(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2012 (27.12.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/175617 Al

(51) International Patent Classification:
C07K 14/00 (2006.01) *A61K 47/48* (2006.01)
C07K 14/78 (2006.01)

(21) International Application Number:
PCT/EP2012/061974

(22) International Filing Date:
21 June 2012 (21.06.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11005155.4 24 June 2011 (24.06.2011) EP

(71) Applicant *(for all designated States except US):* **STICHTING DUTCH POLYMER INSTITUTE** [NL/NL]; Kennispoort John F. Kennedylaan 2, NL-5612 AB Eindhoven (NL).

(72) Inventors; and
(75) Inventors/ Applicants *(for US only):* **ARMANDO HERNÁNDEZ, Garcia** [MX/NL]; Postbus 8038, NL-6700 EK Wageningen (NL). **DE VRIES, Rindert, Jacob** [NL/NL]; Postbus 8038, NL-6700 EK Wageningen (NL). **DE WOLF, Frederik, Anton** [NL/NL]; Postbus 8038, NL-6700 EK Wageningen (NL).

(74) Agent: **RENKEMA, Jaap;** IPecunia Patents B.V., P.O. Box 593, NL-6160 AN Geleen (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** DIBLOCK COPOLYMER

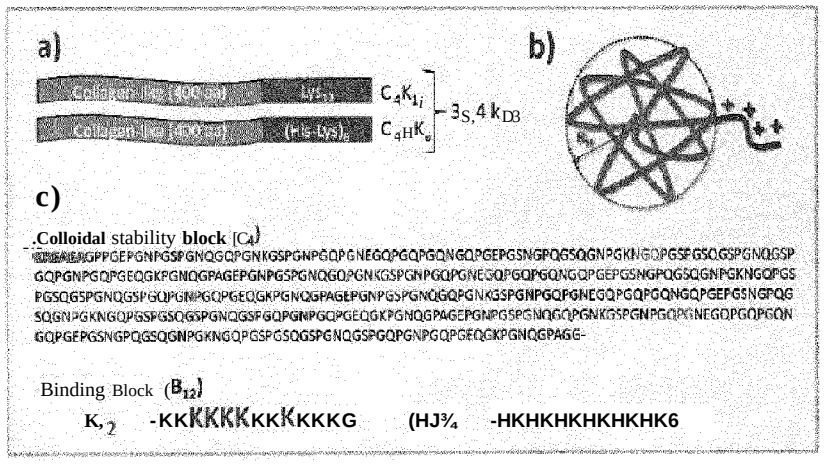


Figure t.

(57) **Abstract:** The invention relates to a diblock copolymer comprising a collagen-like block consisting of a first amino acid sequence of m amino acids and a DNA binding block consisting of a second amino acid sequence of n amino acids, wherein m and n is an integer and wherein the ratio between m and n is at least 20.

WO 2012/175617 A1

DIBLOCK COPOLYMER

The invention relates to a diblock copolymer, its use as a coating for a nucleic acid and a nucleic acid coated by such a copolymer and to the use of the coated nucleic acid in gene therapy, non-viral gene delivery, sequencing by transport through nanopores, nanoelectronics such as nanowire DNA templating artificial virus or DNA origami..

DNA is no longer only a carrier of genetic information, but rather a multi-purpose material that is being explored for both biological and non-biological applications.

Coating DNA with binder moieties is an effective way to modify physical parameters and modulate its interactions with the surrounding environment[1,2,3]. For example, non-viral gene delivery relies on complexation of DNA with cationic lipids and/or polymers to promote passage of DNA across cell membranes. A controlled DNA coating would be advantageous to regulate doses of transfected DNA and also to direct vector shape. Also, a protein coating on ssDNA has been used to facilitate passage through nanopores. Binder moieties could in principle be used for modulating many properties as bending rigidity, thickness, stiffness and stretching. All of these are crucial to other nanotechnological processes involving DNA, such as nanoscale self assembly ("DNA origami"), templating of nanowires, and transport in nano-channels for the stretching, separation and sequencing of DNA[4,5,6,7].

Control of the architecture of the coating using DNA binder moieties is an initial key element needed. Only a few studies so far have focused on understanding the requirements for coating single DNA molecules using rationally designed polymers. Typical for the use of polymeric DNA binders is the formation of random aggregates of uncontrolled size and irregular shape consisting of many nucleic acids molecules per aggregate [8,9, 10]. This limits their applicability when a precisely controlled coating is desired. In this sense, an understanding of the polymer design requirements that control monomolecular DNA coating with regular shape and uniform molecular weight could benefit the entire field of DNA-based nanobiotechnology.

The simplest design for a DNA-coating macromolecule would be a diblock structure containing a DNA binding domain, connected to another domain that should confer colloidal stability. A number of such designs have been published in which the binding domain is a simple flexible polycation. For example, Kostiainen *et al*[3]**.** have conjugated various globular proteins with polycationic blocks and showed that the conjugates coat dsDNA. A number of studies have also been performed on synthetic diblocks consisting of polycationic domain coupled to a large uncharged and hydrophilic flexible domain. A judicious analysis of these studies indicates clearly that the ratio between shielding and cationic block lengths (symmetry) it is important for the final architecture of the DNA molecule-polymer complex. In most of these cases the binding domains were still quite large (on the order of 20 positive charges or

more) leading to the formation of multimolecular aggregates by "bridging" several DNA molecules[11'12'13] and/orto the folding of single DNA molecules (intramolecular aggregation) creating a diverse population of shapes and sizes[14'10]. Intermolecular aggregation can be quite slow in these cases, and is typically enhanced at high DNA and salt concentrations[15].

A few of these studies have claimed monomolecular DNA coating. Radler and coauthors[16] reported controlled coating of single DNA using PEG-PEI (50 positive charges). Borner and coauthors[11], used PEO-PAMAM (20 positive charges) to form a mixed population of single-plasmid toroids with different folded degrees and supercoiled dsDNA. Kataoka and coauthors reported a PEG-poly-L-Lys (19 positive charges) polymer that multi-fold single plasmid DNA[12]. Intramolecular folding/aggregation is presumably driven by polycations bridging neighboring pieces of dsDNA. This type of bridging does not distinguish between inter- or intramolecular bridging (or folding). If intramolecular bridging occurs, it is therefore very likely that over time, or at high DNA concentrations, intramolecular aggregation will also occur. Despite the fact that they used much longer binding blocks (at a roughly similar size of the neutral block), Radler et. al.[16] report neither intermolecular aggregation nor intramolecular folding. This may be related to the fact that they have used DNA that was quite short DNA and concentrations that were quite low.

It is an object of the present invention to provide a diblock copolymer that can coat DNA with a reduced degree of aggregation even at a higher concentration of DNA.

Accordingly, the present invention provides a diblock copolymer comprising a collagen-like block consisting of a first amino acid sequence of m amino acids and a DNA binding block consisting of a second amino acid sequence of n amino acids, wherein m and n each independently stand for an integer and wherein the ratio between m and n is at least 20.

The present invention also provides a diblock copolymer comprising a collagen-like block consisting of a first amino acid sequence and a DNA binding block consisting of a second amino acid sequence, wherein the ratio between the molecular weight of the first amino acid sequence and the molecular weight of the second amino acid sequence is at least 15.

It was surprisingly found that a high asymmetry between shielding-to-binding block lengths in the diblock copolymer according to the present invention results in the prevention of formation of random aggregates of uncontrolled size and shape over time. Decreasing binding block length into a minimal and increasing enough the asymmetry (difference of the length ratio between block lengths) between shielding/binding blocks will lead into a controlled single DNA coated rod assembly; reducing significantly the presence of populations with different shapes and thus homogenizing the size and shape of the coated single DNA molecules. The present invention provides a technology for producing a highly asymmetric neutral-cationic diblock protein diblock copolymer that accurately coats a long single dsDNA, which shows no signs of intermolecular aggregation even at high DNA concentrations.

Preferably, the ratio between m and n is at least 25, more preferably at least 30. Preferably, the ratio between m and n is at most 100, more preferably 80, 60 or 50.

Binding of the diblock copolymer to the DNA template is achieved by a short stretch of basic amino acids, i.e. the binding block having a positive charge. The binding block needs to be short enough to avoid bridging but at the same time it should be long enough to bind DNA strongly. The collagen block has to be large enough to confer colloidal stabilization, while at the same time not constraining the binding. Hence, m is preferably from 300 to 600, preferably from 325 to 550, preferably from 350 to 500, preferably from 375 to 450 or from 390 to 425. Preferably, n is from 10 to 25, preferably from 11 to 20 or from 12 to 15

Preferably, the ratio between the molecular weight of the first amino acid sequence and the molecular weight of the second amino acid sequence is 15-35, more preferably 18-30, more preferably 20-25.

Preferably, the molecular weight of the first amino acid sequence is 30-50 kDa, more preferably 32-40 kDa. Preferably, the molecular weight of the second amino acid sequence is 1-2 kDa, more preferably 1.2-1.8 kDa.

In order to bridge dsDNA, the binding block needs to be at least larger than two dsDNA molecules diameter (>~4.8 nm). The binding block is preferably shorter than around 4.8 nm. This is at least larger than two dsDNA molecules diameter, which allows bridging of the dsDNA.

The protein polymer design according to the present invention can serve as a scaffold for a step-wise addition of extra functional blocks for a tighter control of shape and functionality.

The diblock copolymer may comprise (biofunctional) domains that give specific functionalities, beyond merely covering the DNA. This can be used as modulators of nanotechnological processes involving DNA. An attractive technology to produce biocompatible polymers with a built-in possibility to incorporate biofunctional domains is the use of genetic engineering to produce designed protein-polymers.

Accordingly, the present invention provides a copolymer comprising the diblock copolymer according to the present invention and a functional block comprising a third amino acid sequence attached to the first amino acid sequence or the second amino acid sequence of the diblock copolymer according to the present invention.

The diblock copolymer according to the present invention may have a collagen-like block consisting of the first amino acid sequence represented by SEQ. NO. 1 or a sequence having at least 80%, 90%, 95%, 96%, 97%, 98%, or 99%identity with SEQ ID NO. 1.

The second amino sequence preferably consists of (i) positively charged amino acids selected from the group consisting of lysine, arginine and histidine and (ii) one non-polar amino acid at

the end of the sequence. More preferably, the second amino sequence consists of positively charged amino acids selected from the group consisting of lysine and histidine and (ii) one non-polar amino acid at the end of the sequence. Said non-polar amino acid at the end of the sequence may be chosen from the group consisting of glycine, alanine, valine, leucine, isoleucine, methionine and proline. Preferably, said non-polar amino acid is glycine for its short length.

The second amino acid sequence is preferably attached to the first amino acid sequence with its end having the positively charged amino acid.

The second amino acid sequence may be represented by SEQ. NO. 2 or SEQ. NO. 3.

K_{12} (12 Lys; plus G corresponds to SEQ. N0.2) and $(HK)_6$ (six His-Lys dyads; plus G corresponds to SEQ. NO.3) show a different pH dependence of their charge density. $(HK)_6$ is more preferred of the two, since it coats DNA faster than K12. This is thought to be due to possible intercalation of the cationic aromatic imidazole rings of histidines between the purine nucleic acids

According to another aspect of the present invention, a process for making the diblock copolymer is provided, the process comprising the step of expressing the nucleic acid sequence encoding said copolymer in *Pichia pastoris* using recombinant expression techniques known per se.

Preferably, said step is performed at pH of at least 2.5 and at most 6.0. This range of pH is suitable for growing *Pichia pastoris* cells. The step is preferably performed at pH of at most 5.5, at most 5.0, at most 4.5, at most 4.0 or at most 3.5. A particularly preferred range is pH of 2.5-3.5. This range of pH of 2.5-3.5 gives a protective effect against degradation to the block copolymer of the present invention.

The present invention further provides a double strand nucleic acid coated with the diblock polypeptide according to the present invention. The double strand nucleic acid may be ribonucleotide, deoxy(ribo)nucleotide or a modified form of either type of nucleotide. Preferably, the double strand nucleic acid is DNA, which may be in the form of a vector or a linear DNA.

The present invention further provides use of the coated nucleic acid according to the present invention in gene therapy, non-viral gene delivery, sequencing by transport through nanopores, nanoelectronics such as nanowire DNA templating artificial virus or DNA origami.

The present invention further provides an isolated amino acid sequence SEQ ID NO. 1 or a sequence having at least 80%, 90%, 95%, 96%, 97%, 98% or 99% 99% identity with SEQ ID NO. 1.

The present invention further provides a nucleic acid sequence encoding the amino acid sequence SEQ ID NO.l or a sequence having at least 80%, 90%, 95%, 96%, 97%, 98% or 99%identity with SEQ ID NO. 1.

The present invention further provides a nucleic acid sequence encoding the diblock copolymer according to the present invention.

The skilled person will be aware of the fact that several different computer programs are available to determine the homology between two sequences. For instance, a comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percentage identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. (48): 444-453 (1970)) algorithm which has been incorporated into the GAP program in the Accelrys GCG software package (available at http://www.accelrys.com/products/qcq/), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

In yet another embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the Accelrys GCG software package (available at http://www.accelrys.com/products/gcg/), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity of two amino acid or nucleotide sequence is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:1 1 -17 (1989) which has been incorporated into the ALIGN program (version 2.0) (available at the ALIGN Query using sequence data of the Genestream server IGH Montpellier France http://veqa.iqh.cnrs.fr/bin/aliqn-quess.cqi) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

Based on the present description and possibly in combination with e.g. , Sambrook and Russel, Cold Spring Harbor Laboratory Press, 2001 hereby incorporated by reference, the skilled person can produce the copolymer of the present invention using the recombinant expression techniques described herein.

Figure 1 shows design of diblock copolypeptides for nucleic acid binding. (a) Schematic modular design, green is the colloidal stability collagen-like block and red is the DNA binding block. (b) Model of protein diblock copolymer as a random coil. (c) Amino acid sequence of the colloidal stability block and of the two used binding blocks.

Figure 2 shows molecular mass characterization of protein diblock copolymers. (a) SDS-PAGE image: MM: molecular markers, lane 1: C4K12 harvested medium at the end of the

fermentation, 2: Purified C4K12 1 g/L, 3: Purified C4K12 0.1 g/L, 4: C4HK6 harvested medium at the end of the fermentation, 5: Purified C4HK6 sample 1 g/L 6: Purified C4HK6 sample 0.1 g/L. (b) MALDI-TOF of purified C4K12. The molecular weight for M+ is 38,586, M2+ 19,305 and 2M+ 76,967. (c) MALDI-TOF of purified C4HK6. The molecular weight for M+ is 38,432, M2+ 19,228.

Figure 3 shows electrophoretic mobility of DNA-Protein diblock copolymer complexes on 1% agarose gel. Top row shows N/P ratio. Lane 1: DNA molecular markers, 2: pDNA alone (N/P = 0).

Figure 4(a). shows light scattering intensity for DNA-protein diblock copolymer complexes formed at different N/P. # C4K12 complexes, * C4HK6 complexes.

Figure 4(b) shows DLS intensity and hydrodynamic radius for complexes formed between diblock copolypeptides and pDNA at different N/P values or salt conditions. (a) C4K12-pDNA nanoplexes measured at Θ= 12.8°. (b) C4HK6-pDNA nanoplexes measured at Θ= 12.8°

Figure 5 shows disassembly of DNA-protein nanoplexes by salt addition. a) Light scattered intensity by size distribution of C4K12-pDNA nanoplexes (N/P = 1.1) at different salt concentrations. Appearance of free diblock dynamics of the nanoplexes was followed. b) Light scattered intensity of complexes at different salt concentrations.

Figure 6 shows typical atomic force microscopy (AFM) images and height profiles of rod nanoplexes formed between diblock copolypeptides and 3.7 kbp pDNA a) C4K12-pDNA at N/P = 2.2 b) C4HK6 -pDNA at N/P = 1.5. Scale is 2 μm.

Figure 7 shows atomic Force Microscopy in air of deposited single pDNA molecules complexed with protein diblock copolymers at different N/P ratios. Top row C4K12-pDNA nanoplexes: (a) N/P = 0.95, (b) N/P = 2.2. (c) N/P = 8.3. Bottom row C4(HK)6-pDNA nanoplexes (d) N/P = 0.4, (e) N/P = 1.5, (f) N/P = 8.5. Bar scale 200nm.

Figure 8 shows coated Linear DNA with C4K12 protein diblock copolymer. Scale bar: 200 nm. L is 1258 nm. 0.45 nm height, 24nm width. In mica.

Figure 9 schematically illustrates side-chain stretching.

Figure 10 schematically illustrates coated supercoiled DNS.

Experiments

The choices of the DNA binding block and the collagen-like block were made such that the DNA binding block was shorter (12 amino acids) than the shortest ones reported and the collagen-like block was longer (400 amino acids) than the longest ones reported (see Table 1). Thermodynamically stable single-DNA complexes were obtained. The single-DNA complexes were not only kinetically stable at low concentrations or short times, but were stable in conditions that ultimately will show either intermolecular bridging or intramolecular bridging (folding) if the copolymer according to the present invention is not used.

Table 1

Binding Block		Colloidal Block		[DNA]	Reference
Type	n_+	Type	Mw		
Branched PEI 2 kDa	-50	PEG	20 kDa	<0.33 ng^L	1
polyamidoamine 2.6 kDa	20	PEO	2.7 kDa		2
Poly-L-Lys	17	PEG	12 kDa	33.3 ng/uL	3
Lys -1.5 kDa	12	Collagen-like	36 kDa	43 ng/uL	This study
His-Lys -1.5 kDa	12	Collagen-like	36 kDa	43 ng/uL	This study

1. DeRouchey, J.; Schmidt, C ; Walker, G. F.; Koch, C ; Plank, C ; Wagner, E.; Radler, J. O. Biomacromolecules. 2008, 9, 724-732[16].

2. Hartman, L.; Hafele, S.; Peschka-Suss, R.; Antonietti, M.; Borner, H. G. Chem. Eur. J. 2008, 14, 2025-2033[11].

3. Osada, K.; Oshima, H.; Kobayashi, D.; Doi, M.; Enoki, M.; Yamasaki, Y.; Kataoka, K. JACS. 2010, 132, 12343-12348[14].

Biosynthesis of asymmetric cationic diblock copolypeptides. The two diblock protein copolymers were successfully produced using recombinant yeast strains of *P. pastoris* carrying the artificial genes coding for their amino acid sequences, prepared by molecular cloning. After simple and scalable steps of differential precipitation (see materials and methods) lyophilized diblocks were recovered with a yield up to lg/L of cell-free medium. The precipitation steps were chosen in such a way that they did not precipitate the low concentrations of extracellular *P. pastoris* proteins present in the medium, but only the diblock protein copolymers. Note that biosynthesis of cationic block polypeptides has already been reported in bacterial systems such as *E. coli* , but not in yeast [17'12'18]. As far as we know, this is the first report of the production of cationic block copolypeptides in eukaryotic yeast cells with high yields.

Biosynthesis of highly cationic polypeptides has been difficult since they are generally toxic because they bind to the dominantly negatively charged biomolecules inside the cell, disrupting biological functions.

Molecular characterization of diblock copolypeptides. Purity and molecular weight of C_4-K_{12} and $C_4(HK)_6$ after purification were characterized using SDS-PAGE and MALDI-TOF (figure 2). SDS-PAGE showed well defined single bands for both purified diblocks that corresponded with the dominant bands in the lane of the harvested medium. The molecular weight of the purified polypeptides as determined by SDS-PAGE (98kDa) does not correspond to the expected value (38 kDa). The reduced mobility is caused by the very low content of hydrophobic amino acids in C_4 block that leads to poor binding of SDS. For $C_4(HK)_6$ a second upper band is observed. This may corresponded to a multimer, since the same band is not present in the lane of the SDS-PAGE corresponding to the harvested medium sample.

Accurate experimental molecular weights for the diblocks protein polymers have been determined using MALDI-TOF. For C_4K_{12} three signals are found (19,306, 38,587 and 76,967 Da) that correspond respectively to M2+ M+ 2M+. For $C_4(HK)_6$, two signals are found (19,228 and 38,433 Da) corresponding to M2+ and M+, respectively. The MALDI-TOF sizes are in very good agreement with the expected theoretical sizes of -38,407 and 38,461 Da for C4K42 and $C_4(HK)_6$, respectively. Small deviations are possible, since Glu-Ala repeats can be left when pre-pro signal is removed from the diblock sequence before being secreted. The possible presence of these extra amino acids has no effect on the DNA binding properties, since they are located on the opposite side of the proteins.

Electrophoretic mobility of DNA-polypeptide complexes. Agarose gel mobility retardation was used as an initial assay for the interaction between the two protein diblock copolymers and plasmid DNA. Figure 3 shows that electrophoretic mobility of DNA is reduced when diblock copolypeptides are added. The effect increases at higher binding block to DNA phosphate ratio (N/P), and saturates at about N/P =5. This behavior is similar to that of some cationic polypeptides also carrying a short binding block, where a total retardation was not observed either. In contrast, less asymmetric neutral-cationic diblock polymers completely neutralize DNA charge or even reverse the mobility [11,16]. As a control, we have also verified that there is no mobility shift when DNA is mixed with C_4 block alone. Although the net charge of the resulting complexes was almost neutral at high N/P it never became positive. Presumably, this is because for the highly asymmetric diblock, the coating is not dense enough to neutralize all the charge even at high N/P. Finally, the outside C_4 blocks are weakly negatively charged, and this may also play a role in determining the observed mobility of the complexes. (The predicted isoelectric point of the C_4 block is pi = 4.9).

Light Scattering studies. Light Scattering was used to further study the interaction between DNA and the diblock copolypeptides, since it is a convenient technique for obtaining solution

sizes and molecular weight estimates for the complexes that does not require any labeling. Figure 4(a) shows the absolute intensity of scattered light (or Rayleigh ratio, see materials and methods) Rg of the complexes as a function of the N/P ratio. For both copolypeptides, the scattered intensity increased rapidly at N/P < 1, and then saturates at N/P » 1. The scattered intensity of the fully covered DNA was about a factor 100 larger than that of the naked DNA. For a rough estimate of the degree of coverage, it may be assumed that the scattering at N/P»l is dominated by the complexes and that the contribution of excess free polymer to the scattering can be neglected. Assuming full charge neutralization, the predicted molar mass M of a single coated DNA molecule is

$$M = (1 + \alpha)M_{DNA}$$

$$\alpha = \frac{1}{n_p}\frac{M_p}{M_0}$$

where M_{DNA} and M_p are respectively, the DNA and polymer molar mass, n_p is the number of basepairs covered by a single bound C_4B_{12} polymer, and M_0= 600 Da is molar mass of a single basepair. Assuming n_p = 6bp (one C_4B_{12} polymer per 12 DNA phosphate charges or 6 basepairs), gives a = 10.7. In the limit of zero scattering angle and zero concentration, the predicted absolute scattering intensity R_θ for the complexes is

$$R_g = K_R C M$$

where $C = (1 + \alpha)C_{DNA}$ is the weight concentration of complexes. Literature values for the refractive index increments for dsDNA and the C_4B_{12} polymers are within the same range of dn/dc = 0.16...0.18 ml/g (see materials and methods). Using a single average value of dn/dc = 0.17 gives a value of the Rayleigh constant of $K_R = 2.105\text{-}10^{-5}$ mol $m^2 kg^{-2}$ (see materials and methods). Roughly speaking, both C and M increase by a factor of about 10 by the polymer coating, such that the light scattering increases by a factor of about 100, as observed. The estimated absolute scattering intensity of R_θ =0.319 m^{-1} for charge neutral complexes is indicated as a dashed line in Fig 4a, and lies somewhat above the limiting values for both types of complexes. Consistent with the results of the mobility shift assay, this indicates that at saturation, for both complexes, the majority, but not all of the DNA phosphate charges have been neutralized by complexation with binding domains.

Figure 4(b) shows the effective hydrodynamic radii (R_h) determined using DLS, as a function of N/P. For C4HK6, the size rapidly increases from 75 nm in the absence of polymer, to 150 nm at N/P = 1.5. For C4K12 + DNA, the final size is somewhat smaller, Rh = 135 nm at N/P = 1.5 These values are close to values for similar bulky cationic polypeptides [18]. Finally, for the individual diblocks, in the absence of DNA, the scattered intensity was very low. The hydrodynamic radius for the diblocks was 5.4 nm for C_4K_{12} and 6.1 nm for $C_4(HK)_6$.

It is important to note that the reported effective R_h of the complexes is calculated from a measured diffusion D constant assuming a spherical particle size,

$$D = \frac{k_B T}{6\pi\eta R_h}$$

where $k_B T$ is the thermal energy, and η is the solvent viscosity. A 3.7kb supercoiled plasmid, both with and without adhering diblock polymers, presumably is closer to a rod-like than to a spherical particle. Assuming the diffusion of both naked and coated supercoiled plasmid can be approximated by that of a cylindrical particle with length L and diameter d, the predicted diffusion constant is:

$$D = \frac{Ak_B T}{3\wedge L}$$

$$A = \ln(L/d) + 0.312 + 0.565/(L/d) - 0.1/(L/d)^2$$

This approximation was used by deRouchey et al.[16] for short (rod-like) linear DNA (<lkb) coated with cationic-neutral diblock polymers, to extract an estimated thickness d of the coated DNA, using the known length L. For a very rough estimate of the effective thickness of the (coated) plasmid "cylinder", we use L = 650 nm, which is half the contour length of 3.7kb DNA. Then, for uncoated DNA we find d = 12nm, which is not unreasonable for the diameter of a plectonemic supercoil. For the fully coated supercoils we find d=87 nm (assuming R_h = 135nm for C_4K_{12}) and d=\ 12nm (assuming R_h = 150 nm for $C_4(HK)_6$).

Charge Screening by salt addition. To confirm that complexes are formed due to electrostatic interactions, N/P =1.1 complexes were exposed to increasing concentration of salt (NaCl) and their disassembly was followed using DLS (figure 5). These experiments were carried out at a larger scattering angle of 173°, corresponding to a larger value of the wavevector q (see materials and methods) to allow for a more sensitive detection of the diffusion of the much smaller free polymers. Figure 5.a shows that the scattering intensity decreases rapidly for salt concentrations in between 0.1M and 0.25 M. Figure 5.b shows the size distributions as determined using DLS. The contribution due to diffusion of free diblocks (R_H = 6 nm) starts to become visible at 0.25 M, implying that at this concentration, most diblocks are no longer bound to the DNA. Above 0.4M the scattering intensity is constant, implying that no polymers are bound anymore. These experiments demonstrate that the complexation is clearly electrostatic in nature. This also means that further tuning of the binding strength may be achieved via changes of pH that affect the charge on the binding blocks, and that makes the design can be tailored to specific pH dependency requirements.

Atomic Force Microscopy. AFM studies in air were carried out to determine the morphology of the complexes and to confirm that complexes consist of individual DNA

molecules coated by diblock copolymer. We have found that the collagen-like C_4 blocks rapidly adsorb on the mica and silica surfaces that we have used as a substrate for AFM. Whereas naked DNA does not adhere on on bare mica or silica, all of the complexes did without requiring divalent cations or poly-Lysine. This is advantageous because it this excludes any possible influence of these additives on the morphology of the adhering DNA complexes.

The figure 6 shows typical images of C4-K12 complexes at N/P=1.5 and C4HK6-DNA complexes at N/P = 2.2. For both cases the complexes are well dispersed. They appear as semiflexible single rods of regular shape. The average contour length of the nanoplexes at N/P = 2.2 is 649.4 ± 15 nm for C4K12 and 624.7 ± 29.3 for C4HK6 (N = 20). The height for both complexes is very uniform with an average around 1 nm (1.07 ± 0.1 and 0.99 ± 0.2 (N = 20) for C4K12 and C4HK6, respectively).

A series of images of individual complexes at increasing values of N/P (for both polymers) is shown in Fig. 7. Clearly, for both polymers, complexes become thicker, stiffer and more homogeneously coated at higher N/P. At low N/P, complexes appear thin and less rigid, and an occasional loop of the supercoiled DNA molecules can still be seen. C4HK6-DNA complexes appear thicker at lower N/P as compared to C4K12 complexes: N/P = 0.4 complexes of C4HK6 are similar to N/P = 1 complexes of C4K12. Also, both types of complexes continue to increase in thickness for N/P > 1. Both findings are in agreement with the light scattering and electrophoresis results.

A good test for the absence of intramolecular folding is to compare with complexes of linearized pDNA: this should give a contour length exactly equal to the contour length of the 3.7kb pDNA (L_{DNA} = 1258 nm). A representative image of a C4K12- linearized pDNA complex at N/P = 1 is shown in Figure 7. For this case the contour length of the complexes indeed equals that of the naked DNA, indicating that individual DNA molecules get coated, and that there is no intramolecular folding due to bridging of the binding blocks.

In conclusion, the AFM measurements are completely consistent with the light scattering measurements and indicate that the complexes consist of individual DNA molecules, coated with diblock copolymers: contour lengths for the pDNA complexes are in a very good agreement with the theoretical half contour length for 3.7 kbp DNA (-629 nm, 0.34 nm/bp), and for the complexes with linearized pDNA the contour length equals the full contour length of the naked DNA, as expected. Furthermore, the observed heights are very uniform, which would not have been the case if there had been a spread in the number of dsDNA chains per complex.

Numerical Estimates. DNA has two charges every 0.34 nm (one bp), such that the bare linear charge density ν of dsDNA (elementary charges e per unit length) is $\nu \sim 5.9\ nm^{-1}$. Assuming a single B_{12} block neutralizes 12 DNA phosphate charges, it covers an estimated length of dsDNA of h_{min} = 2 nm. DNA-bound diblock copolypeptides start feeling each other when they are bound at a spacing h along the DNA contour smaller than the coil size of the C_4

block. This occurs at a critical spacing h^* and critical degree of neutralization f^*

$h^* \approx 2R \approx 12\text{nm}$

$f^* \approx h_{min} / h^* \approx 0.17$

where $R \approx 6$ nm is the coil radius of the C_4 block, as determined using DLS. For N/P $> f^*$, electrostatic binding of the diblocks is opposed by steric repulsion of neighbouring C_4 blocks. As is clear from the experimental data, binding continues to increase well beyond N/P $= f^*$, and this will lead to stretching of the C_4 blocks. When the free energy penalty due to side-chain stretching exceeds the binding free energy of the B_{12} block, no more additional polymers are bound, even if not all of the DNA charges have been neutralized. Both the electrophoresis and the static light scattering suggest that this is the case for our diblocks at N/P $\approx$ 1.

The side-chain stretching has been discussed extensively in the literature on so-called bottle-brush polymers, consisting of a main chain covered with regularly spaced side chains [19]. Assuming good solvent conditions, the scaling estimate for the degree of stretching of the grafted side chains (each consisting of N segments of length l) is:

$$\frac{R_s}{R} \propto \Gamma^{1/4}$$

$$\Gamma = N\frac{l}{h}$$

where R_s is the mean-square end-to-end distance of the stretched chain and R is the mean-square end-to-end distance of the unstretched chain. Note that in the absence of a numerical prefactor for this scaling estimate, it can only be used to get order-of-magnitude estimates. Assuming $l = 0.5$ nm and N=400, for the fully charge-neutral complex with $h=h_{min}$ =2nm, we find $\Gamma^{1/4} \approx 3$. This implies that there is only moderate stretching. From dynamic light scattering, we inferred diffusion diameters the coated supercoils of $D \approx 100$ nm. Assuming the diffusion diameter of the coated supercoils is about 4 times the radius R_s of the stretched side-chains (see sketch), this gives $R_s \approx 25$ nm, or a stretching factor $R_s / R \approx 4$, which is of the expected order of magnitude.

A further issue that is highly debated in the literature on bottle-brush polymers, is the stiffening of the main chain, induced by the side chains. Current simulations suggest that at least for relatively short side-chains ($N < 100$) the contribution $l_{P,S}$ of the side chains to the total persistence length of dense bottle-brush polymer is typically on the order of its thickness. According to numerical self-consistent field calculations, very strong stiffening only occurs for densely grafted, extremely long side chains ($N \gg 1000$):

$$l_{p,s} \approx \mu \cdot \frac{l^3}{h^2} N^2$$

This is mainly due to a very small numerical prefactor, $\mu \approx 0.02$. For our case (again

considering the limiting case of assuming full charge neutralization, $h = 2$ nm), the latter equation would imply $1_{p,s}$ « 100 nm for a single dsDNA. This order-of-magnitude estimate seems to be consistent with the images of the fully coated 3.7kb supercoiled plasmid molecules, that are close to rod-like. Considering the entire supercoil as a semiflexible object, its estimated persistence length would increase to about 300nm after coating, about half the total supercoil contour length of about 650nm, such that is should indeed appear as close to rod-like.

Conclusions

Light scattering confirmed that the DNA was successfully coated with the diblock copolymer of the present invention.

Charge screening confirmed that complexes are formed due to electrostatic interactions.

AFM studies confirmed that complexes consist of individual DNA molecules coated by diblock copolymer rather than aggregates.

Materials and Methods

Materials. circular supercoiled plasmid DNA (3.7 kbp) used for binding studies was prepared using a Megaprep kit (Qiagen, Germany) from a recombinant *E. coli* broth culture grown overnight. The purified DNA was dissolved in sterile MQ-water for a final concentration of 232 ng^L (Nanodrop ND-1000 Spectrophotometer; Thermo Scientific, Waltham, MA). Portions were diluted till a final concentration of 116 ng^L in 10 mM acetate buffer at pH 5 and filtrated (0.2μη) for DLS analysis. Restriction enzymes used for molecular cloning of the cationic protein polymers were purchased from New England Biolabs (Ipswich, MA) or Fermentas (ON, Canada). Acetate buffer was prepared from acetic acid and sodium acetate of analytical quality purchased from Sigma (St. Louis, MO).

Cloning of cationic protein polymers. C4 block and binding blocks were built separately and later put together. Two different cationic polypeptides C4-K12 and C4-HK6 were produced. "C4" refers a 4-fold repeat of a collagen-like block of approximately 100 amino acids (SEQ ID NO.l) and "B12" a binding block of twelve basic amino acids length either Lys (K12) (SEQ ID NO. 2) or the dyad His-Lys [$(HK)_6$] (SEQ ID NO.3). The Collagen block (C4) was built it from pMTL23-C2 plasmid by directional ligation. The plasmid pMTL23-C2 was doubly digested with *Van91I+Dralll* restriction enzymes and the C2 fragment was isolated. Then, it was ligated into a previously *Van91I* digested/dephosphorylated pMTL23-C2 plasmid obtaining pMTL23-C4 where the *Dralll* restriction site is lost and only *Van9II* kept in both sites of the duplicated segment. The binding blocks were build using two complementary oligonucleotides (Eurogentec, Belgium) codifying for the particular sequence of each block. They both carried out sticky ends for *Van911 and EcoRI* restriction enzymes. They were annealed by heating at 90 °C for 5-10 min and left at room temperature for 2-3 h. Upon this, they were ligated into the plasmid pMTL23-C4 previously digested with EcoRI + Van9 II, obtaining the plasmids

pMTL23-C4-K12 and pMTL23-C4-HK6. These plasmids were propagated into E. coli after transformation using Z-competentTM E. coli transformation kit (Zymo Research Corporation). Plasmid sequencing and PCR-colony using M13 primers were carried out to verify presence of intact blocks genes in the correct reading frame.

Fragments codifying for the corresponding C4-K12 (SEQ ID NO.1 and SEQ ID NO.2) and C4HK6 (SEQ ID NO.1 and SEQ ID NO.3) cationic protein polymers were obtained from pMTL23-C4-K12 and pMTL23-C4-HK6 through digestion with *XhoI + EcoRI* and ligated into pPIC9 plasmid previously digested with the same restriction enzymes. The resulting plasmids pPIC9-C4-K12 and pPIC9-C4-HK6 were linearized *with SalI* and electroporated into *Pichia pastoris* strain GS115 (Mut^+). The plasmid integrates into the genome through homologous recombination at the *his4* locus providing high genetic stability of the resulting recombinant strains and leaves the cationic protein polymers genes expression under the control of AOX1 promoter enabling growth on methanol. Transformed yeasts were poured into Minimal Dextrose agar medium plates (1.34% Yeast Nitrogen Base, 2% dextrose, $4x10^{-5}$% biotin and 2% agar) and grew for 2-3 days at 30 °C. Some recombinant colonies were re-streaked and grown on plates with fresh medium. The presence of the genes was verified by PCR using specific primers for the AOX1 flanking sequences (5-AOX1: GACTGGTTCCAATTGACAAGC (SEQ ID NO. 4) and 3-AOX1: GCAAATGGCATTCTGACATCC (SEQ ID NO. 5)). Positive recombinant colonies were grown in YPD medium (1% yeast extract, 2% peptone plus 2% dextrose) until reaching an optical density (λ = 600 nm) between 5-10 to be dissolved into 2/3 of 80% glycerol and stored at -80oC in vial ready-to-use for fermentation.

Cationic Protein Polymers Biosynthesis. Fed-batch fermentations using minimal basal salts medium were performed in 2.5-L Bioflo 3000 fermenters (New Brunswick Scientific, Edison, NJ). The methanol fed-batch phase for protein production lasted two to three days. A homemade semiconductor gas sensor-controller was used to monitor the methanol level in the off-gas and to maintain a constant level of 0.2%> (w/v) methanol in the broth. The pH was maintained at 3.0 throughout the fermentation by addition of ammonium hydroxide. At the end of the fermentation, the cells were separated from the broth by centrifugation for 15 min at 10000 × g (room temperature or 4 °C) in an SLA-3000 rotor (Thermo Scientific, Waltham, MA), and the supernatant was microfiltered (Pall Corporation, Port Washington, NY) to avoid possible proteolysis by remaining cells and immediately stored at 4 °C for subsequent purification.

Cationic Protein Polymers Purification. All centrifugations were done for 30 min at 20000 x g at 4 °C, interchangeably in a Sorvall SLA-1500 or SLA-3000 rotor. First, medium salts impurities were removed from the cell-free broth solution by precipitation by adjusting the pH to 8.0 with NaOH 1M followed by centrifugation. After this the diblock copolypeptides were selectively precipitated from the solution meanwhile *P. pastoris* proteins stayed in solution by

adding ammonium sulphate till a saturation of 45% m/v and incubating over night at 4 °C. The pellet was resuspended in an equal volume (relative to the cell-free broth) of Milli-Q water and precipitation was repeated once at 4°C for 1-2 hrs. The pellet was resuspended in 0.2 volumes (relative to the cell-free broth) of Milli-Q water and sodium chloride (50 mM) and acetone (40% v/v) were added. After centrifugation the pellet was ruled out and acetone concentration was raised to 80% (v/v) and solution was centrifuged to separate the pellet containing the pure diblock copolypeptide. The o/n air-dried pellet was resuspended in Milli-Q water, extensively desalted by dialysis against Milli-Q water and lyophilized.

SDS-PAGE and Mass Spectrometry Analysis. SDS-PAGE was carried out using the NuPAGE Novex system (Invitrogen, Carlsbad, CA) with 10% Bis-Tris gels, MES SDS as running buffer and SeeBlue Plus2 prestained molecular mass markers. Gels were stained with Coomassie SimplyBlue SafeStain (Invitrogen). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) was carried out in an Ultraflex mass spectrometer (Bruker, Billerica, MA). Proteins samples (lmg/mL) were prepared by the dried droplet method. The matrix was made up of 5 mg/mL 2.5-dihydroxyacetophenone, 1.5 mg/mL diammonium hydrogen citrate, 25% (v/v) ethanol, and 1% (v/v) trifluoroacetic acid on a 600 μη AnchorChip target (Bruker). Measurements were made in the positive, linear mode. An external mass calibration was done based on Protein Calibration Standard II (Bruker).

Agarose Gel Mobility Retardation Assay. 46 ng of pDNA dissolved in 10 mM acetate buffer pH 5 were mixed with different volumes of C4-B12 lg/L solution dissolved in the same buffer for a final volume of 15 μL. After 30-60 min at room temperature they were poured in a 1% agarose gel and run at 90v for 60 min using TAE IX buffer. Bands were visualized using ethidium bromide.

Dynamic Light Scattering (DLS). Light scattering measurements were performed with a Zetasizer NanoZS apparatus (Malvern Instruments, UK) equipped with a 4mW He-Ne ion laser working at a wavelength of 633nm. DNA-cationic protein complexes were formed at room temperature mixing 5.6 μL of DNA (116ng^L) dissolved in buffer together with certain volume of desalted and purified C4-B12 (1-10 mg/mL) dissolved in the same buffer necessary to obtain the particular N/P ratio to be analyzed. DNA-diblock copolypeptide premixed solutions were filled up till 15 μL with 10 mM acetate buffer pH 5 and mixed. DNA, protein and buffer solutions were previously filtrated using 0.2μη cut-off filters for all the complexing experiments. The resulting solutions were left to equilibrate for 1-2 hrs at room temperature before DLS measurements. The intensity of scattered light by the DNA-cationic protein polymers complexes previously formed was determined from an average of 5 correlation measurements followed by 60 seconds at angle of 12.8 degrees and/or 173 degrees at 25°C. Effective hydrodynamic radius was automatically gotten from the apparent diffusion coefficient calculated by the apparatus.

Atomic Force Microscopy. DNA-diblock copolypeptide complex solutions (prepared the same way as for DLS experiments) were diluted 10 or 20 times with filtrated Milli-Q water. Immediately 3-5 μL of the resulting solution was added onto a clean hydrophilic lxl cm silicon wafer and left for 1-2 min. Then it was washed with 500 μL of filtrated Milli-Q water to remove salts and not absorbed particles followed by tissue removal of water excess and a slow drying by nitrogen stream. Samples were analyzed using a Digital Instruments NanoScope V equipped with a silicon nitride probe with a spring constant of 0.32 N/m in ScannAsyst mode (Veeco, NY, USA). Images were recorded between 0.488-0.965 Hz and 384-1024 samples/line. Image processing was done with NanoScope Analysis software. Contour length and long axis length measurements were performed with ImageJ software.

REFERENCES

1. Hegner M, Smith S. B., Bustamante, C. PNAS, 1999.
2. Hall AR, van Dorp S, Lemay SG, Dekker C. Nanoletters, 2009.
3. Kostiainen MA, Szilvay GR, Lehtinen J, Smith DK, Linder MB, Urtti A, Ikkala O. ACS nano, 2007, 1.
4. Douglas S. M.; Dietz, H.; Liedl, T.; Hogberg. B.; Graf. F.; Shih, W. Nature. 2009, 459.
5. Maubach G, Born D, Csaki A, Fritzsche W. Small, 2005, 1.
6. Keren K, Berman RS, Buchstab E, Sivan U, Braun E. Science, 2003.
7. Smeets RMM (Smeets, R. M. M.)l, Kowalczyk SW (Kowalczyk, S. W.)l, Hall AR (Hall, A. R.)l, Dekker NH (Dekker, N. H.)l, Dekker C (Dekker, C.)l nanoletters, 2009, 9.
8. Walker, G. F.; Fella, C ; Pelisek, J.; Fahrmeir, J.; Boeckle, S.; Ogris, M.; Wagner, E. molecular therapy, 2004, 11.
9. Martin, E. M.; Rice, K. G. Peptide-guided gene delivery. AAPS journal, 2007, 9.
10. J.K.W. Lama, Y. Mab, S.P. Armesb, A.L. Lewisc, T. Baldwina and S. Stolnika, J. control release, 2004.
11. Hartmann, L.; Hafele, S.; Peschka-Suss, R.; Antonietti, M.; Borner, H. G. Chem. Eur. J. 2008, 14, 2025-2033.
12. Numata, K.; Subramanian, B.; Currie, H. A.; Kaplan, D. L. Biomaterials. 2009. 30, 5775-5784.
13. Grayson, A. C ; Doody, A. M.; Putnam, D. PharmRes. 2006, 23, 1868-1876.
14. Osada, K.; Oshima, H.; Kobayashi, D.; Doi, M.; Enoki, M.; Yamasaki, Y.; Kataoka, K. JACS. 2010, 132, 12343-12348.
15. Mann, A.; Richa, R.; Ganguli, M. J Control Release. 2008, 125, 252-262.
16. DeRouchey, Walker, G. F.; Wagner, E.; Radler, J. O. J. phys. Chem B. 2006, 110, 4548-4554.

17. Canine, B. F.; Wang, Y.; Hatefi, A. J Control Release. 2008, 129, 117-123.

18. Zeng, J.; Wang. S. Biomaterials. 2005, 26, 679-686.

19. Hsu, H.; Wolfgang, P.; Kurt, B. Macromolecular theory and simulations, 2007.

SEQ ID NO. 1.

GPPGEPGNPGSPGNQGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGSNGPQGSQG
NPGKNGQPGSPGSQGSPGNQGSPGQPGNPGQPGEQGKPGNQGPAGEPGNPGSPGNQG
QPGNKGSPGNPGPGNEGQPGQPGQNGQPGEPGSNGPQGSQGNPGKNGQPGSPGSQGSP
GNQGSPGQPGNPGQPGEQGKPGNQGPAGEPGNPGSPGNQGQPGNKGSPGNPGQPGNE
GQPGQPGQNGQPGEPGSNGPQGSQNPGKNGQPGSPGSQGSPGNQGSPGQPGNPGQPGE
QGKPGNQGPAGEPGNPGSPGNQGQPGNKGSPGNPGQPGNEGQPGQPGQNGQPGEPGS
NGPQGSQGNPGKNGQPGSPGSQGSPGNQGSPGQPGNGQPGEQGKPGNQGPAGG

SEQ. ID NO. 2

KKKKKKKKKKKKG

SEQ. NO. 3

HKHKHKHKHKHKG.

SEQ.NO. 4

GACTGGTTCCAATTGACAAGC

SEQ.NO.5

GCAAATGGCATTCTGACATCC

CLAIMS

1. A diblock copolymer comprising a collagen-like block consisting of a first amino acid sequence of m amino acids and a DNA binding block consisting of a second amino acid sequence of n amino acids, wherein m and n each independently stand for an integer and wherein the ratio between m and n is at least 20.

2. The copolymer according to claim 1, wherein m is from 300 to 600, preferably from 325 to 550, preferably from 350 to 500, preferably from 375 to 450 or from 390 to 425.

3. The copolymer according to claim 1 or 2, wherein n is from 10 to 25, preferably from 11 to 20 or from 12 to 15.

4. The copolymer according to any one of the preceding claims, wherein the first amino acid sequence is represented by SEQ ID NO. 1 or a sequence having at least 80%, 90%, 95%, 96%, 97%, 98% or 99% identity with SEQ ID NO. 1.

5. The copolymer according to any one of the preceding claims, wherein the second amino acid sequence is represented by SEQ ID NO. 2 or SEQ ID NO. 3, or a sequence having at least 80%, 90%, 95%, 96%, 97%, 98%, 99% identity with SEQ ID NO.2 or SEQ ID NO. 3.

6. A process for making the diblock copolymer according to any one of the preceding claims, the process comprising the step of expressing the nucleic acid sequence encoding said copolymer in *Pichia pastoris* using recombinant expression techniques known per se.

7. The process according to claim 6, wherein said step is performed at pH of at least 2.5 and at most 6.0, preferably at most 5.0, at most 4.5, at most 4.0 or at most 3.5.

8. A copolymer comprising the diblock copolymer according to any one of the preceding claims and a functional block comprising a third amino acid sequence attached to the first amino acid sequence or the second amino acid sequence of the diblock copolymer according to the present invention.

9. A double strand nucleic acid coated with the diblock polypeptide according to any one of claims 1-5 or with the copolymer according to claim 8.

10. Use of the coated nucleic acid of claim 9 in gene therapy, non-viral gene delivery, sequencing by transport through nanopores, nanoelectronics such as nanowire DNA templating artificial virus or DNA origami.

11. A nucleic acid sequence encoding the diblock copolymer according to any one of claims 1-5 or the copolymer of claim 8.

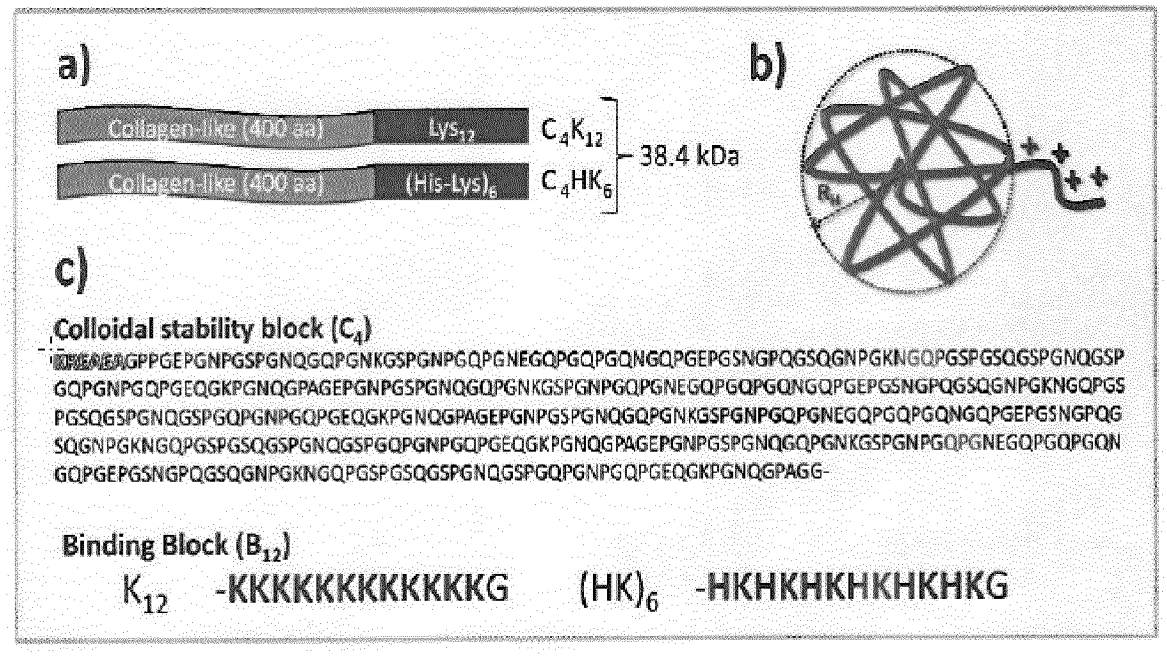


Figure 1.

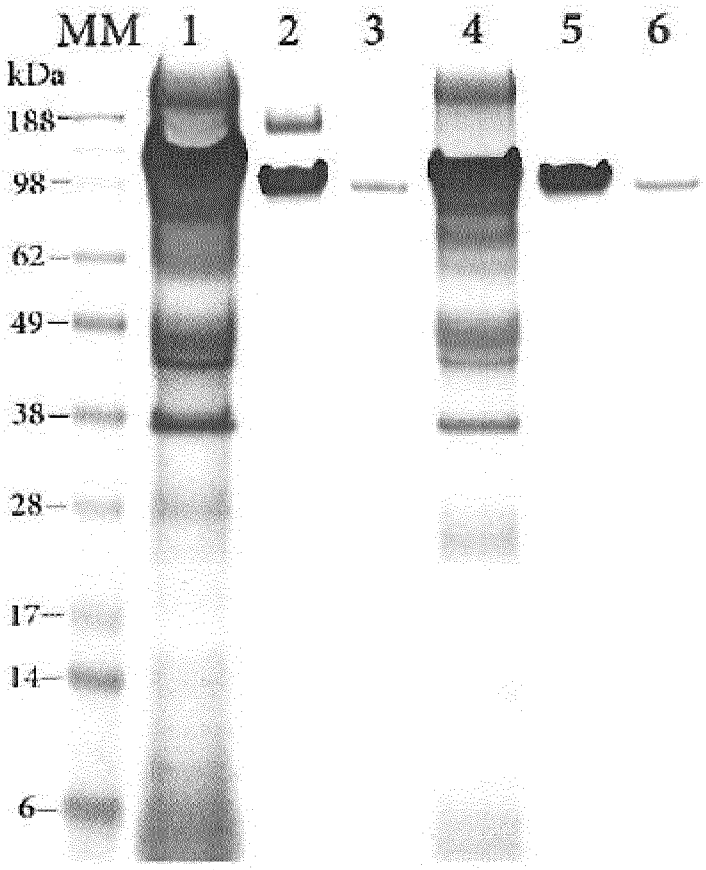


Figure 2a.

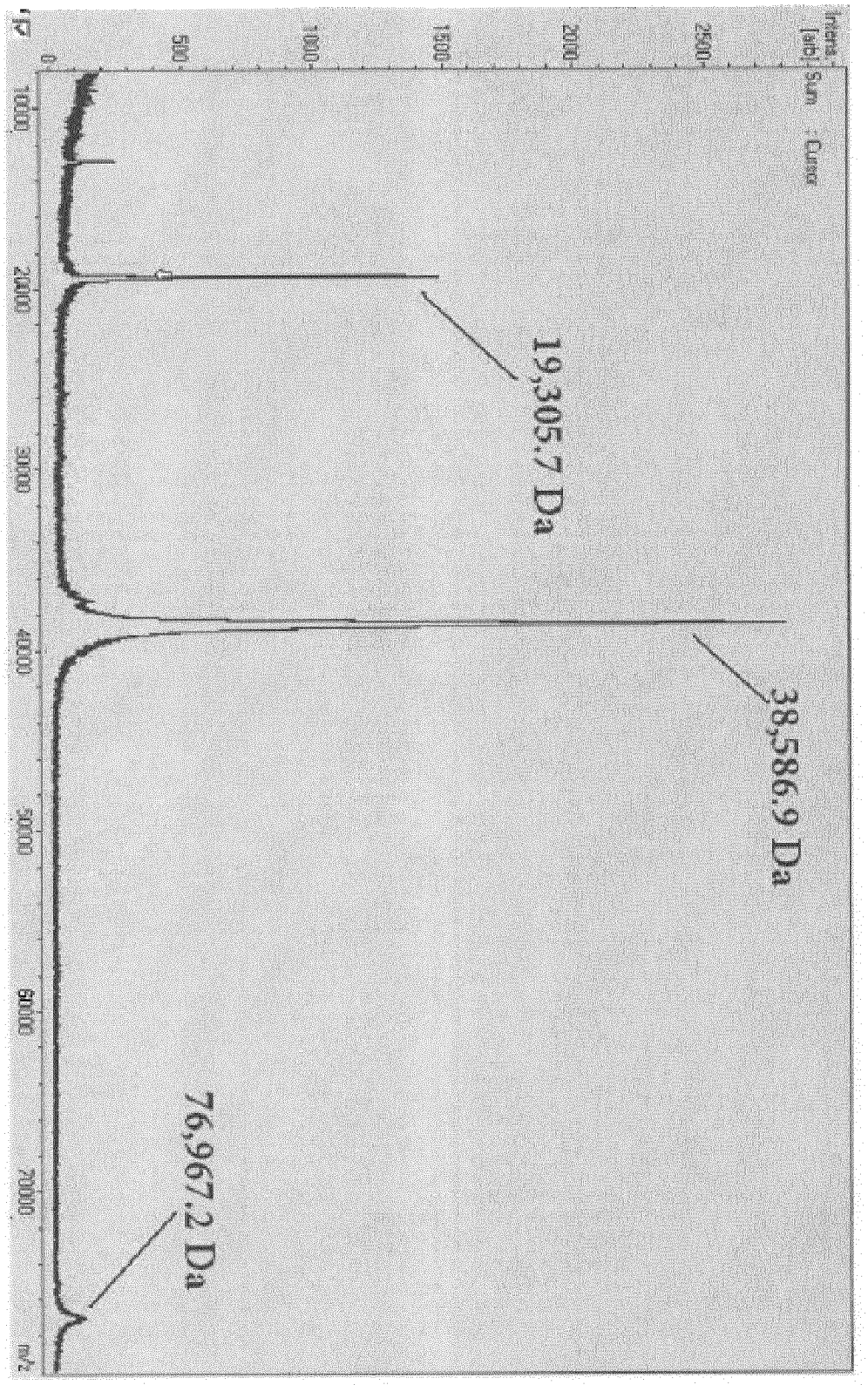


Figure 2b.

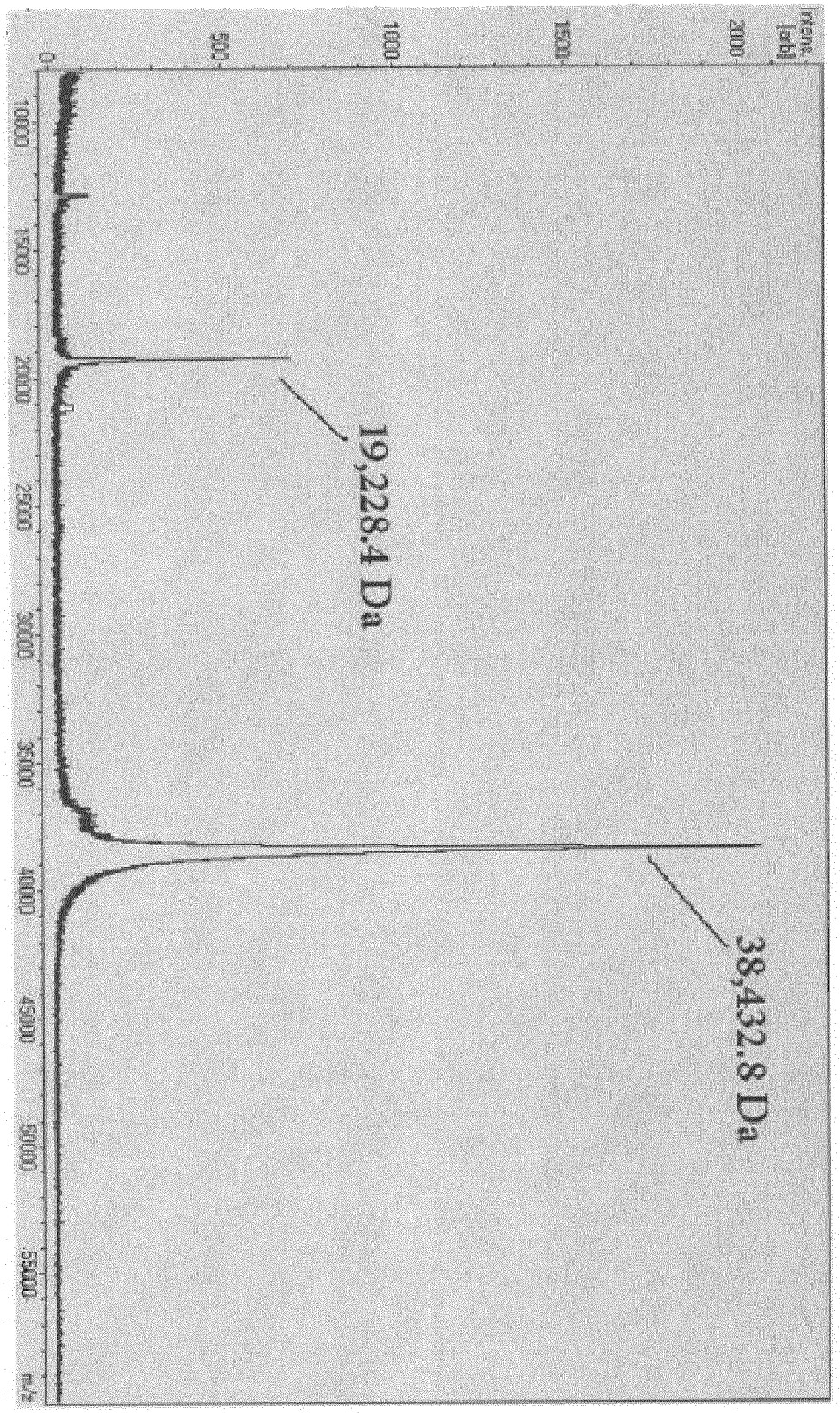

Intens.
[a.u.]
2000
1500
1000
500
0
10000
15000
20000
25000
30000
35000
40000
45000
50000
55000
m/z
19,228.4 Da
38,432.8 Da

Figure 2c.

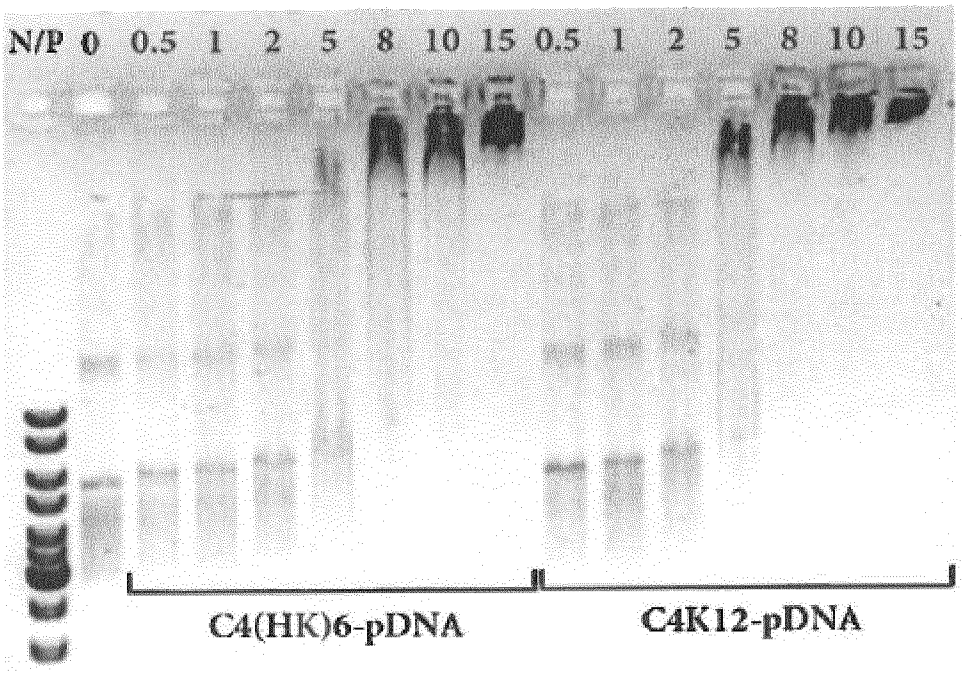


Figure 3.

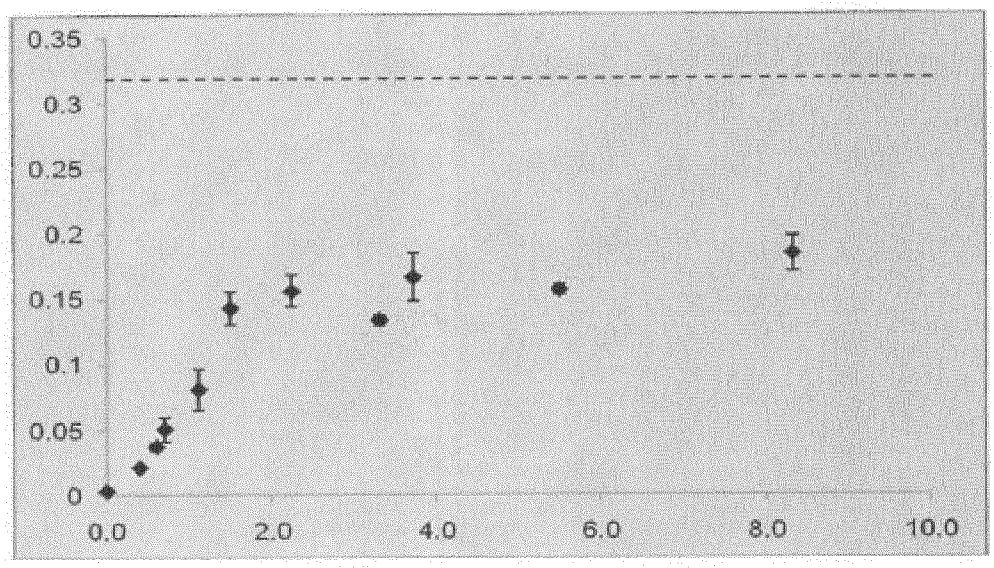


Figure 4-1.

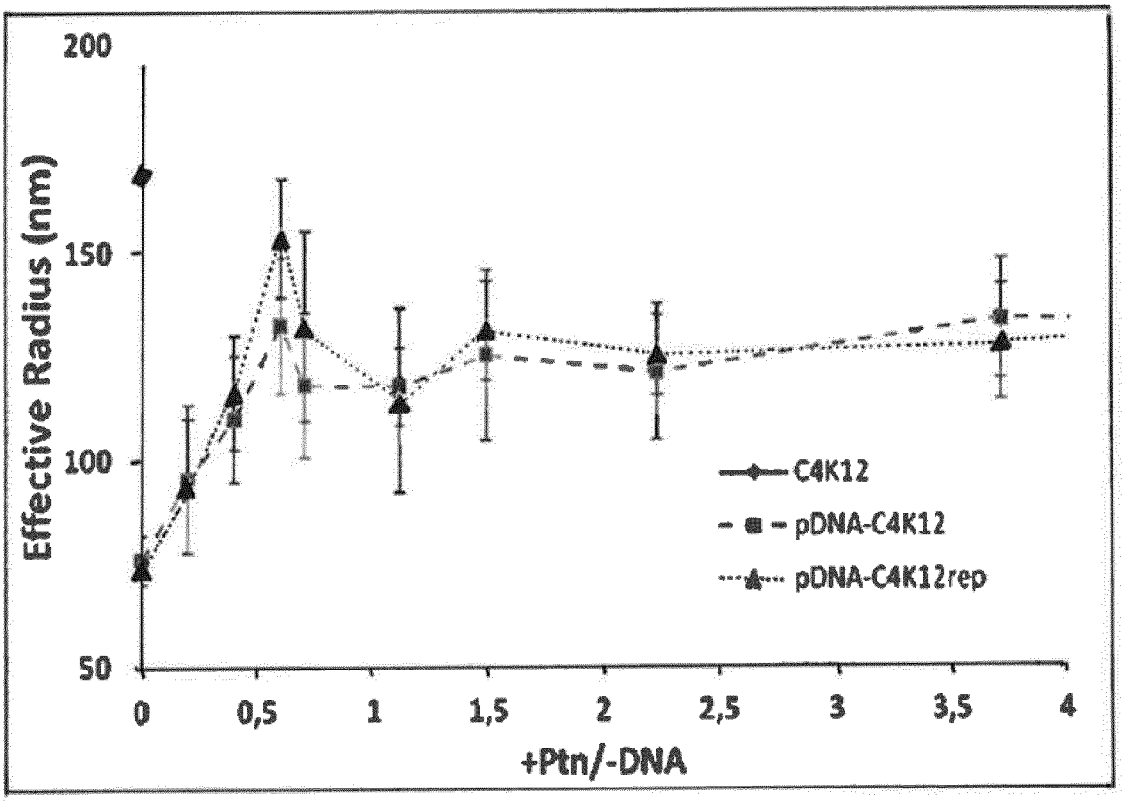
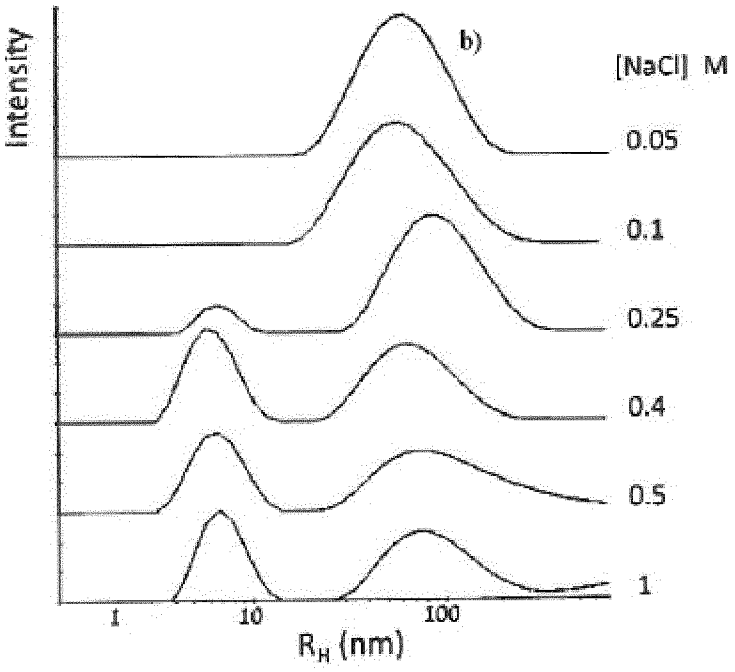
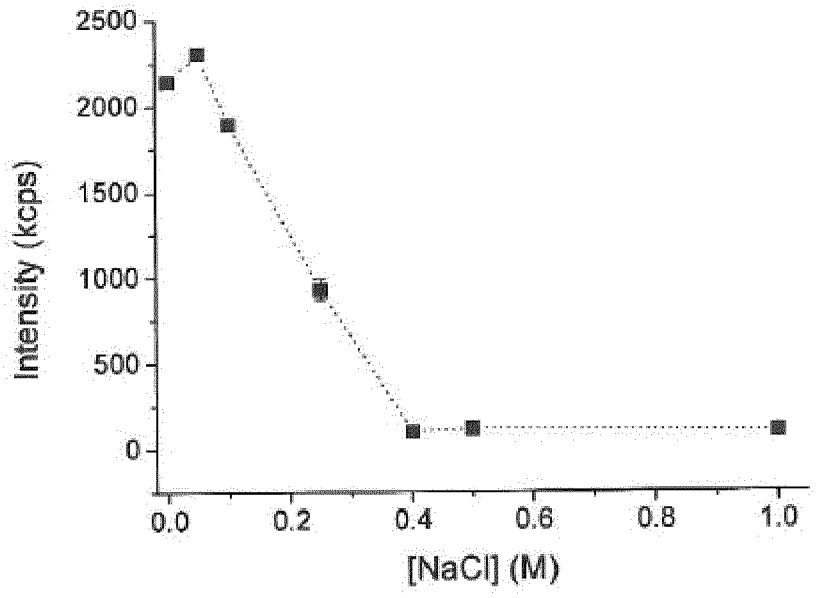

200
150
100
50
Effective Radius (nm)
0
0,5
1
1,5
2
2,5
3
3,5
4
+Ptn/-DNA
C4K12
pDNA-C4K12
pDNA-C4K12rep

Figure 4(a).


b)
[NaCl] M
0.05
0.1
0.25
0.4
0.5
1
Intensity
1
10
100
R_H (nm)

(a)


2500
2000
1500
1000
500
0
Intensity (kcps)
0.0
0.2
0.4
0.6
0.8
1.0
[NaCl] (M)

(b)

Figure 5.

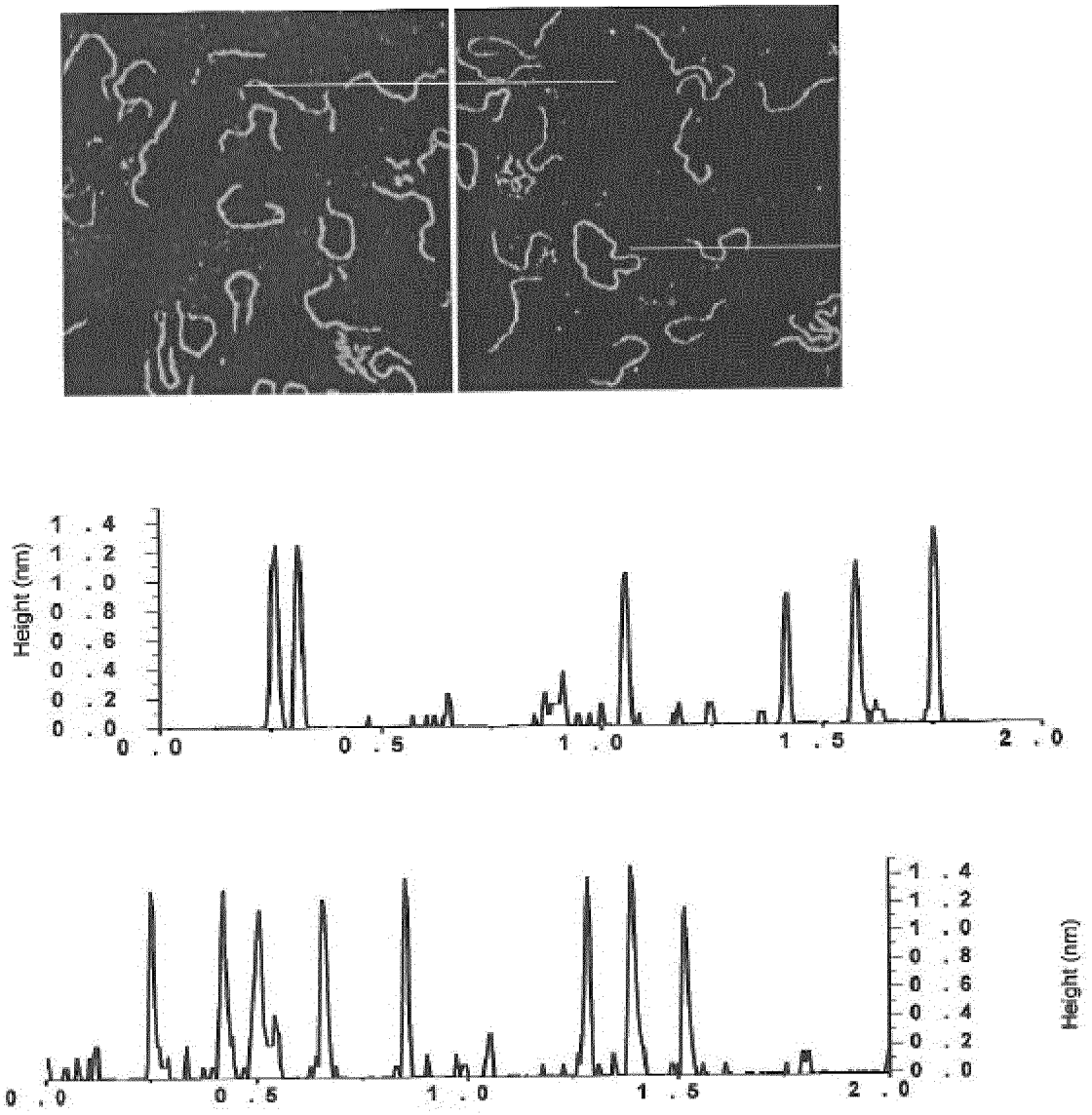


Figure 6.

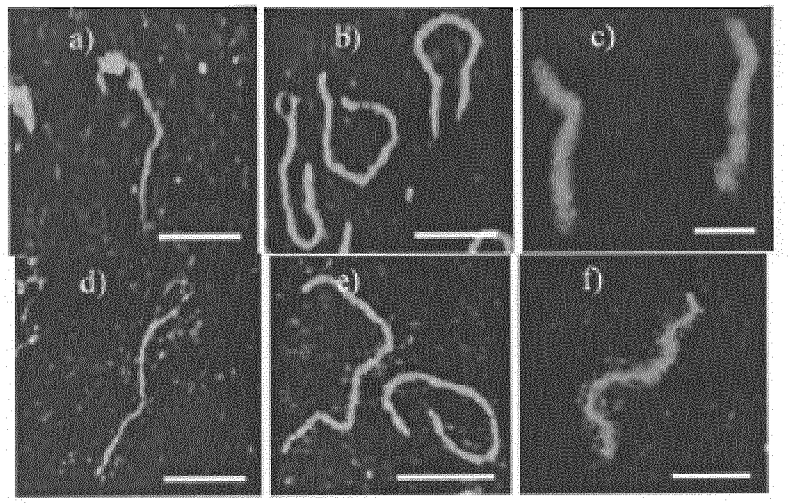


Figure 7.

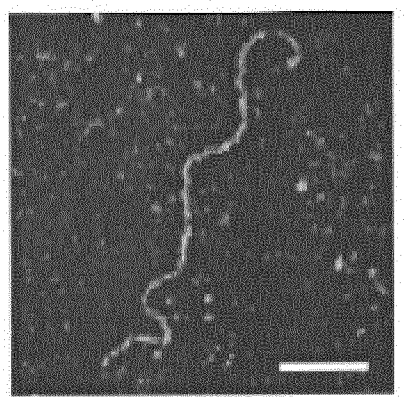


Figure 8.

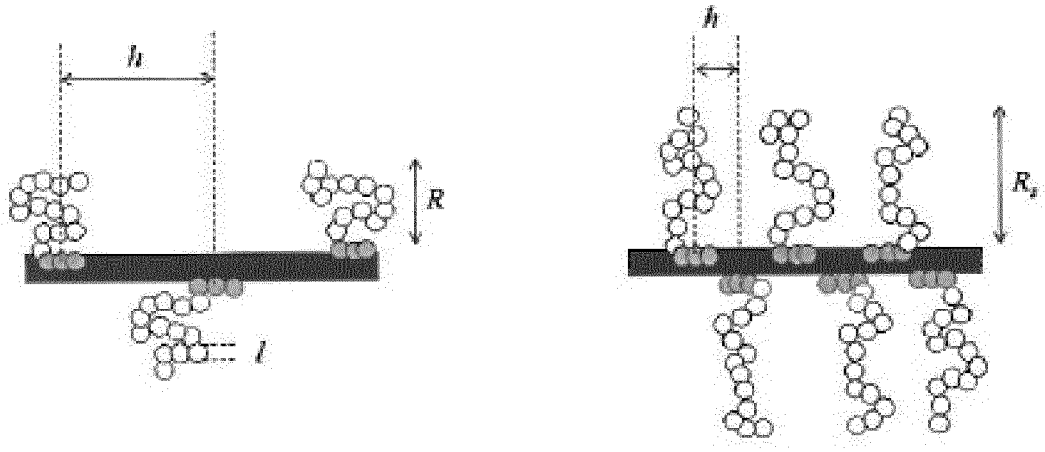


Figure 9.

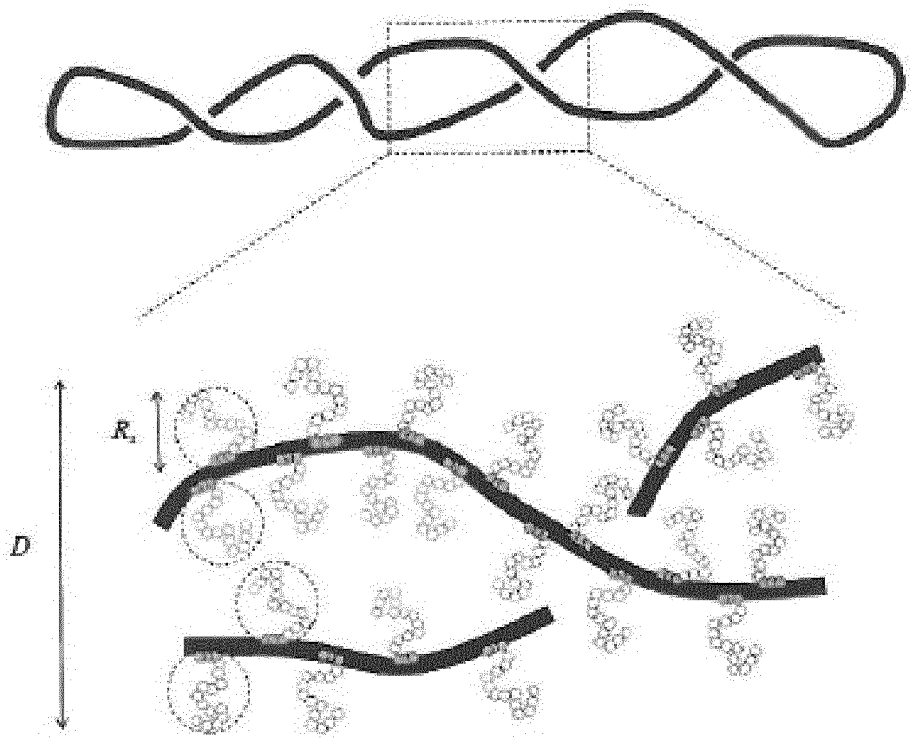


Figure 10. Coated supercoiled DNS.

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/061974

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/00 C07K14/78 A61K47/48
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal , WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/006133 A2 (TUFTS COLLEGE [US] ; NUMATA KEIJI [US] ; KAPLAN DAVID L [US]) 13 January 2011 (2011-01-13) page 3; claims 1-79 ; figures 1,2	1-11
X	NUMATA K ET AL: "Bioengineered silk protein-based gene delivery systems" , BIOMATERIALS, ELSEVIER SCIENCE PUBLISHERS BV. , BARKING, GB, vol . 30, no. 29, 1 October 2009 (2009-10-01) , pages 5775-5784, XP026470004, ISSN : 0142-9612 , DOI : 10. 1016/J . BIOMATERIALS.2009.06.028 [retrieved on 2009-07-04] cited in the application abstract; figures 2,8	1-11

-/- .

X Further documents are listed in the continuation of Box C.

X See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" documentwhich may throw doubts on priority claim(s) orwhich is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 July 2012	30/07/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Schmidt, Harald

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/061974

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/151327 A1 (STICHTING DI ENST LANDBOUWKUNDI [NL] ; DE WOLF FREDERI K ANTON [NL] ; WERT) 17 December 2009 (2009-12-17) page 35; cl aims 1-4	1
X	US 6 903 200 B1 (CH0U MIN-YUAN [TW] ET AL) 7 June 2005 (2005-06-07) exampl e 6	1
A	US 2003/040496 A1 (CHANDLER LOIS ANN [US] ET AL) 27 February 2003 (2003-02-27) cl aims 1-27	1-11
A	FERD0US A ET AL: "COMB-TYPE COPOLYMER: STABI LIZATION OF TRI PLEX DNA AND POSSIBLE APPLICATION I N ANTIGENE STRATEGY", JOURNAL OF PHARMACEUTICAL SCI ENCES, AMERICAN PHARMACEUTICAL ASSOCIATION, WASHINGTON, US, vol . 87, no. 11, 1 November 1998 (1998-11-01) , pages 1400-1405 , XP000783389 , ISSN: 0022-3549 , D0I : 10. 1021/JS980066G abstract; f i gure 1	1-11
A	W0 01/00708 A1 (PLANK CHRISTIAN [DE] ; STEMBERGER AXEL [DE] ; SCHERER FRANZ [DE]) 4 January 2001 (2001-01-04) cl aims 1-15	1-11
A	LE0NG K W ET AL: "DNA-polycati on nanospheres as non-vi ral gene del i very vehi cl es", JOURNAL OF CONTROLLED RELEASE, ELSEVI ER, AMSTERDAM, NL, vol . 53, no. 1-3 , 30 Apri l 1998 (1998-04-30) , pages 183-193 , XP004121269 , ISSN: 0168-3659 , DOI : 10. 1016/50168-3659 (97)00252-6 abstract; f i gure 1	1-11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2012/061974

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011006133 A2	13-01-2011	AU 2010271238 A1	09-02-2012
		EP 2451953 A2	16-05-2012
		US 2012171770 A1	05-07-2012
		WO 2011006133 A2	13-01-2011
WO 2009151327 A1	17-12-2009	EP 2133364 A1	16-12-2009
		WO 2009151327 A1	17-12-2009
US 6903200 B1	07-06-2005	TW 499433 B	21-08-2002
		US 6903200 B1	07-06-2005
US 2003040496 A1	27-02-2003	NONE	
W0 0100708 A1	04-01-2001	AT 265488 T	15-05-2004
		AU 776715 B2	16-09-2004
		AU 5222800 A	31-01-2001
		CA 2377207 A1	04-01-2001
		ES 2219346 T3	01-12-2004
		JP 2003503370 A	28-01-2003
		US 2003026840 A1	06-02-2003
		US 2009239939 A1	24-09-2009
		WO 0100708 A1	04-01-2001