Journal of Peptide Science
J. Peptide Sci. **11**: 17–28 (2005)
Published online 7 July 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.577

JOURNAL OF PEPTIDE SCIENCE

Efficient synthesis and comparative studies of the arginine and N^{ω},N^{ω}-dimethylarginine forms of the human nucleolin glycine/arginine rich domain

SOTIR ZAHARIEV,[a]* CORRADO GUARNACCIA,[a] FRANCESCO ZANUTTIN,[a] ALESSANDRO PINTAR,[a] GENNARO ESPOSITO,[b] GORDANA MARAVIĆ,[a] BERNARD KRUST,[c] ARA G. HOVANESSIAN[c] and SÁNDOR PONGOR[a]

[a] Protein Structure and Bioinformatics Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), Area Science Park, Padriciano 99, 34012 Trieste, Italy
[b] Dipartimento di Scienze e Tecnologie Biomediche, Universita' di Udine, 33100 Udine, Italy
[c] Unité de Virologie et Immunologie Cellulaire, URA 1930 CNRS, 75724 Paris Cedex 15, France

Received 7 January 2004; Accepted 23 February 2004

Abstract: The Gly- and Arg-rich *C*-terminal region of human nucleolin is a 61-residue long domain involved in a number of protein–protein and protein–nucleic acid interactions. This domain contains 10 aDma residues in the form of aDma-GG repeats interspersed with Phe residues. The exact role of Arg dimethylation is not known, partly because of the lack of efficient synthetic methods. This work describes an effective synthetic strategy, generally applicable to long RGG peptides, based on side-chain protected aDma and backbone protected dipeptide Fmoc-Gly-(Dmob)Gly-OH. This strategy allowed us to synthesize both the unmodified (N61Arg) and the dimethylated (N61aDma) peptides with high yield (~26%) and purity. As detected by NMR spectroscopy, N61Arg does not possess any stable secondary or tertiary structure in solution and N^{ω},N^{ω}-dimethylation of the guanidino group does not alter the overall conformational propensity of this peptide. While both peptides bind single-stranded nucleic acids with similar affinities ($K_d = 1.5 \times 10^{-7}$ M), they exhibit a different behaviour in ssDNA affinity chromatography consistent with the difference in pK_a values. It has been previously shown that N61Arg inhibits HIV infection at the stage of HIV attachment to cells. This study demonstrates that Arg-dimethylated *C*-terminal domain lacks any inhibition activity, raising the question of whether nucleolin expressed on the cell-surface is indeed dimethylated.

Keywords: backbone protection; dimethylarginine; HIV; nucleolin; RGG box

INTRODUCTION

Nucleolin, one of the most abundant proteins in the nucleoli of vertebrate cells, is highly conserved among eukaryotes and carries out multiple functions in ribosome biogenesis, transcription, repression, cell proliferation and growth. It is also involved in embryogenesis, differentiation, systemic autoimmune disorders and shuttling between cytoplasm and nucleus (for reviews see [1–3]).

The *C*-terminal RGG box motif (residues 646–706, ChNu) contains ten Arg that are N^{ω},N^{ω}-dimethylated (aDma) and clustered in nine aDma-Gly-Gly repeats interspersed with six Phe residues (Figure 1). This region controls nucleolar localization, is able to perturb RNA base-stacking and secondary structure [4] and participates in the formation and maturation of preribosomal particles. It also mediates or interacts directly with high affinity and specificity with a subset of ribosomal proteins and nucleic acids [2], in particular four-stranded DNA stabilized by G-quartets (G4 DNA). ChNu is the target of a potent pseudopeptide inhibitor of HIV entry [5] and for cancer therapy [6].

Five different forms of methylated Arg (Figure 2) have been identified: Mma, aDma, sDma and recently δMma [7] and Tma [8]. In extracts of human cells over 50 proteins have been found to be methylated at the Arg residues and the predominant (>90%) product of Arg methylation is aDma, produced by type I *N*-methyltransferases (from PRMT1-1 to PRMT1-4 and PRMT1-6) which transfer methyl groups from S-adenosylmethionine [9].

The exact biological role of Arg methylation, believed to be an irreversible post-translational protein

Abbreviations: As recommended in *J. Pept. Sci.* 2003; **9**: 1–8 and URL: http://www.chem.qmw.ac.uk/iupac/AminoAcid/with the following additions and variations: aDma, N^{ω},N^{ω}-dimethylarginine; BSA, N,O-*bis*-(trimethylsilyl)acetamide; Dmob, 2,4-dimethoxybenzyl; EDT, ethanedithiol; Hmb, 2-hydroxy-4-methoxybenzyl; IE, ion exchange; Mma, N^{ω}-methylarginine; Mts, 2,4,6-trimethylbenzene-1-sulfonyl, mesitylene-2-sulfonyl; Pbf, 2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl; PIP, piperidine; PITC, phenylisothiocyanate; sDma, N^{ω},$N^{\omega'}$-dimethylarginine; TA, thioanisole; TIPS, triisopropylsilane; Tma, trimethylarginine; Tmob, 2,4,6-trimethoxybenzyl; Tms, trimethylsilyl; Tris, tris(hydroxymethyl)amino methane; TSP, 3-trimethylsilyl-[$^2H_4$]-propionate sodium salt. Amino acid symbols denote the L-configuration. All solvent ratios and percentages are volume/volume unless stated otherwise.

* Correspondence to: Dr Sotir Zahariev, ICGEB, Padriciano 99, 34012 Trieste, Italy; e-mail: sotir@icgeb.org

646 706

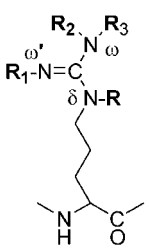
↓-----**N6**----↓----**N5**---↓---**N4**--↓---**N3**--↓---**N2**--↓------**N1**-----↓

H-EGGFGGRGGGRGGFGGRGGGRGGRGGFGGRGRGGFGGRGGFRGGRGGGGDHKPQGKKTKFE-NH_2 **(N61Arg-b)**

H-EGGFGGRGGGRGGFGGRGGGRGGRGGFGGRGRGGFGGRGGFRGGRGGGGDHKPQGKKTKFE-NH_2 **(N61Arg)**

H-EGGFGG**R**GGG**R**GGFGG**R**GGG**R**GG**R**GGFGG**R**G**R**GGFGG**R**GGF**R**GG**R**GGGGDHKPQGKKTKFE-NH_2 **(N61aDma)**

Figure 1 Sequences of the 61 amino acid long peptide amides, corresponding to the *C*-terminal region of human nucleolin containing either R = arginine (N61Arg-b and N61Arg) or **R** = N^{ω},N^{ω}-dimethylarginine (N61aDma). N61Arg-b was synthesized by classical stepwise strategy but was obtained in low yield. In the successful automated stepwise synthesis of N61Arg and N61aDma, a Gly-Gly dipeptide building block protected at the amide bond with the 2,4-dimethoxybenzyl (Dmob) group was used (underlined). For a solid-phase segment assembly attempt the peptide sequences separated by arrows and marked from N1 to N6 were synthesized; N3 and N4 are in fact identical.



i. Mma: R_1=CH_3, R=R_2=R_3=H

ii. sDma: R_1=R_2=CH_3, R=R_3=H

iii. aDma: R_2=R_3=CH_3, R=R_1=H

iv. δMma: R=CH_3; R_1=R_2=R_3=H

v. Tma: undetermined
R=R_1=R_2=CH_3, R_3=H or
R=R_2=R_3=CH_3, R_1=H or
R_1=R_2=R_3=CH_3, R=H:

Figure 2 Different forms of methylated arginine: Mma, sDma and aDma were found in human cells. δMma was identified in yeast proteins. The presence of Tma was demonstrated in a protein from parasitic protozoan *Trypanosoma brucei*, but the position (R – R_3) of methylation is undetermined (three position isomers).

modification in eukaryotes, is still poorly understood [9–11]. Apart from influencing protein–protein interactions, Arg methylation has been also implicated in cell proliferation, signal transduction and protein trafficking [12]. Our recent studies have demonstrated that Arg methylation, although reducing the capacity for hydrogen bonding formation, does not affect the nucleic acid binding. However, aDma containing peptides perturb the helical structure of RNA to a lower extent than the Arg-containing counterparts [13].

High level expression of recombinant ChNu, and of Arg-rich peptides in general, is difficult because of their high basicity on one hand, which makes them toxic to cells, and their susceptibility to proteolytic cleavage on the other hand. Prokaryotes lack PRMT1 and can therefore express only unmethylated proteins. Enzymatic *in vitro* methylation of model RGG-containing peptides or proteins is not quantitative and produces a mixture of products [14,15]. Moreover, ChNu peptide amide is a challenging sequence for SPPS both because of the length and of the high Gly content (34 out of 61 residues, Figure 1); repetitive peptides are known for poor recovery due to low coupling yields [16]. Lastly, Arg-rich peptides are prone to incomplete cleavage and severe adduct formation in the crude stage [17].

Here, an efficient strategy is described for the synthesis of Arg-Gly-Gly- (N61Arg) and aDma-Gly-Gly- rich (N61aDma) 61-amino acid residue peptide amides, corresponding to ChNu (Figure 1). In our hands, stepwise (product N61Arg-b) and convergent approaches to the SPPS of ChNu suffered from low coupling efficiency and solubility/purification problems. To overcome these problems a backbone protected dipeptide Fmoc-Gly-(Dmob)Gly-OH was prepared and successfully used as a building block in the automated SPPS of N61Arg and N61aDma. The purified polypeptides were employed in a set of experiments for further elucidation of the role played by Arg methylation.

MATERIALS AND METHODS

Materials

The TBTU and PyBOP C-activating reagents were from Calbiochem-Novabiochem AG (Switzerland) while HATU was from Applied Biosystems (USA). Anhydrous DMF, NMP and PIP were from Biosolve Ltd (The Netherlands); BSA, DIPEA, TFA, EDT, TA, TIPS, Tms-Br and HOBt from Fluka Chemie AG (Switzerland). All other reagents and solvents were reagent grade and were purchased from Fluka, Sigma-Aldrich (USA), Biosolve Ltd (Netherlands), Alexis (USA) and Advanced ChemTech (USA). Fmoc-protected amino acids were obtained from Novabiochem (Switzerland), Inbios (Italy), Bachem AG (Switzerland) or Applied Biosystems. Side-chain protecting groups were: Boc (Lys), Trt (Gln, His), Bu^t (Ser,Thr), OBu^t (Asp, Glu), Pmc or Pbf (Arg), Mts or Tos (aDma).

Peptide Synthesis

Manual peptide synthesis [18] was carried out in 6 or 20 ml polypropylene syringes (International Sorbent Technology Limited, UK) equipped with polyethylene frits, mounted on a solid phase extraction vacuum chamber (Macherey-Nagel, Germany) for the solvent washes and subjected to continuous nitrogen bubbling.

Automatic peptide synthesis of N61Arg-b was performed on a MilliGen 9050, (Milford, USA) synthesizer. N61Arg and N61aDma were synthesized automatically an a Protein Technologies PS3 peptide synthesizer (Rainin Instrument Co. Inc., USA).

The resins used, generally with substitution 0.16–0.34 mmol/g, were: TentaGel Sieber amide resin (Calbiochem-Novabiochem AG, 0.16 mmol/g) loaded with Fmoc-Glu(OBut)-OH for *C*-terminal amidated peptides N61Arg and N61aDma; TentaGel S-Trt-Gly-Fmoc (Fluka, 0.18 mmol/g) or 2-chlorotrityl-resin (Alexis, USA, 1.47 mmol/g) loaded with Fmoc-Gly-OH (0.34 mmol/g) according to the published procedure [19] for peptides in Table 1. Resin loading was determined by spectrophotometric quantitation [20] of Fmoc groups after cleavage with PIP in DMF and was in agreement with the loading as determined by AAA. The synthesis scales used were 0.05 mmol and 0.1 mmol.

Peptides were synthesized according to standard solid-phase Fmoc chemistry [21] with TBTU or PyBOP as peptide coupling reagent [22] using 4 equivalents of TBTU(PyBOP)/HOBt/Fmoc-Xaa-OH/DIEA (0.95/1.00/1.00/1.90). The coupling time was 45 min. Single coupling was performed throughout the synthesis unless stated otherwise. Coupling efficiency after each cycle of manual SPPS was monitored by the Kaiser ninhydrin test [23]. If required, double couplings were applied using HATU or/and PyBOP as coupling(s) reagents.

Fmoc-deprotection was performed with 20% PIP in DMF (or NMP) for 7 min at a flow rate of 3 ml/min (MilliGen 9050 synthesizer) or for 2 × 5 min with 5 ml 20% PIP/NMP each one (PS3 peptide synthesizer). After peptide chain assembly on the resin, machine-assisted removal of the *N*-terminal amino group protection was performed. The resin was washed with DMF, PrOH(2), DMF, DCM and petroleum ether (fraction 40°–60 °C). Peptide-resins swollen in DCM were placed in polypropylene syringes equipped with polyethylene frits, and cleaved/deprotected: those containing less than five Pbf-side chain protected Arg residues, with TFA/H_2O/TIPS (20 ml/g resin) for 1–12 h at room temperature. Peptides N61Arg, N61Arg-b and N61aDma were deprotected with Tms-Br/thioanisole, for 3.5 h according published procedures [17] with modifications for long peptides [24].

The resin was removed by filtration, washed with small portions of TFA and dried *in vacuo* at room temperature. Peptides were precipitated and washed with diethyl ether, dissolved in 5% acetic acid (25 ml/g peptide-resin), extracted with diethyl ether (6 × 10 ml) and freeze dried. N-side chain protected peptides were cleaved from 2-chlorotrityl resins with HFIP (6 resin volumes) for 1 h and the procedure was repeated twice. The Dmob-amide bond protecting group was cleaved under the same acidic conditions used for deprotection of Hmb [25].

Analytical RP-HPLC was carried out on a Gilson HPLC system, using Zorbax 300 SB 5C18 (150 × 4.6 mm I.D.) columns. Samples were eluted with a linear gradient from A = 0.1% TFA in water to B = 0.1% TFA in MeCN. Preparative RP-HPLC was performed on a Waters RCM with PrepPak Cartridge Delta-Pak 300 15RP18 (100 × 25 mm I.D.) columns at a flow rate of 7 ml/min or on a Waters Prep LC universal base module with a PrepPak Cartridge Delta-Pak 300 15RP18 (100 × 40 mm I.D) column at a flow rate of 18 ml/min. Samples were injected manually and eluted from the column with a gradient slope from 0.4% to 0.6% B/min. Pure fractions, according to analytical RP-HPLC [26] or ESI-MS analysis, were pooled and freeze-dried. IE-HPLC was carried out on AKTA Basic 10 (Amersham Pharmacia, Sweden) using a Shodex IEC SP-825 column (75 × 8 mm I.D., Millipore, MA, USA) equilibrated in 20 mM Na-phosphate buffer pH 8.5. Elution was carried out with a linear NaCl gradient from 0.05 to 0.60 M NaCl at a flow rate of 1 ml/min.

Synthesis of the Gly-Gly Building Block

H-(Hmb)Gly-OH (**3a**) and H-(Dmob)Gly-OH (**3b**) were prepared (Figure 3) according to published procedures [25,27]. After crystallization both **3a**/**3b** always contained occluded Gly. Therefore the reaction mixture from $NaBH_4$ reduction was stirred overnight at pH ~10 with 0.3 equivalents of Dde-OH, acidified, diluted with 1 volume water and extracted with

Table 1 Analytical Data for N^{ω},N^{ω}-Dimethylarginine (aDma)-Containing Peptides from ChNu Sequence

Entry	Sequence	MW calc./exp.	Purity (%) crude	Yield(%) crude/purif.
N2	Fmoc-GF**R**[b] (Tos)GG**R** (Tos)GG-OH	1349.56/1349.7	98	83/67
N3[a]	Fmoc-**R**(Tos)GGFGG**R**(Tos)G-OH[c]	1349.56/1349.6	92	77/—[c]
N5	Fmoc-FGG**R**(Tos)GGG**R**(Tos)GG-OH	1463.67/1464.1	96	82/62
N6	Fmoc-E(But)GGFGG**R**(Tos)GGG**R**(Tos)GG-OH	1763.00/1763.6	96	85/60
N2a	Fmoc-GF**R**(Mts)G(Hmb)G**R**(Mts)GG-OH	1542.83/1543.4	97[d]	86[d]/70
N3a[a]	Fmoc-**R**(Mts)G(Hmb)GFGG**R**(Mts)G-OH	1542.83/1542.7	93[d]	89[d]/61
N5a	Fmoc-FGG**R**(Mts)G(Hmb)GG**R**(Mts)GG-OH	1656.94/1657.4	94[d]	78[d]/60
N6a	Fmoc-E(But)GGFG(Hmb)G**R**(Mts)-GG(Hmb)G**R**(Mts)GG-OH	2093.43/2094.2	94[d]	77[d]/59
N3b[a]	Fmoc-**R**(Mts)G(Dmob)GFG(Dmob G)**R**(Mts)G-OH	1708.05/1708.7	97	87/63
N3c[a]	Fmoc-**R**(Mts)G(AcHmb)GFGG**R**(Mts)G-OH	1584.87/1585.2	89	77/58

[a] Peptides **N3** and **N4** are the same compound and correspond to identical repeats in the sequence of ChNu.
[b] **R** is N^{ω},N^{ω}-Dimethylarginine (aDma).
[c] Hardly soluble in solvents for peptide synthesis; limited solubility in HFIP and HFIP/$CHCl_3$ mixtures.
[d] After Fmoc deprotection and Fmoc reintroduction with Fmoc-OSu (see Results and Discussion).

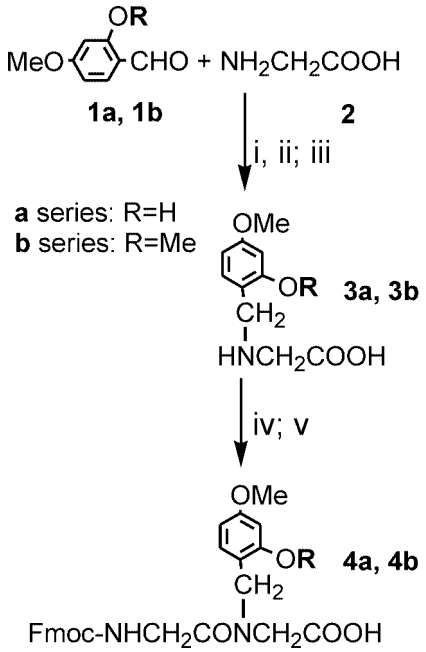


Figure 3 Synthesis of amide bond protected glycyl-glycine building blocks. Dipeptide **4b** was used in this study for successful automated SPPS of N61Arg and N61aDma. Reagents and conditions: (i) **1**/**2**/$NaBH_4$ = 1.1/1.0/1.0 equivalents in THF/H_2O (1 : 1), 4°–10 °C, 2 h. (ii) pH ~ 10, Dde-OH (0.3 equivalents), r.t., 12 h; (iii) pH 3–4, EtOAc extraction, crystallization of **3a** from water phase [25,27]; (iv) BSA(3 eq.)/DCM, rt, 0.5 h; (v) Fmoc-Gly-OPfp/DIEA, 1.0/0.2 eq. in DCM, rt, 12 h or Fmoc-Gly-Cl/DIEA, 1.0/1.0 eq. in DCM, 0 °C for 0.5 h, then 1.5 hr at rt.

EtOAc (10 mmol scale, 3 × 30 ml). The pH was adjusted to 5 and **3a**/**3b** were crystallized as described [25,27]. The yields of **3a**/**3b** were 60%–65%, the purity >99% (TLC or HPLC after precolumn derivatization with Fmoc-Cl). Dipeptides **4a**/**4b** were prepared in 30 mmol scale from *in situ* silylated **3a** or **3b** [28] and Fmoc-Gly-Cl or Fmoc-Gly-OPfp and then purified to homogeneity (>98%) by flash chromatography on SiO_2 ($CHCl_3$/MeOH/AcOH 96/4.5/0.5) with a yield ~80%.

Pilot Synthesis of Gly/Arg- Rich Peptides

Protected dimethylarginines Fmoc-aDma(Mts)-OH and Fmoc-aDma(Tos)-OH were synthesized according to Székely *et al.* [29]. All peptides in Table 1 were prepared by manual synthesis [18]. **N2**–**N6** were prepared by stepwise synthesis, **N2a**–**N6a, N3c** using preformed **4a**, **N3b** using preformed **4b**. Protected peptides (with the exception of **N3**) were analysed by RP-HPLC and LC-MS, and purified by preparative RP-HPLC.

Synthesis of ChNu Peptides

Automated stepwise SPPS of N61Arg-b (0.1 mmol scale) was performed on a MilliGen 9050 synthesizer using Sieber amide resin and a standard Fmoc-protocol. SPPS of N61Arg and N61aDma (0.1 mmol scale) was performed on a PS3 peptide synthesizer using the same resin. **4b** was used for coupling at 14 positions (Figure 1). Crude peptides were cleaved, deprotected and recovered as described in the Peptide Synthesis section, and purified by a two-step procedure using IE-HPLC and RP-HPLC.

Amino Acid Analysis

Peptide-resins or peptides were hydrolysed with 5.5 N HCl containing 2% phenol in tubes sealed *in vacuo* for 24 h at 110 °C. AAA was performed by RP-HPLC after derivatization of the amino acid mixture with phenylisothiocyanate on a PicoTag system (Waters). Mma, sDma, aDma [30] were derivatized separately with PITC. Phenylthiocarbamoyl-derivative of Mma was eluted before Pro, sDma derivative after Pro, and derivative of aDma, after sDma, before and partially overlapped with NH_3.

ESI-MS

Mass spectrometry analysis was performed using an API 150EX electrospray single quadrupole mass spectrometer (Applied Biosystems, Foster City, USA). Samples were analysed in positive mode, by direct syringe infusion (300 µl/h). The samples were either HPLC fractions dissolved in 0.1% TFA in water/MeCN, solid-phase extracted from solutions containing other buffer components using ZipTipC18 (Millipore Corporation, Bedford, USA) and eluted with 0.1% TFA in water/MeCN, or RP-HPLC purified samples, lyophilized and re-dissolved in a mixture of 0.1% HCOOH in MeCN/water (1 : 1). The molecular mass of the peptides was calculated by deconvolution of the ion spectra using the software BioMultiview (Applied Biosystems).

LC-MS

Analytical LC-MS was carried out on a Waters system configured with two model 510 pumps and a Waters automated gradient controller operated at 1 ml/min. The flow was then split (1 : 20) to a Zorbax 300 SB 3.5C18 (150 × 1.0 mm I.D.) column. Samples were eluted with a linear gradient from A = 0.1% TFA in water to B = 0.1% TFA in MeCN, and the column effluent diverted directly to a Applied Biosystems Sciex API 150EX mass spectrometer.

NMR

Samples for NMR spectroscopy were prepared by dissolving the freeze-dried material (N61Arg and N61aDma) in potassium phosphate buffer (0.1 M, pH 6.0) containing 10% D_2O (Aldrich, Gillingham, UK), to give a final peptide concentration of ~0.5 mM. 1D ^{1}H-NMR spectra were acquired at 298 K on a Bruker AVANCE spectrometer operating at a ^{1}H frequency of 500.13 MHz and equipped with a triple resonance probe and z-gradients. The carrier frequency was placed on the solvent resonance and spectra were recorded with a sweep width of 8000 Hz and 8k points. The solvent suppression was achieved either by low power irradiation during the pre-acquisition delay or using a 3-9-19 pulse sequence [31]. Data were zero-filled and processed using XWIN-NMR (Bruker). Spectra were referenced to TSP.

DNA Binding Assays

Double filter binding assays [32] were carried out as described [13]. Briefly, the synthetic DNA (5′-GATCTCGCATCACGT-GACGAAGATC-3′) oligonucleotide (MWG-Biotech, Germany)

was labelled using [γ^{32}P]-ATP and T4 polynucleotide kinase (New England Biolabs, Beverly, USA). Samples for filter binding assays were prepared by incubating the nucleic acid substrates at a constant concentration (1 nM) with increasing amounts of N61Arg or N61aDma (final concentration from 0.2 to 100 μM). All binding experiments were carried out in a 20 mM sodium phosphate buffer (pH 7.2) at 25 °C on nitrocellulose and DEAE membranes in a 96-well dot-blot apparatus (Schleicher & Schuell, Dassel, Germany). The amounts of the free oligonucleotide bound to the DEAE membrane and that of the oligonucleotide/peptide complex bound to the nitrocellulose membrane were detected and quantified using a Cyclone Phosphoimager system (PerkinElmer, Wellesley, USA). The estimated error in the [DNA_{bound}]/[DNA_{total}] ratio was ±0.03. Non specific adsorption on the membranes was not detectable.

For ssDNA affinity chromatography, 0.5 ml of ssDNA agarose (Amersham Pharmacia Biotech, Uppsala, Sweden) was packed in a HR 5/5 column (Amersham Pharmacia Biotech), the column was equilibrated with 20 column volumes of 50 mM Na-phosphate pH 7.4 and after loading the polypeptide N61Arg or N61aDma, was washed with 5 column volumes; the elution step was performed in the same buffer with a gradient of NaCl (from 0 to 1 M in 40 column volumes) at a flow rate of 0.5 ml/min.

Helicase Activity Assay

The assay was performed as described [33,34]. Briefly, the substrate was an M13 phage plasmid annealed either with a 17-mer or a 41-mer labelled oligonucleotide. In the assay 100 pmol of polypeptide N61Arg or N61aDma and 0.5 pmol of the substrate were used in the presence or in the absence of ATP. After incubation the reaction mix was separated by native 12% PAGE and the gel visualized by autoradiography.

In vitro Methylation of N61Arg

The human PRMT1 coding sequence was amplified from the EST plasmid clone 27 498 (I.M.A.G.E. consortium) and cloned into NdeI and EcoRI sites of pET-25(b)+ vector (Novagen, Darmstadt, Germany). 6×His-tagged protein was expressed in *E. coli* HMS174(DE3) cells (Novagen,). At $OD_{600} \sim 1.0$ expression was induced with 1 mM IPTG and carried out for 4 h at 30 °C. Cells were harvested and stored at −80 °C. The cleared bacterial extract was first applied onto Ni-NTA resin (Qiagen S.p.A., Milan, Italy) according to the manufacturer's instructions. Partially purified PRMT1 was then desalted on a HiPrep 26/10 column (Amersham Biosciences Europe GMBH, Milan, Italy) using buffer containing 20 mM Bis-Tris, pH 7.5, 1 mM EDTA and 1 mM DTT and applied to a Shodex IEC QA-825 (75 × 8 mm I.D., Millipore, Bedford, MA, USA). The protein was eluted by NaCl gradient, concentrated in Centriprep® YM-10 concentrators (Millipore) and stored at 4 °C. The protein concentration was determined spectrophotometrically using a PRMT1 calculated extinction coefficient at 280 nm.

The *in vitro* methylation of N61Arg was carried out using 250 μM S-adenosyl-methionine (SAM), 20 μg of N61Arg peptide and 10 μg of recombinant PRMT1 in 50 mM Na-MOPS, pH 7.2, 300 mM NaCl, 2 mM EDTA for a total volume of 150 μl. After 3 h and 72 h incubation at 30 °C the reaction was stopped by the addition of 2% TFA and analysed by LC-MS.

Assay of HIV Infection in HeLa CD4+ Cells

Human HeLa-CD4-LTR-*LacZ* expressing or not expressing CCR5 were referred to as HeLa P4-C5 and HeLa P4, respectively. These HeLa cells (provided by Drs P. Charneau and O. Schwartz; Institut Pasteur, Paris) were cultured in Dulbecco's modified Eagle's medium (Gibco) supplemented with G418 sulfate (500 μg/ml) for the HeLa P4 cells and with G418 sulfate (500 μg/ml)/hygromycin B (300 μg/ml) (Calbiochem-Novabiochem Corp., La Jolla, CA) for the HeLa P4-C5 cells [35]. The T-lymphocyte tropic (X4) HIV-1 LAI isolate was propagated and purified as described previously [36]. The macrophage-tropic (R5) HIV-1 Ba-L isolate [37] was provided by the AIDS Program, NIAID, National Institutes of Health. The MOI of the HIV-1 for different infections was 1.

HIV-1 LAI infection was monitored indirectly in HeLa-CD4-LTR-*lac* Z cells (HeLa P4 cells) containing the bacterial *lac* Z gene under the control of HIV-1 LTR [35]. HIV-1 entry and replication result in the activation of the HIV-1 LTR leading to the expression of β-galactosidase. At 48 h post-infection, the cells were lysed and assayed for β-galactosidase activity using the chlorophenol-red-β-D-galactopyranoside as a substrate. The optical density was measured at 570 nm. The HIV-1 Ba-L infection was monitored in HeLa-CD4-LTR-*lac* Z cells but expressing also CCR5 (HeLa P4-C5 cells) as above.

RESULTS AND DISCUSSION

General Strategy and Synthesis of the Gly-Gly Building Block

As a first approach to the synthesis of N61Arg and N61aDma a solid-phase segment assembly (convergent synthesis) was used due to the highly repetitive character of the Gly/Arg rich region. For example, fragments **N3** and **N4** have the same sequence, which can be found also in **N5** and **N6** (Figure 1, Table 1); in addition, four GFGGRG repeats can be identified within the ChNu sequence. Although the synthesis of these repeats looks straightforward [38], during the manual parallel synthesis of the protected peptides **N2** to **N6**, we observed incomplete coupling at several positions, which leads to truncated peptides that are very difficult to separate from the target product. Moreover, the protected peptide **N3** was poorly soluble in all the most commonly used solvents.

This outcome suggested the need for a reversible 'structure-breaking' protecting group [39]. Recently, two strategies were explored to prevent aggregation when using Fmoc methodology in SPPS: (i) incorporation of pseudoproline dipeptide [40] and (ii) reversible protection of the peptide bond with the Hmb group [25,41] or its analogues (Tmob, Dmob). When introduced before the 'difficult' part of the sequence, both groups have the ability to disrupt the secondary structure (β-sheet) formation and to increase the solvation of the growing peptide chain for at least six amino acid residues from the point of insertion.

Thus analytically pure H-(Hmb)Gly-OH was prepared using a modification of the published procedures (Figure 3). Dipeptide **4a** was prepared in solution with a purity of >90%–95% and isolated with a yield of >70% from the reaction of Fmoc-Gly-Cl [42] or Fmoc-Gly-OPfp with *in situ* silylated H-(Hmb)Gly-OH according to a modified method of Muller *et al.* [28]. H-(Dmob)Gly-OH and H-(Tmob)Gly-OH were obtained in a similar way. With respect to the yield and purity of the products, Dmob-protected Gly was found to be superior to the Tmob analogues. Ten grams of dipeptide **4b** was prepared (Figure 3) and purified by SiO_2 column chromatography. In a model SPPS experiment, the dipeptide was shown to be free of Fmoc-Gly-OH and AcOH.

Pilot Synthesis of Gly/Arg Rich Peptides

Compound **4a** was purified (Figure 3) and used for the parallel manual synthesis of peptides **N2a-N6a** (Table 1). To increase the coupling efficiency in the synthesis of these peptides, TBTU/HOBt [27,43] was used in place of pentafluorophenyl activated esters [41]. Small portions from the peptide resins were cleaved with HFIP and analysed. According to ESI-MS analysis, the main peak (>80%, HPLC) corresponded to the correct sequence but was additionally O-acylated on the Hmb group with the amino acid used in the last coupling step. Peptide-resins were therefore simultaneously N- and O-deprotected with PIP/DMF and then their *N*-termini were selectively reacylated with Fmoc-OSu. The desired peptides **N2a**–**N6a** were obtained in excellent yield and purity (Table 1). Peptides **N3c**, carrying an acetylated Hmb protection [44], were prepared for the convergent synthesis of N61aDma, but the occurrence of acetylation at the Arg guanidino group, as revealed by ESI-MS (not shown), diminished the homogeneity of the crude product (Table 1). This and other discussed shortcomings [27,45] of Hmb pointed to the need of permanently O-blocked versions of Hmb-amide bond protection.

Dmob and Tmob groups have been used in the protection of Asn/Gln carboxamide [46], reduced peptide bond [47] or to improve the solubility of protected peptides [39,46,48]. Moreover, both groups may be cleaved under acidic, mild oxidative or reducing conditions [49], compatible with the Fmoc methodology for SPPS.

4b was therefore used for the preparation of **N3b** (Table 1) and demonstrated its efficacy in peptide solvation (both in SPPS and RP-HPLC) without the disadvantages of Hmb protection.

Peptide Synthesis of N61Arg and N61aDma

N61Arg and N61aDma were successfully synthesized by a standard stepwise automated Fmoc protocol for SPPS using preformed **4b** in 14 positions (Figure 1). Amino

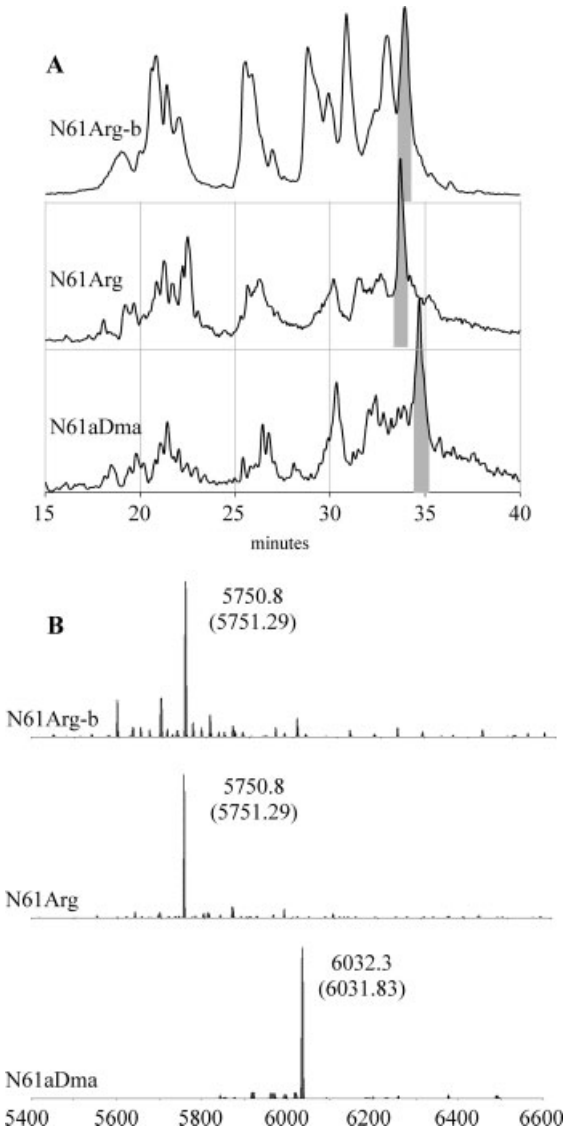


Figure 4 Panel (A) LC-MS analysis of the three crude polypeptides N61Arg-b, N61Arg and N61aDma. In panel (B) the deconvoluted mass spectra and the MW (expected mass in parenthesis) corresponding to the relative products (shaded peaks in A) are reported.

acids/dipeptide were used in 10% molar excess with respect to the coupling reagent (PyBOP). The cleavage/deprotection conditions (Tms-Br/TA/EDT, 3.5 h) of peptide-resins N61Arg and N61aDma were optimized according the published procedure for long peptides [17,24] in order to minimize the formation of *N*-terminal pyroglutamic acid, which represents the main product after 12 h of deprotection. The RP-HPLC chromatograms of crude N61Arg and N61aDma are shown in Figure 4; both products were purified by cation exchange chromatography with a yield of ~31%–35% (Figure 5), then desalted/purified by RP-HPLC to homogeneity (>95%) (Figure 6A) with a yield of ~26%. For comparison, an automated stepwise synthesis of N61Arg-b by standard Fmoc protocol [21]

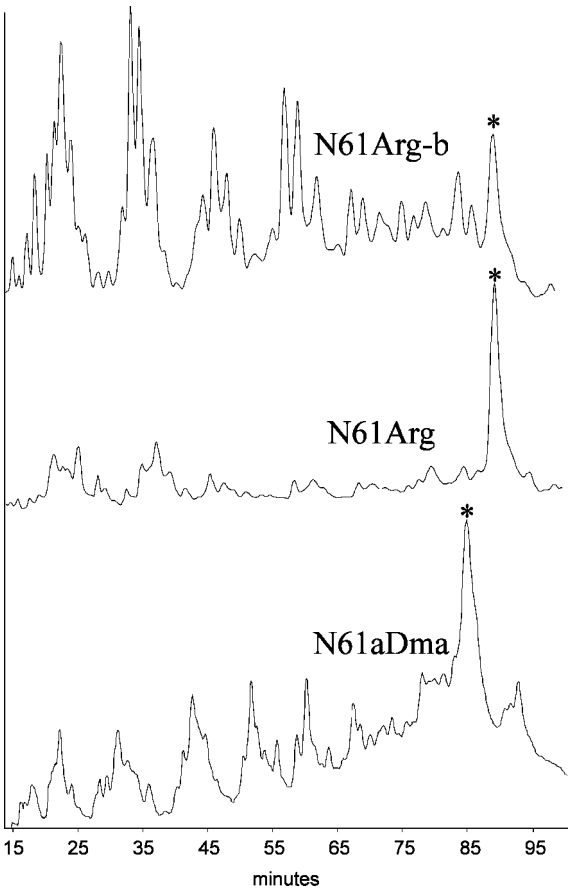


Figure 5 IE-HPLC separation of the three crude polypeptides N61Arg-b, N61Arg and N61aDma. The asterisk indicates the peak containing the correct product. For chromatographic conditions, see Materials and Methods.

for SPPS gave a final yield of ~4%–6% (Figures 4 and 5) and the separation of N61Arg-b from desGly truncated peptides turned out to be impossible. The purity and identity of the peptides were confirmed by ESI-MS (Figure 4B), amino acid analysis, SDS-PAGE (Figure 6B) and PAGE in acetic acid/urea (not shown). LC-MS analysis of the main peaks from IE-HPLC of N61Arg and N61aDma showed the presence of very low amounts of desGly truncated impurities (Figure 4B). It seems that this deletion cannot be completely avoided, unless preformed protected triglycyl-building blocks are applied for coupling on all three positions containing this sequence (Figure 1). Alternatively, solid-phase segment assembly (convergent synthesis) of ChNu's, using Dmob-amide bond protection as a solubilizing group, could be applied.

NMR

In the spectrum of N61Arg in H_2O/D_2O solution, the large majority of the backbone NH signals are clustered between 8.1 and 8.6 ppm, while the aromatic protons of the six Phe residues give rise to the sharp multiplets between 7.15 and 7.35 ppm. The remaining isolated

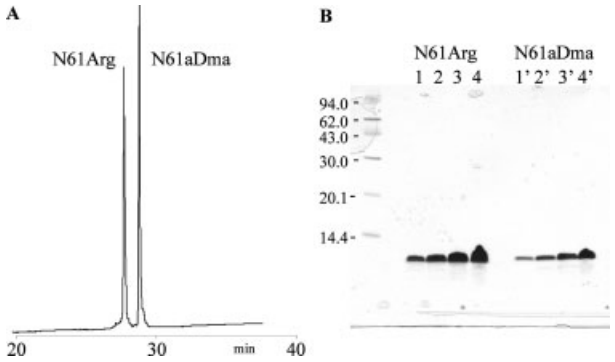


Figure 6 (A): Analytical RP-HPLC separation of purified, coinjected N61Arg (25 μg) and N61aDma (30 μg). (B): 15% acrylamide PAGE in Tris-glycine buffer: lanes from 1 to 4 were loaded, respectively, with 1.6, 3.2, 6.4 and 12.8 μg of N61Arg; lanes from 1′ to 4′ were loaded, respectively, with 0.8, 1.6, 3.2 and 6.4 μg of N61aDma. MW markers (MW in kDa) are shown on the left.

resonances between 6.8 and 8.6 ppm can be assigned to the peptide *C*-terminal NH_2 protons, to the NH_2 amide protons of Gln^{699} side chain, and to the remaining backbone NH protons of the non repetitive, *C*-terminal region (residues 695–706). The additional, broad peak at ~6.6 ppm can be assigned to the guanidino protons of the Arg side chains. In the aliphatic region the spectrum is dictated by the overlapping resonances of the 34 Gly (3.8–4.0 ppm) Hα protons, the 10 Arg Hα protons (~4.3 ppm) and the Hδ protons of the Arg side chains (3.11 ppm).

The spectrum of N61aDma is nearly identical to that of N61Arg, with a few notable exceptions. The broad resonance at 6.6 ppm splits into two sharper peaks at 6.66 and 6.68 ppm, while an additional broad peak at ~5.7 ppm appears. In the aliphatic region the Hδ resonance of aDma side chain is shifted from 3.11 to 3.17 ppm and the aDma methyl group resonances appear at ~2.9 ppm.

With 34 Gly, 10 Arg and 6 Phe over 61 residues, the nucleolin *C*-terminal region is highly repetitive in its sequence and therefore very difficult to analyse by NMR methods in the absence of uniform or specific isotopic labelling. However, the lack of chemical shift dispersion for the backbone NH and Hα protons, the overlap of the Phe aromatic ring resonances, and the fact that all chemical shift values that can be assigned are very close to those expected for random coil peptides show that N61Arg does not possess any stable secondary or tertiary structure in aqueous solution. As the NMR spectrum of N61aDma is nearly identical to that of N61Arg, it can be concluded that N,N-dimethylation of the guanidino group does not alter the overall conformational propensity of this peptide. In support of the NMR findings, the circular dichroic spectra of N61Arg and N61aDma (not shown) are identical, with a strong negative band at 200 nm and a slightly positive

band at 220 nm, typical of random coil structures. It should be borne in mind, however, that certain polypeptides, that are apparently unstructured when isolated, can adopt a well defined conformation in the presence of specific target molecules or other domains of the same protein [50]. This conformational plasticity has been suggested to work as an energy saving expedient [51]. This might be the case also for the nucleolin *C*-terminal region. A different conformation of this region in the context of the entire protein cannot be ruled out, as well as a conformational switch induced by the interaction of the *C*-terminal region with its target protein or nucleic acid.

At a local level N,N-dimethylation is known to affect the guanidino group physico-chemical properties, such as the pK_a, hydrophobicity, steric hindrance [52] and hydrogen bonding capabilities [13]. While the chemical shift of aDma methyl groups is consistent with that reported for positively charged substituted guanidines, the small change in the chemical shift of the Hδ resonance, together with the different behaviour of guanidino protons in regard to solvent exchange, suggest that subtle changes both in the charge distribution and resonance energy are occurring upon N,N-dimethylation of the guanidino moiety. At this stage, however, it is difficult to predict how these changes are going to affect the interaction of aDma containing peptides with target proteins or nucleic acids.

DNA Binding Assays

Arg residues have a strong tendency to form hydrogen bonds, especially double hydrogen bonded structures [13]. In the interaction of Arg side chain with nucleic acids, various patterns between the guanidino group and the phosphates can be formed.

Dimethylation drastically reduces the number of hydrogen bonds that Arg can form, and some of the patterns with double H-bonds are excluded [13]. Furthermore, dimethylation increases the hydrophobicity, the accessible surface and the molecular volume of the guanidino moiety, but does not alter the net charge of the residue [52]. Therefore the role of Arg dimethylation on the interaction between the nucleolin *C*-terminal domain and nucleic acids was investigated.

Filter-binding assays (Figure 7) revealed no appreciable differences in the binding affinity towards a single-stranded nucleic acid substrate between the Arg-dimethylated and the unmethylated peptides (with an approximate $K_d = 1.5 \times 10^{-7}$ M for both N61Arg and N61aDma). This value is comparable to those collected with recombinant *C*-terminal nucleolin ($K_d = \sim 2 \times 10^{-7}$ M) [4]. Also examined were the binding properties of an intermediate peptide produced during the stepwise synthesis of N61Arg and corresponding to residues 669–706 of human nucleolin. This peptide contains six Arg involved in five RGG sequences. The

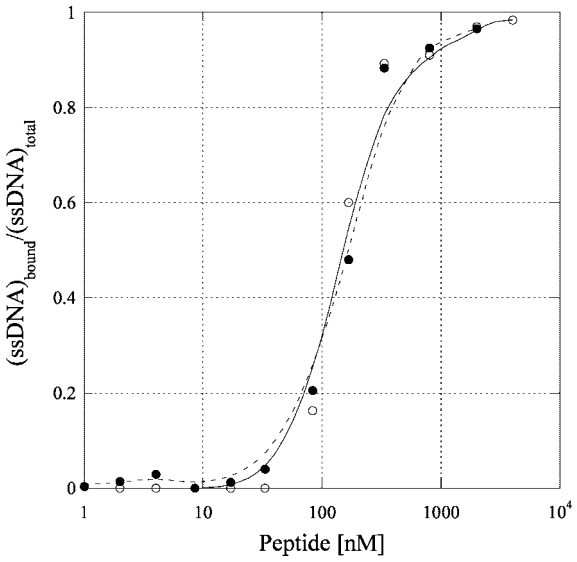


Figure 7 Double filter binding assay of N61Arg (●) and N61aDma (○) in the presence of ssDNA.

assay indicates an approximate ten-fold lower affinity for ssDNA with respect to N61Arg and a five-fold higher affinity with respect to peptides containing four RGG-repeats (data not shown), suggesting that the nucleic acid binding capacity increases with the number of the RGG repeats. These findings confirm that the non-specific binding of RGG peptides to the nucleic acid is primarily ionic in nature and are in agreement with our previously published results on shorter Gly/Arg-rich peptides [13].

To confirm the ionic nature of the DNA/RGG repeat interaction, the affinity for nucleic acid substrates was also tested by affinity chromatography on a ssDNA

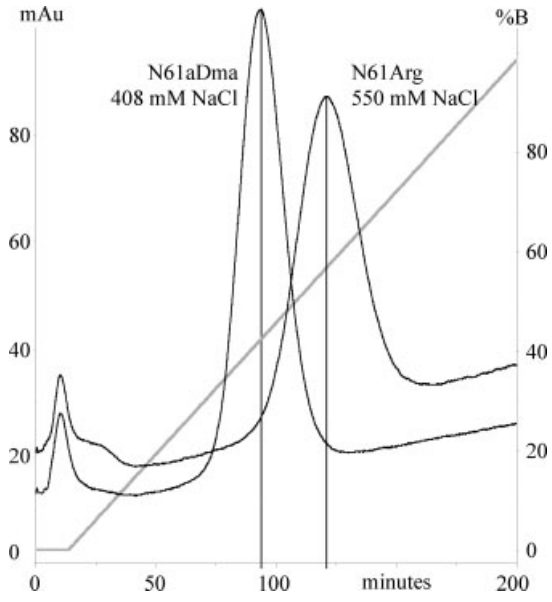


Figure 8 ssDNA-agarose affinity chromatography of N61Arg and N61aDma.

agarose column challenging the binding of N61Arg and N61aDma by increasing the ionic strength with a salt gradient. The lower ionic strength (Figure 8) required to elute N61aDma (408 mM NaCl) compared with N61Arg (550 mM NaCl) can be explained by the reduced p*K*a value of the dimethylated Arg side chain, which is further confirmed by the ion-exchange chromatography step performed for the purification, where N61aDma again elutes before N61Arg (Figure 5). The fact that asymmetric guanidino group alkylation renders the Arg molecule less basic was also previously reported and discussed [52].

Helicase Activity

Nucleolin has been previously identified as the human DNA helicase IV. The ATP-dependent duplex unwinding activity was attributed to the Gly/Arg-rich domain [53]. The helicase activity of N61Arg and N61aDma was tested using the same experimental set-up in the presence or in the absence of ATP but no unwinding activity was detected (not shown). This result confirms the findings of Ginisty *et al.* [1] who could not detect any helicase activity with different nucleolin samples or recombinant fragments comprising the *C*-terminal domain [1]. The absence of helicase activity seems also to be consistent with our previous data about a stabilization effect exerted by RGG containing peptides on the DNA double helix as demonstrated by DNA melting studies [13]. An indication in this sense is supported also by the work of Hanakahi *et al.* [54] showing how nucleolin accelerates annealing of complementary or partially mismatched oligonucleotides and that this annealing activity can be localized to the *C*-terminal region.

Methylation Assay

The *C*-terminal domain of nucleolin is rich in RGG clusters, which are asymmetrically methylated in the cells by the protein Arg methyltransferase 1 (PRMT1) giving ten aDma residues. In order to investigate whether N61Arg can be used as a substrate for PRMT1, the methylation assay was performed using recombinant human PRMT1. The results clearly show that N61Arg is efficiently methylated *in vitro*. The MS analysis of the reaction product revealed the presence of a vast number of mono- and dimethylated species that appeared throughout the methylation process (Figure 9). During the short 3 h reaction entities were detected that contained only a couple of dimethylated Arg. The number of dimethylated Arg increased progressively with time, leading to the formation of traces of fully dimethylated N61Arg (N61aDma) within the 72 h reaction (Figure 9).

Assay of HIV Infection in HeLa CD4$^+$ Cells

The pentameric pseudopeptide 5[Kψ(CH_2N)PR]-TASP (referred to as HB-19) is a potent inhibitor of HIV infection [35]. HB-19 binds specifically the cell-surface-expressed nucleolin and thereby blocks the attachment of HIV particles to target CD4+ cells [55]. By the use of various deletion constructs of nucleolin, the RGG domain of nucleolin was designated as the domain that binds HB-19. Consequently, a synthetic non-methylated peptide corresponding to the last *C*-terminal 63 amino acids of nucleolin was shown to inhibit HIV infection at the stage of HIV attachment to cells, thus suggesting that this domain could be functional in the HIV anchorage to target cells [5]. This work compares the anti-HIV activities of N61Arg and N61aDma on HIV-1 infection of HeLa-CD4-LTR-*lacZ* cells (HeLa P4 and HeLa P4-C5 cells). HIV entry and replication in these cells result in the activation of HIV LTR, leading to the expression of the *lacZ* gene. Therefore, the β-galactosidase activity could be measured in cell extracts to monitor HIV entry into cells. The value of β-galactosidase activity obtained in the presence of the HIV-replication inhibitor AZT is referred to as the background value in a given experiment. Another control for inhibition of HIV infection was obtained by the nucleolin-binding HB-19 pseudopeptide. Consistent with a previous report [5], N61Arg inhibited the T-tropic HIV-1 LAI infection of HeLa P4 cells in a dose-dependent manner with a 50% inhibitory concentration (IC_{50}) value of 0.5 μM. A complete inhibition of HIV infection was observed at 5 μM of N61Arg, a value similar to those obtained with AZT and HB-19 (Figure 10A). The inhibitory effect of N61Arg was not restricted to T lymphocyte-tropic HIV-1 isolates, since it inhibited also the macrophage-tropic HIV-1 Ba-L isolate in a dose-dependent manner with an IC_{50} value of 0.7 μM (Figure 10B). In contrast, the N61aDma peptide had no significant effect on infection of cells with both cell types (Figure 10). These results demonstrate that the Arg dimethylated synthetic peptide lacks any inhibition activity. This difference raises the question whether the cell-surface expressed nucleolin is indeed dimethylated. It should be noted that two subspecies of nucleolin could be resolved by two-dimensional gel isoelectric focusing: cytoplasmic/surface nucleolin with pI values at about 4.5, and nuclear nucleolin with pI values between pH 4 and 6 [56]. The difference between the pI values could be accounted for by the possibility that, unlike nuclear nucleolin, the RGG domain of surface nucleolin does not contain aDma.

CONCLUSIONS

The *C*-terminal region of human nucleolin, ChNu, contains several tandem aDma-Gly-Gly repeats. The

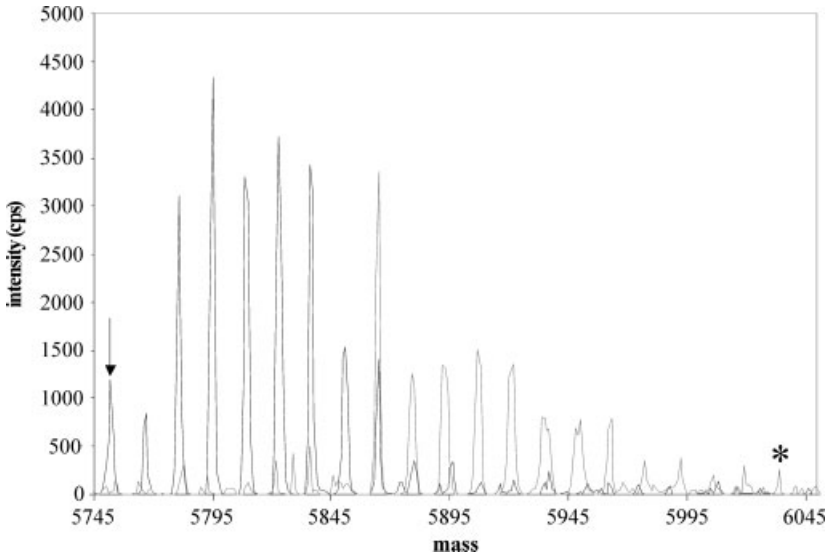


Figure 9 N61Arg is efficiently methylated *in vitro* by recombinant PRMT1. Deconvoluted mass spectra of N61Arg after 3 h (blue spectrum) and 72 h (red spectrum) of methylation reaction. The peaks are regularly spaced at multiples of 14 Da (+ one methyl group) indicating the presence of various intermediate species of the methylation reaction. The arrow indicates the starting product N61Arg, the asterisk indicates the presence of a peak with mass corresponding to the fully dimethylated polypeptide.

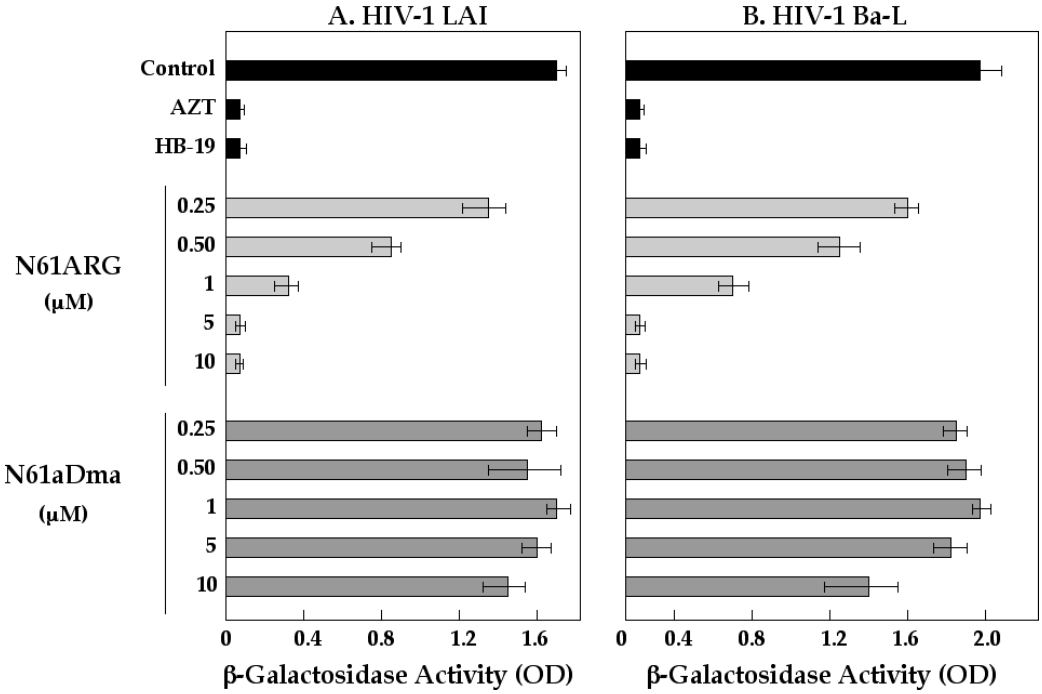


Figure 10 N61Arg but not N61aDma inhibits HIV infection. Inhibition of HIV infection by the synthetic N61Arg and N61aDma peptides was assayed in HeLa P4 cells infected by HIV-1 LAI (panel A) and in HeLa P4C5 cells infected by the HIV-1 Ba-L (panel B) isolate. HIV entry and replication result in the activation of the HIV-1 LTR promoter controlling the expression the lacZ gene (corresponding to ß-galactosidase). Cells were infected in the presence of AZT (5 mM), HB-19 (1 mM), or N61Arg and N61aDma at 0.125 0.25, 0.50, 1, 5 and 10 mM. The ß-galactosidase activity was measured at 48 h post-infection (OD 570 nm). Each point represents the mean ± S.D. of triplicate samples. Note that at 5 mM N61Arg the inhibition is as efficient as with AZT (which gives the background value in these cells).

investigation of the biological function of this region has been hampered on one hand by the impossibility of obtaining pure, fully methylated ChNu by enzymatic methods from the recombinant unmethylated polypeptide, and on the other by the difficulties in the synthesis of peptides containing multiple GG repeats. The synthetic strategy described here, based on the Fmoc-Gly-(Dmob)Gly-OH dipeptide as a building block

and the direct introduction of aDma, overcomes these obstacles. The advantages are a decrease in the number of coupling steps (from 60 to 45), a higher solubility of the growing peptide chain, higher coupling yields and elimination of Arg acylation, a side reaction connected to the Hmb backbone protection strategy. Therefore, this approach is of general applicability in the synthesis of peptides with repeating Xaa-Gly dipeptides at multiple locations.

Thus it was possible to prepare in high yield and purity two peptide amides corresponding to the 61-residue *C*-terminal region of human nucleolin, containing either unmodified Arg or aDma, and to study their structure and function in a series of comparative assays.

NMR and CD data show that Arg N,N-dimethylation does not change the overall conformational preference of ChNu and that this region is essentially unstructured in solution as an isolated polypeptide.

It was shown that both N61Arg and N61aDma bind to single-stranded nucleic acid substrates with similar affinities in filter binding assays, but elute at different ionic strength from a ssDNA-sepharose column, suggesting that the interaction with nucleic acids is mainly dictated by ionic interactions. Although controversal reports [1,53] exist about the helicase activity of ChNu, no DNA unwinding activity could be detected for either N61Arg or for N61aDma.

N61Arg is particularly suitable for the kinetic analyses of Arg methylation in substrates with multiple sites, which would add data to the scarce information available on the catalytic mechanism of PRMT1. To the best of our knowledge, this is the first ESI-MS analysis of *in vitro* enzymatic Arg methylation which gives information on the distribution of methyl groups in long RGG repeats and suggests that quantitative methylation of such polypeptides is not practicable by enzymatic methods.

While most abundant in the nucleoli, nucleolin is also found at the cell surface, and it has been shown that the N61Arg peptide inhibits HIV entry into HeLa cells blocking the virus attachment to the target cell. On the contrary, the dimethylated analogue N61aDma lacks any inhibition activity, which brings up the question of whether nucleolin is indeed present in its methylated form at the cell surface.

In this respect, it is of interest to note that the expression of surface nucleolin is highly regulated by activation conditions of cell growth and proliferation, whereas the expression of nuclear nucleolin remains somewhat constant. Post-translational modifications of surface and nuclear nucleolin, such as the modification of Arg residues to aDma, could account for trafficking of nucleolin toward distinct targets in the cell, i.e. to the nucleus or the plasma membrane.

Acknowledgements

The help and advise of Professor Arturo Falaschi (ICGEB) is gratefully acknowledged.

REFERENCES

1. Ginisty H, Sicard H, Roger B, Bouvet P. Structure and functions of nucleolin. *J. Cell. Sci.* 1999; **112**: 761–772.
2. Srivastava M, Pollard HB. Molecular dissection of nucleolin's role in growth and cell proliferation: new insights. *FASEB J.* 1999; **13**: 1911–1922.
3. Tuteja R, Tuteja N. Nucleolin: a multifunctional major nucleolar phosphoprotein. *Crit. Rev. Biochem. Mol. Biol.* 1998; **33**: 407–436.
4. Ghisolfi L, Joseph G, Amalric F, Erard M. The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J. Biol. Chem.* 1992; **267**: 2955–2959.
5. Nisole S, Said EA, Mische C, Prevost MC, Krust B, Bouvet P, Bianco A, Briand JP, Hovanessian AG. The anti-HIV pentameric pseudopeptide HB-19 binds the *C*-terminal end of nucleolin and prevents anchorage of virus particles in the plasma membrane of target cells. *J. Biol. Chem.* 2002; **277**: 20 877–20 886.
6. Xu X, Hamhouyia F, Thomas SD, Burke TJ, Girvan AC, McGregor WG, Trent JO, Miller DM, Bates PJ. Inhibition of DNA replication and induction of S phase cell cycle arrest by G-rich oligonucleotides. *J. Biol. Chem.* 2001; **276**: 43 221–43 230.
7. Zobel-Thropp P, Gary JD, Clarke S. δ-N-Methylarginine is a novel posttranslational modification of arginine residues in yeast proteins. *J. Biol. Chem.* 1998; **273**: 29 283–29 286.
8. Pelletier M, Xu Y, Wang X, Zahariev S, Pongor S, Aletta JM, Read LK. Arginine methylation of a mitochondrial guide RNA binding protein from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 2001; **118**: 49–59.
9. Gary JD, Clarke S. RNA and protein interactions modulated by protein arginine methylation. *Progr. Nucleic Acid Res. Mol. Biol.* 1998; **61**: 65–131.
10. Aletta JM, Cimato TR, Ettinger MJ. Protein methylation: a signal event in post-translational modification. *Trends Biochem. Sci.* 1998; **23**: 89–91.
11. McBride AE, Silver PA. State of the arg: protein methylation at arginine comes of age. *Cell* 2001; **106**: 5–8.
12. Lee J, Bedford MT. PABP1 identified as an arginine methyltransferase substrate using high- density protein arrays. *EMBO Report* 2002; **3**: 268–273.
13. Raman B, Guarnaccia C, Nadassy K, Zakhariev S, Pintar A, Zanuttin F, Frigyes D, Acatrinei C, Vindigni A, Pongor G, Pongor S. N^{ω}-Arginine dimethylation modulates the interaction between a Gly/Arg-rich peptide from human nucleolin and nucleic. *Nucleic Acids Res.* 2001; **29**: 3377–3384.
14. Liu Q, Dreyfuss G. *In vivo* and *in vitro* arginine methylation of RNA-binding proteins. *Mol. Cell. Biol.* 1995; **15**: 2800–2808.
15. Tang J, Gary JD, Clarke S, Herschman HR. 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.* 1998; **273**: 16 935–16 945.
16. Ottl J, Musiol HJ, Moroder L. Heterotrimeric collagen peptides containing functional epitopes. Synthesis of single-stranded collagen type I peptides related to the collagenase cleavage site. *J. Pept. Sci.* 1999; **5**: 103–110.
17. Fischer PM, Retson KV, Tyler MI, Howden ME. Application of arylsulphonyl side-chain protected arginines in solid-phase peptide synthesis based on 9-fluorenylmethoxycarbonyl amino protecting strategy. *Int. J. Pept. Protein Res.* 1992; **40**: 19–24.
18. van Abel RJ, Tang YQ, Rao VS, Dobbs CH, Tran D, Barany G, Selsted ME. Synthesis and characterization of indolicidin, a tryptophan-rich antimicrobial peptide from bovine neutrophils. *Int. J. Pept. Protein Res.* 1995; **45**: 401–409.

19. Barlos K, Chatzi O, Gatos D, Stavropoulos G. 2-Chlorotrityl chloride resin. Studies on anchoring of Fmoc-amino acids and peptide cleavage. *Int. J. Pept. Protein Res.* 1991; **37**: 513–520.
20. Benz H. The role of solid-phase fragment condenzation (SPFC) in peptide synthesis. *Synthesis* 1994; 337–358.
21. Fields CG, Lloyd DH, Macdonald RL, Otteson KM, Noble RL. HBTU activation for automated Fmoc solid-phase peptide synthesis. *Pept. Res.* 1991; **4**: 95–101.
22. Schnolzer M, Alewood P, Jones A, Alewood D, Kent SB. *In situ* neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *Int. J. Pept. Protein Res.* 1992; **40**: 180–193.
23. Kaiser E, Colescott RL, Bossinger CD, Cook PI. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* 1970; **34**: 595–598.
24. Sparrow JT, Monera OD. Improvements to the TMSBr method of peptide resin deprotection and cleavage: application to large peptides. *Pept. Res.* 1996; **9**: 218–222.
25. Johnson T, Quibell M, Sheppard RC. N,O-bisFmoc derivatives of N-(2-hydroxy-4-methoxybenzyl)-amino acids: useful intermediates in peptide synthesis. *J. Pept. Sci.* 1995; **1**: 11–25.
26. Rivier J, McClintock R, Galyean R, Anderson H. Reversed-phase high-performance liquid chromatography: preparative purification of synthetic peptides. *J. Chromatogr.* 1984; **288**: 303–328.
27. Nicolas E, Pujades M, Bacardit J, Giralt E, Albericio F. A new approach to Hmb-backbone protection of peptides: synthesis and reactivity of Nα-Fmoc-Nα-(Hmb)amino acids. *Tetrahedron Lett.* 1997; **38**: 2317–2320.
28. Muller B, Besser D, Kleinwachter P, Arad O, Reissmann S. Synthesis of N-carboxyalkyl and N-aminoalkyl functionalized dipeptide building units for the assembly of backbone cyclic peptides. *J. Pept. Res.* 1999; **54**: 383–393.
29. Székely Z, Zakhariev S, Guarnaccia C, Antcheva N, Pongor S. A highly effective method for synthesis of N-substituted arginines as building blocks for Boc/Fmoc peptide chemistry. *Tetrahedron Lett.* 1999; **40**: 4439–4442.
30. Patthy A, Bajusz S, Patthy L. Preparation and characterization of N$^{\gamma}$-mono-, di- and trimethylated arginines. *Acta Biochim. Biophys. Acad. Sci. Hung.* 1977; **12**: 191–196.
31. Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* 1992; **2**: 661–665.
32. Wong I, Lohman TM. A double-filter method for nitrocellulose-filter binding: application to protein–nucleic acid interactions. *Proc. Natl Acad. Sci. USA* 1993; **90**: 5428–5432.
33. Tuteja N, Rahman K, Tuteja R, Falaschi A. DNA helicase IV from HeLa cells. *Nucleic Acids Res.* 1991; **19**: 3613–3618.
34. Tuteja N, Tuteja R, Ochem A, Taneja P, Huang NW, Simoncsits A, Susic S, Rahman K, Marusic L, Chen J, Zhang J, Wang S, Pongor S, Falaschi A. Human DNA helicase II: a novel DNA unwinding enzyme identified as the Ku autoantigen. *EMBO J.* 1994; **13**: 4991–5001.
35. Nisole S, Krust B, Dam E, Bianco A, Seddiki N, Loaec S, Callebaut C, Guichard G, Muller S, Briand JP, Hovanessian AG. The HB-19 pseudopeptide 5[KΨ(CH$_2$N)PR]-TASP inhibits attachment of T lymphocyte- and macrophage-tropic HIV to permissive cells. *AIDS Res. Hum. Retroviruses* 2000; **16**: 237–249.
36. Callebaut C, Nisole S, Briand JP, Krust B, Hovanessian AG. Inhibition of HIV infection by the cytokine midkine. *Virology* 2001; **281**: 248–264.
37. Gartner S, Markovits P, Markovitz DM, Kaplan MH, Gallo RC, Popovic M. The role of mononuclear phagocytes in HTLV-III/LAV infection. *Science* 1986; **233**: 215–219.
38. Choma CT, Robillard GT, Englebretsen DR. Synthesis of hydrophobic peptides: An Fmoc "Solubilising Tail" method. *Tetrahedron Lett.* 1998; **39**: 2417–2420.
39. Blaakmeer J, Tijsse-Klasen T, Tesser GI. Enhancement of solubility by temporary dimethoxybenzyl-substitution of peptide bonds. Towards the synthesis of defined oligomers of alanine and of lysyl-glutamyl-glycine. *Int. J. Pept. Protein Res.* 1991; **37**: 556–564.
40. Mutter M, Nefzi A, Sato T, Sun X, Wahl F, Wohr T. Pseudo-prolines (ΨPro) for accessing 'inaccessible' peptides. *Pept. Res.* 1995; **8**: 145–153.
41. Johnson T, Quibell M, Owen D, Sheppard RC. A reversible protecting group for the amide bond in peptides. Use in the synthesis of 'difficult sequences'. *J. Chem. Soc. Chem. Commun.* 1993; **4**: 369–372.
42. Carpino LA, Cohen BJ, Stephens KE, Sadat-Aalaee SY, Tien J-H, Langridge DC. [(9-Fluorenylmethyl)oxy]-carbonyl (Fmoc) amino acid chlorides. Synthesis, characterization, and application to the rapid synthesis of short peptide segments. *J. Org. Chem.* 1986; **51**: 3732–3734.
43. Zeng W, Regamey PO, Rose K, Wang Y, Bayer E. Use of Fmoc-N-(2-hydroxy-4-methoxybenzyl)amino acids in peptide synthesis. *J. Pept. Res.* 1997; **49**: 273–279.
44. Quibell M, Packman LC, Johnson T. Synthesis of the 3-repeat region of human Tau-2 by the solid phase assembly of backbone amide-protected segments. *J. Am. Chem. Soc.* 1995; **117**: 11 656–11 668.
45. Sampson WR, Patsiouras H, Ede NJ. The synthesis of 'difficult' peptides using 2-hydroxy-4-methoxybenzyl or pseudoproline amino acid building blocks: a comparative study. *J. Pept. Sci.* 1999; **5**: 403–409.
46. Weygand F, Steglich W, Bjarnason J, Akhtar R, Khan M. Leicht abspaltung schutzgruppen für säureamidfunktionen. *Tetrahedron Lett.* 1966; **29**: 3483–3487.
47. Sasaki Y, Abe J. Protection of Ψ(CH$_2$NH) peptide bond with 2,4-dimethoxybenzyl group in solid-phase peptide synthesis. *Chem. Pharm. Bull.* 1997; **45**: 13–17.
48. Narita M, Ishikawa K, Nakano H, Isokawa S. Tertiary peptide bond containing-oligo(Leu)s. Conformational studies in solution of oligo (L-leucine)s with L-proline residue and glycyl-N-(2, 4-dimethoxybenzyl)-L-leucine sequence. *Int. J. Pept. Protein Res.* 1984; **24**: 14–24.
49. Watson DJ, Dowdy ED, Li W-S, Wang J, Polniaszek R. Electronic effects in the acid-promoted deprotection of N-2,4-dimethoxybenzyl maleimide. *Tetrahedron Lett.* 2001; **42**: 1827–1830.
50. Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 1999; **293**: 321–331.
51. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R. Extended disordered proteins: targeting function with less scaffold. *Trends Biochem. Sci.* 2003; **28**: 81–85.
52. Kennedy KJ, Lundquist JTt, Simandan TL, Kokko KP, Beeson CC, Dix TA. Design rationale, synthesis, and characterization of non-natural analogs of the cationic amino acids arginine and lysine. *J. Pept. Res.* 2000; **55**: 348–358.
53. Tuteja N, Huang NW, Skopac D, Tuteja R, Hrvatic S, Zhang J, Pongor S, Joseph G, Faucher C, Amalric F, Falaschi A. Human DNA helicase IV is nucleolin, an RNA helicase modulated by phosphorylation. *Gene* 1995; **160**: 143–148.
54. Hanakahi LA, Bu Z, Maizels N. The *C*-terminal domain of nucleolin accelerates nucleic acid annealing. *Biochemistry* 2000; **39**: 15 493–15 499.
55. Nisole S, Krust B, Callebaut C, Guichard G, Muller S, Briand JP, Hovanessian AG. The anti-HIV pseudopeptide HB-19 forms a complex with the cell-surface-expressed nucleolin independent of heparan sulfate proteoglycans. *J. Biol. Chem.* 1999; **274**: 27 875–27 884.
56. Hovanessian AG, Puvion-Dutilleul F, Nisole S, Svab J, Perret E, Deng JS, Krust B. The cell-surface-expressed nucleolin is associated with the actin cytoskeleton. *Exp. Cell Res.* 2000; **261**: 312–328.

