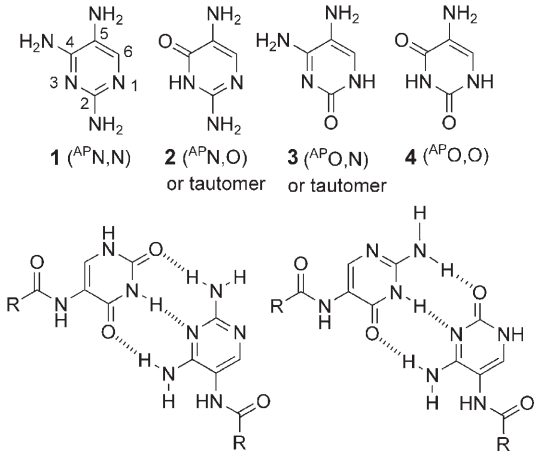
VIP Base Pairing

DOI: 10.1002/ange.200603209

Mapping the Landscape of Potentially Primordial Informational Oligomers: Oligodipeptides Tagged with 2,4-Disubstituted 5-Aminopyrimidines as Recognition Elements**

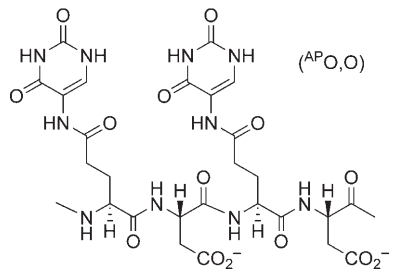
Gopi Kumar Mittapalli, Yazmin M. Osornio, Miguel A. Guerrero, Kondreddi Ravinder Reddy, Ramanarayanan Krishnamurthy, and Albert Eschenmoser**

The canonical purine bases of the natural nucleic acids occupy a central position in the chemist's reasoning of the problem of life's origin. Their formation from aqueous ammonium cyanide under a variety of potentially prebiotic conditions[2] and the fundamental function these bases fulfill as recognition elements in contemporary life's genetic system constitute facts deemed to be intimately related to each other. The generational relationship the other type of canonical nucleobases, the pyrimidines, have to HCN, is more complex. Although their formation in reaction mixtures of oligomerizing HCN has been experimentally observed,[3] the major part of the pyrimidine derivatives formed under such conditions (observed after hydrolysis of the reaction mixtures) are derivatives that bear a hydroxy or amino group at position 5 of the pyrimidine ring.[3b,4] The hydroxy groups supposedly result from replacement of an amino group in the course of the hydrolytic process. Importantly, a mechanistic analysis of the chemistry involved in these processes shows the formation of the canonical pyrimidines from HCN to require a reductive step, whereas that of 5-aminopyrimidines does not, as neither does the formation of the canonical purines.[5] These facts and relationships, together with the specific opportunities that an amino group in position 5 of the pyrimidine ring offers for backbone tagging, led us to investigate the properties of the HCN-derivable 5-aminopyrimidines **1**–**4** (Scheme 1) with respect to their potential to act as recognition elements in one of the dipeptide-based oligomer systems described in the preceding Communication (see Scheme 2).[1]



Scheme 1. The family of the 5-amino-2,4-disubstituted pyrimidines (**1**–**4**) that could form two informational pyrimidine–pyrimidine base pairs as 5-aminoacyl derivatives.

Chemical syntheses of all four of these 5-aminopyrimidines are known.[6] Among them, 5-aminouracil and 5-aminocytosine have been used as a substitute of thymine and cytosine, respectively, in the chemistry of DNA oligonucleotides in various contexts. In these cases, the heterocycle is attached to the sugar backbone through the conventional a C1–N1 nucleosidic bond.[7] Our own interest focused on the potential of 2,4-disubstituted 5-aminopyrimidines to become attached as recognition elements to oligomer backbones through the extra amino group and to form two tridentate informational pyrimidine–pyrimidine base pairs, a possibility



Scheme 2. Structure type of the 5-aminouracil-tagged AspGlu oligodipeptide system; end groups are either free NH_2 and CO_2H or CH_3CONH and $CONH_2$.

[*] Dr. G. K. Mittapalli, Dr. Y. M. Osornio, Dr. M. A. Guerrero, Dr. K. R. Reddy, Prof. Dr. R. Krishnamurthy, Prof. Dr. A. Eschenmoser
Department of Chemistry and
The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-9573
E-mail: rkrishna@scripps.edu
Eschenmoser@org.chem.ethz.ch
Prof. Dr. A. Eschenmoser
Laboratorium für Organische Chemie
Eidgenössische Technische Hochschule
HCI Hönggerberg, 8093 Zürich (Switzerland)
Fax: (+41) 1-632-1043

[**] Chemistry of α-Aminonitriles, Part 46. For Part 45, see the preceding paper (reference [1]). This work was supported by the Skaggs Research Foundation. G.K.M., Y.M.O, M.A.G, and K.R.R. are Skaggs Postdoctoral Fellows.

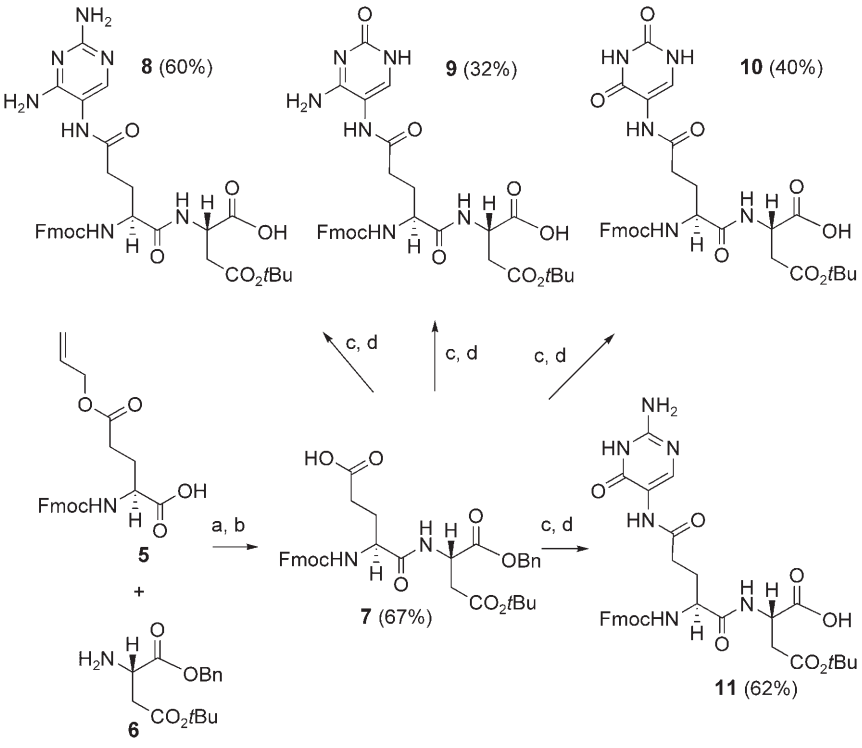
Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

that 5-aminopyrimidines connected to a backbone through a conventional nucleosidic bond to N1 lack. Such a pair of 5-aminopyrimidines has the potential to act as a functional alternative to the two canonical Watson–Crick base pairs (Scheme 1).[8]

The extra amino group in position 5 of all four 5-aminopyrimidines displays pronounced nucleophilicity and is known to react (by mechanistically understandable reasons) regioselectively with acylating agents.[6e,9] Herein, we describe the pairing properties of oligodipeptides built of L-Asp-L-Glu dipeptide monomers in which the γ-carboxyl function of the glutamic acid residue is tagged with the recognition element by such acylation (Scheme 2). The choice of oligodipeptides derived from glutamic and aspartic acids as oligomer backbones served the purpose of comparing base-pairing behavior in this series with data described in the preceding paper.[1] Again, the procedures used in the synthesis of the oligodipeptides were chosen without regard to generational constraints of etiological nature. The four building blocks (**8**–**11**), suitably protected for the solid-support oligodipeptide synthesis, were prepared from a common intermediate, the AspGlu dipeptide **7**, which is accessible from the known glutamic acid and aspartic acid derivatives **5**[10] and **6**,[11] respectively (Scheme 3).[12] Once again, as in the case of the triazine derivatives, all synthetic manipulations could be conducted without the need for protecting the amino or oxo groups on the pyrimidine nucleus.

Table 1 summarizes the observations made on the base-pairing capabilities of L-Asp-L-Glu oligodipeptides tagged on the Glu residue with the four 5-aminopyrimidine bases. Most remarkably, the homododecamer tagged with the 2,4-dioxopyrimidine nucleus (^{AP}OO) was found to pair strongly with complementary DNA sequences (Table 1, entries 1–6 and Figure 1a,c), whereas the analogous homododecamer bearing 2,4-diaminopyrimidine (^{AP}NN) tags does only poorly so (Table 1, entries 14–16 and Figure 1b). This contrasting cross-pairing behavior of the two complementary homo-oligomers towards DNA is reminiscent of the equally contrasting behavior of the two corresponding oligomers that are tagged with the complementary 2,4-dioxo- and 2,4-diaminotriazines as nucleobases described in the preceding communication, except that there it was the diamino-substituted base that showed the strong pairing and the dioxo derivative the one that paired only very weakly. Not surprisingly, therefore, the combination of the two 5-aminopyrimidine-tagged homo-oligodipeptide dodecamers AspGlu(^{AP}O,O)$_{12}$ and AspGlu(^{AP}N,N)$_{12}$ results in only very weak intrasystem pairing (Table 1, entry 17), which is reminiscent of the (even weaker) intrasystem pairing between the strands with identical backbone but tagged with the corresponding triazine heterocycles described in the preceding paper.[1] Interestingly, as well as not unexpectedly, the intersystem combination of two complementary homobasic oligo-(Asp-Glu)-dipeptide dodecamers, one tagged with the 2,4-dioxopyrimidine nucleus and the other with 2,4-diaminotriazine (both bases representing in their series the ones that cross-pair strongly with DNA), showed reasonably strong intersystem cross-pairing (Table 1, entries 7, 8, and 13 and Figure 1d), whereas the inverse combination (both bases representing those that, in their series, pair with DNA weakly) showed no pairing at all (Table 1, entry 18). Even though an interpretation of the remarkable antipodal pairing behavior of oxo/oxo and the amino/amino members in two families of heterocycles may, to some extent, be complicated by differences in backbone constitution in some of the pairing experiments carried out so far, we believe that this conspicuously diverging behavior of the members in the two series' of alternative nucleobases touches upon an essential aspect concerning the chemical factors that co-determine base-pairing strength in nucleic acids (see below).

While both the 2,4-dioxo and the 2,4-diamino member of the 5-aminopyrimidine family do not present any uncer-



Scheme 3. Synthesis of the four 5-aminopyrimidine-tagged AspGlu-dipeptide building blocks used in the solid-support synthesis of the oligomers. a) 0.5 M EDCI (1.0 equiv), 0.5 M HOBt (1.0 equiv), DMF, RT, 4 h; b) 0.12 M $PhSiH_3$ (4.0 equiv), [$(Ph_3P)_4Pd$] (0.05 equiv), CH_2Cl_2, RT, 5 h; c) 0.54 M HBTU (1.5 equiv), 0.36 M HOBt (1.0 equiv), 0.72 M 5-aminopyrimidine (**1**, **2**, **3**, or **4**; 2.0 equiv), DMF, RT−35 °C, 36–60 h; d) 10% Pd/C, H_2 (1 atm), MeOH/DMF/HCO_2H (4.8:4.8:0.4), 0 °C, 2–3.5 h. EDCI = 1-ethyl-3-(3′-dimethylaminopropyl)carbodiimide, HOBt = *N*-hydroxybenzotriazole, DMF = *N*,*N*-dimethylformamide, HBTU = *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate, Fmoc = 9-fluorenylmethoxycarbonyl.

Table 1: T_m data of 5-aminopyrimidine-tagged AspGlu oligodipeptides.

Entry	Pairing system[a]	T_m(UV) (λ [nm])[b]	T_m(CD) (λ [nm])[b]	Comments[c]
1	^{H_2NOC}AspGlu(^{AP}OO)$_{12}$ + poly-d(A)	61.1 (255)		
2	+ poly-r(A)	40.8 (255)		
3	+ d(A)$_{12}$	51.4 (255)	50 (251)	
4	+ r(A)$_{12}$	35.5 (255)	44 (250)	
5	+ d(D)$_{12}$	41.9 (270)		
6	+ t(D)$_{12}$	54.3 (275)	55 (275)	
7	+ ^{H_2NOC}AspGlu(^{T}NN)$_{12}$	16.0 (270)		
8	+ ^{H_2NOC}AspAsp(^{T}NN)$_{12}$	15.5 (270)		
9	^{H_2NOC}AspGlu(^{AP}OO)$_{16}$ + poly-d(A)	64.7 (255)		
10	+ poly-r(A)	44.0 (255)		
11	+ d(A)$_{16}$	54.7 (255)		
12	+ r(A)$_{16}$	41.7 (255)		
13	+ ^{H_2NOC}AspGlu(^{T}NN)$_{16}$	24.4 (255)		1:1 duplex by Job plot (UV)
14	^{H_2NOC}AspGlu(^{AP}NN)$_{12}$ + poly-d(T)	23.3 (250)	26 (275)	
15	+ poly-r(U)	$<$10 (250)		
16	+ d(T)$_{12}$	$<$10 (250)		
		$<$10 (250)		pH 8
17	+ ^{H_2NOC}AspGlu(^{AP}OO)$_{12}$	15.5 (255)		melting is hypochromic; 1:1 duplex by Job plot (UV)
		22.7 (255)		50 + 50 μM
18	+ ^{HOOC}AspGlu(^{T}OO)$_{12}$	$<$0 (240)		
19	^{H_2NOC}AspGlu(^{AP}OO)$_4$(^{AP}ON)$_4$(^{AP}OO)$_4$ + d(A$_4$^{iso}G$_4$A$_4$)	48.8 (260)	45 (285)	1:1 duplex by Job plot (UV)
		57.0 (260)		after 2 days at 4 °C
20	+ d(A$_4$G$_4$A$_4$)	$<$0 (250)		
21	^{H_2NOC}AspGlu(^{AP}OO)$_4$(^{AP}NO)$_4$(^{AP}OO)$_4$ + d(A$_4$G$_4$A$_4$)	$\approx$35 (260)		
22	+ d(A$_4$^{iso}G$_4$A$_4$)	$<$0 (260)		
23	^{H_2NOC}AspGlu[(^{AP}ON)(^{AP}NO)]$_6$ (self-pairing)	$<$0 (275)		
		27.0 (275)	26 (236)	100 μM
		32.7 (275)		100 μM, after 15 days at 4 °C
		38.0 (275)		100 μM, after 43 days at 4 °C
24	^{H_2NOC}AspGlu(^{AP}NO)$_6$(^{AP}ON)$_6$ (self-pairing)	$<$0 (280)		
		$\approx$10 (280)		100 μM
		28.0 (280)		100 μM, after 3 days at 4 °C
		32.0 (280)	25 (310)	100 μM, after 7 days at 4 °C

[a] Oligodipeptide sequences are written from the COOH terminus with every second amino acid residue tagged with the heterocycle; Asp = aspartyl; Glu = glutamyl; (^{AP}O,O) = 5-amino-2,4-dioxopyrimidine; (^{AP}N,N) = 2,4,5-triaminopyrimidine; (^{AP}N,O) = 2,5-diamino-4-oxopyrimidine; (^{AP}O,N) = 4,5-diamino-2-oxopyrimidine; (^{T}N,N) = (2,4-diamino)triazin-6-yl; A = adenine; D = 2,6-diaminopurine; T = thymine; U = uracil. [b] Measurements were made at the indicated wavelength (nm), the pairing systems (each strand $c \approx 5$ μM) in phosphate buffer solution (1 M NaCl, 10 mM NaH_2PO_4, 0.1 mM Na_2EDTA; pH 7.0) except when stated otherwise. T_m values [°C] are derived from maxima of the first derivative of the heating curve (software Kaleidagraph). [c] The majority of the curves show hysteresis.

tainty with regard to their constitution, each of the two isomeric oxo-amino members, (^{AP}ON) and (^{AP}NO), can exist as two NH tautomers (not considering phenolic tautomers), a dichotomy directly relevant to the question of both specificity and strength of the two bases' pairing behavior in oligomers. All our ^{1}H and ^{13}C NMR spectroscopic observations on these two bases themselves and on the corresponding 5-amino-acylated monomer derivatives point in each case to the presence of one single tautomer in dimethyl sulfoxide (DMSO) solution.[14]

Cross-pairing experiments listed in Table 1, entries 19–21 and illustrated in Figure 2 were designed to determine the relative accessibility of the respective tautomers of the two isomeric oxo-amino bases for base pairing in aqueous solution: The dodecamer AspGlu[(^{AP}OO)$_4$ (^{AP}ON)$_4$ (^{AP}OO)$_4$], containing in the center of the sequence four units bearing the ^{AP}ON nucleus (cytosine analogue), is found to pair unambiguously and strongly with the DNA sequence d(A$_4$^{iso}G$_4$A$_4$) (Table 1, entry 19 and Figure 2a,b), but not with the corresponding DNA sequence containing G instead of ^{iso}G (Table 1, entry 20 and Figure 2a). In contrast, the T_m curves observed for the isomeric dodecamer AspGlu-[(^{AP}OO)$_4$(^{AP}NO)$_4$(^{AP}OO)$_4$] containing the ^{AP}NO nucleus (isocytosine analogue) combined with d(A$_4$G$_4$A$_4$) (Table 1, entry 21) are ambiguous (see Figure 2a,c), yet nevertheless point to a stronger interaction of ^{AP}NO with G as compared with an ^{iso}G-containing partner sequence. We concluded that the pairing mode of ^{AP}ON with ^{iso}G is Watson–Crick (and not reverse Watson–Crick,[16] Scheme 4). As ^{iso}G, contrary to G, can engage in base pairing not only in one but two tautomeric forms,[17] it remains undecided as to whether ^{AP}ON engages in its cross-pairing with isoguanine the position 1 NH or its (presumably less favored[18]) position 3 NH tautomer (see the Supporting Information). Importantly, we observe clean intrasystem self-pairing in each of the two self-complementary sequences AspGlu [(^{AP}ON)(^{AP}NO)]$_6$ and AspGlu

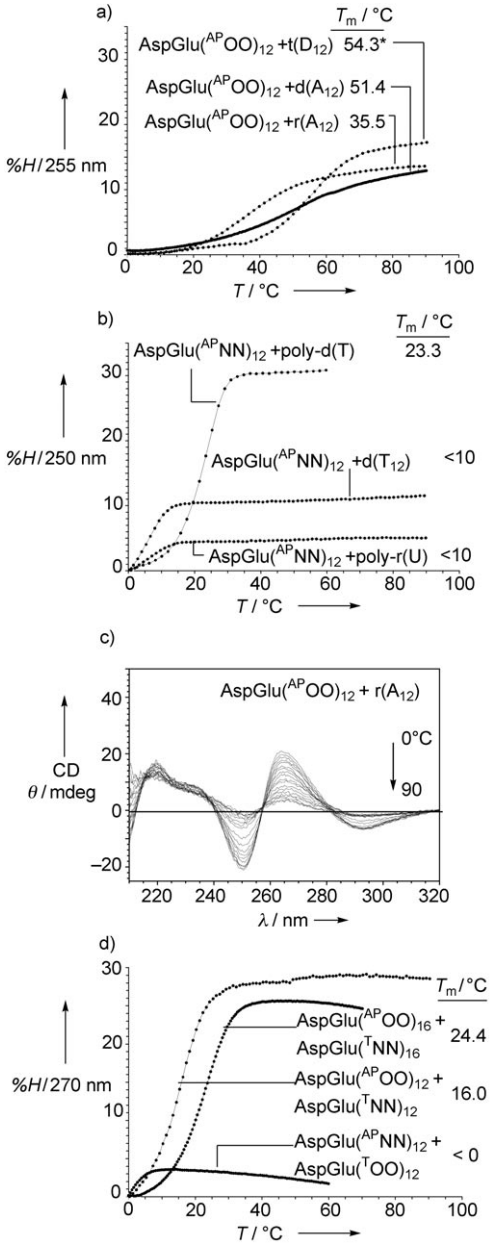


Figure 1. Experiments documenting inter- and intrasystem pairing. a) T_m(UV) curves of the duplex formed from strong cross-pairing of AspGlu(^{AP}OO)$_{12}$ with complementary DNA (d), RNA (r), and TNA (t) sequences; b) Weak cross-pairing of AspGlu(^{T}NN)$_{12}$ with complementary d and r sequences; c) Temperature-dependent CD spectrum of the duplexes formed by AspGlu(^{T}OO)$_{12}$ with RNA (corresponding CD spectrum with DNA in the Supporting Information); d) Temperature-dependent T_m(UV) curves of intrasystem pairing in oligodipeptide backbones (Table 1, entries 7, 13, and 18); For a Job plot of AspGlu-(^{AP}OO)$_{12}$ + AspGlu(^{AP}NN)$_{12}$ (Table 1, entry 17) showing the 1:1 ratio of the pairing partners in the homoduplex, see the Supporting Information. Measurements were made with strand concentration of 5 μM each (1:1) in NaCl (1 M), aqueous NaH_2PO_4 (10 mM), Na_2EDTA (0.1 mM); pH 7.0. No self-pairing was observed for individual partner stands; hysterisis was observed in some UV cooling curves. CD temperature increments in 5 °C steps; * at 275 nm. %*H* = percentage hyperchromicity.

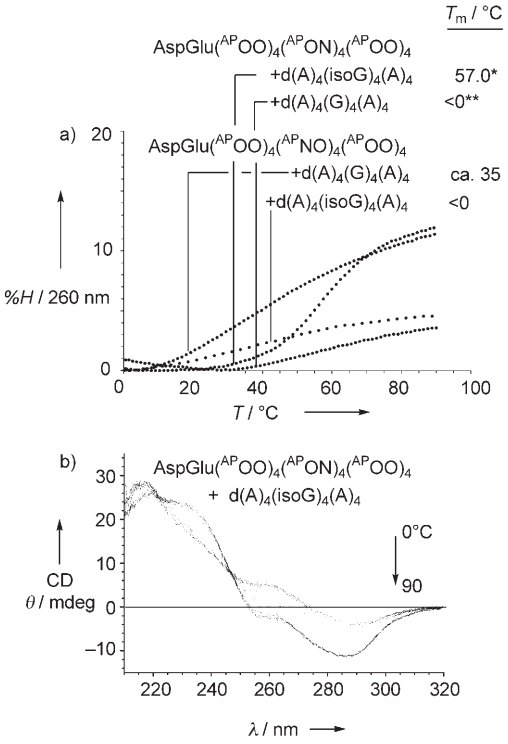
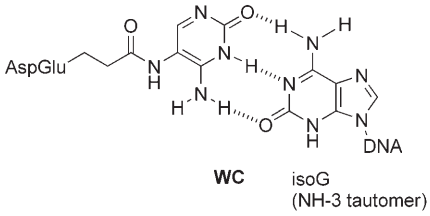


Figure 2. Intersystem pairing of oligodipeptide sequences containing the isomeric oxo-aminopyrimidine members. a) Comparison of the T_m(UV) curves of the duplexes formed from cross-pairing with corresponding complementary d sequences; b) Temperature-dependent CD spectra documenting the unambiguous pairing of (^{AP}ON) with isoG (corresponding CD spectra of ambiguous behavior of (^{AP}NO) with G is shown in the Supporting Information). Measurements were made with the oligodipeptide sequence and the complementary DNA (≈10 μM; 1:1) in 1 M NaCl, aqueous 10 mM NaH_2PO_4, 0.1 mM Na_2EDTA, pH 7.0. CD temperature increments in 5 °C steps. * after 2 days at 4 °C; ** at 250 nm.



Scheme 4. One possible mode of Watson–Crick (WC) base pairing that is available for the 2-oxo-4-amino member with isoguanine (all possible base-pairing combinations of the two isomeric oxo-amino members when pairing with guanine and isoguanine are depicted in the Supporting Information).

(^{AP}ON)$_6$ (^{AP}NO)$_6$ (Table 1, entries 23 and 24 and Figure 3), therefore one of the two isomeric oxo-aminopyrimidine bases must be able to engage with its position 3 NH tautomer in base pairing. The formulation for the (^{AP}ON)-(^{AP}NO) base pair depicted in Scheme 1 is the variant that we surmise to be the more probable of the two possible formulations. The apparent need of one of the two pyrimidine bases (probably ^{AP}NO) to adjust in this self-pairing to the pairing partner by tautomerization may well contribute to the huge hysteresis

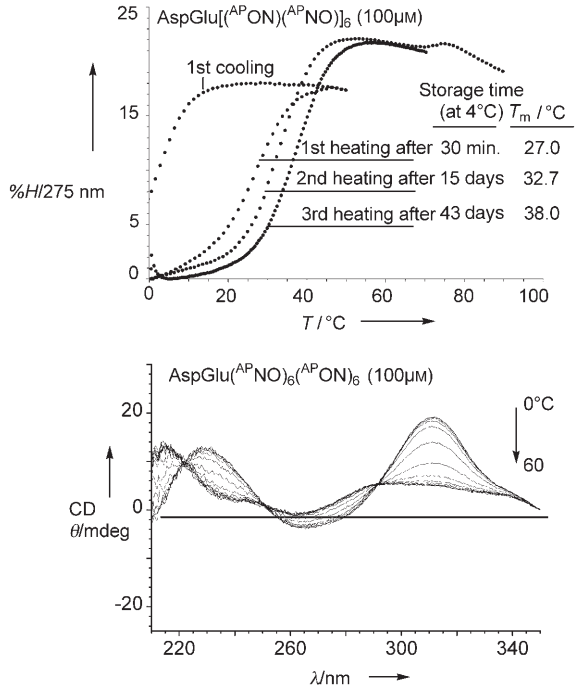


Figure 3. a) Storage-time dependence of the T_m(UV) of sequence containing the isomeric oxo-aminopyrimidine members (Table 1, entry 23), first-heating curves of the self-pairing duplexes ($c \approx 100$ μM); sample stored at 4 °C; b) Temperature-dependent CD spectrum of the self-pairing block sequence (Table 1, entry 26); T_m(CD) = 25 °C ($c \approx 100$ μM); CD temperature increments in 5° steps. Measurements were in 1 M NaCl, 10 mM aqueous NaH_2PO_4, 0.1 mM Na_2EDTA; pH 7.0.

observed for the cooling versus the heating curves in UV spectroscopic T_m determination of these two duplexes, as well as for the remarkable finding that the T_m values of these two duplexes is dependent on sample history: the T_m value derived from the heating curve in Table 1, entry 23 increases parallel to the storage time of the duplex at 4 °C (Figure 3). The isomeric block sequence (Table 1, entry 24) behaves similarly. The self-pairing of the latter demonstrates antiparallel strand orientation in the duplex.

The deviation in pairing behavior of the 2,4-diamino-5-acylaminopyrimidine base from what would be expected on the basis of the perfectly "normal" pairing of the corresponding 2,4-dioxo analogue, in conjunction with the strikingly reciprocal behavior of the 2,4-dioxo and 2,4-diamino members in the triazine family described in the preceding paper, may be considered to represent, besides its possible etiological relevance, the main chemical message of this and the preceding paper. As also shown in the previous communication, one of the two bases undergoes pairing with the complementary canonical base with normal strength, whereas the other does not. In the previous communication it was the dioxo member that deviated from the norm, in this case it is the diamino member. In this case, as previously, the most accessible chemical property that parallels this deviation in pairing behavior relative to the standard of the canonical nucleobases is the pK_a value: although in the previous case a 2,4-dioxotriazine (pK_a of 6-methyl-2,4-dioxo-1,3,5-triazine = 7.2[19]) is a distinctly stronger acid than a 2,4-dioxopyrimidine (pK_a of deoxyuridine = 9.3[20]), in the present case, a 5-acylamino-2,4-diaminopyrimidine (pK_a of the 5-formylamino derivative = 6.03[21]) is a stronger base than deoxyribofuranosyl-2,6-diaminopurine (pK_a = 4.4[22]) or, for that matter, the canonical deoxyadenosine (pK_a = 3.8[23]). Figure 4 juxtaposes the so-far-observed examples of normal versus deviating pairing behavior with the corresponding (known or estimated) pK_a values of pairing partners. The juxtaposition points to a correlation between ΔpK_a values of pairs of complementary bases and their pairing strength in the sense that the smaller the ΔpK_a value of a pair of complementary 2,4-dioxo and 2,4-diamino pairing partners (as compared with the standard difference of about 5 pK_a units for the canonical pair of bases), the weaker the pairing (in aqueous solution at neutral pH). The trend of this correlation points in the opposite direction to the one that is found in the literature concerning ΔpK_a values and relative strengths of hydrogen bonds in nonaqueous media.[24]

We would expect a rationalization of this difference to refer, besides considering the influence of factors such as nucleo- and electrophilicities of pairing centers, primarily to differences in nucleobase–water[25] interactions in single as compared with double strands. The need for collecting further facts demands an extension of our studies to experiments that will eliminate the remaining uncertainties regarding the influence of differences in backbone structure on the relative strength of base pairing and will, furthermore, provide the opportunity for directly comparing relative pairing strength in aqueous solutions with that in nonaqueous solvents.[26]

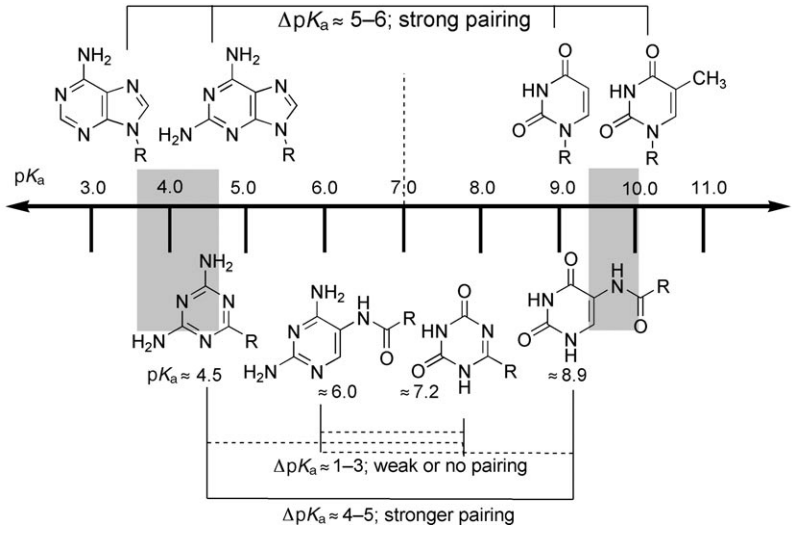


Figure 4. Correlation between base-pairing strength with ΔpK_a values of pairs of complementary bases in aqueous solution at neutral pH.

From an etiological point of view, the findings described in this and the preceding paper in conjunction with our previous

work on the etiology of nucleic acid structure can be interpreted to indicate that it may have been mainly the structure of the recognition elements and not so much the structure of the oligomer backbone that had been critical in nature's choice of the molecular basis of a genetic system. Although a variety of backbone alternatives of generational complexity and base-pairing capability similar to that of RNA would have been available for nature's choice, there seems to be a distinct scarcity of potential natural alternatives to the two pairs of Watson–Crick nucleobases. Our observations indicate that 2,4-disubstituted 1,3,5-triazines and 2,4-disubstituted-5-aminopyrimidines, two families of heterocycles deemed to be of generational simplicity comparable with that of the canonical nucleobases, yet offering chemically wider opportunities for backbone tagging, are clearly functionally inferior to the family of Watson–Crick bases by reasons that seem intrinsically chemical in nature. The findings provide a chemical illustration of the view that the canonical nucleobases represent a functional optimum with respect to informational base pairing in aqueous solution. Our observations, however, should not be interpreted as excluding the possibility that functionally less-than-optimal recognition elements, such as the 5-aminopyrimidines, may have played a role in the self-organization of organic matter.

Received: August 7, 2006
Published online: November 17, 2006

Keywords: 5-aminopyrimidines · base pairing · DNA · nucleic acid hybridization · oligonucleotides

[1] 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.

[2] a) J. Oro, A. Kimball, *Biochem. Biophys. Res. Commun.* **1960**, *2*, 407; b) J. Oro, A. Kimball, *Arch. Biochem. Biophys.* **1962**, *96*, 293; c) R. A. Sanchez, J. P. Ferris, L. E. Orgel, *J. Mol. Biol.* **1967**, *30*, 223; d) M. Levy, S. L. Miller, J. Oro, *J. Mol. Evol.* **1999**, *49*, 165; e) S. L. Miller, L. E. Orgel, *The Origins of Life on the Earth*, Prentice-Hall, Englewood Cliffs, NJ, **1974**.

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

[4] a) J. P. Ferris, P. C. Joshi, J. G. Lawless, *Biosystems* **1977**, *9*, 81; b) J. P. Ferris, P. C. Joshi, E. H. Edelson, J. G. Lawless, *J. Mol. Evol.* **1978**, *11*, 293.

[5] For alternative (potentially prebiotic) formation of canonical pyrimidine, see, for example: a) J. P. Ferris, R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1968**, *33*, 693; b) M. P. Robertson, S. L. Miller, *Nature* **1995**, *375*, 772.

[6] For 5-aminocytosine, see: a) G. Andresen, L.-L. Gundersen, M. Lundmark, F. Rise, S. Sundell, *Tetrahedron* **1995**, *51*, 3655; b) A. Holy, *Collect. Czech. Chem. Commun.* **1979**, *44*, 2846; c) C. O. Johns, *Am. Chem. J.* **1911**, *45*, 79. For 5-aminoisocytosine, see: d) T. B. Johnson, C. O. Johns, *Am. Chem. J.* **1905**, *34*, 554; e) J. L. Kelly, E. W. Mclean, *J. Heterocycl. Chem.* **1981**, *18*, 671. For 2,4,5-triaminopyrimidine, see f) D. J. Brown, *J. Appl. Chem.* **1957**, *7*, 109. 5-Aminouracil is commercially available.

[7] a) J.-W. Chern, D. S. Wise, W. Butler, L. B. Townsend, *J. Org. Chem.* **1988**, *53*, 5622; b) M. A. Sofan, A. E.-S. Abdel-Megied, M. B. Pedersen, E. B. Pedersen, C. Nielsen, *Synthesis* **1994**, 516; c) E. Ferrer, M. Wiersma, B. Kazimierczak, C. W. Müller, R. Eritja, *Bioconjugate Chem.* **1997**, *8*, 757; d) V. S. Rana, K. N. Ganesh, *Nucleic Acids Res.* **2000**, *28*, 1162; e) M. J. Storek, A. Suciu, G. L. Verdine, *Org. Lett.* **2002**, *4*, 3867.

[8] For an early constitutional analysis of alternative nucleobase pairs, see: a) J. A. Piccirilli, T. Krauch, S. E. Moroney, S. E. Benner, *Nature* **1990**, *343*, 33; b) S. A. Benner, *Acc. Chem. Res.* **2004**, *37*, 784.

[9] a) T. Otzen, E. G. Wempe, B. Kunz, R. Bartels, G. L. -Yvetot, W. Hänsel, K.-J. Schaper, J. K. Seydel, *J. Med. Chem.* **2004**, *47*, 240; b) C. Bamford, K. Al-Lamee, *J. Chem. Soc. Chem. Commun.* **1993**, 1580; c) D. J. Brown, K. Mori, *Aust. J. Chem.* **1985**, *38*, 467. An amino group in position 5 of a pyrimidine is, together with the NH group in position 1, part of a vinylogous hydrazine system, whereas amino groups at position 2 or 4 are part of a guanidine or an amidine system, respectively.

[10] a) P. Wipf, A. Cunningham, R. L. Rice, J. S. Lazo, *Bioorg. Med. Chem.* **1997**, *5*, 165. b) The procedure used in the preparation for the D enantiomer was used: K. L. Webster, A. B. Maude, M. E. O'Donnell, A. P. Mehorotra, D. Gani, *J. Chem. Soc. Perkin Trans. 1* **2001**, 1673.

[11] S. Nakabayashi, C. D. Warren, R. W. Jeanloz, *Carbohydr. Res.* **1988**, *174*, 279.

[12] All compounds were characterized by ^{1}H and ^{13}C NMR and mass spectral data following purification by column chromatography on silica gel. Oligodipeptides were synthesized on an Expedite 8909 Nucleic Acid Synthesizer (Perseptive Biosystems) by using a modified PNA protocol, purified by HPLC (ion exchange) to a minimal purity of 95%, and checked by MALDI-TOF-MS (see the Supporting Information). RNA and DNA oligonucleotides were purchased from commercial sources; TNA sequences were available from our previous work.[13]

[13] K.-U. Schoening, P. Scholtz, X. Wu, S. Guntha, G. Delgado, R. Krishnamurthy, A. Eschenmoser, *Helv. Chim. Acta* **2003**, *86*, 1259.

[14] The similarity between the UV spectra of cytosine (λ_{max} = 267 nm, ε = 6100)[15a] and 1-methyl-cytosine (λ_{max} = 273 nm, ε = 8100)[15a] or 1-ribofuranosylcytosine (λ_{max} = 270 nm, ε = 8800)[15b] closely corresponds to the similarity between the UV spectra of the 5-aminocytosine (λ_{max} = 292 nm, ε = 3800)[15c] and of its 1-ribofuranosyl derivative (λ_{max} = 298 nm, ε = 6200)[15b] in buffer solution pH 7. This indicates that the 5-aminocytosine nucleus prefers the position 1 NH tautomer in aqueous solution. In the 5-aminoisocytosine series we observe UV spectra of quite different structure for the free base **2** (λ_{max} = 287 nm, ε = 4100; 240 nm (shoulder), ε = 6400; ε_{220nm} = 10,500) and its N3 methyl derivative (λ_{max} = 308 and 242 nm, ε = 7300 and 6800; ε_{220nm} = 3500; both spectra in aqueous phosphate buffer solution, pH 7). The absorption maxima of the UV spectrum of the 5-formyl amino derivative of 3-methyl-isocytosine (for its preparation and for the X-ray structure analysis of the corresponding *N,N*-dimethylformamidine derivative, see the Supporting Information) are hypsochromically shifted (λ_{max} = 295 and 235 nm, ε = 8800 and 7000; ε_{220nm} = 5200) relative to the maxima of the 3-methylisocytosine derivative. The spectrum is, however, of the same type as for the free N-3-methylated base. These findings point to the conclusion that 5-aminoisocytosine disfavors its NH(3)-tautomer in aqueous solution. The conclusion remains tentative as the necessary UV comparison with the 1-methylisocytosine derivative is lacking.

[15] a) A. R. Katritzky, A. J. Warring, *J. Chem. Soc.* **1963**, 3046; b) J. Fox, D. V. Praag, *J. Org. Chem.* **1961**, *26*, 526; c) S. F. Mason, *J. Chem. Soc.* **1954**, 2071.

[16] See, for example, 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.

[17] a) G. Groebke, J. Hunziker, W. Fraser, Ling Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1998**, *81*, 375; b) C. Roberts, R. Bandaru, C. Switzer, *J. Am. Chem. Soc.* **1997**, *119*, 4640. In both the homo-DNA and the pyranosyl-RNA series, it has been shown that isoguanine can form strong Watson–Crick pairs with guanine. This demonstrates the capability of the isoguanine nucleus to act in base pairing as the position 3 NH tautomer (see references [16] and [17a]). It should be noted, however, that the generation of such a guanine–isoguanine pair from guanine and the isoguanine position 3 NH tautomer might well be indistinguishable from that of an identical pair from guanine and the phenolic isoguanine tautomer (see the discussion on pages 438–444 in reference [17a]).

[18] We surmise that, in this guanine–cytosine-like pairing interaction, it is the 2-amino-4-oxo member that takes the role of the guanine-analogue; the argument being that, if the 2-oxo-4-amino member would fulfill that role, it would have to do so as its position 3 NH tautomer, in which the location of two amino functions at a common C–C double bond should be electronically unfavorable (local four-center, six-electron system).

[19] J. Jonas, J. Gut, *Collect. Czech. Chem. Commun.* **1962**, *27*, 716.

[20] K. Nakanishi, N. Suzuki, F. Yamakazi, *Bull. Chem. Soc. Jpn.* **1961**, *34*, 53.

[21] B. Roth, J. Z. Strelitz, *J. Org. Chem.* **1969**, *34*, 821.

[22] Spectroscopically determined from the pH-dependent UV spectrum. For details, see the Supporting Information of reference [1].

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

[24] For a review of factors that determine base-pairing strength in nonaqueous medium, see: S. C. Zimmermann, P. S. Corbin, *Struct. Bonding (Berlin)* **2000**, *96*, 63. For a correlation between ΔpK_a values and the formation of short hydrogen bonds in the solid state, see: T. Steiner, I. Majerz, C. C. Wilson, *Angew. Chem.* **2001**, *113*, 2728; *Angew. Chem. Int. Ed.* **2001**, *40*, 2651, footnote 14 as well as references therein. The correlation deduced from observations in nonaqueous media that indicates that hydrogen bonds are stronger the smaller the ΔpK_a value between donor and acceptor ("pK_a match") has recently been proposed to explain the fact that, as a rule, RNA duplexes have higher melting temperatures than corresponding DNA duplexes; see: P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya, J. Chattopadhyaya, *J. Am. Chem. Soc.* **2004**, *126*, 2862, and references therein.

[25] J. D. Dunitz, *Science* **1994**, *264*, 670.

[26] Studies along these lines are being carried out in collaboration with J. Rebek (TSRI). See also in this context: K. S. Jeong, T. Tjivikua, J. Rebek, Jr., *J. Am. Chem. Soc.* **1990**, *112*, 3215.

