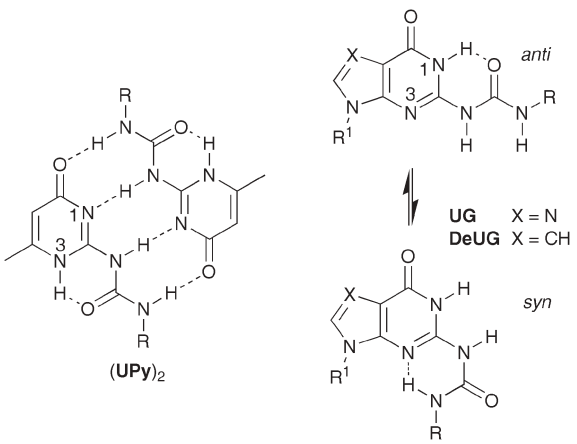
Self-complementary purines by quadruple hydrogen bonding†

Alisha M. Martin, Roslyn S. Butler, Ion Ghiviriga, Rachel E. Giessert, Khalil A. Abboud and Ronald K. Castellano*

Received (in Austin, TX, USA) 21st July 2006, Accepted 16th August 2006
First published as an Advance Article on the web 7th September 2006
DOI: 10.1039/b610239e

The first discrete, self-complementary, quadruply hydrogen-bonded complexes based on the 2,6-diaminopurine (DAP) scaffold have been prepared; regioselective urea formation at the C(2) amino group of the heterocycle allows intermolecular dimerization (K_{dim} ~ 1–1.6 × 10^3 M^{-1} in $CDCl_3$) through a DADA hydrogen bonding motif.

Nucleobases are readily available heterocycles for the construction of self-complementary, quadruply hydrogen-bonded complexes, among which the pyrimidines have been extensively explored.[1,2] For example, the UPy unit of Meijer *et al.* (derived from isocytosine) has found applications that span materials science due to its exceedingly high dimerization constant in organic solution (K_{dim} ~ 10^7 M^{-1} in $CDCl_3$) *via* an accessible DDAA[3] hydrogen bonding arrangement.[1,4] Surprisingly little work has considered urea-functionalized purines in this vein, although the bicycles boast additional sites for functionalization, an expanded π-surface[5] and untapped mechanisms to control assembly strength/dynamics *via* the modulation of ring electronics remote from the dimer hydrogen bonding interface.



To this end, Zimmerman and co-workers have only recently shown that urea-functionalized guanine (UG) and 7-deazaguanine (DeUG), unlike UPy, are tautomerically stable (preferring the N(1)–H tautomer), and that the *anti*/*syn* conformational equilibrium can be controlled by atomic mutation in the fused ring (UG prefers *anti*; DeUG prefers *syn*).[6,7] With a preorganized ADDA edge and free from competing tautomeric/conformational equilibria, DeUG is particularly well-suited to forming tight heterodimeric (complementary) complexes (K_{assoc} > 10^7 M^{-1}) with DAAD partners.[6] Likewise, both UG and DeUG are designed to only weakly self-associate and do so by various DA motifs in $CDCl_3$ (K_{assoc} for UG ~ 230 M^{-1}; K_{dim} for DeUG = 880 M^{-1}).[6] Herein, we present the first discrete, self-complementary, quadruply hydrogen-bonded complexes based on purines using the tautomerically-stable 2,6-diaminopurine (DAP) platform (Fig. 1). Upon regioselective urea formation at the C(2) amino group of DAP, two intramolecularly hydrogen-bonded, low-energy (*anti*) conformers of the ureidodiaminopurine (UDAP),[8] $\mathbf{1}^{N1}$ and $\mathbf{1}^{N3}$ (interconverted through a single C–N bond rotation), are accessible,[9] where the latter is preorganized for intermolecular dimerization *via* a DADA hydrogen bonding motif.

Synthetic considerations and X-ray crystallography guided the choice of substituents shown in Fig. 1; the synthesis of **1a** illustrates the design and execution (Scheme 1). The route begins from commercially available (or routinely prepared[10]) 6-chloro-2-aminopurine (**2**), derived from guanine. Substituted benzyl substituents were selected for the N(9) position, primarily due to their known convenient installation by N-alkylation[11] and decent organic solubility-imparting properties. Standard alkylation affords **3a** at room temperature, its yield somewhat diminished by the unavoidable formation of the N(7) regioisomer.[12] Subsequent displacement of the 6-chloro group with ammonia

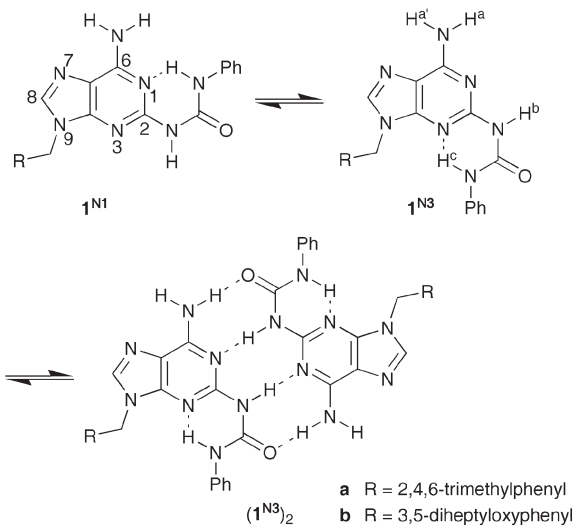
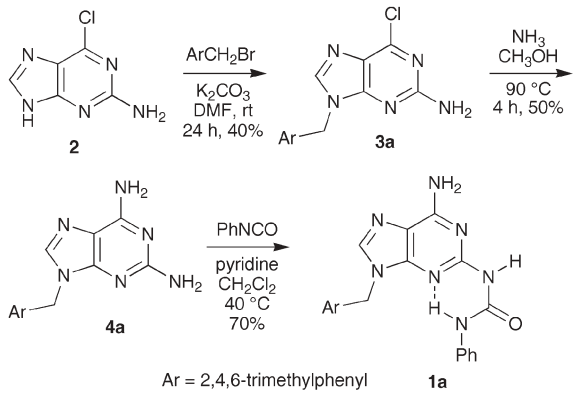


Fig. 1 Intramolecular hydrogen bonding and intermolecular dimerization of ureidodiaminopurine (UDAP) **1**.

Department of Chemistry, University of Florida, P.O. Box 117200, Gainesville, FL 32611, USA. E-mail: castellano@chem.ufl.edu; Fax: +1 352 846 0296; Tel: +1 352 392 2752

† Electronic supplementary information (ESI) available: Synthesis, characterization, copies of ^{1}H NMR spectra for all new compounds, additional NMR data for **1a** (gHMBC, NOESY and VT NMR), and dimerization data for **1a** and **1b**. See DOI: 10.1039/b610239e



Scheme 1 Representative UDAP synthesis; preparation of **1a**.

provides **4a**. We next found that while most isocyanates react unusually sluggishly with **4a**,[13] aryl isocyanates appear to enjoy a significant reactivity advantage. Hence, treatment of **4a** with phenylisocyanate in the presence of pyridine exploits the differential nucleophilicity of its C(2) and C(6) amino groups,[14] forming **1a** regioselectively. Functionalized aryl isocyanates are equally effective in this reaction. Compound **1b**, bearing alkoxy substituents, was prepared similarly (ESI†) and offers a modest increase in organic solubility.

Of additional consideration at the design stage was the steric compatibility of the N(9) benzyl and urea phenyl substituents in the desired $\mathbf{1}^{N3}$ conformation.[15] We were pleased to find that the X-ray structure (Fig. 2) of model compound **5** (ESI†) indeed shows that the desired mode of intramolecular hydrogen bonding to N(3) is accessible through a planar arrangement (N(3)···N(12) = 2.74 Å).‡ Also observed in the solid state is a near edge-to-face relationship between the two aromatic substituents (angle between the least-squares planes of the aromatic rings = 86.3°); while the rings are slightly offset with respect to one another in this orientation (center-to-center distance = 5.42 Å), they are in close contact (closest carbon–carbon distance = 3.67 Å).[16] Although this type of interaction is expected to be relatively weak in organic solution,[16] it does inspire strategies to achieve conformational control in the UDAP monomers. Consistent with the solid state data for **5**, the chemical shift of its urea H^c proton in $CDCl_3$ (~5 mM) is significantly deshielded to δ 11.4 (relative to TMS); H^b appears at $\sim\delta$ 7.2, a value that correlates with the limiting upfield chemical shift (also for H^b) of monomeric **1** ($\delta_{monomer}$), as discussed below.

The solution-phase dimerization of **1a** and **1b** could be analyzed through routine ^{1}H NMR measurements in $CDCl_3$; representative data for **1a** are shown in Fig. 3. A 2-D NMR (gHMBC) experiment in $CDCl_3$ at ~2 mM was first used to assign the urea proton chemical shifts (ESI†). The peak at δ ~12 arises from intramolecularly hydrogen-bonded H^c, while the H^b signal, deshielded due to intermolecular hydrogen bonding, appears at δ ~9. Dilutions were performed at 25 °C from ~5 to ~0.1 mM; five selected spectra are shown in Fig. 3. The H^b resonance was monitored as it moved upfield throughout the series[17] (the H^a signal could not be followed due to peak broadness and overlap) and the data was fitted to a non-linear binding equation using standard software.[18] The dimerization constant (K_{dim}) emerges as 1100 ± 360 M^{-1} for **1a** (1200 ± 200 M^{-1} for the run shown) and 1600 ± 380 M^{-1} for **1b**, similar values that are the average of four independent runs in each case. Furthermore, the downfield and upfield limiting chemical shifts for H^b of **1** appear from the calculations (δ_{dimer} = 10.1 ± 0.1; $\delta_{monomer}$ = 7.2 ± 0.1), and are consistent with variable temperature (VT) NMR experiments (*vide infra*) and model compound **5** (*vide supra*). Finally, while dimer formation necessarily requires that **1a** (and **1b**) adopts the $\mathbf{1}^{N3}$ conformation, direct evidence that this conformation is populated in solution comes through a NOESY spectrum. Key NOEs are identified between the N(9) trimethylbenzyl substituent and both the urea phenyl and H^c protons (ESI†).

Importantly, VT ^{1}H NMR studies performed with **1a** in $CDCl_3$ confirm that dimerization by *quadruple* hydrogen bonding is the predominant mode of assembly in the concentration range studied (ESI†). Upon cooling a ~2 mM solution from 55 to −55 °C, the

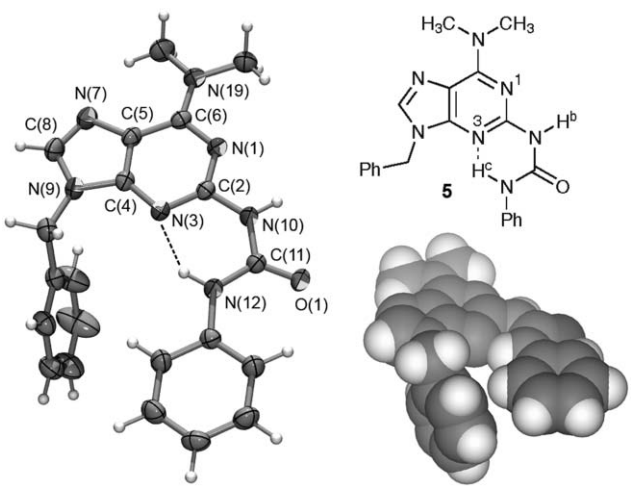


Fig. 2 Left: Intramolecular hydrogen bonding of model ureidopurine **5** in the solid state (ellipsoids drawn at the 50% probability level). Right: The side view of a CPK representation shows the extent of contact achievable for the two phenyl rings in the N(3) hydrogen-bonded conformer.

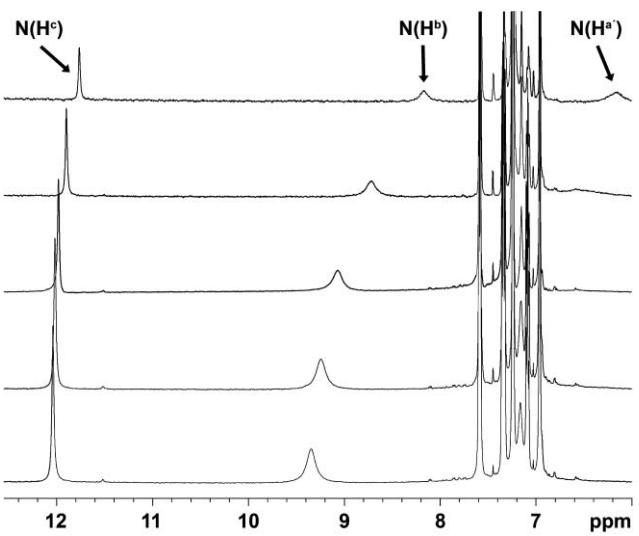


Fig. 3 ^{1}H NMR spectra (500 MHz) of **1a** in $CDCl_3$ (25 °C) at the following concentrations (from bottom to top): 4.7, 2.9, 1.5, 0.59 and 0.16 mM. Data at five additional concentrations was used in the calculation of K_{dim} (ESI†). See Fig. 1 for the atom labelling scheme.

chemical shift of H^c moves downfield from δ 11.8 to δ 12.4. H^b, on the other hand, shifts even more substantially from δ 8.1 to δ 10.2 ($\sim\delta_{dimer}$ for H^b calculated from dilution studies), in accordance with its intermolecular hydrogen bonding. Also significantly, the amino protons N($H^{a/a'}$) decoalesce at $\sim$5 °C. The chemical shift of proton $H^{a'}$ appears at δ $\sim$5.9 and remains there as the temperature is lowered,[19] while H^a, participatory in dimer formation, moves downfield from δ 8.4 (5 °C) to δ 9.3 (−55 °C).

Given that the UDAP derivatives dimerize somewhat more weakly (by 10-fold) than the most comparable DADA (self-complementary) quadruple hydrogen bonding system (the alkyl ureidotriazines of Meijer and co-workers[9,20]), there is additional optimization to do and subtleties to be understood. We would expect the K_{dim} of **1** to be affected by (a) the interaction of the urea and N(9) substituents and (b) the use of aryl rather than the more commonly employed alkyl ureas.[9,21] We initiated explorations of "(a)" by preparing the N(7)-alkylated regioisomer of **1a** (ESI†). We were surprised to find that this regioisomer does not dimerize by quadruple hydrogen bonding in $CDCl_3$;[22] hence, interaction between the N(9) and urea substituents appears to be an important (and modifiable) parameter in these systems. Along the same lines, we studied the intrinsic $\mathbf{1}^{N1}$ *vs.* $\mathbf{1}^{N3}$ conformational preference by computation. When a substituent in the N(9) position is too small to interact appreciably with the phenylurea group, such as methyl (**1**, where R = H), computation (MP2/6–31G*//HF/6–31G*) shows that the $\mathbf{1}^{N1}$ and $\mathbf{1}^{N3}$ conformers are essentially isoenergetic, $\sim$0.55 kcal mol^{-1} in favour of the desired $\mathbf{1}^{N3}$ conformer in the gas phase. This leaves the monomer conformational equilibrium, and also likely K_{dim}, easily perturbed. We are currently developing the synthetic chemistry to test "(b)".

To conclude, the first self-complementary, quadruply hydrogen bonding purines have been prepared from the readily-available 2,6-diaminopurine scaffold. The monomers are routinely synthesized and should offer unique handles (and bioinspired strategies) for the control of association strength through their multiple substitution sites, some remote from the hydrogen bonding interface. Their expanded π-surfaces will facilitate their association into stacked assemblies and more complex architectures. Explorations in these directions are currently under way.

We are grateful to the University of Florida and the Research Corporation (Research Innovation Award (RI-1198) to R. K. C.) for financial support. R. S. B. and R. E. G. were funded through University of Florida Alumni Fellowships. K. A. A. wishes to acknowledge the National Science Foundation and the University of Florida for funding the X-ray equipment. We are grateful to Prof. Blake R. Peterson (Penn State) for providing a copy of *Associate 1.6* for the calculation of dimerization constants and Prof. Adrian Roitberg (UF) for computational time.

Notes and references

‡ Crystal data for **5**: $C_{21}H_{21}N_7O$ (M = 387.45), monoclinic, space group $P21/n$, radiation type = Mo-Kα, λ = 0.71073 Å, a = 9.0629(6), b = 19.8629(13), c = 11.3182(7) Å, α = γ = 90, β = 106.862 (1)°, V = 1949.9(2) $Å^3$, Z = 4, μ = 0.087 mm^{-1}, D_c = 1.320 g cm^{-3}, $F(000)$ = 816, T = 173(2) K, 4406 independent reflections (R_{int} = 0.0337), final R indices (272 parameters) [$I > 2\sigma(I)$] were R_1 = 0.0374, wR_2 = 0.0966 (using 3194 reflections), GOF = 1.063. Refinement was done using F^2. CCDC 615504. For crystallographic data in CIF or other electronic format see DOI: 10.1039/b610239e

1 R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. Hirschberg, R. F. M. Lange, J. K. L. Lowe and E. W. Meijer, *Science*, 1997, **278**, 1601–1604; F. H. Beijer, R. P. Sijbesma, H. Kooijman, A. L. Spek and E. W. Meijer, *J. Am. Chem. Soc.*, 1998, **120**, 6761–6769; F. H. Beijer, H. Kooijman, A. L. Spek, R. P. Sijbesma and E. W. Meijer, *Angew. Chem., Int. Ed.*, 1998, **37**, 75–78; S. H. M. Söntjens, R. P. Sijbesma, M. H. P. van Genderen and E. W. Meijer, *J. Am. Chem. Soc.*, 2000, **122**, 7487–7493.
2 V. G. H. Lafitte, A. E. Aliev, P. N. Horton, M. B. Hursthouse, K. Bala, P. Golding and H. C. Hailes, *J. Am. Chem. Soc.*, 2006, **128**, 6544–6545.
3 This nomenclature is used throughout the paper: D = hydrogen bond donor; A = hydrogen bond acceptor.
4 L. Brunsveld, B. J. B. Folmer, E. W. Meijer and R. P. Sijbesma, *Chem. Rev.*, 2001, **101**, 4071–4097; A. W. Bosman, L. Brunsveld, B. J. B. Folmer, R. P. Sijbesma and E. W. Meijer, *Macromol. Symp.*, 2003, **201**, 143–154; R. P. Sijbesma and E. W. Meijer, *Chem. Commun.*, 2003, 5–16 and references therein.
5 D. W. Guo, R. P. Sijbesma and H. Zuilhof, *Org. Lett.*, 2004, **6**, 3667–3670.
6 T. Park, S. C. Zimmerman and S. Nakashima, *J. Am. Chem. Soc.*, 2005, **127**, 6520–6521; T. Park, E. M. Todd, S. Nakashima and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2005, **127**, 18133–18142; H. C. Ong and S. C. Zimmerman, *Org. Lett.*, 2006, **8**, 1589–1592; T. Park and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2006, **128**, 11582–11590.
7 For multimeric hydrogen-bonded assemblies of guanines (*e.g.* G-quartets) see: J. T. Davis, *Angew. Chem., Int. Ed.*, 2004, **43**, 668–698 and references therein.
8 Not shown are the two lowest energy *syn*-conformers (non-hydrogen-bonded), calculated at the molecular mechanics level (Amber* and MM3* in $CHCl_3$) to be $\sim$5 kcal mol^{-1} higher in energy (for R = H in Fig. 1).
9 These features have been discussed for various urea-functionalized heterocycles. See ref. 1 and F. H. Beijer, R. P. Sijbesma, J. Vekemans, E. W. Meijer, H. Kooijman and A. L. Spek, *J. Org. Chem.*, 1996, **61**, 6371–6380.
10 *Eur. Pat.*, EP 543095, 1993.
11 E. Camaioni, S. Costanzi, S. Vittori, R. Volpini, K. N. Klotz and G. Cristalli, *Bioorg. Med. Chem.*, 1998, **6**, 523–533.
12 The regioisomeric mixture depends to some extent on the size of the electrophile, as the equivalent alkylation with 3,5-diheptyloxybenzyl bromide yields the desired N(9) isomer in over 80% yield (ESI†).
13 Poor reactivity has also been observed with guanine derivatives (see ref. 6). In our experience, the nucleophilicity of the amino group on C(2) is very sensitive to the C(6) substituent. With a deactivating –Cl group in this position (*i.e.* **3a**), even reactivity with aryl isocyanates is poor. Strongly activating groups, such as $N(CH_3)_2$ (preparation of **5**), see conversion to the urea in minutes at room temperature.
14 S. Porcher and S. Pitsch, *Helv. Chim. Acta*, 2005, **88**, 2683–2704.
15 Ribose groups may be accommodated but have yet to be completely explored due to their chemical lability during our initial studies (for similar observations, see ref. 6): A. M. Martin and R. E. Giessert, unpublished work.
16 E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem., Int. Ed.*, 2003, **42**, 1210–1250.
17 The upfield shift of H^c with dilution demonstrates weakening, through induction, of the intramolecular hydrogen bond upon dimer decomplexation (see ref. 6 for related observations).
18 B. R. Peterson, Associate 1.6, *PhD thesis*, University of California, Los Angeles, 1994.
19 This demonstrates that the Hoogsteen edge of the molecule is not engaged in measurable hydrogen bonding at these concentrations and temperatures (see ref. 6).
20 Comparisons between multiply (>3) hydrogen-bonded systems are not straightforward. For a discussion, see: O. Lukin and J. Leszczynski, *J. Phys. Chem. A*, 2002, **106**, 6775–6782.
21 Although the arylurea H^c proton is more acidic (the pK_a (DMSO) for urea is 26.7 and for diphenylurea is 19.6; see: F. G. Bordwell, *Acc. Chem. Res.*, 1988, **21**, 456–463), its carbonyl is a correspondingly weaker base.
22 While H^c is intramolecularly hydrogen-bonded (δ 11.5), H^b is only modestly deshielded (δ 7.8). Weaker DA-type hydrogen bonding may be operative due to a preferred $\mathbf{1}^{N1}$ conformation and/or distortion at the C(6) amino group.