Journal of Pharmaceutical Analysis 11 (2021) 638–645

Contents lists available at ScienceDirect

Journal of Pharmaceutical Analysis

journal homepage: www.elsevier.com/locate/jpa

Original article

Acid-base and lipophilic properties of peptide nucleic acid derivatives

Pramod Thakare [a], Francesca Vasile [a], Maura Vallaro [b], Sonja Visentin [b,**], Giulia Caron [b], Emanuela Licandro [a], Silvia Cauteruccio [a,*]

[a] Department of Chemistry, University of Milan, 20133, Milan, Italy
[b] Molecular Biotechnology & Health Sciences Department, University of Turin, 10135, Turin, Italy

ARTICLE INFO

Article history:
Received 17 March 2020
Received in revised form
20 July 2020
Accepted 21 July 2020
Available online 2 August 2020

Keywords:
Peptide nucleic acids
Ionization properties
Potentiometric pH titration
Immobilized artificial membrane chromatography
Lipophilicity

ABSTRACT

The first combined experimental and theoretical study on the ionization and lipophilic properties of peptide nucleic acid (PNA) derivatives, including eleven PNA monomers and two PNA decamers, is described. The acidity constants (pK_a) of individual acidic and basic centers of PNA monomers were measured by automated potentiometric pH titrations in water/methanol solution, and these values were found to be in agreement with those obtained by MoKa software. These results indicate that single nucleobases do not change their pK_a values when included in PNA monomers and oligomers. In addition, immobilized artificial membrane chromatography was employed to evaluate the lipophilic properties of PNA monomers and oligomers, which showed the PNA derivatives had poor affinity towards membrane phospholipids, and confirmed their scarce cell penetrating ability. Overall, our study not only is of potential relevance to evaluate the pharmacokinetic properties of PNA, but also constitutes a reliable basis to properly modify PNA to obtain mimics with enhanced cell penetration properties.

1. Introduction

Ionization and lipophilicity have a pivotal role in governing not only drug-receptors interactions but also absorption, distribution, metabolism, excretion and toxicity of drugs. Ionization impacts pH-dependent solubility, lipophilicity, permeability, stability and protein binding. For instance, aqueous solubility and membrane permeability both depend on the charge state of the molecules; therefore, the pK_a value of drugs governs oral bioavailability and absorption in the different compartments of the human body [1]. The ionization properties of the nucleobases also play a crucial role in the self-assembly process of DNA and RNA, which are mainly governed by stacking and hydrogen-bonding interactions [2,3]. In particular, intermolecular hydrogen bonding between nucleobases in nucleic acids duplexes is essential in determining the tertiary structure and functions of natural DNA and RNA in cellular systems [3]. The proton accepting and donating ability of the nucleobases in forming these hydrogen bonds and the corresponding base-pairing strength were found to be strictly associated with the ionization constants (pK_a) of the complementary base-pairing partners [4]. Thus, the evaluation of pK_a values of nucleobases in nucleosides and nucleotides is an important tool to disclose their hydrophobic and hydrophilic natures as well as their biological behavior [5]. Different experimental techniques have been employed to measure the pK_a values of unmodified nucleobases and properly modified DNA and RNA monomeric building blocks, such as UV absorption spectroscopy [6,7], fluorescent methods [8], ^{1}H NMR shift experiments [9–13], potentiometric pH titration [14] or pH-dependent high-field NMR measurements [15,16]. Computational methods have also been used to estimate pK_a values of nucleobases in gas- and solution-phases [17–25]. Both experimental and theoretical studies showed that the nature of sugar backbone and the position in the DNA and RNA strand significantly affect the pK_a values of nucleobases [26,27].

On the other hand, lipophilicity is a key physicochemical property with a major role as a predictor of eventual compound success as a drug. It is commonly measured by its distribution behavior in a biphasic system, mostly 1-octanol/water [28]. Because of the limits of this traditional system to mimic the electrostatic interactions between drugs and membranes, immobilized artificial membrane (IAM) chromatography, which combines membrane simulation with rapid measurements, is nowadays

Peer review under responsibility of Xi'an Jiaotong University.
* Corresponding author.
** Corresponding author.
E-mail addresses: silvia.cauteruccio@unimi.it (S. Cauteruccio), sonja.visentin@unito.it (S. Visentin).

https://doi.org/10.1016/j.jpha.2020.07.007

considered as a valid alternative to log P/log D measurements [29–31]. IAM stationary phases are based on phosphatidylcholine, which is the major component of biological membranes, and the determination of chromatographic retention coefficients of the analytes on such stationary phases is assumed as direct measures of their phospholipophilicity, i.e., the affinity that the analytes have for phospholipids [32]. Furthermore, IAM chromatography has been applied not only to small molecules but also to peptides of pharmaceutical interest [33]. Notably, the potential of IAM indices to predict passive transport through biological barriers can be of particular relevance in near future to estimate pharmacokinetic properties of compounds with potential pharmacological applications. Overall, to the authors' best knowledge, no study on the ionization properties and lipophilicity of peptide nucleic acid monomers and oligomers has so far been reported [34].

Peptide nucleic acids (PNAs), introduced by Nielsen et al. [35] in 1991, are synthetic DNA/RNA mimics in which the neutral *N*-(2-aminoethyl)glycine (*aeg*) unit replaces the sugar phosphate backbone, and the pseudopeptide chain is covalently linked to nucleobases through a carboxymethylene spacer (Fig. 1).

Since the distance between two nucleobases in PNA is comparable to the spacing between bases in DNA or RNA, single-stranded PNAs bind to complementary DNA, RNA and PNA sequences following the Watson-Crick base-pairing rules [36]. Unlike natural nucleic acids, the PNA backbone is neutral, thus providing strong hybridization properties with their complementary DNA/RNA targets in a high sequence specific manner [36]. Indeed, they are capable of forming PNA-DNA, PNA-RNA, and PNA-PNA duplexes or more complex structures (e.g., triplexes, quadruplexes, or hairpins) that are generally of higher stability and are endowed with greater sequence discrimination than the corresponding DNA-DNA or DNA-RNA duplexes [37–39]. Moreover, the nonstandard backbone makes PNAs resistant to enzymatic degradation by nucleases and proteases [40]. Thus, PNAs are ideal candidates for therapeutic and biomedical applications, especially within gene therapy and diagnostics [41,42]. However, it is found that PNA derivatives should fully overcome limitations in the efficiency of delivery to the nuclei of the desired cells to gain a major relevance in drug discovery. Indeed, these compounds often show poor pharmacokinetics which significantly limits their potential clinical use [43,44].

Since artificial systems like PNA may have different ionization and lipophilicity characteristics from those of DNA, to fully exploit their potential as drug candidates, there is a need for characterization of PNA derivatives in terms of ionization and lipophilicity. The aim of such characterization is to determine the pK_a and lipophilicity values of monomers used to build PNA oligomers and to rationalize the poor internalization properties of these compounds in order to design more permeable PNA derivatives. For these reasons, in this paper we report a systematic study on the ionization and lipophilic properties of eleven PNA monomers **1**–**11** (Fig. 2) that could be divided into three main groups: (1) four commercially available PNA monomers **1**–**4**, in which the amino group of aminoethylglycine portion is protected with an *N*-*tert*-butyloxycarbonyl (*N*-Boc) group, and the exocyclic amino groups of adenine, cytosine and guanine are protected with an *N*-benzyloxycarbonyl (*N*-Z) protecting group; (2) four PNA monomers **5**–**8** containing the free terminal amino group on the *aeg*-backbone, obtained by Boc deprotection of **1**–**4**; (3) three fully deprotected PNA monomers **9**–**11** in which the free exocyclic amino groups of adenine, cytosine and guanine are also present. The study of PNA monomers with different level of amine groups protection should allow to attribute the correct pK_a value to each potential ionizable groups in the monomers.

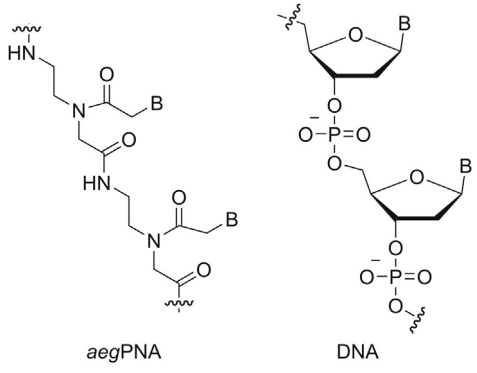


Fig. 1. Structure of peptide nucleic acid (PNA) and DNA.

Monomers **5**–**11** were synthesized for the first time and, in this paper, they were fully characterized by NMR experiments also in terms of the estimation of their equilibrium mixture of the *trans*- and *cis*-rotamers. Next, the ionization properties of monomers **1**–**11** and the model PNA decamer **12** (Fig. 3) were investigated through the measurement of their pK_a values by means of automated potentiometric pH titration using the Sirius T3 apparatus. Finally, the lipophilic properties of monomers **1**–**11** and the model PNA decamers **12** and **13** (Fig. 3) were evaluated by IAM chromatographic technique.

2. Experimental

2.1. Materials

PNA monomers **1**–**4** were purchased from ASM Research Chemicals GmbH (Hannover, Germany), and used as received. Other chemicals and solvents were obtained from commercial sources and used without further purification. The homothymine decamer **12** was synthesized according to the literature [45]. The details of the synthetic procedures of the PNA monomers **5**–**11** and the PNA decamer **13** (Scheme S1) as well as their characterization data can be found in the Supplementary data.

2.2. NMR study

NMR experiments were performed at 303 K on a Bruker Avance 600 MHz spectrometer (Milan, Italy). The NMR experiments were carried out in 600 μL of deuterated DMSO at a concentration range of 30–50 mM. In some cases, in proton spectra, the water signal at 3.3 ppm was reduced using a presaturation pulse (58 db) during relaxation delay. ^{13}C experiments were performed using attached proton test sequence (Figs. S1–S7). For quantitative ^{13}C analysis, the sequence with inverse gated decoupling (which provide proton decoupling without nuclear Overhauser effect (NOE) was applied using 100 s as relaxation delay. All proton and carbon chemical shifts were assigned unambiguously using bi-dimensional experiments (COSY, NOESY, HSQC and HMBC) and the assignments are reported in Tables S1–S7. NOESY experiments were performed with a mixing time of 400 and 700 ms; HMBC spectra were acquired applying 8 Hz as long range coupling constant $^{1}H–^{13}C$.

2.3. pK_a measurements

To measure pK_a values, a potentiometric approach was applied. All measurements were performed using the Sirius T3 apparatus (Sirius Analytical Instruments Ltd., Forest Row, UK) equipped with an Ag/AgCl double junction reference pH electrode (Forest Row, UK) and a turbidity sensing device (Forest Row, UK). All tests were performed in triplicate using standardized 0.5 M KOH and 0.5 M

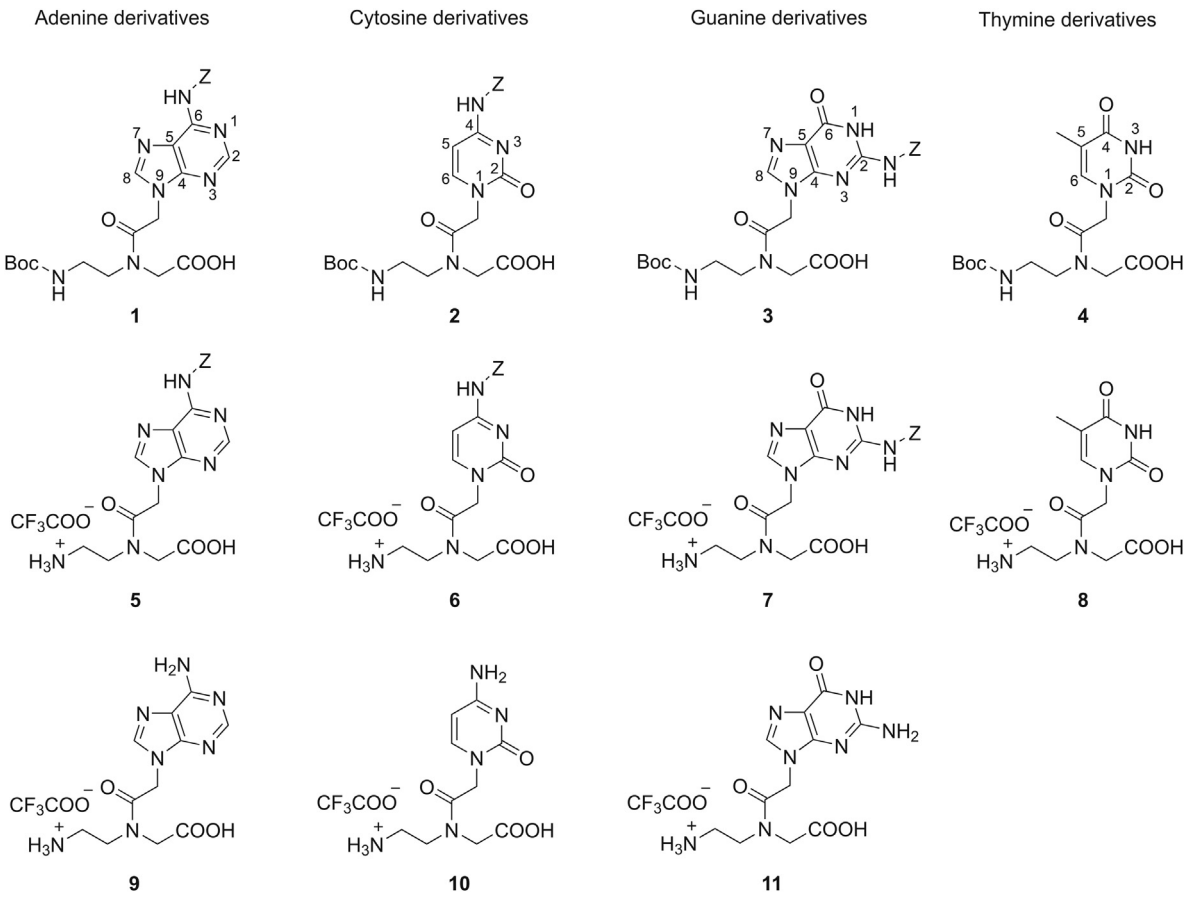


Fig. 2. Chemical structures of PNA monomers **1**–**11** (Boc: *tert*-butyloxycarbonyl; Z: benzyloxycarbonyl).

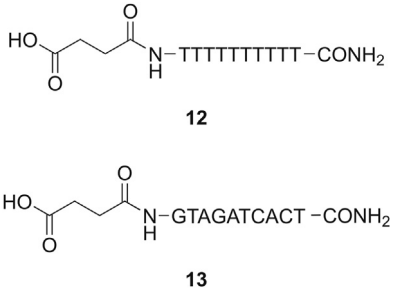


Fig. 3. Structure of PNA decamers **12** and **13**.

HCl as titration reagents. The titration experiments were conducted in 0.15 M KCl solution either in water or in co-solvent/water mixtures under nitrogen atmosphere at a temperature of (25 ± 1) °C. Aqueous pK_a values were extrapolated using the Yasuda–Shedlovsky method [46].

Shortly, in this pH-metric method, pK_a was measured by titrating a solution of the sample with acid and base, and the results were obtained by a complex computational process. The pH of each point in the titration curve was calculated using equations that contain pK_a, and the calculated points were fitted to the measured curve by manipulating the pK_a value. The pK_a that provided the best fit was taken to be the measured pK_a. This sophisticated method enables high precision in the determination of multiprotic compounds with pK_a in the 2–12 pH range [46].

2.4. Computational studies

The pK_a values were calculated with MoKA version 2.5.6 (Molecular Discovery Ltd.).

2.5. Chromatographic indexes

The applied method has already been described elsewhere [47]. The analyses were performed at 30 °C with 20 mM ammonium/acetate at pH 7.0 in mixture with acetonitrile at various percentages. The stationary phase was IAM.PC.DD.2 (Regis Technology, 10 cm × 4.6 mm, 10 μm, packing 300 Å pore size). The flow rate was 1.0 mL/min, and the injection volume was 10 μL. Wavelength of detection was 260 nm.

Chromatographic retention data at a given amount of cosolvent, expressed as log k^{IAM} (the logarithm of the retention factor), were calculated with the following expression:

$$\log k^{IAM} = \log [(t_r - t_0)/t_0]$$

where t_r and t_0 are the retention time of the drug and a non-retained compound (citric acid), respectively (data not shown). All log k^{IAM} values are the average of at least three measurements. To avoid that the experimental measurements were affected by

retention changes due to column aging, the retention times of five gold standard compounds (caffeine, carbamazepine, ketoprofen, theobromine and toluene) was checked daily.

The indexes log k_W^{IAM} were calculated by an extrapolation method. Log k^{IAM} values were determined at least three different acetonitrile percentages (φ) in the mobile phases (from 10% to 40%, *V/V*) and the intercept values of the linear relationships ($R^2 \geq 0.98$ for the monomers and $R^2 \geq 0.80$ for the oligomers) between log k and φ values were assumed as log k_W^{IAM} values. An example of extrapolation is reported in Fig. S8.

An HPLC Varian ProStar instrument (Leini, Torino, Italy) equipped with a 410 autosampler, a PDA 335 LC detector and Galaxie Chromatography Data System version 1.9.302.952 was used to perform all the measurements.

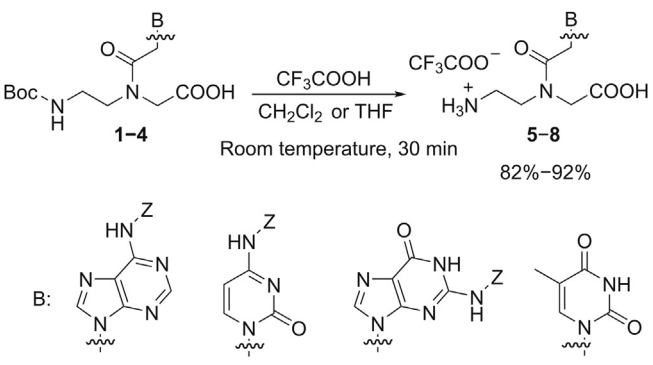
3. Results and discussion

3.1. Synthesis and NMR studies of PNA monomers 5–11

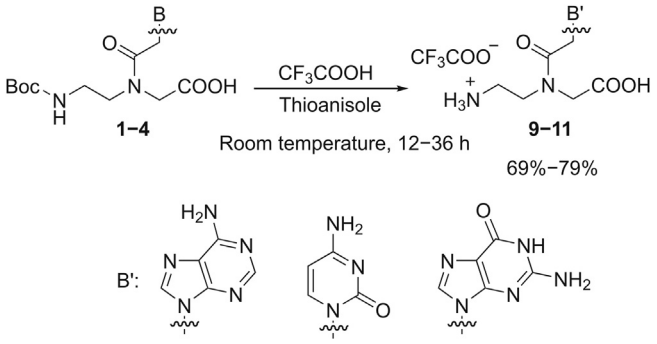
Monomers **5**–**8** were synthesized in 82%–92% yield through the selective removal of the *N*-Boc group from the corresponding commercially available PNA monomers **1**–**4**, using a mixture of trifluoroacetic acid (TFA) and dichloromethane or tetrahydrofuran at room temperature (Scheme 1).

To obtain the fully deprotected monomers **9**–**11**, the simultaneous cleavage of *N*-Boc and *N*-Z groups was performed by treating monomers **1**–**4** with a mixture of TFA and thioanisole at room temperature, according to the well-known "push-pull" mechanism [48], in which the combination of TFA and thioanisole efficiently ensures the removal of *N*-Z group under mild conditions. The fully deprotected monomers **9**–**11** were obtained in 69%–79% yield (Scheme 2).

Monomers **5**–**11** were isolated as TFA salts after crystallization with MeOH and diethyl ether, and they were fully characterized by means of NMR and MS techniques. In particular, NMR experiments were used to characterize all PNA monomers, obtain additional information about the protonation of amine and carboxylic groups, and analyze their conformational features. Indeed, the monomers exist as mixture of two amide rotamers, namely, the *E*- and *Z*-rotameric forms [49], that slowly interconvert on the NMR timescale at 303 K. Thus, for each compound, we performed a complete ^{1}H and ^{13}C NMR analyses, and we identified two sets of signals corresponding to the two rotamers that, in solution, interchange between major and minor conformers. The assignments are reported in Tables S1–S7 for both major and minor forms (indicated with M and m, respectively). In order to distinguish between the two rotamers, we made use of NOESY spectra that yield through space correlations and they can be used to estimate the molecule



Scheme 1. Boc-deprotection from PNA monomers **1**–**4**.



Scheme 2. Simultaneous Boc- and Z-deprotection from PNA monomers **1**–**4**.

conformations in solution [50,51]. Making use of bidimensional NOESY experiments, we were able to assign the NOE contacts typical of the two rotamers: the NOE contact between protons 10 and 12 (Fig. 4) was considered diagnostic for the presence of *E*-rotamer, while the interaction between protons 10 and 14 identified the *Z*-rotamer.

From the integration of their ^{1}H spectra, we can evaluate the rotamer population distribution for each monomer as reported in the Supplementary data. Fig. 4 shows NOESY experiment for monomer **8**. We assigned the resonances relative to the more intense signals (70% populated) to the *E*-rotamer while the *Z*-rotamer was 30% populated. We observed that *E*-rotamer was the major conformer for all compounds with the exception of monomer **10** that existed as 50:50 mixture of two amide isomers. NMR analysis of PNA demonstrated that the range of *E*- and *Z*-rotamers' population was between 70:30 and 50:50, and it depended on the nature of the nucleic base.

^{13}C NMR experiments were used in order to verify and quantify the presence of TFA. In a ^{13}C NMR spectrum, the area under the signal is not simply proportional to the number of carbons giving rise to the signal, because the NOE from proton decoupling is not equal for all the carbons. To exclude the contribution of NOE, we acquired monodimensional ^{13}C spectra using inverse gated decoupling sequence applying 100 s as relaxation delay in order to allow the relaxation of all ^{13}C before the acquisition of the next experiment scan. High-resolution spectra and accurate quantification using peak area integration permitted accurate measurements of relative ratios between *E*- and *Z*-rotamers and that of the ratio between PNA and TFA. We observed that for each analyzed PNA monomer (monomers **8** and **9**) the ratio of monomer–TFA was 1:1.

3.2. Ionization properties

Potentiometry, as implemented in the automatic Sirius T3 instrument, was applied to measure pK_a values of monomers **1**–**11** (Table 1). Most of the compounds under study are multiprotic molecules bearing both acidic and basic centers. In a multiprotic molecule it is crucial to identify the acidic/basic nature of different pK_a values. For instance, for monomer **11** we measured three pK_a values: 2.90, 8.67 and 9.79. In order to establish the acidic/basic nature of these pK_a values, two combined strategies can be used: predicted values and pK_a determination in the presence of a co-solvent. There are a number of commercial software and online websites that provide pK_a calculations. The methods' performances strongly depend on the chemical nature of the investigated compounds and on the availability of experimental values for similar structures. Here, we used MoKa [52,53] after a preliminary validation through the prediction of the pK_a and evaluation of the most stable tautomer of the standard free nucleobases (i.e., adenine, cytosine, guanine and thymine, see Fig. S9). Since predictions are in

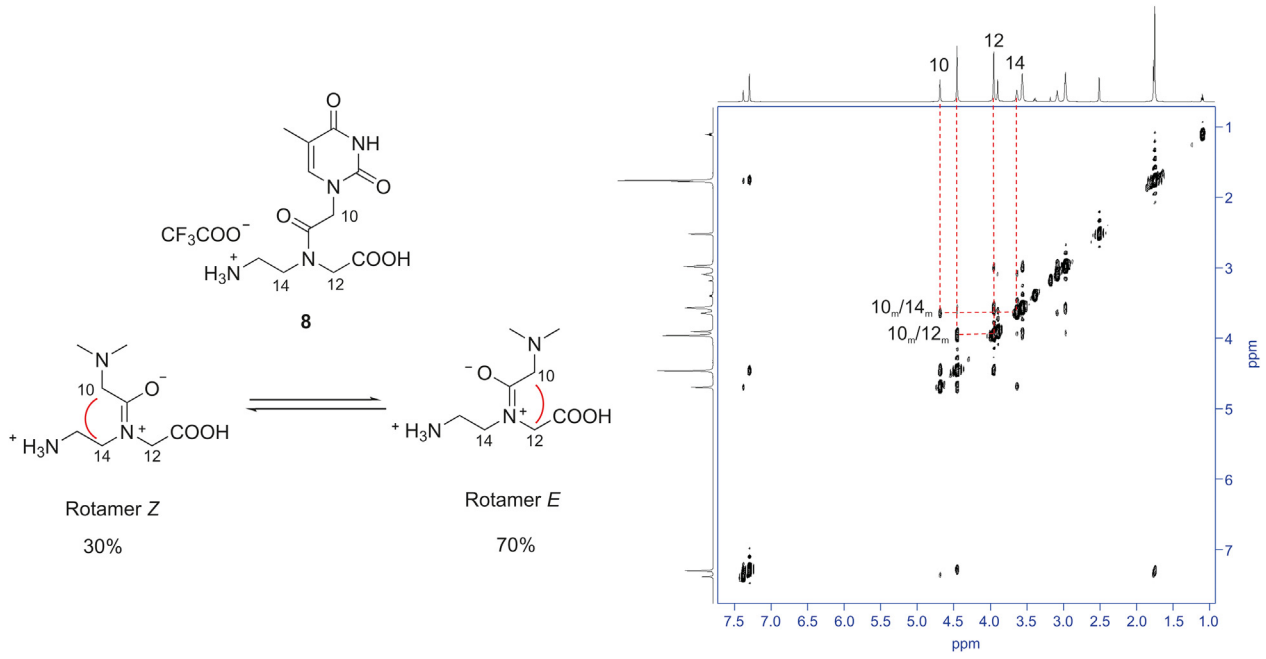


Fig. 4. 2D-NOESY experiment obtained for monomer **8** in DMSO-d6. Different nuclear Overhauser effect (NOE) cross peaks are observed, as evidenced in the spectrum: a NOE contact between protons 10–12 suggested that *E*-rotamer is the most populated, while the NOE contact between protons 10-14 indicated that *Z*-rotamer is the less populated.

excellent agreement with the experimental values [25], we decided to use MoKa also for predicting pK_a values of PNA monomers (an example of MoKa predictions is provided in Fig. S10). Thus, as far as monomer **11**, MoKa predictions indicated two acidic and one basic centers (Table 1 and Fig. 5A).

Reasonably, the COOH moiety has a pK_a of 2.90. Calculated values cannot be used to assign the second acidic pK_a. The calculation is unreliable because the values of 8.67 and 9.79 are too close. The second strategy to evaluate the acidic/basic nature of pK_a consists in measuring pK_a in water and methanol as co-solvent. As the dielectric constant of the solvent mixture decreases by increasing the amount of the co-solvent, the pK_a of an acid increases, whereas for a base the reverse is true. The variation of pK_a in the presence of the co-solvent for monomer **11** (Table 1 and

Table 1
Predicted and experimental pKa for peptide nucleic acid (PNA) monomers **1**–**11**. Standard deviation is reported.

Monomer[a]	Nucleobase	MoKa pK_a	MoKa pK_a nature	pK_a water	Slope[b] ($\times 10^{-2}$)	pK_a nature	Dominant species at pH 7.0
2	Cytosine	3.07 ± 0.34	Acidic	3.62 ± 0.04	0.72	Acidic	Monoanion
		9.17 ± 0.60	Acidic	10.27 ± 0.03	0.68	Acidic	
6	Cytosine	2.56 ± 0.34	Acidic	2.84 ± 0.01	0.51	Acidic	Zwitterion
		9.17 ± 0.60	Acidic	8.89 ± 0.02	−0.54	Basic	
		9.99 ± 0.33	Basic	10.39 ± 0.02	0.84	Acidic	
10	Cytosine	2.44 ± 0.34	Acidic	2.27 ± 0.02	1.02	Acidic	Zwitterion
		3.18 ± 0.57	Basic	4.15 ± 0.01	−0.67	Basic	
		9.99 ± 0.33	Basic	9.21 ± 0.01	−0.28	Basic	
1	Adenine	2.59 ± 0.34	Acidic	3.35 ± 0.02	1.40	Acidic	Monoanion
		9.94 ± 0.60	Acidic	10.31 ± 0.02	1.44	Acidic	
5	Adenine	2.08 ± 0.34	Acidic	2.77 ± 0.02	1.14	Acidic	Zwitterion
		9.65 ± 0.60	Acidic	9.10 ± 0.04	−0.19	Basic	
		10.25 ± 0.33	Basic	10.73 ± 0.07	0.95	Acidic	
9	Adenine	2.08 ± 0.34	Acidic	2.45 ± 0.07	0.74	Acidic	Zwitterion
		4.12 ± 0.53	Basic	3.89 ± 0.01	−0.76	Basic	
		9.97 ± 0.33	Basic	9.10 ± 0.01	−0.10	Basic	
3[c]	Guanine	2.59 ± 0.34	Acidic	3.19 ± 0.01	1.14	Acidic	Monoanion
		7.48 ± 0.53	Acidic	8.24 ± 0.08	1.17	Acidic	
7[c]	Guanine	2.27 ± 0.34	Acidic	2.69 ± 0.06	0.79	Acidic	Zwitterion
		7.48 ± 0.53	Acidic	7.95 ± 0.01	0.86	Acidic	
		10.23 ± 0.33	Basic	9.44 ± 0.01	−0.01	Basic	
11[c]	Guanine	2.08 ± 0.34	Acidic	2.90 ± 0.05	1.15	Acidic	Zwitterion
		9.42 ± 0.53	Acidic	8.67 ± 0.01	−0.20	Basic	
		10.11 ± 0.33	Basic	9.79 ± 0.01	1.13	Acidic	
4	Thymine	3.10 ± 0.34	Acidic	3.78 ± 0.03	0.96	Acidic	Monoanion
		7.90 ± 0.48	Acidic	9.62 ± 0.01	1.47	Acidic	
8	Thymine	2.59 ± 0.34	Acidic	2.54 ± 0.01	0.93	Acidic	Zwitterion
		7.90 ± 0.48	Acidic	8.96 ± 0.01	−0.08	Basic	
		9.99 ± 0.33	Basic	9.94 ± 0.01	1.04	Acidic	

[a] Range 2–9 for acids; range 3–10 for bases.
[b] The slope value was used to assess the acidic/basic nature of pK_a values (i.e., negative value for bases, positive values for acids).
[c] The most abundant tautomer calculated by MoKa was > 50.

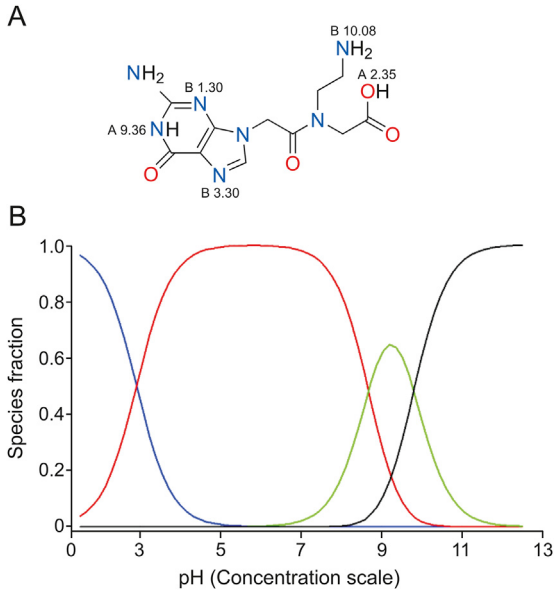


Fig. 5. Monomer **11**. (A) MoKa output; (B) the ionization profile (in red the zwitterionic species).

Fig. S11) clearly shows that the pK_a of the second acidic center is 9.79 (positive slope), whereas 8.67 is a basic center (negative slope).

Fully protected monomers **1**–**4** and Boc-deprotected monomers **5**–**8** show a lower number of ionizable centers than monomers **9**–**11** and were used to confirm the acidic/basic nature of pK_a discussed above. For instance, we compared the ionization behavior of cytosine derivatives **2**, **6**, and **10** by experimental measurements in both water and water/methanol mixtures. In particular, for the fully protected monomer **2** we observed the presence of two acidic groups corresponding to pK_a values of 3.62 and 10.27. The first proton (pK_a 3.62) of monomer **2** was mainly released from the −COOH group on the *aeg* backbone, while the second constant (pK_a 10.27) was presumably due to proton loss from the −NH(Z) group of cytosine, as also suggested by Moka simulations. Monomer **6** displayed three experimental pK_a values: a basic center (pK_a 8.89) related to the free aliphatic amino group on the backbone, and two acidic centers with pK_a values of 2.84 and 10.39 related to the −COOH group on the *aeg* backbone and the −NH present in the Z group of cytosine, respectively. These data were in reasonable agreement with the values determined for monomer **2** and with MoKa predictions. A major change of the ionization profile was found for the fully deprotected monomer **10** when compared to monomer **6**. Indeed, as the consequence of the removal of Z group from the exocyclic $-NH_2$ of cytosine, the acidic center involving the −NH in the Z group was lost, and a weak basic moiety ($pK_a = 4.15$) appeared. This basic center can be associated with the nitrogen atom N(3) of the cytosine ring. For adenine derivatives **1**, **5** and **9**, the evolution of the ionization behavior when passing from the most protected monomer **1** to the fully deprotected monomer **9** was similar to that described for the cytosine derivatives (more details in the Supplementary data).

Slightly different ionization properties were found for the guanine derivatives **3**, **7** and **11**. As expected, the fully protected monomer **3** showed two acidic centers with pK_a of 3.19 and 8.24, corresponding to the loss of proton from the −COOH group on the *aeg* backbone and the NH(1) of the guanine ring, respectively. The removal of the Boc protecting group from the *aeg* backbone from monomer **3** provided monomer **7**, in which a strong basic center (pK_a 9.44) appeared owing to the free aliphatic amine group. Fully deprotected monomer **11** showed the acidic center (pK_a 2.90), the basic center (pK_a 8.67), and a slightly acidic nitrogen (pK_a 9.79) related to the NH(1) of the guanine ring. Finally, the two thymine derivatives **4** and **8** differed by the protection of the aliphatic amino group on the backbone. As the consequence of the Boc-deprotection, monomer **8** showed a basic center (pK_a 8.96) due to the free amino on the *aeg*, besides the two acidic centers (pK_a 2.54 and 9.94 for −COOH and −NH(1), respectively) similar to those found in monomer **4**.

Overall, three main findings are obtained from ionization data of PNA monomers. Firstly, data shown in Table 1 support that partially deprotected monomers **5**–**8** and fully deprotected monomers **9**–**11** are present as zwitterions at physiological pH (the ionization profile of monomer **11** is shown in Fig. 5B, red line). The third ionization center has a weak acidic nature (pK_a 8–10) in monomers **5**–**8** and **11** and a weak basic center (pK_a 3.9–4.1) in monomers **9** and **10**; thus these sites have a poor relevance on the ionization profile. Secondly, the 2-ethyl amino group on the *aeg* backbone of monomers **5**–**11** is characterized by pK_a values ranging from 8.67 to 9.44. For the free $-NH_2$ of the alanine we found a basic pK_a value of 9.72 using the same potentiometric equipment. This means that the 2-ethyl amino group shows slightly lower basicity than that of $-NH_2$ in α-aminoacids. Finally, our data suggest that nucleobases do not significantly alter their pK_a values when included in PNA monomers, although the basicity of adenine N(1) center and cytosine N(3) center slightly decreases.

Once we obtained a complete characterization of the ionization properties of monomers **1**–**11**, we determined the pK_a values of decamer **12** that had been selected as model PNA oligomer for this study. In particular, decamer **12**, endowed with a terminal −COOH group useful for the potentiometric titration, was synthesized on solid phase according to the literature [45]. Taking into account its structure, decamer **12** is expected to show eleven acidic pK_a values, being one around 4.00 for the −COOH group and the others ≥9.0 related to NH(1) of the thymine ring. We then submitted decamer **12** to potentiometric measurements and the obtained results confirmed our expectations ($pK_{a1} = 5.08$, $pK_{a2} = 8.51$, $pK_{a3} = 9.56$, $pK_{a4} = 9.56$, $pK_{a5} = 10.22$, $pK_{a6} = 10.23$ and pK_{a7}–$pK_{a11} > 10.3$). The solubility of decamer **13** was too low for potentiometric measurements.

3.3. Lipophilicity

As discussed above, IAM chromatography offers a promising alternative to octanol/water partitioning as a tool to mimic specific interactions of compounds with membrane phospholipids. Lipophilicity data of the monomers **1**–**11** and the two model PNA decamers **12** and **13** expressed as log k_W^{IAM} are reported in Table 2.

The fully protected PNA monomers **1**, **2** and **3** show very similar positive log k_W^{IAM} (1.43–1.76), and they are the highest values obtained in the PNA monomers series. Thus, monomers **1**–**3** are the most lipophilic systems. As expected, this is due to the simultaneous presence of the hydrophobic butyloxycarbonyl protecting group on the *aeg* backbone and the benzyloxycarbonyl protecting group on the nucleobases. As the consequence of the Boc deprotection, monomers **4**, **6** and **7** have shorter IAM retention time, with log k_W^{IAM} values ranging from −0.49 to 0.97, that is about half of those obtained for monomers **1**–**3**. Negative log k_W^{IAM} values were obtained for the fully deprotected monomers **8**–**11** (from −0.68 to −0.75), which are the most hydrophilic ones. These data seem to be in agreement with the results obtained from a systematic study on the retention properties of nucleobases in IAM chromatographic systems [54]. Indeed, a set of structurally congeneric purines possess weaker IAM retention, and their log k_W^{IAM} values are close

Table 2
Lipophilicity data of PNA monomers **1**–**11** and PNA decamers **12** and **13** (for monomers **1**–**11** SD < 0.08, for decamers **12** and **13** SD < 0.20).

Compound	Nucleobase	Log k_W^{IAM}
2	Cytosine	1.43
6	Cytosine	0.72
10	Cytosine	−0.75
1	Adenine	1.43
5	Adenine	0.62
9	Adenine	−0.68
3	Guanine	1.76
7	Guanine	0.97
11	Guanine	−0.73
4	Thymine	−0.49
8	Thymine	−0.73
12	–	−1.00
13	–	−0.97

or smaller than 0, except for those with specific H-bond and/or electrostatic interactions. Since similar values were also obtained for PNA monomers, the presence of the *aeg* backbone does not significantly affect the lipophilic properties of the nucleobases.

Finally, we evaluated the lipophilicity of two model PNA decamers **12** and **13**, containing ten thymine monomers and all four nucleobases (GTAGATCACT), respectively, in order to verify the influence of the nucleobases' nature on the absorption properties of PNAs. Oligomers **12** and **13** have very similar lipophilic properties since they display negative log k_W^{IAM} values of −1.00 and −0.97, respectively. These are the shortest IAM retention time obtained in this study, which suggests that PNA decamers are slightly more hydrophilic than the single PNA monomers. Moreover, the nature of the nucleobases seems not to influence the lipophilicity of the oligomer. The poor affinity of PNA decamers towards IAM surface, which is a good model of the membrane lipid barrier, could rationalize the well-documented limited pharmacokinetic properties of PNA oligomers. Indeed, commercial drugs with good properties generally (at least when passive diffusion is known to be the dominant permeation mechanism) show positive log k_W^{IAM} values (e.g., caffeine 0.31, dexamethasone 1.71, and diltiazem 2.83) [47,55], while PNA decamers 12 and 13 display negative values. However, very high log k_W^{IAM} values should also be avoided since they are potentially connected with undesired effects such as toxicity.

4. Conclusion

In this paper we reported the synthesis, the NMR analysis of the conformational properties, the ionization and the lipophilicity features of eleven PNA monomers **1**–**11** along with the study of the ionization and lipophilicity properties of two model PNA decamers **12** and **13**. From the conformational point of view, NMR study confirms that PNA monomers exist as a mixture of *E*- and *Z*-rotamer in 70:30 or 50:50 ratio, depending on the nature of the nucleic base. Moreover, NMR experiments were also used to determine the ratio between the cationic monomer and the anion CF_3COO^- in monomers **5**–**11**, thus confirming their molecular formula.

The multiprotic nature of PNA derivatives imposed to set up a complex strategy to measure pK_a values and assign their acidic/basic nature. Therefore, we combined a sophisticated potentiometric method to measure the ionization constants with a convenient computational tool to guide the pK_a assignment. Findings supported that acid-base properties of single nucleobases are maintained in the PNA decamer **12**. The lipophilicity properties of PNA monomers and oligomers were investigated by means of IAM chromatography. Results show that the investigated PNA derivatives show poor affinity towards phospholipid-based membranes.

Overall, we provide for the first time an insight into the physicochemical properties responsible for the modulation of the pharmacokinetic behavior of PNA derivatives. Our data in fact support the well-known inefficient cell penetration of PNA that restricts its clinical application. In this scenario, our study represents an important milestone since it delivers efficient property-based drug design tools required to discover new drug-like PNA derivatives.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Acknowledgments

Pramod Thakare thanks the University of Milan for the Ph.D. fellowship. Giulia Caron, Maura Vallaro and Sonja Visentin acknowledge the financial support from the University of Turin (Ricerca Locale ex-60%, Bando2019).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpha.2020.07.007.

References

[1] D.T. Manallack, R.J. Prankerd, E. Yuriev, et al., The significance of acid/base properties in drug discovery, Chem. Soc. Rev. 42 (2013) 485–496.
[2] G.A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures, Springer-Verlag Berlin Heidelberg, Berlin, 1991.
[3] V.A. Bloomfield, D.M. Crothers, I. Tinoco Jr., Nucleic Acids: Structures, Properties and Functions, University Science Books, Sausalito, California, 2000.
[4] P. Acharya, P. Cheruku, S. Chatterjee, et al., Measurement of nucleobase pKa values in model mononucleotides shows RNA–RNA duplexes to Be more stable than DNA–DNA duplexes, J. Am. Chem. Soc. 126 (2004) 2862–2869.
[5] R. Krishnamurthy, Role of pKa of nucleobases in the origins of chemical evolution, Acc. Chem. Res. 45 (2012) 2035–2044.
[6] H. Sigel, Acid-base properties of purine residues and the effect of metal ions: quantification of rare nucleobase tautomers, Pure Appl. Chem. 76 (2004) 1869–1886.
[7] J.C. González-Olvera, J. Martínez-Reyes, E. González-Jasso, et al., Determination of pKa values for deprotonable nucleobases in short model oligonucleotides, Biophys. Chem. 206 (2015) 58–65.
[8] J.L. Wilcox, P.C. Bevilacqua, A simple fluorescence method for pKa determination in RNA and DNA reveals highly shifted pKa's, J. Am. Chem. Soc. 135 (2013) 7390–7393.
[9] A. Domínguez-Martín, S. Johannsen, A. Sigel, et al., Intrinsic acid–base properties of a hexa-2′-deoxynucleoside pentaphosphate, d(ApGpGpCpCpT): neighboring effects and isomeric equilibria, Chem. Eur J. 19 (2013) 8163–8181.
[10] A. Mucha, B. Knobloch, M. Jeżowska-Bojczuk, et al., Comparison of the acid–base properties of ribose and 2′-deoxyribose nucleotides, Chem. Eur J. 14 (2008) 6663–6671.
[11] G. Kampf, L.E. Kapinos, R. Griesser, et al., Comparison of the acid–base properties of purine derivatives in aqueous solution. Determination of intrinsic proton affinities of various basic sites, J. Chem. Soc., Perkin Trans. 2 (2002) 1320–1327.
[12] N.A. Corfù, H. Sigel, Acid-base properties of nucleosides and nucleotides as a function of concentration, Eur. J. Biochem. 199 (1991) 659–669.
[13] R. Tribolet, H. Sigel, Self-association and protonation of adenosine 5′-monophosphate in comparison with its 2′- and 3′-analogues and tubercidin 5′-monophosphate (7-deaza-AMP), Eur. J. Biochem. 163 (1987) 353–363.
[14] H. Sigel, E.M. Bianchi, N.A. Corfù, et al., Acid–base properties of the 5′-triphosphates of guanosine and inosine (GTP4– and ITP4–) and of several related nucleobase derivatives, J. Chem. Soc., Perkin Trans. 2 (2001) 507–511.
[15] P. Acharya, S. Acharya, A. Földesi, et al., Tandem electrostatic effect from the first to the third aglycon in the trimeric RNA owing to the nearest-neighbor interaction, J. Am. Chem. Soc. 125 (2003) 2094–2100.
[16] S. Acharya, P. Acharya, A. Földesi, et al., Cross-modulation of physicochemical character of aglycones in dinucleoside (3′→5′) monophosphates by the nearest neighbor interaction in the stacked state, J. Am. Chem. Soc. 124 (2002) 13722–13730.
[17] K.N. Rogstad, Y.H. Jang, L.C. Sowers, et al., First principles calculations of the pKa values and tautomers of isoguanine and xanthine, Chem. Res. Toxicol. 16

(2003) 1455–1462.
[18] J.Y. Hee, H. Sungu, C.D. Soo, DFT calculation of site-specific acid dissociation constants of purine nucleobases, Chem. Lett. 36 (2007) 1496–1497.
[19] Y.H. Jang, W.A. Goddard, K.T. Noyes, et al., pKa values of guanine in Water: density functional theory calculations combined with Poisson–Boltzmann Continuum–Solvation model, J. Phys. Chem. B 107 (2003) 344–357.
[20] A.K. Chandra, M.T. Nguyen, T. Zeegers-Huyskens, Theoretical study of the interaction between thymine and water. Protonation and deprotonation enthalpies and comparison with uracil, J. Phys. Chem. 102 (1998) 6010–6016.
[21] A.K. Chandra, D. Michalska, R. Wysokiński, et al., Theoretical study of the acidity and basicity of the cytosine tautomers and their 1:1 complexes with water, J. Phys. Chem. 108 (2004) 9593–9600.
[22] A.K. Chandra, M.T. Nguyen, T. Zeegers-Huyskens, Theoretical study of the protonation and deprotonation of cytosine. Implications for the interaction of cytosine with water, J. Mol. Struct. 519 (2000) 1–11.
[23] Y. Podolyan, L. Gorb, J. Leszczynski, Protonation of nucleic acid bases. A comprehensive post-Hartree–Fock study of the energetics and proton affinities, J. Phys. Chem. 104 (2000) 7346–7352.
[24] C.L. Tang, E. Alexov, A.M. Pyle, et al., Calculation of pKas in RNA: on the structural origins and functional roles of protonated nucleotides, J. Mol. Biol. 366 (2007) 1475–1496.
[25] V. Verdolino, R. Cammi, B.H. Munk, et al., Calculation of pKa values of nucleobases and the guanine oxidation products guanidinohydantoin and spiroiminodihydantoin using density functional theory and a polarizable continuum model, J. Phys. Chem. B 112 (2008) 16860–16873.
[26] S. Acharya, J. Barman, P. Cheruku, et al., Significant pKa perturbation of nucleobases is an intrinsic property of the sequence context in DNA and RNA, J. Am. Chem. Soc. 126 (2004) 8674–8681.
[27] S. Chatterjee, W. Pathmasiri, O. Plashkevych, et al., The chemical nature of the 2′-substituent in the pentose-sugar dictates the pseudoaromatic character of the nucleobase (pKa) in DNA/RNA, Org. Biomol. Chem. 4 (2006) 1675–1686.
[28] J.A. Arnott, S.L. Planey, The influence of lipophilicity in drug discovery and design, Expet Opin. Drug Discov. 7 (2012) 863–875.
[29] A. Taillardat-Bertschinger, P.-A. Carrupt, F. Barbato, et al., Immobilized artificial membrane HPLC in drug Research, J. Med. Chem. 46 (2003) 655–665.
[30] F. Tsopelas, T. Vallianatou, A. Tsantili-Kakoulidou, Advances in immobilized artificial membrane (IAM) chromatography for novel drug discovery, Expert Op. Drug Discovery 11 (2016) 473–488.
[31] C. Giaginis, A. Tsantili-Kakoulidou, Alternative measures of lipophilicity: from octanol–water partitioning to IAM retention, J. Pharm. Sci. 97 (2008) 2984–3004.
[32] F. Barbato, G. di Martino, L. Grumetto, et al., Prediction of drug-membrane interactions by IAM–HPLC: effects of different phospholipid stationary phases on the partition of bases, Eur. J. Pharm. Sci. 22 (2004) 261–269.
[33] D. Gussakovsky, H. Neustaeter, V. Spicer, et al., Peptide retention time prediction for immobilized artificial membrane phosphatidylcholine stationary phase: method development and preliminary observations, ADMET & DMPK 6 (2018) 190–199.
[34] O. Plashkevych, S. Chatterjee, D. Honcharenko, et al., Chemical and structural implications of 1′,2′- versus 2′,4′- conformational constraints in the sugar moiety of modified thymine nucleosides, J. Org. Chem. 72 (2007) 4716–4726.
[35] P.E. Nielsen, M. Egholm, R.H. Berg, et al., Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide, Science 254 (1991) 1497–1500.
[36] M. Egholm, O. Buchardt, L. Christensen, et al., PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules, Nature 365 (1993) 566–568.
[37] P. Wittung, P.E. Nielsen, O. Buchardt, et al., DNA-like double helix formed by peptide nucleic acid, Nature 368 (1994) 561–563.
[38] P.E. Nielsen, M. Egholm, O. Buchardt, Evidence for (PNA)2/DNA triplex structure upon binding of PNA to dsDNA by strand displacement, J. Mol. Recogn. 7 (1994) 165–170.
[39] V.V. Demidov, E. Protozanova, K.I. Izvolsky, et al., Kinetics and mechanism of the DNA double helix invasion by pseudocomplementary peptide nucleic acids, Proc. Natl. Acad. Sci. Unit. States Am. 99 (2002) 5953–5958.
[40] S. Karkare, D. Bhatnagar, Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino, Appl. Microbiol. Biotechnol. 71 (2006) 575–586.
[41] C. Sharma, S.K. Awasthi, Versatility of peptide nucleic acids (PNAs): role in chemical biology, drug discovery, and origins of life, Chem. Biol. Drug Des. 89 (2017) 16–37.
[42] A. Gupta, A. Mishra, N. Puri, Peptide nucleic acids: advanced tools for biomedical applications, J. Biotechnol. 259 (2017) 148–159.
[43] A. Dragulescu-Andrasi, S. Rapireddy, G. He, et al., Cell-permeable peptide nucleic acid designed to bind to the 5′-untranslated region of E-cadherin transcript induces potent and sequence-specific antisense effects, J. Am. Chem. Soc. 128 (2006) 16104–16112.
[44] W.B. Kauffman, S. Guha, W.C. Wimley, Synthetic molecular evolution of hybrid cell penetrating peptides, Nat. Commun. 9 (2018), 2568.
[45] G. Prencipe, S. Maiorana, P. Verderio, et al., Magnetic peptide nucleic acids for DNA targeting, Chem. Commun. (2009) 6017–6019.
[46] A. Avdeef, Physicochemical profiling (solubility, permeability and charge state), Curr. Top. Med. Chem. 1 (2001) 277–351.
[47] G. Ermondi, M. Vallaro, G. Caron, Learning how to use IAM chromatography for predicting permeability, Eur. J. Pharm. Sci. 114 (2018) 385–390.
[48] Y. Kiso, K. Ukawa, T. Akita, Efficient removal of N-benzyloxycarbonyl group by a 'push–pull' mechanism using thioanisole–trifluoroacetic acid, exemplified by a synthesis of Met-enkephalin, J. Chem. Soc., Chem. Commun. (1980) 101–102.
[49] P.C. Meltzer, A.Y. Liang, P. Matsudaira, Peptide nucleic acids: synthesis of thymine, adenine, guanine, and cytosine nucleobases, J. Org. Chem. 60 (1995) 4305–4308.
[50] F. Vasile, M. Civera, L. Belvisi, et al., Thermodynamically–weighted conformational ensemble of cyclic RGD peptidomimetics from NOE data, J. Phys. Chem. B 120 (2016) 7098–7107.
[51] F. Vasile, G. Tiana, Determination of structural ensembles of flexible molecules in solution from NMR data undergoing spin diffusion, J. Chem. Inf. Model. 59 (2019) 2973–2979.
[52] F. Milletti, L. Storchi, G. Sforna, et al., New and original pKa prediction method using grid molecular interaction fields, J. Chem. Inf. Model. 47 (2007) 2172–2181.
[53] F. Milletti, L. Storchi, G. Sforna, et al., Tautomer enumeration and stability prediction for virtual screening on large chemical databases, J. Chem. Inf. Model. 49 (2009) 68–75.
[54] H. Luo, C. Zheng, Y.-K. Cheng, The retention properties of nucleobases in alkyl C8-/C18- and IAM-chromatographic systems in relation to log Pow, J. Chromatogr. B 847 (2007) 245–261.
[55] F. Tsopelas, T. Vallianatou, A. Tsantili-Kakoulidou, The potential of immobilized artificial membrane chromatography to predict human oral absorption, Eur. J. Pharm. Sci. 81 (2016) 82–93.