Gene Transfection Agents

Gemini Surfactants: New Synthetic Vectors for Gene Transfection

Anthony J. Kirby, Patrick Camilleri, Jan B. F. N. Engberts, Martin C. Feiters, Roeland J. M. Nolte, Olle Söderman, Mark Bergsma, Paul C. Bell, Matthew L. Fielden, Cristina L. García Rodríguez, Philippe Guédat, Andreas Kremer, Caroline McGregor, Christele Perrin, Gaël Ronsin, and Marcel C. P. van Eijk*

Keywords:
DNA · drug delivery · gene therapy · surfactants · transfection

The superior surfactant properties of cationic gemini surfactants are applied to the complex problem of introducing genes into cells. Of almost 250 new compounds tested, of some 20 different structural types, a majority showed very good transfection activity in vitro. The surfactant is shown to bind and compact DNA efficiently, and structural studies and calculations provide a working picture of the "lipoplex" formed. The lipoplex can penetrate the outer membranes of many cell types, to appear in the cytoplasm encapsulated within endosomes. Escape from the endosome—a key step for transfection—may be controlled by changes in the aggregation behavior of the lipoplex as the pH falls. The evidence suggests that DNA may be released from the lipoplex before entry into the nucleus, where the new gene can be expressed with high efficiency.

[*] Prof. A. J. Kirby, Dr. P. Guédat, Dr. C. McGregor, Dr. G. Ronsin
University Chemical Laboratory
Cambridge CB2 1EW (UK)
Fax: (+44) 1223-336-362
E-mail: ajk1@cam.ac.uk

Dr. P. Camilleri, Dr. A. Kremer, Dr. C. Perrin
Glaxo-SmithKline Pharmaceuticals
Science Complex 1, New Frontiers Science Park North
Third Avenue, Harlow, Essex CM19 5AW (UK)

Prof. J. B. F. N. Engberts, M. Bergsma, Dr. P. C. Bell, Dr. M. L. Fielden
Department of Chemistry, University of Groningen
Nijenborgh 4, 9747 AG Groningen (The Netherlands)

Dr. M. C. Feiters, Prof. Dr. R. J. M. Nolte, Dr. C. L. García Rodríguez
Department of Organic Chemistry, University of Nijmegen
Toernooiveld 1, 6525 ED Nijmegen (The Netherlands)

Prof. O. Söderman, Dr. M. C. P. van Eijk
Division of Physical Chemistry 1
Chemical Center, University of Lund
P.O.B. 124, 22100 Lund (Sweden)

1. Introduction

This Minireview describes an extensive interdisciplinary investigation of gemini surfactants as vectors for gene transfection in vitro. It reflects a highly successful collaboration involving five independent groups from three European countries.[1] The long-term objective is the development of synthetic vectors for gene therapy, which in its simplest form requires the introduction of a missing or defective gene into the cell nucleus. It is known that some cationic surfactants can support transfection, and that gemini surfactants possess superior surfactant properties. We find that geminis are also—in many cases—superior transfection agents.

1.1. Gemini Surfactants

Gemini surfactants[2] are a relatively new class of amphiphilic molecules containing two head groups and two aliphatic chains, linked by a rigid[3] or flexible[4] spacer, as illustrated in Figure 1. They typically show greatly enhanced surfactant properties relative to the corresponding monovalent (single chain, single head group) compounds—surface activity can be increased 1000-fold. This makes them of special interest for biological and especially biomedical applications, where it is

  DOI: 10.1002/anie.200201597

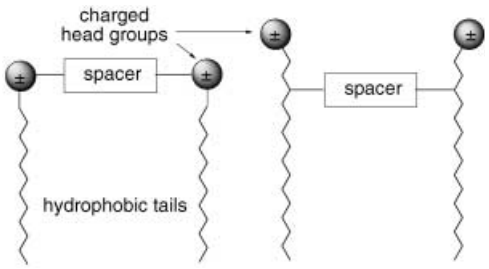


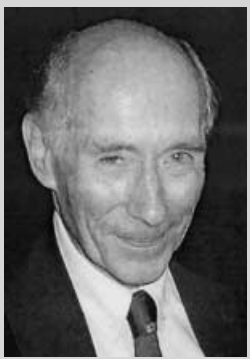
Figure 1. The basic structure of gemini (twinned) surfactants.

essential to optimize the safety profile of any foreign compound: the first and simplest step is to minimize its concentration in vivo. Since using less compound to achieve the same effect also has clear economic advantages, recent years have seen a rapidly growing number of patents describing a whole range of applications of gemini surfactants.

1.2. Gene Therapy and Transfection

The simplest form of gene therapy involves the introduction of a missing or defective gene into the cell nucleus. Once in place it can be expressed, using the natural machinery of the cell, to produce proteins needed to correct a specific pathological condition. A major problem for the routine use of gene therapy in the treatment of disease is the efficient introduction of DNA into the cell nucleus (transfection). The research effort in this field is correspondingly intense.[5–8]

There is broad agreement[9] that the most significant barriers to efficient transfection are the cell and nuclear membranes, and there is intense interest in the development of practical techniques that will allow DNA to cross these membranes. Potentially the most promising approach involves chemical or biological modification, or packaging, of the genetic material. Currently the most efficient method is the incorporation of therapeutic DNA into engineered viruses,[10–12] which have the inherent advantage of specific biological "keys" that allow them passage across the various membranes and into the nucleus. However, their DNA carrying-capacity is limited, and a cloud hangs over their therapeutic use, because of their unpredictable immune response and concerns about safety.[13,14] As a result the search for alternative vectors has intensified.



Tony Kirby studied chemistry at the University of Cambridge, where he obtained his PhD in 1962. He then spent a postdoctoral year with William P. Jencks at Brandeis University. He has been coordinator of three European networks on catalytic antibodies (1993–1996), gemini surfactants (1997–2001), and artificial nucleases (2000–2004). He retired in October 2002 as Professor of Bioorganic Chemistry at the University of Cambridge, and he is now free to concentrate on research and on some exciting new collaborations, most recently in Brazil.

2. The Delivery Problem

Gene transfection can be regarded as a special problem in drug delivery. It is typically dependent on such factors as solubility (which will be pH-dependent for ionizable compounds), how the drug is transported through the system, how long it survives before it is metabolized, and how easily it can penetrate various physical barriers. It is special in that the "drug" to be delivered is a specific piece of DNA, with success depending absolutely on effective penetration of the cell and nuclear membranes. The problems are illustrated schematically in Figure 2.

2.1. Steps Leading to Transfection and Protein Expression

The essential first step is the complexation and consequent compaction (Section 4.2) of the nucleic acid polyanion by the cationic surfactant (Figure 2, step 1). In the absence of a suitable vector no significant level of transfection is observed. This step is discussed in detail in Section 4.3 in the light of our new results, which show that lipoplexes are readily formed from gemini surfactants with various forms of DNA (and RNA), at lower than critical aggregation concentrations of surfactant.

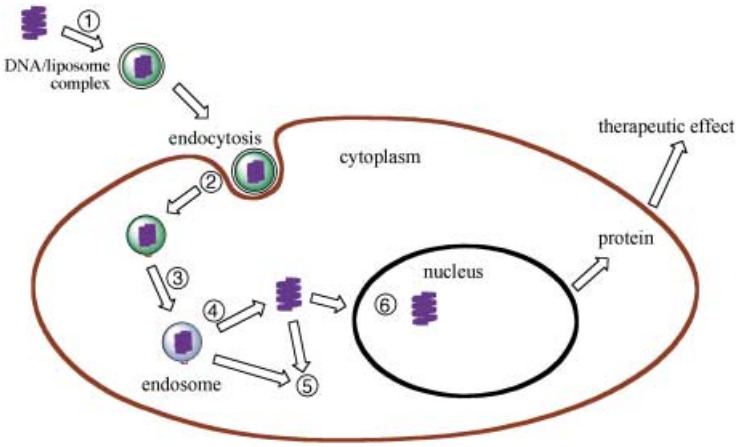


Figure 2. Simplified summary of the key steps involved in liposome-mediated transfection, leading to the expression of a new protein. (Steps 1–6: see text for details.)

The initial passage through the cell membrane mediated by synthetic surfactants is presumed to occur by endocytosis (Figure 2), as demonstrated by Zabner et al. using electron microscopy.[15] Endocytosis involves a controlled invagination of the cell membrane, allowing the DNA/liposome complex to be enveloped by the cell membrane. The membrane then "buds off" (step 2) to form a new vesicle inside the cell. Such vesicles can combine to form endosomes (step 3), which develop increasingly powerful hydrolytic capabilities, with the internal pH falling sharply, until they eventually merge with lysosomes. Only a fraction of the complexed DNA escapes from the endosome: the rest is eventually digested either in the late endosome or after fusion with a lysosome. (Lyso-

somes deploy powerful acid hydrolases—nucleases, peptidases, glycohydrolases, etc.—to provide an efficient recycling capability.) DNA within the liposome complex is protected from nuclease action, but will inevitably be degraded unless it is released from the endosome, perhaps still complexed to the vector. This key step (4) is discussed in detail in Section 4.4. Release of the DNA from the DNA/liposome complex presumably occurs at or before this stage because passage through the nuclear membrane (step 6) appears—at least from the circumstantial evidence from our systems—to involve uncomplexed DNA (see Section 3.2). Passage through the nuclear membrane has to compete with the rapid degradation of uncomplexed DNA by cytoplasmic nucleases (step 5).

Thus the strategic points at which vector design can be expected to have significant effects on the efficiency of gene delivery are the formation of the lipoplex (Figure 2, step 1), its passage through the cell membrane (step 2), and the release of the lipoplex from the endosome and its subsequent dissociation (step 4). These processes are discussed in Section 4.

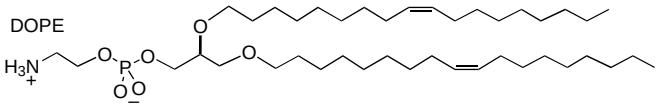
2.2. Principles of Vector Design

Essential requirements for a vector are thus that it should bind DNA sufficiently strongly and rapidly, readily penetrate the target cell and perhaps its nucleus, and eventually release DNA from the lipoplex inside the target cell at the right time and in the correct place. It should also be nontoxic, nonimmunogenic, and biodegradable, though it must be reasonably stable in biological fluids to survive long enough to carry out its functions. The first, key step in the whole process (Figure 2, step 1) is the compaction of the extended, high-molecular-mass, negatively charged DNA into a dense, positively charged (possibly neutral) particle small enough to be taken up by the cell. This generally requires a chemical species bearing multiple positive charges to replace the monovalent counterions of DNA. (Something as simple as calcium phosphate has been used,[16] but gives very low transfection efficiency.)

Many macromolecular and supramolecular systems have been developed to carry the positive charge. These include cationic polyelectrolytes such as DEAE-dextran,[17] polylysine,[18,19] polyethyleneimine,[20] polynorbornane,[21] and polyamine dendrimers.[22,23] The supramolecular systems of particular interest are those that form amphiphile aggregates, most commonly liposomes (or vesicles).[8] Synthetic cationic surfactants are involved in 18% of current clinical trials based on gene therapy, and this proportion increases year by year.[24]

Cationic amphiphiles can compact and stabilize DNA by a combination of attractive electrostatic interactions and hydrophobic interactions between the apolar hydrocarbon tails.[25] The DNA in the aggregate (lipoplex) is protected from endogenous nucleases, while the hydrophobic elements of the aggregate may also promote escape from the endosome by fusion or aggregation with the endosomal membrane.

This step—escape from the endosome—appears to be a major obstacle to efficient transfection using these multivalent cationic vectors, since generally only a small fraction of the compacted DNA escapes (the rest being digested in the lysosome). Various methods have been devised to promote endosomal escape, in particular the use of additional "helper lipids" such as DOPE that are capable of changing the



morphology of the lipoplex from vesicular to inverted hexagonal in the endosome, and thus the interaction with the endosomal membrane. The similarities between these structures and some of the amphiphilic transfection agents discussed below suggest that suitable compounds could play both roles, dispensing with the need for additional helper lipids.

3. Application of Gemini Surfactants

We have developed a new class of transfection agents, designed to combine the cationic character necessary for binding and compacting DNA, with the superior surface activities of gemini surfactants.[26,27] In the basic design (Figure 1) the central "spacer" bears pairs of identical long-chain hydrophobic tails and cationic head groups derived from positively charged α-amino acids and/or amine-linked carbohydrates. All three components are based on natural metabolites, to minimize potential problems with toxicity. The design allows an enormous range of structural variation.

This basic structure offers several opportunities to take advantage of the "chelate effect",[28] by which multiple binding interactions can deliver binding constants and derived effects much greater than the sum of individual binding interactions. (The general principle operates in many different systems, ranging from metal-ion binding and intramolecular reactions to antibiotic activity.[28,29]) Our basic design proved to be remarkably successful: of some 250 new compounds with the basic structure shown in Figure 1 the majority (Figure 3) showed good to very good activity in standard in vitro transfection assays. Thus it is of great interest to identify the structural, physicochemical, and pharmacological factors that are responsible for this activity.

3.1. Structures

The basic gemini design (Figure 1) lends itself to an almost unlimited range of potential structures (we have made and tested only some 20 different structural types), allowing extensive structure–activity studies aimed at identifying features necessary for optimum transfection activity. The central spacer can be derived from any available system with twofold symmetry, or can be custom-made by linking together two molecules of interest. To minimize potential problems

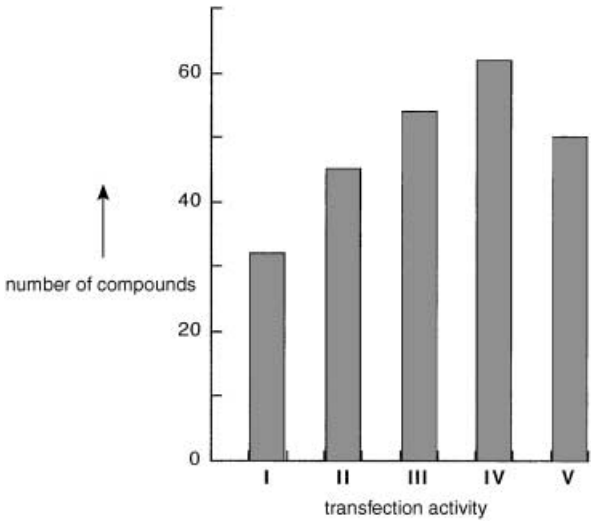


Figure 3. *In vitro* transfection activities of all gemini surfactants produced in this investigation. Activities are defined in broad ranges, from I (low) to V (very high). See text for details.

with toxicity our structures are based on naturally occurring subunits—fatty acids, α-amino acids, lipids, and carbohydrates—to make them generally readily biodegradable. Carbohydrate-based systems are of special interest for potential cell targeting.

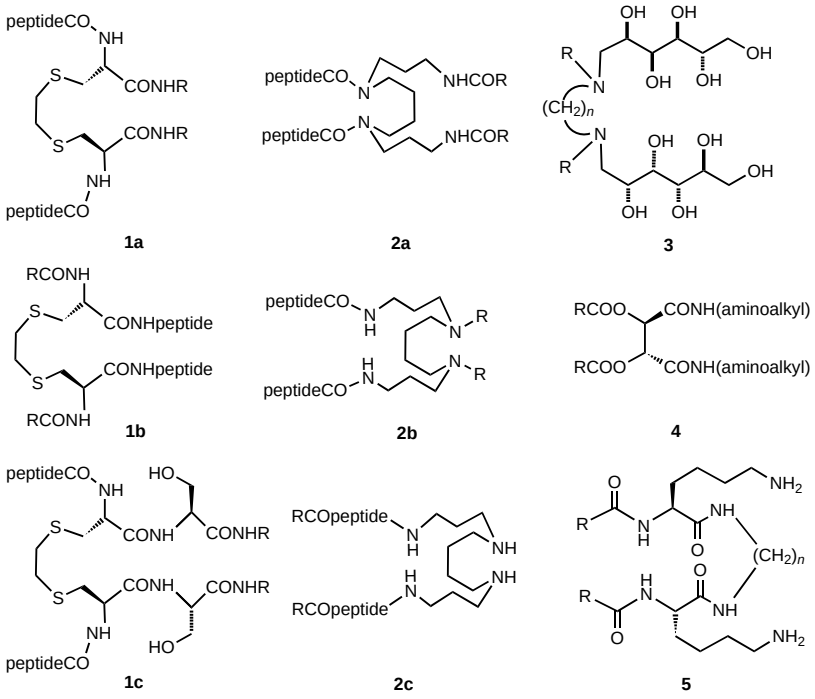
Of the twenty different systems, five were investigated in more depth (Scheme 1): gemini vectors **1** (based on a synthetic cysteine dimer),[30] **2** (based on spermine), **3** and **5** (based on simple symmetrical diamines), and **4** (derived from symmetrical dicarboxylic acids). The polar head groups are in most cases oligopeptides based on naturally occurring α-amino acids, though some of the most effective transfectants are those with amine-linked carbohydrate-based head groups. The hydrophobic tails are derived from naturally occurring fatty acids. Simple changes of connectivity allow two or more series of structures based on each spacer, and these can be multiplied by including branching points based, for example, on α-amino acids such as lysine, with functionality in the side chain. Potential future developments include the use of head groups or spacers designed to bind selectively to cell-surface receptors, and of structural features introduced to control the release of gemini–DNA complexes from the endosome following endocytosis.

3.2. Structure–Activity Correlations

The assisted path of a gene from outside an organism to the nucleus of a particular cell depends on so many variables that simple relationships between transfection efficiency and vector structure are not to be expected. Experimental investigations typically use relatively simple systems, working in vitro with cell lines known to be easy to transfect, and genes coding for proteins that are not native and thus readily detected. The most effective vectors at this level can then be challenged with more testing systems: those that show most promise will make the much bigger jump to in vivo trials. The most effective vectors in vitro of the many gemini structures summarized in Scheme 1 are now reaching the in vivo stage.

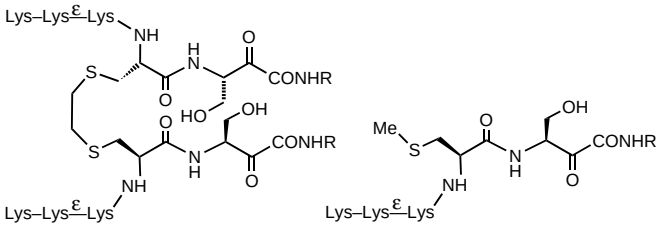
For an efficient primary measure of transfection efficiency we developed a 96-well-plate assay, using a luciferase reporter gene.[31] This assay is both convenient and sensitive, with no background in normal animal cells. Levels of transfection were classified into five broad groups, ranging from I (no activity) to V (most active). (Activities are based on a minimum of eight parallel runs for each compound. These biological systems do not give the high reproducibility familiar to most physical scientists, and our comparisons of activity in plasmid DNA delivery by our gemini surfactants are deliberately based on this broad classification scheme.) To allow informed comparisons with results obtained in other laboratories with other types of vectors, we routinely ran parallel "control" experiments with the commercial nonviral transfection agent lipofectamine 2000/+. (These agents fall in the middle of class V of active transfectants in our experiments.)

The most striking effect is the greatly enhanced activity afforded by the complete gemini surfactant structure. Systems such as **6** with oligolysine head groups including one or



Scheme 1. Structural features of the main classes of gemini surfactants showing excellent gene transfection efficiencies. The R groups are generally long hydrocarbon chains (C_{12}–C_{18}), though in a few cases also steroidal. Structure **3** contains a reduced glucose head group: variants with two of the methylene groups in the spacer replaced by oxygen atoms show particularly high gene-transfer efficacy.

more ε linkages (as opposed to the natural peptide α linkages) are among the most active transfection agents of their type (Scheme 2, R = oleyl). Activity disappears almost completely for the surfactant monomer **7**, and similarly for the compound **6** (R = methyl), which lacks the long aliphatic



6 **7**

Scheme 2. Structure of compounds **6** and **7**. R is defined under Scheme 1.

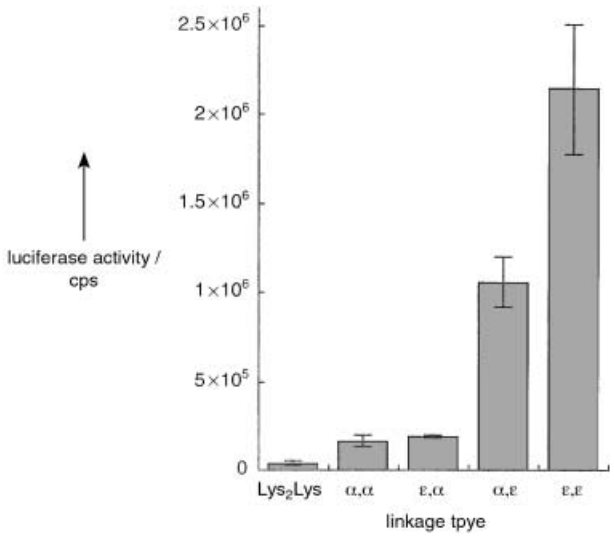


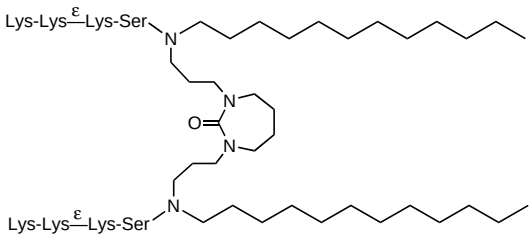
Figure 4. Effect of varying trilysine linkage; data are shown for compounds (e.g. **6**) with an oleyl (C18) side chain.

chain that makes it amphiphilic. All three compounds have the hydrophobic chain linked through a serine residue. This effective extension of the spacer was found empirically to improve transfection activity. Similarly, in most, but not all, series of geminis the oleyl chain R gives the highest activity.[32]

The widest scope for structural variation lies in the polar head group (Scheme 1). Of special interest are oligopeptide sequences corresponding to nuclear localization signals (NLS), which offer a key to allow much larger molecules to be transported through nuclear membranes. We find that typical NLS sequences offer no advantage over the best, shorter peptide sequences in our systems **2** based on spermine and conclude that—at least in the systems tested—the lipoplexes involved in the initial stages of the delivery process (Figure 2) have dissociated before the (presumably unassisted) passage of the DNA through the nuclear membrane.

We found another clear-cut correlation in the cysteine-based series **1a** when the hydrophobic chain R is the preferred oleyl group and the peptide head group is varied. As a general rule lysine is the preferred basic α-amino acid when compared with histidine and arginine; we found no advantage in using units longer than a tetrapeptide. Three lysine units can be attached to an amino group in five different ways, and we found a strong dependence on the connectivity, and thus the spacing, of the positive charges along the trilysine chain (Figure 4).

It should be emphasized that such well-defined correlations are the exception rather than the rule, even when closely similar structures are compared. Selected examples of all the main classes of gemini surfactant structures shown in Scheme 1 support highly efficient transfection, but a wide range of activities is found in all series (Figure 3). Apparently minor changes in structure can have major effects on biological activity. Thus compound **8**, based on cyclospermine with the head group Lys-Lys-ε-Lys-Ser (that is, with one ε linkage between the inner two lysine units), is one of the most active gemini transfectants, whereas the corresponding structure with *both* lysine units ε-linked (the optimal combination in systems **1a**, as indicated in Figure 4) is one of the



8

least active. Similarly, compound **5** (RCO = oleyl, $n = 6$) based on a hexamethylenediamine spacer is highly active in transfection, whereas the same structure with a saturated 18-membered chain is inactive. The generally (but not invariably) superior properties of the oleyl tails confer good (class IV) activity on the otherwise inactive system with $n = 4$. Notably, the oleyl chains in the natural, commercially available amine and alcohol are a mixture of geometrical isomers, *cis*/*trans* ≈ 80:20. We made the compounds **1a** (R = oleyl, peptide = ε-linked trilysyl), based on one of the most effective transfection agents, with exclusively *cis* and *trans* double bonds, and found no significant difference in activity compared with various mixtures of isomers.

The unsurprising but unhelpful conclusion is that the complex series of processes involved in transfection involve multiple interactions, which depend on the unique, combined effects of the head group, spacer, and hydrophobic tail of a particular vector. Clear-cut structure–activity relationships are thus not generally to be expected, except perhaps in series involving minor changes in a single structural parameter. Of the various classes of gemini surfactant summarized in Scheme 1, structures **3** with the carbohydrate-based head groups showed the highest activity: the majority of these compounds showed very high activity and all showed at least some activity (Figure 5). However, high activity was observed in at least some cases in almost all the series tested.

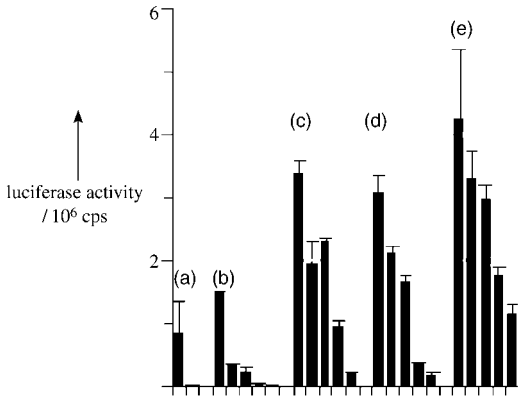


Figure 5. Transfection efficiencies (for luciferase expression in CHO cells) of a series of cationic gemini surfactant vectors, compared with lipofectamine 2000 (a). The surfactants tested were glucosyl derivatives **3** with $n = 6$ (b) and with a -$(CH_2)_2O(CH_2)_2O(CH_2)_2$- spacer (c). The results for the corresponding two mannosyl derivatives are shown in (d) and (e). (Gemini surfactant concentrations: 4, 8, 10, 20, and 30 μM).

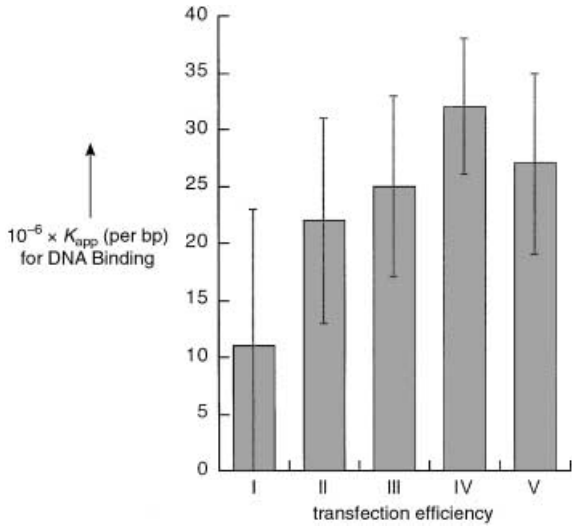
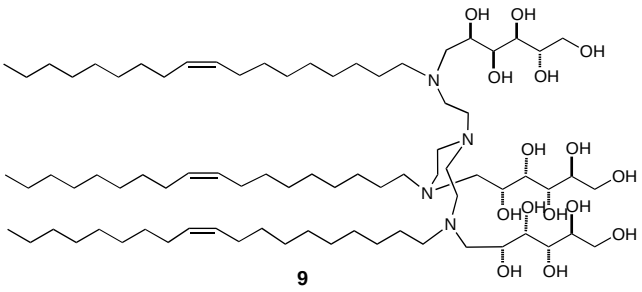


Figure 6. Relationship between transfection efficiency and the strength of DNA binding for 46 gemini surfactants with the general structure **1**.[34]

Since gemini surfactants are generally much more efficient transfection agents than the corresponding monomeric structures, a key basic question is whether the effectiveness of the gemini structure can be improved further by increasing the multiplicity of the structural elements. We have no comprehensive answer to this question, but tests on two compounds based on the most reliable carbohydrate-based structure **3** proved decisively negative. Thus both compound **9** and the corresponding system with saturated C_{16} chains showed no significant transfection activity.



9

4. Structure of the Lipoplex

Of the complex series of noncovalent interactions involved in transfection only the first, the formation of the lipoplex, can be readily studied in isolation. We can measure the strength of the binding of DNA and examine the structures of the aggregates formed by the surfactant in the presence and in the absence of DNA. Subsequently, we attempt to relate this information to the transfection process and thus to apply it to the design of improved vectors.

4.1. DNA Binding and Transfection Activity

We know that the efficiency of transfection is not likely to show a simple dependence on the strength of DNA binding by the amphiphile. The initial binding process must be strong enough for the lipoplex to form rapidly and to survive the subsequent passage through various solutions and membranes. Yet it must also eventually dissociate efficiently, presumably at or after the point of release from the endosome, thus the binding must not be too strong. A series of measurements on 46 compounds with general structure **1** (Scheme 1) gave results consistent with this picture (Figure 6). The binding constants, in the order of 10^7 per base pair, are clearly lower for less efficient vectors, but this trend of efficiency increasing with the binding constant is not maintained for the most efficient transfectants, consistent with the idea that there is an optimal level of binding.[34]

4.2. Aggregation States

To complement the synthetic and transfection studies, the aggregation behavior of selected surfactants was investigated by physical and calculation methods. Although gemini surfactants are active in transfection at well below typical aggregation concentrations, their aggregation behavior must be at least indirectly relevant to lipoplex formation. In practice aggregation behavior observed by transmission electron microscopy (TEM) generally shows no clear correlation with transfection efficiency. An interesting exception is a series of compounds with the general structure **1c** (Scheme 1), with saturated tails (R = dodecyl) linked to the cysteine-based spacer through a serine residue, and short head groups made up of 1–3 basic α-amino acids. Five compounds with this structure all bound DNA well, but the two with single α-amino acids as head groups were poor transfectants. These two formed characteristic fibrillar aggre-

gates when absorbed from water onto carbon/Formvar-coated grids, whereas the three highly active compounds with di- or tripeptide head groups did not.[35] More typically, the series of tartaric acid based gemini transfectants **4**, compounds of moderate transfection efficiency studied by TEM, formed variously short ribbons, ribbons with a tendency to twist, or no aggregates at all.[27]

For the same series of compounds **4**, a circular dichroism titration with λ-phage DNA showed that the gemini surfactant:DNA stoichiometry in the lipoplex is dictated by charge complementarity.[31] A more detailed CD study suggests that the immediate hydration shell in the region of the backbone phosphate groups is not significantly changed on complexation.[27] Conductivity studies showed that some of these compounds have very low critical aggregation concentrations ($\leq 0.24\ mol\,m^{-3}$).[36] The compaction of DNA by cationic gemini surfactants (or any other cationic surfactant) presumably involves an initial interaction of a small cluster of surfactant molecules with DNA (especially for compounds with singly charged head groups). Further interactions, strengthened by the combined (chelate) effect of electrostatic and hydrophobic attraction, enhance compaction even more. This process will depend on some of the same factors as self-association and should thus be favored by a low critical aggregation concentration.

The interaction of a range of cationic gemini surfactants with bacteriophage T4 DNA was studied by means of fluorescence microscopy.[36] Upon addition of surfactant, the DNA undergoes a transition from a random coil to a globule with an intermediate coexistence region. The state behavior of a DNA–gemini surfactant system depends on spacer length, valency, head group size, and tail length. A series of simple alkanediyl-α,ω-bis(dimethylalkylammonium bromide) gemini surfactants ($Me_2N(CH_2)_{12}NH(CH_2)_s$-$(CH_2)_{12}NMe_2$, [12-*s*-12]), with fixed tail length and variable spacer lengths, showed a minimum in compaction efficiency at $s = 6$, as a result of the competition between entropy loss and enthalpy gain. This occurs at roughly the same spacer length as a maximum (at $s = 5$) in the critical aggregation concentration. In comparison with a single-tailed divalent surfactant [12-3-1] it was shown that the gemini equivalent [12-3-12] was more efficient in compacting DNA. A series of gemini surfactants **5** based on cationic peptides with α,ω-diamino alkyl spacers showed similar behavior with changing spacer length. Additionally, two surfactants **4**, based on diastereomers of tartaric acid, with hexadecanoic acid tails and α,ω-diaminopropyl and spermidine head groups, respectively, showed effects of head group size that depended strongly on entropy effects.[36]

The most detailed studies concentrated on gemini surfactant **3** (R = oleyl, $n = 6$) because of its relatively simple molecular structure and because it is one of the most active transfection agents discovered in this study. This compound undergoes a proton-induced vesicle-to-micelle transition as the pH value is decreased to below 7. This behavior was investigated in more detail by using self-consistent-field theory and was explained in terms of significant protonation of the two amine centers in combination with the flexible sugar head groups.[37] A high-level molecular-dynamics simulation for the neutral surfactant shows a 46-Å thick bilayer with significant interdigitation of the alkyl chains (Figure 7). The dimensions of the bilayer, and the degree of interdigitation, closely match those obtained from small-angle X-ray-scattering (SAXS) experiments (Section 4.3).[37]

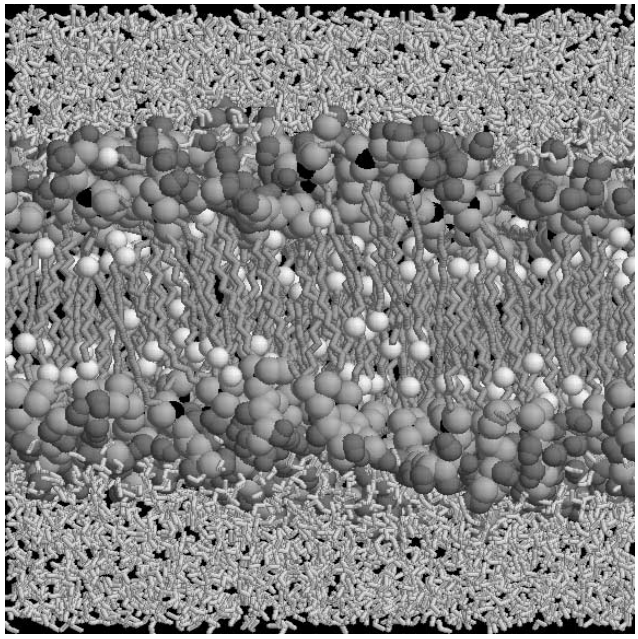


Figure 7. Molecular dynamics simulation for the neutral gemini surfactant **3** (R = oleyl, $n = 6$) in water.[36]

The morphology of the aggregates formed by compounds with the general structure **5** and of their complexes with λ-phage DNA, were also investigated by means of transmission and cryo-scanning electron microscopy (TEM and cryo-SEM).

4.3. SAXS Studies of Selected Gemini Surfactants and Lipoplexes

Selected gemini surfactants were studied as concentrated dispersions and as lipoplexes formed with salmon-sperm DNA[*].[38] All dispersions studied showed at least some long-range order, which changed on addition of DNA, although the diffraction peaks were often broad. Higher-order diffraction peaks were observed in some cases. Assignments of packing were made by using these higher-order diffraction peaks and/or correlation with the dimensions of the extended molecular structures, obtained from molecular modeling by Quanta/Charm or estimated from Corey–Pauling–Koltun (CPK) models. Selected results are shown in Table 1.

The results for most tartaric acid and lysine-based surfactants are consistent with lamellar packing, which is expanded upon binding of DNA. When the most active

[*] These experiments were carried out at the Small Angle X-ray Scattering station of the Dutch-Belgian beamline (DUBBLE) at the ESRF in Grenoble, France, with financial support from the NWO.

Table 1: Molecular Dimensions and Packings Derived from SAXS experiments.

Compound (RCO, headgroup)	Activity	Estimated length [Å][a]	Obs. spacing/ Free surfactant	Obs. spacing/ Lipoplex
4 ($C_{15}H_{31}$, 3-aminopropyl)	I	*24*: 29 Å	lam., $d = 53$ Å	lam., $d = 64$–65 Å
4 ($C_{15}H_{31}$, lysine amide)	III	*26*: 31 Å	lam., $d = 57$ Å	?, $d = 61$ Å
5 (dodecyl, $n = 2$)	I	*19*: 23 Å	lam., $d = 36$ Å	lam., $d = 45$ Å
5 (oleyl, $n = 4$)	IV	*25*: 30 Å	lam., $d = 46$ Å	lam., $d = 52$–55 Å
5 (oleyl, $n = 6$)	V	*25*: 30 Å	lam., $d = 45$ Å	lam., $d = 52$ Å
5 (dodecyl, $n = 8$)	III	*25*: 30 Å	lam., $d = 44$ Å	lam., $d = 51$–54 Å
3 (R = oleyl, $n = 6$)	IV/V	*26*: 31 Å	lam., $d = 49$ Å	ext. lam., $d = 59$ Å
pH 7.0				cd. lam., $d = 57$ Å
pH 6.5				cd. lam., $d = 55$ Å
pH 5.5				hex., $a = 58$ Å

[a] *No. of atoms* and estimated maximum length of the molecule in its extended conformation. *lam. = lamellar; ext. = extended; cd. = condensed; hex. = hexagonal; d = lamellar spacing; a = hexagonal or columnar spacing.*

compound of this structural type **5** (RCO = oleyl, $n = 6$) was studied in more detail, it was found to exhibit temperature-dependent transitions between expanded and condensed lamellar phases.

The structure of the lipoplex formed from DNA and the sugar-based cationic gemini surfactant **3** (R = oleyl, $n = 6$), which exhibits excellent transfection efficiency, was investigated over the pH range 8.8–3.0 by both SAXS and cryo-TEM. Uniquely, three well-defined morphologies of the lipoplex were observed as the pH was reduced (Figure 8): a lamellar phase (between pH 8.80–7.97), with dimensions consistent with the molecular dynamics simulation results described above,[37] a condensed lamellar phase (from pH 7.49–6.00), and an inverted hexagonal (H_{II}) columnar phase (from pH 5.75–3.81). Molecular modeling suggested a correlation between the observed lipoplex morphologies and physical behavior, and specific structural features in the surfactant.

These results suggest that key factors for future surfactant design are: a spacer of six methylene groups, the presence of two nitrogen atoms that can be protonated in the physiological pH range, two unsaturated alkyl tails, and hydrophilic sugar head groups. Assuming that the mechanism of transfection by synthetic cationic surfactants involves endocytosis, we suggest that the efficacy of gemini surfactant **3** (R = oleyl, $n = 6$) as a gene delivery vehicle can be explained by this unprecedented observation of a pH-induced formation of the inverted hexagonal phase of the lipoplex in the endosomal pH range. This change in morphology could lead to destabilization of the endosome through fusion of the lipoplex with the endosomal wall, resulting in release of DNA into the cytoplasm.

We suppose that the acid-induced change in morphology of the lipoplex formed from **3** (R = oleyl, $n = 6$) and salmon-sperm DNA from condensed lamellar to hexagonal is driven by the close association of mostly doubly protonated surfactant molecules with the phosphate groups of the DNA. The increase in pH observed to accompany this event can be explained by the exposure and subsequent protonation of unprotonated amine groups that were initially internalized in the condensed bilayers. In earlier studies on these gemini surfactants[31,39] a proton-induced vesicle-to-micelle transition was observed and explained in terms of significant protonation of the second amine center. This would result in a greater extent of counterion association and increased hydration, leading to an increase in the head group size. Thus the formation of micelles is favored over that of bilayers, in line with the shape-structure concept.[40]

The change from a lamellar to an inverted hexagonal phase observed here for the lipoplex formed from **3** (R = oleyl, $n = 6$) requires a decrease in head-group size, and can be rationalized by a strong association between the doubly charged head group of **3** and phosphate moieties of the DNA, leading to 1) local charge neutralization and 2) dehydration of both phosphate and head groups, which results in an effective reduction in head group size. The occurrence of the morphological change at pH 5.45 is consistent with a vesicular pK_a value for **3** (R = oleyl, $n = 6$) at around this pH value. Thus we suggest that DNA is a template for the H_{II} columnar phase, a result of "specific" association of pairs of phosphate groups with the doubly charged gemini surfactant, as opposed to the "atmospheric" DNA association with the singly charged species observed in the lamellar phases. The lipoplex of **3** (R = $C_{18}H_{37}$, $n = 6$) the saturated analogue of **3** (R = oleyl, $n = 6$) does not undergo pH-induced morphological changes: it shows a lower but still good activity in transfection (class IV).

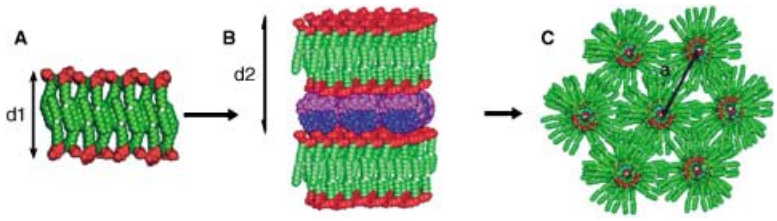


Figure 8. Space-filling molecular models of aggregates of **3** and DNA based on morphology and geometry obtained by SAXS/EM studies: a) free vesicles, d1 = 48.7 Å; b) initial lamellar phase of lipoplex, d2 = 59.8 Å; c) columnar H_{II} structure of lipoplex, $a = 57.8$ Å. Green: alkyl tails; red: sugar head groups; blue and lilac: complementary DNA strands.

4.4. Escape from the Endosome

It has been suggested that the escape of DNA from the endosome mediated by synthetic surfactants depends, in some cases at least, on the formation of a fusogenic inverted hexagonal phase.[41] This phase can be induced by the shape of the cationic surfactant molecule[42] or of a "helper" lipid,[43] or by the interaction of the lipoplex with anionic lipids.[44] The pH-induced formation of an inverted hexagonal (H_{II}) phase for the lipoplex of **3** (R = oleyl, $n = 6$) and DNA in the endosomal pH range, as evidenced by the SAXS and cryo-SEM results described, may facilitate its fusion with the endosomal membrane, an important step toward release of the DNA into the cytoplasm.

This observation of the pH-induced and DNA-templated formation of the hexagonal phase of the lipoplex led us to propose the detailed mechanism for the transfection by **3** (R = oleyl, $n = 6$) outlined in Figure 9. The initial passage through the cell membrane is presumed to be by endocytosis (Figure 2).[44a]

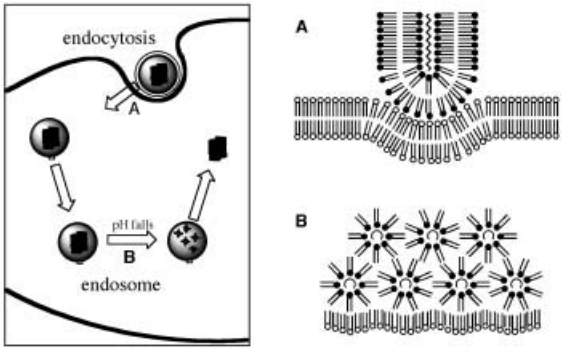


Figure 9. Proposed mechanism for transfection mediated by **3** (R = oleyl, $n = 6$) adapted from Xu and Szoka.[44a] White and black head groups represent zwitterionic/anionic membrane components and cationic gemini surfactants, respectively. **A**: onset of endocytosis of the lipoplex at the cell membrane; **B**: formation of a DNA-templated fusogenic hexagonal phase at low pH value, allowing escape from the endosome. (The Box is a section from Figure 2.)

Although it is highly significant for the mechanism of transfection that such a fundamental change in the morphology of the lipoplex can occur simply on encounter with anionic membranes, the question arises as to why it does not occur when the lipoplex interacts with either the cell or, initially, the endosomal membrane. We propose that the factor that triggers the escape of the DNA from the endosome is the gradual decrease in pH that this organelle undergoes after formation.[45] Our finding that the inverted hexagonal phase can be formed from unadulterated cationic lipids, simply by lowering the pH, is significant for the mechanism of transfection since it provides a reasonable explanation for why escape occurs from the endosome under these circumstances.

5. Summary and Outlook

1. The gemini surfactant structure, as realized in a wide range of systems based on cationic head groups, gives rise to a high proportion of compounds that show excellent activity in the transfection of a wide variety of cell types. Recent (unpublished) work has extended the scope of the process to transfection with oligoribonuceotides and to a number of cell lines that are normally difficult to transfect. These gemini surfactants generally show very low toxicity, and transfection efficiencies can be as high as 90%.
2. These compounds bind strongly to DNA, and there is evidence to support the logical conclusion that the strength of binding goes through an optimum.
3. Binding to DNA can be assumed to involve an important electrostatic contribution. The spacing between the two NH^+ centers in the most efficient transfectants (**3** (R = oleyl, $n = 6$) and **1a** (R = oleyl, peptide = ε-linked trilysyl)) of two of the most interesting series of gemini surfactants, is closely similar, at about 10 Å. This spacing of the NH^+ groups complements the spacing of the anionic phosphodiester groups on either side of the minor groove of DNA.
4. The α NH^+ groups of **1a** (R = oleyl, peptide = ε-linked trilysyl) and **3** (R = oleyl, $n = 6$) have low pK_a values (about 7.5 versus 5.8 and 8.3 in free solution), much lower than those of ε NH^+ groups, which will be fully protonated at all pH values of interest.
5. It is proposed that the key to the exceptional efficacy of **3** (R = oleyl, $n = 6$) in gene transfection is its ability to form a lamellar lipoplex, which changes to the fusogenic inverted hexagonal phase at a critical (endosomal) pH value. a) The spacing of the ammonium centers allows the DNA to template the morphology of the complex into the fusogenic inverted H_{II} columnar phase. b) The second of the two amine nitrogen atoms in the head group has a vesicular pK_a value in the endosomal pH region; its protonation causes a morphological change at a critical pH value. c) Its unsaturated alkyl chains reduce T_c (the gel to liquid crystal main phase transition temperature) to below physiological temperatures, thus increasing the susceptibility of the aggregate to morphological change. d) The hydrophilic sugar head groups increase aqueous solubility but do not obstruct localized ammonium phosphate interactions.
6. Finally, gemini surfactants can confidently be expected to show other sorts of biological activity. For example, many cationic surfactants are active bactericidal agents, and a group of gemini surfactants **2a** (Scheme 2) based on spermine with cholic acid as the hydrophobic tail turn out to be very effective against a broad range of bacteria.[46]

The insight arising from this work offers potential guidelines for surfactant and polymer design, and underlines the need to concentrate attention not just on DNA release from a lipoplex (for example, by using surfactants with chemically labile moieties), but also on the possibility of engineering morphological transformations to take place under critical cellular conditions such as pH change. Of course these are

complex systems, but eventually rules will emerge—most likely from work with simple systems in vitro. This belief is widely held, as witnessed by the hundreds of patents appearing annually that describe the applications of cationic lysosomes to gene transfection. Much of this work tells only the beginning of the story, because structure–activity relationships in vivo and in vitro are not necessarily the same.[41c]

This work is a contribution from the European Network on Gemini Surfactants, and was supported by the European Commission as part of its Training and Mobility of Researchers (TMR) Programme.

Received: September 18, 2002 [M1597]

[1] ENGEMS—the European Network on Gemini Surfactants (1997–2001, web site www.ch.cam.ac.uk/misc/ENGEMS/)—brought together groups from the UK, the Netherlands, and Sweden under the aegis of the Framework 4 Training and Mobility of Researchers (TMR) Program of the European Commission. The compounds discussed have been patented.

[2] The name gemini was coined in 1991 by Menger, who has recently reviewed the chemistry of these remarkable compounds: F. M. Menger, J. S. Keiper, *Angew. Chem.* **2000**, *112*, 1980 – 1996; *Angew. Chem. Int. Ed.* **2000**, *39*, 1906 – 1920.

[3] F. M. Menger, C. A. Littau, *J. Am. Chem. Soc.* **1993**, *115*, 10083 – 10090; R. Zana, M. Bennraou, R. Rueff, *Langmuir* **1991**, *7*, 1072 – 1075; R. Zana, *Adv. Colloid Interface Sci.* **2002**, *97*, 205 – 253; R. Zana, *J. Colloid Interface Sci.* **2002**, *248*, 203 – 220.

[4] R. Zana, Y. Talmon, *Nature* **1993**, *362*, 228 – 230; S. Karaborni, K. Esselink, P. A. J. Hilbers, B. Smit, J. Karthauser, N. M. van Os, R. Zana, *Science* **1994**, *266*, 254 – 256.

[5] L. Felgner, T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, M. Danielsen, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7413 – 7417; D. D. Lasic, N. S. Templeton, *Adv. Drug Delivery Rev.* **1996**, *20*, 221 – 266.

[6] J.-P. Behr, *Bioconjugate Chem.* **1994**, *5*, 382 – 389.

[7] J.-Y. Legendre, F. C. J. Szoka, *Pharm. Res.* **1992**, *9*, 1235 – 1242.

[8] A. D. Miller, *Angew. Chem.* **1998**, *110*, 1862 – 1880; *Angew. Chem. Int. Ed.* **1998**, *37*, 1768 – 1785.

[9] F. Labat-Moleur, A.-M. Steffan, C. Brisson, H. Perron, O. Feugeas, P. Furstenberger, F. Oberling, E. Brambilla, J.-P. Behr, *Gene Ther.* **1996**, *3*, 1010 – 1017.

[10] A. D. Miller, *Nature* **1992**, *357*, 455 – 460.

[11] E. Marshall, *Science* **1995**, *270*, 1751.

[12] B. Abdallah, L. Sachs, B. A. Demeneix, *Biol. Cell* **1995**, *85*, 1 – 7.

[13] S. Lehrman, *Nature* **1999**, *401*, 517 – 518.

[14] H. Pearson, *Nature* **2000**, *403*, 9.

[15] J. Zabner, A. J. Fasbender, T. Moninger, K. A. Poellinger, M. J. Welsh, *J. Biol. Chem.* **1995**, *270*, 18997 – 19007.

[16] F. L. Graham, A. J. Van der Eb, *Virology* **1973**, *54*, 536 – 539.

[17] K. Ohtani, M. Nakamura, S. Saito, K. Nagata, K. Sugamura, Y. Hinuma, *Nucleic Acids Res.* **1989**, *17*, 1589.

[18] J. P. Lemaitre, B. Bayard, B. Lebleu, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 648 – 652.

[19] P. L. Felgner, *Adv. Drug Delivery Rev.* **1990**, *5*, 163 – 187.

[20] O. Boussif, F. Lezoualc'h, M. A. Zanta, M. D. Mergny, D. Scherman, B. Demeneix, J. P. Behr, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7297 – 7301.

[21] E. Serres, P. Vicendo, E. Perez, T. Noel, I. Rico-Lattes, *Langmuir* **1999**, *15*, 6956 – 6960.

[22] J. Haensler, F. C. Szoka, *Bioconjugate Chem.* **1993**, *4*, 372 – 379.

[23] J. F. Kukowska-Latallo, A. U. Bielinska, C. Chen, M. Rymaszewski, D. A. Tomalia, J. R. Baker, Jr, in *Self-Assembling Complexes for Gene Delivery: from Laboratory to Clinical Trial* (Eds.: A. V. Kabanov, P. L. Felgner, L. W. Seymour), Wiley, New York, **1998**, pp. 241 – 253.

[24] Y. Barenholz, *Curr. Opin. Colloid Interface Sci.* **2001**, *6*, 66 – 77.

[25] D. Matulis, I. Rouzina, V. Bloomfield, *J. Am. Chem. Soc.* **2002**, *124*, 7331 – 7342.

[26] a) K. Jennings, I. Marshall, H. Birrell, A. Edwards, N. Haskins, O. Södermann, A. J. Kirby, P. Camilleri, *Chem. Commun.* **1998**, 1951 – 1952; b) G. Ronsin, C. Perrin, P. Guedat, A. Kremer, P. Camilleri, A. J. Kirby, *J. Chem. Soc. Chem. Commun.* **2001**, 2234 – 2235.

[27] P. J. J. A. Buynsters, C. L. García Rodríguez, E. L. Willighagen, N. A. J. M. Sommerdijk, A. Kremer, P. Camilleri, M. C. Feiters, R. J. M. Nolte, B. Zwanenburg, *Eur. J. Org. Chem.* **2002**, 1397 – 1406.

[28] R. Breslow, *Recl. Trav. Chim. Pays-Bas* **1994**, *113*, 493 – 498; M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem.* **1998**, *110*, 2908 – 2953; *Angew. Chem. Int. Ed.* **1998**, *37*, 2754 – 2794.

[29] K. C. Nicolaou, R. Hughes, S. Y. Cho, N. Winssinger, H. Labischinski, R. Endermann, *Chem. Eur. J.* **2001**, *7*, 3824 – 3843.

[30] Gemini surfactants structurally related to **1**, derived from the oxidative dimerization of cysteine-based monomers, have been reported by the Strasbourg group to show interesting transfection capabilities; see: E. Dauty, J.-S. Remy, T. Blessing, J.-P. Behr, *J. Am. Chem. Soc.* **2001**, *123*, 9227 – 9234.

[31] M. L. Fielden, C. Perrin, A. Kremer, M. Bergsma, M. C. Stuart, P. Camilleri, J. B. F. N. Engberts, *Eur. J. Biochem.* **2001**, *268*, 1269 – 1279.

[32] C. McGregor, C. Perrin, M. Monck, P. Camilleri, A. J. Kirby, *J. Am. Chem. Soc.* **2001**, *123*, 6215 – 6220.

[33] M. A. Zanta, P. BelguiseValladier, J. P. Behr, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 91 – 96.

[34] C. McGregor, Ph.D. Thesis, University of Cambridge, **2001**.

[35] K. H. Jennings, I. C. B. Marshall, M. J. Wilkinson, A. Kremer, A. J. Kirby, P. Camilleri, *Langmuir* **2002**, *18*, 2426 – 2429.

[36] L. Karlsson, M. C. P. van Eijk, O. Söderman, *J. Colloid Interface Sci.* **2002**, *252*, 290 – 296.

[37] M. Johnsson, A. Wagenaar, J. B. F. N. Engberts, *J. Am. Chem. Soc.* **2003**, *125*, 757 – 760; M. C. P. van Eijk, M. Bergsma, S.-J. Marrink, *Eur. Phys. J. E* **2002**, *7*, 317 – 324.

[38] P. C. Bell, M. Bergsma, I. P. Dolbnya, W. Bras, M. C. Stuart, A. E. Rowan, M. C. Feiters, J. B. F. N. Engberts, *J. Am. Chem. Soc.* **2003**, *125*, 1551 – 1558.

[39] M. Bergsma, M. L. Fielden, J. B. F. N. Engberts, *J. Colloid Interface Sci.* **2001**, *243*, 491 – 495.

[40] J. N. Israelachvili, S. Marcelja, R. G. Horn, *Q. Rev. Biophys.* **1980**, *13*, 121 – 200.

[41] a) H. H. Ellens, D. P. Siegel, D. Alford, P. L. Yeagle, L. Boni, L. J. Lis, P. J. Quinn, J. Bentz, *Biochemistry* **1989**, *28*, 3692 – 3703; b) D. P. Siegel, R. M. Epand, *Biophys. J.* **1997**, *73*, 3089 – 3111; c) S. Audouy, D. Hoekstra, *Mol. Membr. Biol.* **2001**, *18*, 129 – 143; d) V. Oberle, I. Zuhorn, S. Audouy, U. Bakowsky, J. Smisterová, J. B. F. N. Engberts, D. Hoekstra in *Targeting of Drugs* (Eds.: G. Gregoriades, B. McCormack), IOS Press, **2000**, pp. 146 – 155.

[42] J. Smisterová, A. Wagenaar, M. C. A. Stuart, E. Polushkin, G. Ten Brinke, R. Hulst, J. B. F. N. Engberts, D. Hoekstra, *J. Biol. Chem.* **2001**, *276*, 47615 – 47622.

[43] I. Koltover, T. Salditt, J. O. Rädler, C. R. Safinya, *Science* **1998**, *281*, 78 – 81.

[44] a) Y. Xu, F. C. Szoka, Jr., *Biochemistry* **1996**, *35*, 5616 – 5623; b) I. M. Hafez, P. R. Cullis, *Adv. Drug Delivery Rev.* **2001**, *47*, 139 – 148; c) I. M. Hafez, N. Maurer, P. R. Cullis, *Gene Ther.* **2001**, *8*, 1188 – 1196.

[45] B. Tycko, F. R. Maxfield, *Cell* **1982**, *28*, 643 – 651.

[46] G. Ronsin, A. J. Kirby, S. Rittenhouse, G. Woodnutt, P. Camilleri, *J. Chem. Soc. Perkin Trans. 2* **2002**, 1302 – 1306.

