. Peptide Res.*, **55**, 348–358.
7. Valentini, S.R., Weiss, V.H. and Silver, P.A. (1999) Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation. *RNA*, **5**, 272–280.
8. Rajpurohit, R., Paik, W.K. and Kim, S. (1994) Effect of enzymic methylation of heterogeneous ribonucleoprotein particle A1 on its nucleic-acid binding and controlled proteolysis. *Biochem. J.*, **304**, 903–909.
9. Yang, T.H., Tsai, W.H., Lee, Y.M., Lei, H.Y., Lai, M.Y., Chen, D.S., Yeh, N.H. and Lee, S.C. (1994) Purification and characterization of nucleolin and its identification as a transcription repressor. *Mol. Cell. Biol.*, **14**, 6068–6074.
10. Kiledjian, M. and Dreyfuss, G. (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J.*, **11**, 2655–2664.
11. Cobianchi, F., Karpel, R.L., Williams, K.R., Notario, V. and Wilson, S.H. (1988) Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1. Large-scale overproduction in *Escherichia coli* and cooperative binding to single-stranded nucleic acids. *J. Biol. Chem.*, **263**, 1063–1071.
12. Shen, E.C., Henry, M.F., Weiss, V.H., Valentini, S.R., Silver, P.A. and Lee, M.S. (1998) Arginine methylation facilitates the nuclear export of hnRNP proteins. *Genes Dev.*, **12**, 679–691.
13. Tang, J., Gary, J.D., Clarke, S. and Herschman, H.R. (1998) PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity and regulation. *J. Biol. Chem.*, **273**, 16935–16945.
14. Liu, Q. and Dreyfuss, G. (1995) *In vivo* and *in vitro* arginine methylation of RNA-binding proteins. *Mol. Cell. Biol.*, **15**, 2800–2808.
15. Carpenter, J.E. and Weinhold, F. (1988) Analysis of the geometry of the hydroxymethyl radical by the different hybrids for different spins natural bond orbital procedure. *J. Mol. Struct.*, **169**, 41–62.
16. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheesman, J.R., Zakrzewski, V.G., Montgomery, J.A., Statmann, R.E., Burant, J.C. et al. (1998) *Gaussian 98, Revision A.3*. Gaussian Inc., Pittsburgh, PA.
17. Foresman, J.B. and Frisch, A. (1996) *Exploring Chemistry with Electronic Structure Methods*. Gaussian Inc., Pittsburgh, PA.
18. Glendenning, E.D., Reed, A.E., Carpenter, J.E. and Weinhold, F. (1992) *NBO Version 3.1*.
19. Bernstein, F.C., Koetzle, T.F., Williams, J.B., Meyer, E.F., Jr, Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) The protein data bank. A computer-based archival file for macromolecular structures. *J. Mol. Biol.*, **112**, 535–542.
20. Nádassy, K., Wodak, S.J. and Janin, J. (1999) Structural features of protein-nucleic acid recognition sites. *Biochemistry*, **38**, 1999–2017.
21. McDonald, I.K. and Thornton, J.M. (1994) Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.*, **238**, 777–793.
22. Momany, F., McGuire, R., Burgess, A.W. and Scheraga, H.A. (1975) Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.*, **79**, 2361–2381.
23. Szekely, Z., Zakhariev, S., Guarnaccia, C., Antcheva, N. and Pongor, S. (1999) A highly effective method for synthesis of N<sup>ω</sup>-substituted arginines as building blocks for Boc/Fmoc peptide chemistry. *Tetrahedron Lett.*, **40**, 4439–4442.
24. Wong, I. and Lohman, T.M. (1993) A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc. Natl Acad. Sci. USA*, **90**, 5428–5432.
25. Shimoni, L. and Glusker, J.P. (1995) Hydrogen bonding motifs of protein side chains: descriptions of binding of arginine and amide groups. *Protein Sci.*, **4**, 65–74.
26. Calnan, B.J., Tidor, B., Biancalana, S., Hudson, D. and Frankel, A.D. (1991) Arginine-mediated RNA recognition: the arginine fork. *Science*, **252**, 1167–1171. [Erratum (1992) *Science*, **255**, 665]
27. Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*, **257**, 76–80.
28. Ghisolfi, L., Joseph, G., Amalric, F. and Erard, M. (1992) The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J. Biol. Chem.*, **267**, 2955–2959.
29. Serin, G., Joseph, G., Ghisolfi, L., Bauzan, M., Erard, M., Amalric, F. and Bouvet, P. (1997) Two RNA-binding domains determine the RNA-binding specificity of nucleolin. *J. Biol. Chem.*, **272**, 13109–13116.
30. Saenger, W. (1988) In Cantor, R.C. (ed.), *Principles of Nucleic Acid Structure*. Springer-Verlag, New York, NY.