Oligopeptides

DOI: 10.1002/anie.200904188

Mapping the Landscape of Potentially Primordial Informational Oligomers: Oligo-dipeptides Tagged with Orotic Acid Derivatives as Recognition Elements**

*Xuejun Zhang and Ramanarayanan Krishnamurthy**

Dedicated to Professor Albert Eschenmoser on the occasion of his 84th birthday

In the context of experimentally mapping the landscape of potential primordial informational oligomers,[1] Eschenmoser, Krishnamurthy, and co-workers have jointly investigated the base-pairing properties of oligomers derived from dipeptides tagged with non-canonical recognition elements, 2,4-disubstituted triazines and 5-aminopyrimidines.[2] The criteria for choosing these non-natural candidates were[1,2] a) their perceived 'generational simplicity' under prebiotic conditions, and b) more importantly, their potential to become attached to backbones (overcoming the 'nucleosidation problem'[3]). Widespread base pairing was observed, however, with an unanticipated result: whereas base-pairing was strong in the diaminotriazine-tagged aspartic-glutamic (Asp-Glu) oligodipeptide backbone series, the corresponding dioxotriazine-tagged series exhibited (very) weak base pairing. Furthermore, the exact opposite behavior was observed for the diamino- and dioxo-5-aminopyrimidine-tagged oligomers. Such contrasting base-pairing behavior revealed the existence of a correlation between the magnitude of the ΔpK_a of the complementary partners and the strength of the base pairing, independent of the nature of the backbones to which they were attached; thereby implying that the constitution of recognition elements may have played a more influential role than the structure of the backbone in nature's choice of an informational system.[2b] As a continuation of this work, we report herein the base-pairing properties of oligo-dipeptides tagged with orotic acid (2,4-dioxopyrimidine-6-carboxylic acid; $^{6\text{-cP}}$OO, **1**) and its complementary base-pairing partner, 2,4-diaminopyrimidine-6-carboxylic acid ($^{6\text{-cP}}$NN, **2**) (Figure 1a). The results are consistent with the previous postulate[2b] and, in unison, provide a general rationale for correlating the pK_a of a recognition element with its base-pairing propensity.

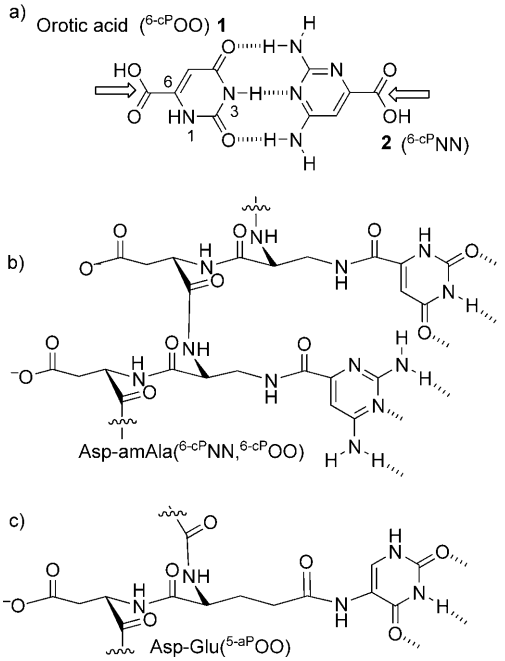

Figure 1. a) The tridentate Watson–Crick-type base pair between orotic acid ($^{6\text{-cP}}$OO, **1**) and its 2,4-diamino counterpart ($^{6\text{-cP}}$NN, **2**); b) Idealized conformation of Asp-amAla oligo-dipeptide backbone tagged with $^{6\text{-cP}}$OO and $^{6\text{-cP}}$NN—derived from the type of conformational reasoning[7] applied in the case of c). c) Asp-Glu oligo-dipeptide tagged with 2,4-dioxo-5-aminopyrimdine ($^{5\text{-aP}}$OO).[2b]

Orotic acid is formed as one of the products of an oligomerization reaction of HCN, and is unique from the standpoint of contemporary biology as the precursor in the de novo synthesis of pyrimidine nucleotides.[4] It offers an alternate way of tagging a backbone through its 6-carboxylic acid moiety, and can function as a recognition element through a Watson–Crick-type tridentate hydrogen bonding with its corresponding 2,4-diamino counterpart (Figure 1a). Moreover, the pK_a values of the closely related N(1)-methylorotamide (9.1)[5a] and 2,4-diaminopyrimidine-6-carboxamide (4.8)[5c] indicate that they could form a base pair.[2] The choice of the backbone was influenced by previous studies wherein 5-aminopyrimidines were tagged to Asp-Glu oligo-dipeptide scaffolds through an amide bond[2] (Figure 1c). However, in orotic acid the CO_2H group is on the heterocycle, necessitating a diaminoacid, 2,3-diaminopropanoic acid,[6] as the tagging unit. We focused on the oligo-dipeptide, derived from alternating units of L-aspartic acid and L-3-aminoalanine

[*] Dr. X. Zhang, Prof. Dr. R. Krishnamurthy
Department of Chemistry, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-9573
E-mail: rkrishna@scripps.edu

[**] We are extremely grateful for the discussions, encouragement and support from Prof. Albert Eschenmoser in carrying out this work. We thank Dr. Jian Xie for help with LC/MS measurements. This work was supported by NASA Astrobiology: Exobiology & Evolutionary Biology (grant NNX07AK18G) and by The Skaggs Research Foundation.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200904188.

(Asp-amAla), tagged with $^{6\text{-cP}}$OO and $^{6\text{-cP}}$NN as recognition elements (Figure 1b). On the basis of qualitative conformational analysis,[7] we anticipated these oligomers would form base pairs since their overall backbone/base-pairing axes are similar to that of Asp-Glu oligo-dipeptide tagged with 2,4-dioxo-5-aminopyrimidine ($^{5\text{-aP}}$OO; Figure 1c), which does exhibit strong base pairing.[2] The effect on base pairing when orotic acid is incorporated through its N1-position at selective locations within a PNA_2:NA triplex, has been studied.[8]

Since the goal of the study was to assess the base-pairing properties of the oligomers, we prepared the required Fmoc-protected (Fmoc = 9-fluorenylmethoxycarbonyl) monomers for the machine assisted synthesis of $^{6\text{-cP}}$OO-tagged Asp-amAla oligo-dipeptides.[9] The ion-exchange HPLC purification was challenging. In the case of the hexamer we obtained homogeneous material, whereas in the case of the dodecamer we could only isolate mixtures containing the 11- and 12-mer; for the hexadecamer we isolated a mixture of 14-, 15-, and 16-mer (see Table S1 in the Supporting Information),[9] and used them in this study. The cross-pairing of Asp-amAla($^{6\text{-cP}}$OO) oligo-dipeptides with adenine containing DNA and RNA sequences were investigated using temperature dependant UV and CD spectroscopy; the results are summarized in Figure 2 (and Table S2, entries 1–16 in the Supporting Information). Strikingly, all of the data consistently point to no (or very weak) base-pairing interactions (Figure 2a,b), which is supported by the negligible effects that the variations in concentration of oligomers, salt, and length of the pairing strands have on the base-pairing behavior. That there is interaction was shown by a Job plot (Figure 2c) which, not surprisingly, shows a triplex with a 2:1 ratio of $^{6-\text{cP}}$OO-tagged dodecapeptide with poly-r(A).

We checked if the weak pairing behavior of the orotamide unit was consistent with a correlation between the ΔpK_a and the base-pairing strength.[2] The pK_a value of the orotamide unit in dipeptide **S21**[9] was determined to be 6.6,[9] lower than that of either orotic acid ($pK_{a2} = 9.45$)[10a] or of *N*(1)-methyl orotamide ($pK_a = 9.1$)[5a], but in agreement with the pK_a of methyl orotate (7.93)[5a]. The difference in pK_a values between deoxyadenosine (3.8)[10a] and an orotamide derivative (6.6) is less (ΔpK_a ca. 3) when compared with the difference (ΔpK_a ca. 5) between deoxyadenosine and deoxythymidine (pK_a ca. 9.5)[5a].[11] These results agree with the correlation, 'the smaller the ΔpK_a of base-pairing partners, the weaker the base-pairing strength'.[2b] That this pK_a value of 6.6 of the orotamide unit in **S21** correlates with deprotonation of its N1–H proton (and not N3–H proton) was established by the similarity of the change in its UV spectrum (λ_{max} shifts from 280 nm to 300 nm by deprotonation)[9] with that of the known[5b] N(3)-methyl orotic acid (a behavior not shown by N(1)-methyl derivative). Similar bathochromic shifts were also observed for the temperature dependent UV spectrum of Asp-amAla($^{6\text{c-P}}$OO)$_{12}$ either alone, or in the presence of its pairing partner, indicating a temperature facilitated deprotonation of the orotamide unit within the oligomer.[9] All these point to the orotamide unit existing in its anionic form at N1, which—and not the change in the backbone—seems to be contributing to its weak base pairing.

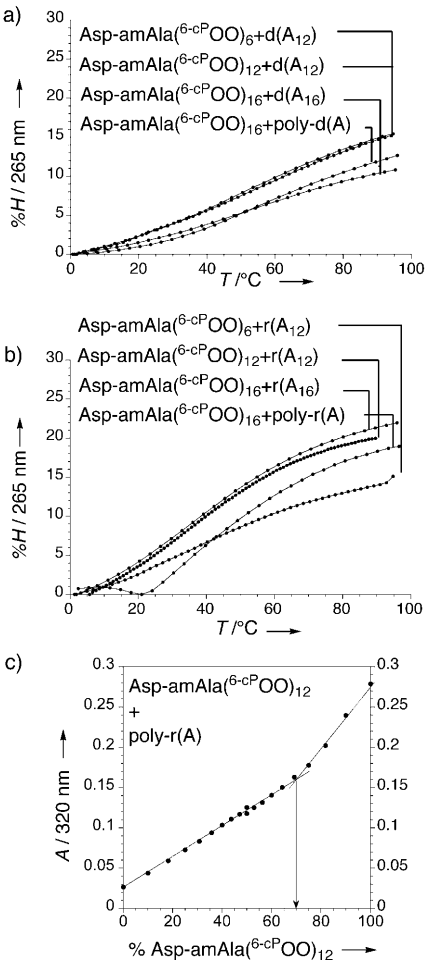

Figure 2. Weak base pair interactions of Asp-amAla($^{6\text{-cP}}$OO) oligomers. UV spectroscopic T_m curves of Asp-amAla($^{6\text{-cP}}$OO) oligomers documenting a) no pairing with DNA oligomers and b) very weak pairing with RNA oligomers. c) Job plot showing the 2:1 ratio of the pairing partners in the homo-duplex formed at 0°C. Measurements were made with total concentration of approximately 10 μM (1:1) in 1 M NaCl, 10 mM aq. NaH_2PO_4, 0.1 mM Na_2EDTA, pH 7.0. No self pairing was observed for individual partner strands.

This interpretation finds support in observations that the $^{5\text{-aP}}$OO-tagged ($pK_a = 8.9$) Asp-Glu oligo-dipeptide, which exhibits base pairing, has the same overall backbone/base-pairing axes (Figure 2c).[2b]

We turned our attention to the $^{6\text{-cP}}$NN-tagged oligo-dipeptides. Being aware that the pK_a value of the heterocycle has an effect on its base-pairing disposition, we measured the pK_a of 2,4-diaminopyrimidine-6-carboxamide derivative **S22**[9] and found it to be 4.7,[9] consistent with that known for orotamide.[5c] Therefore, an oligomer tagged with $^{6\text{-cP}}$NN should form base pairs when partnered with thymine or uracil (pK_a ca. 9.5–9.8)[10a], which is in agreement with the correlation of the ΔpK_a and the base-pairing strength.[2b] The required oligo-dipeptides Asp-amAla($^{6\text{-cP}}$NN)$_{6,12\&16}$ were prepared from Fmoc-protected monomers and purified without any difficulty by ion-exchange HPLC methods (see Table S1 in the Supporting Information).[9] Their base-pairing behavior was investigated with thymine- (DNA) and uracil-containing

(RNA) sequences (Figure 3; see Table S2, entries 17–38 in the Supporting Information). Whereas weak base pairing of the hexamer $(^{6\text{-cP}}NN)_6$ was observed, the 12-mer and 16-mer showed relatively strong base pairing with a preference towards RNA over DNA (Figure 3a). The effects of varying the length of sequence, concentration of oligomers, and concentration of salt on the UV spectroscopic T_m values confirmed duplex formation, which was further corroborated by a 1:1 stoichiometry from a Job plot (Figure 3d).[9] Surprisingly, thymine-containing α-threofuranosyl nucleic acid (TNA) sequences paired stronger indicating that $^{6\text{-cP}}$NN-tagged oligo-dipeptide backbones had more affinity for the structurally least flexible system (TNA) and less for the most flexible system (DNA), which is the reverse of what was observed in the previous study;[2a] the CD spectra of these duplexes suggest that their overall shape is largely similar (Figure 3b). The Asp-amAla($^{6\text{-cP}}$NN) oligomers were found to degrade at pH 4 (RT, 40 min), whereas the Asp-amAla-($^{6\text{-cP}}$OO) oligomers were stable (by HPLC). This suggests that the source of instability is the $^{6\text{-cP}}$NN heterocycle and not the peptide backbone.

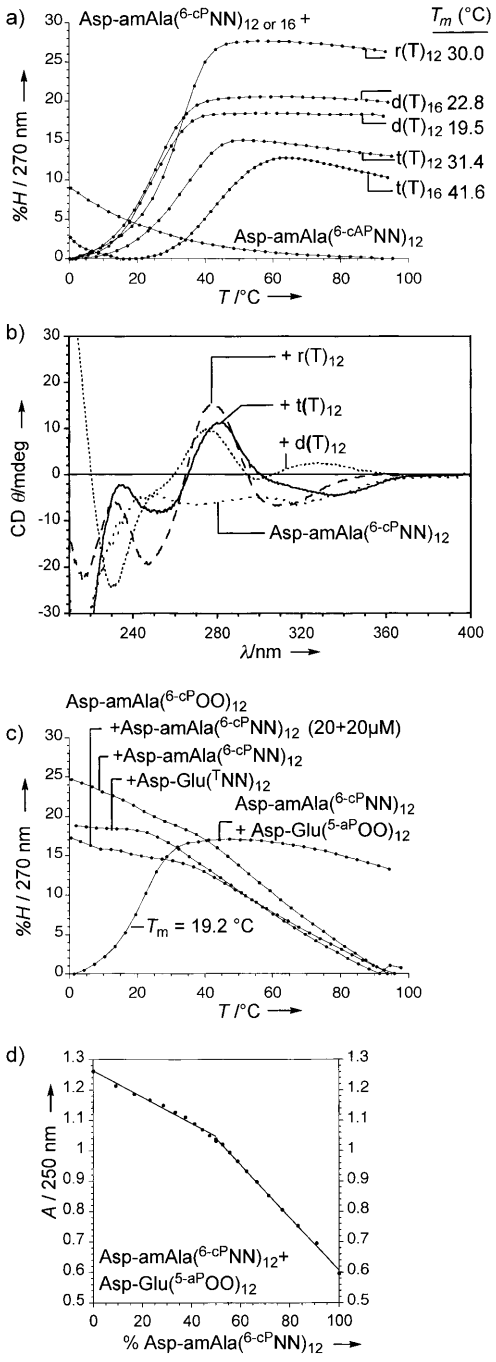

Figure 3. Base-pairing behavior of Asp-amAla($^{6-cP}$NN) oligo-dipeptides. a) UV spectroscopic T_m curves demonstrating strong pairing with (r) RNA, (d) DNA, and (t) TNA sequences containing thymine. b) CD spectral comparison of Asp-amAla$(^{6-cP}NN)_{12}$ and the corresponding duplexes formed with $r(T)_{12}$, $d(T)_{12}$, and $t(T)_{12}$. c) UV spectroscopic T_m curves documenting pairing behavior in the inter- and intra- oligo-dipeptide systems. d) Job plot showing the 1:1 ratio of the pairing partners in the intra-system homo-duplex formed at 0°C (see Table S2, entry 31 in the Supporting Information). Measurement conditions: see caption for Figure 3.

The divergent base-pairing behavior of the $^{6\text{-cP}}$OO and $^{6\text{-cP}}$NN nuclei is consistent with earlier outcomes.[2] Extrapolation of the correlation of ΔpK_a and base-pairing strength is found to be valid for the inter- and intra-system combinations in the all-oligo-dipeptide-backbone: no (or weak) base pairing is expected, and observed (Figure 3c), between oligo-dipeptides containing either 2,4-diaminotriazine,^{T}NN ($pK_a = 4.5$)[2a] or $^{6\text{-cP}}$NN ($pK_a = 4.7$), and $^{6-cP}$OO ($pK_a = 6.6$), where the ΔpK_a is approximately 2, (see Table S2, entry 11, 29, and 37 in the Supporting Information). Whereas for $^{6\text{-cP}}$NN ($pK_a = 4.7$) and its partner $^{5\text{-aP}}$OO ($pK_a = 8.9$)[2b], where the ΔpK_a is approximately 4, stronger base pairing is expected and detected (see Table S2, entries 31 and 38 in the Supporting Information). Thus, when juxtaposed with the canonical nucleobases, and with potentially natural alternative heterocycles, in the context of the correlation between the ΔpK_a and the base-pairing strength, the dissimilar base-pairing behavior of orotoamide and its 2,4-diamino counterpart blend in (Figure 4).

The above weak base-pairing properties of $^{6\text{-cP}}$OO and $^{6\text{-cP}}$NN dismiss the likelihood that they could have been members of the landscape of potentially primordial informational systems.[2a] However, these results, in combination with others,[2,12,13] may be used—in the limited context of pK_a values of the heterocycles—to elicit chemical reasons as to why canonical nucleobases represent an "optimum" with respect to their structure and function, and ask: "why nature chose these nucleobases"? The canonical nucleobases have pK_a values which are 2–3 units away from the physiological pH (ca. 7) of the medium (Figure 4). The alternative heterocycles which base pair weakly have pK_a values closer to the pH (ca. 7) of the medium, whereas relatively stronger base-pairing heterocycles have pK_a values more removed (Figure 4). In other words, the smaller the difference between the pK_a value of the heterocycle and the pH of the aqueous medium ($pK_a - pH < 2$), the weaker the base pairing.[14] Among the alternative heterocycles and nucleobases—which can become charged at physiological pH—the ones with a pK_a value lower than pH 7 should be protonated, whereas the ones with a pK_a value higher than pH 7 should be deprotonated. In these cases, the smaller the difference between the pK_a value of the molecule and the pH of the

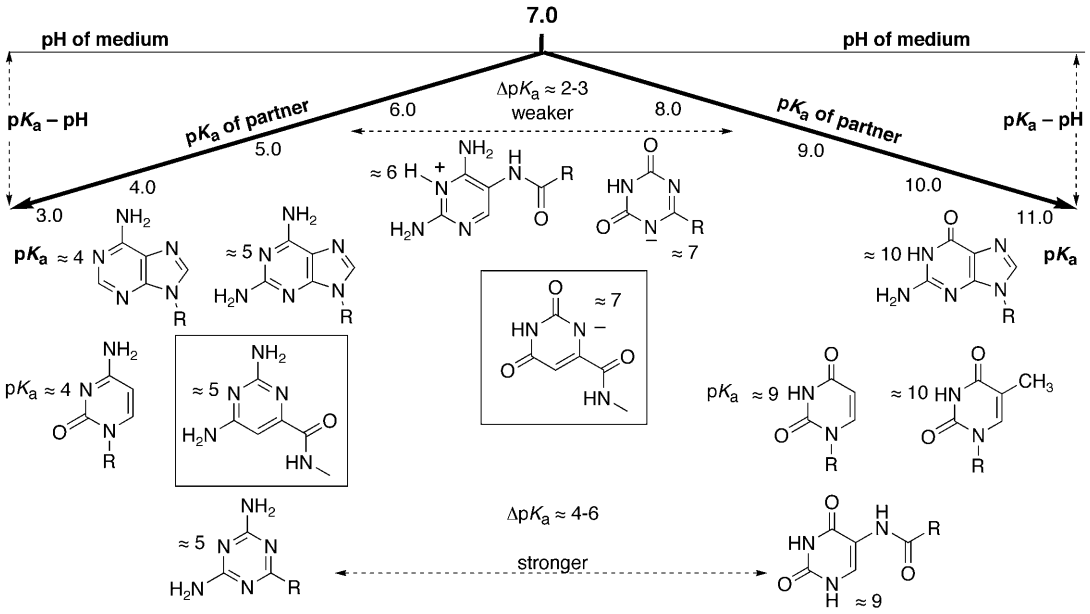

Figure 4. Correlation of the pK_a value of the base-pairing partners with their base-pairing strength (aq. neutral conditions). The pK_a values below 7 are for the conjugate acid of the hetereocycles. The canonical nucleobases and strongly base-pairing (potentially natural) heterocycles have pK_a values at least 2–3 units away from the pH of the medium, whereas the weakly pairing heterocycles have pK_a values which are closer.

medium, the greater the degree of ionization of the heterocycle. However, if it were the other way around the exact opposite is expected, as exemplified by xanthine (pK_a = 5.5)[15] which exists as a monoanion at physiological pH.[15d] It pairs (relatively) more strongly with adenine when the pH is lowered from 7.5 to 5.5,[15b,c] as a consequence of the reversal of the extent of ionization when the pK_a value of the molecule and the pH of the medium become equal. This situation is akin to hydrogen-bond formation in (nonpolar) organic solvents, giving rise to the "pK_a match" rule[16]—the smaller the ΔpK_a between the hydrogen-bond donor and hydrogen-bond acceptor, the stronger the association between them—the exact opposite of what has been observed in aqueous medium, both herein and previously.[2] Notably, in all cases, the common factor seems to be the magnitude of neutral character of the partners involved.[17]

Whereas the pK_a values of the canonical nucleobases are such that they remain neutral under physiological conditions, it is not so for the phosphodiester backbone (pK_{a1} ca. 2). In arguing "Why Nature chose Phosphates",[18a] Westheimer has emphasized the "importance of being ionized",[18b] pointing out that under physiological conditions a phosphodiester is ionized—contributing to the kinetic stability of DNA/RNA and enabling their retention within a bilayer membrane—making it suitable for its primary function as a linker in an informational polymer. Comparing the base-pairing properties of alternative heterocycles and the canonical nucleobases, we are led to the conclusion that the exact opposite is true, that is, "the importance of being **not** ionized" under physiological conditions. The canonical nucleobases, by virtue of being not ionized—are hydrophobic, minimizing their interactions with water and maximizing the stacking interactions among them.[19] This, in turn, reinforces the hydrogen bond between appropriate partners. The alternative heterocycles which are ionized in aqueous medium become less hydrophobic, leading to greater interaction with water molecules, and thus weak (or no) base pairing. The significance of pK_a values of canonical nucleobases being less than 4 and greater than 9, correlating to their neutral forms and their ability to form Watson–Crick base pairs, has been pointed out in the context of structural studies.[20] This dichotomous and simultaneous ionization and non-ionization behavior of two different portions of the same biomolecule, serves the purposes of RNA/DNA being impermeable to the cell membrane and kinetically stable to hydrolysis, while possessing the capacity to form informational base pairs.

The work presented herein, along with previous results,[2] suggests that a balance between pK_a of the heterocycle and pH of the medium—manifesting as the correlation between the magnitude of ΔpK_a of complementary heterocycles and their base-pairing strength—could be used as a rough barometer of base-pairing propensity, all other things being equal; and could be invoked to explain the effect of nucleobase structural variations on duplex stability[12,13,21,22] (with exceptions[23]). Such an optimal interplay between pK_a and pH, underscores[2] the role of the physicochemical properties of canonical nucleobases, under near neutral aqueous conditions, in contributing to their functional superiority (optimization of base-pairing strength)—a status conferred by their constitutional uniqueness.

Received: July 28, 2009
Published online: September 18, 2009

Keywords: DNA · nucleic acids · oligonucleotides · oligopeptides · RNA

[1] a) A. Eschenmoser, *Origins Life Evol. Biospheres* **2004**, *34*, 277; b) A. Eschenmoser, *Tetrahedron* **2007**, *63*, 12821.

[2] a) G. K. Mittapalli, K. R. Reddy, H. Xiong, O. Munoz, B. Han, F. De Riccardis, R. Krishnamurthy, A. Eschenmoser, *Angew. Chem.* **2007**, *119*, 2522; *Angew. Chem. Int. Ed.* **2007**, *46*, 2470; b) G. K. Mittapalli, Y. M. Osornio, M. Guerrero, K. R. Reddy, R. Krishnamurthy, A. Eschenmoser, *Angew. Chem.* **2007**, *119*, 2530; *Angew. Chem. Int. Ed.* **2007**, *46*, 2478.

[3] Recent elegant work has demonstrated that this "nucleosidation problem" can be circumvented: M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239.

[4] a) J. P. Ferris, P. C. Joshi, *Science* **1978**, *201*, 361; b) A. B. Voet, A. W. Schwartz, *Origins Life Evol. Biospheres* **1982**, *12*, 45; c) S. Miyakawa, H. J. Cleaves, S. L. Miller, *Origins Life Evol. Biospheres* **2002**, *32*, 209.

[5] a) J. Jonas, J. Gut, *Collect. Czech. Chem. Commun.* **1962**, *27*, 716; b) J. J. Fox, N. Yung, I. Wempen, *Biochim. Biophys. Acta* **1957**, *23*, 295; c) B. Roth, J. Z. Strelitz, *J. Org. Chem.* **1969**, *34*, 821.

[6] 2,3-Diaminoacids have been identified in Murchison meteorite: a) U. J. Meierhenrich, G. M. Munoz Caro, J. H. Bredehoft, E. K. Jessberger, W. H.-P. Thiemann, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9182; in experiments simulating chemistry of interstellar medium: b) G. M. Munoz Caro, U. J. Meierhenrich, W. A. Schutte, B. Barbier, A. Arcones Segovia, H. Rosenbauer, W. H.-P. Thiemann, A. Brack, J. M. Greenberg, *Nature* **2002**, *416*, 403; and in hydrolysis of dilute HCN solutions: c) K. E. Nelson, M. Levy, S. L. Miller, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3868.

[7] A. Eschenmoser, *Chimia* **2005**, *59*, 836.

[8] a) R. E. Hudson, F. Wojciechowski, *Can. J. Chem.* **2005**, *83*, 1731; b) R. E. Hudson, F. Wojciechowski, *Nucleosides Nucleotides Nucleic Acids* **2005**, *24*, 1123.

[9] See the Supporting Information for additional details.

[10] R. M. C. Dawson, D. C. Elliot, W. H. Elliot, K. M. Jones, *Data for Biochemical Research*, Oxford, Clarendon Press, **1959**.

[11] In the DNA series, $d(T_{12})/d(A_{12})$, ΔpK_a ca. 6, $T_m = 41.6\,^{\circ}C$ (1 M NaCl, 10 mM NaH_2PO_4, 0.1 mM Na_2EDTA, pH 7.0).

[12] a) F. Seela, C. Wei. G. Becher, M. Zulauf, P. Leonard, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 289; b) F. Seela, X. Peng, H. Li, P. Chittepu, K. I. Shaikh, J. He, Y. He, I. Mikhailopulo, *Collect. Symp. Ser.* **2005**, *7*, 1; c) X. Peng, H. Li, F. Seela, *Nucleic Acids Res.* **2006**, *34*, 5987; d) I. Luyten, P. Herdewijn, *Eur. J. Med. Chem.* **1998**, *33*, 515.

[13] a) N. J. Leonard, A. G. Morrice, M. A. Sprecker, *J. Org. Chem.* **1975**, *40*, 356; b) S. A. Benner, A. Ricardo, M. A. Carrigan, *Curr. Opin. Chem. Biol.* **2004**, *8*, 672; c) A. A. Henry, F. Romesberg, *Curr. Opin. Chem. Biol.* **2003**, *7*, 727; d) A. T. Krueger, H. Lu, A. H. F. Lee, E. T. Kool, *Acc. Chem. Res.* **2007**, *40*, 141; e) T. R. Battersby, M. Albalos, M. J. Friesnhahn, *Chem. Biol.* **2007**, *14*, 525; f) B. D. Heuberger, C. Switzer, *ChemBioChem* **2008**, *9*, 2779.

[14] Exceptions are found for nucleobases that form base pairs upon protonation, for example, cytosine self pairing: a) K. Gehring, J. L. Leroy, M. Gueron, *Nature* **1993**, *363*, 561; 8-amino-G self pairing: b) M. Hattori, J. Frazier, H. T. Miles, *Biopolymers* **1975**, *14*, 2095; c) A. E. Engelhart, T. H. Morton, N. V. Hud, *Chem. Commun.* **2009**, 647; $A(H^+)$–C and $A(H^+)$–G base pairing: d) T. Brown, G. A. Leonard, E. D. Booth, G. Kneale, *J. Mol. Biol.* **1990**, *212*, 437.

[15] a) K. Groebke, J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmerman, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1998**, *81*, 375; b) R. Eritja, D. M. Horowitz, P. A. Walker, J. P. Ziehler-Martin, M. S. Boosalis, M. F. Goodman, K. Itakura, B. E. Kaplan, *Nucleic Acids Res.* **1986**, *14*, 8135; c) S. C. Jurczyk, J. Horlacher, K. G. Devined, S. A. Benner, T. R. Battersby, *Helv. Chim. Acta* **2000**, *83*, 1517; d) E. Kulikowska, B. Kierdaszuk, D. Shugar, *Acta Biochim. Pol.* **2004**, *51*, 493. This purine, as opposed to pyrimidines, is able to base pair despite being ionized at physiological pH, which is in line with observations that charged pyrimdines tend to "unstack" more than charged purines; page 132 of reference [20a].

[16] a) Y. Kyogoku, R. C. Lord, A. Rich, *Proc. Natl. Acad. Sci. USA* **1967**, *57*, 250; b) S. C. Zimmerman, P. S. Corbin, *Struct. Bonding (Berlin)* **2000**, *96*, 63; c) P. Gilli, L. Pretto, V. Bertolasi, G. Gilli, *Acc. Chem. Res.* **2009**, *42*, 33. For a discussion of how this pK_a match rule explains stronger duplexes in RNA versus DNA, see d) P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya, J. Chattopadhyaya, *J. Am. Chem. Soc.* **2004**, *126*, 2862.

[17] A striking example is the W–C pairing between guanine (G; $pK_a = 9.5$) and isoguanine (isoG; $pK_a = 9.0$) where the ΔpK_a is less than one; however, pK_a–pH is approximately 2–3, enabling the partners to be neutral; a) R. Krishnamurthy, S. Pitsch, M. Minton, C. Miculka, N. Windhab, A. Eschenmoser, *Angew. Chem.* **1996**, *108*, 1619; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1537; b) reference [15a]; c) reference [13f].

[18] a) F. H. Westheimer, *Science* **1987**, *235*, 1173; b) B. Davis, *Arch. Biochem. Biophys.* **1958**, *78*, 497.

[19] a) P. O. P. Ts'o, *Basic Principles in Nucleic Acid Chemistry, Vol. I* (Ed.: P. O. P. Ts'o), Academic Press, New York, **1974**, chap. 6; b) M. Eigen, D. Pörschke, *J. Mol. Biol.* **1970**, *53*, 123; c) A. Pohorille, S. K. Burt, R. D. MacElroy, *J. Am. Chem. Soc.* **1984**, *106*, 402; d) P. Cieplak, P. A. Kollman, *J. Am. Chem. Soc.* **1988**, *110*, 3734.

[20] a) W. Saenger, *Principles of Nucleic Acid Structure* (Ed.: C. R. Cantor), Springer, New York, **1984**; b) M. Roitzsch, B. Lippert, *J. Am. Chem. Soc.* **2004**, *126*, 2421.

[21] The weaker pairing of 2,6-diaminopurine—thymine (ΔpK_a ca. 4) compared to guanine—cytosine (ΔpK_a ca. 5) fits with this correlation (alternative explanations: a) W. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008; b) A. R. Leach, P. Kollman, *J. Am. Chem. Soc.* **1992**, *114*, 3675). In *p*-RNA series: c) S. Pitsch, S. Wenderborn, R. Krishnamurthy, A. Holzner, M. Minton, M. Bolli, C. Miculka, N. Windhab, R. Micura, M. Stanek, B. Jaun, A. Eschenmoser, *Helv. Chim. Acta* **2003**, *86*, 4270; and in the DNA series see also references [13e and 13f]. For additional examples of ΔpK_a—base-pairing-strength correlation see Table S3 in the Supporting Information.[9]

[22] For the size-expanded (x) nucleobases, the low pK_a value (7.2) of dxG may also account for the observations: a) that C–dxG base pairs (ΔpK_a ca. 3) are less stable than the inverse dxC–G (predicted ΔpK_a ca. 5) combination and b) why the increase in UV spectroscopic T_m values—in going from AT only sequences to mixed ATGC sequences—is larger in the DNA series when compared with corresponding sequences in an xDNA series: a) H. Liu, S. R. Lynch, E. T. Kool, *J. Am. Chem. Soc.* **2004**, *126*, 6900; b) J. Gao, H. Liu, E. T. Kool, *Angew. Chem.* **2005**, *117*, 3178; *Angew. Chem. Int. Ed.* **2005**, *44*, 3118.

[23] For example, halogenated heterocycles in reference [12]; see also reference [14].

