Rebek Imides and Their Adenine Complexes: Preferences for Hoogsteen Binding in the Solid State and in Solution

Ronald K. Castellano,[a] Volker Gramlich,[b] and François Diederich*[a]

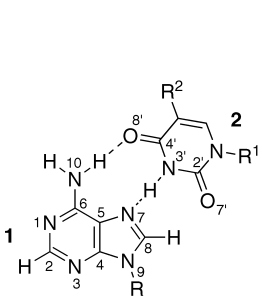
In memory of Professor Donald J. Cram

Abstract: Rebek imides (**3**), formed from Kemp's triacid, were developed in the mid-1980's as model receptors for adenine derivatives. We report here the first account of their hydrogen-bonding preferences upon binding 9-ethyladenine (**1a**) in the solid state. Structural analysis begins with simple imides **3a–e** that form discrete dimers, while bisimide **4** forms ribbon-like structures in the crystalline phase. The hydrogen-bonding interface within each of the representative assemblies features short intermolecular $N(3)_{imide} \cdots O(8^*)_{imide^*}$ distances (ca. 2.95 Å), indicative of two-point hydrogen bonding. Imides **3f–h** could be co-crystallized with **1a**; single-crystal X-ray analysis of the resulting complexes reveals hydrogen-bonding geometries nearly identical to those observed in nucleobase complexes of adenine and pyrimidine derivatives. Imides **3f** and **3g** form 2:1 ternary assemblies with **1a**; the complex of the former, $(\mathbf{3f})_2 \cdot \mathbf{1a}$, displays both Watson–Crick- and Hoogsteen-type hydrogen bonding, whereas the complex of the latter, $(\mathbf{3g})_2 \cdot \mathbf{1a}$, shows the Hoogsteen motif *and* imide hydrogen bonding to N(3) of the purine base ($N(3)_{adenine} \cdots N(3'')_{imide} = 3.07(1)$ Å). Imide **3h** forms a 1:1 complex with **1a** ($\mathbf{3h} \cdot \mathbf{1a} \cdot CHCl_3$) and displays Hoogsteen binding exclusively. All of the $\mathbf{3} \cdot \mathbf{1a}$ assemblies show $C_{adenine} \cdots O_{imide}$ distances (3.38–3.75 Å) that are consistent with C-H···O hydrogen bonding. Base-pairing preferences for the Rebek imides are further explored in solution by ^{1}H NMR one-dimensional NOE experiments and by computational means; in all cases the Hoogsteen motif is modestly favored relative to its Watson–Crick counterpart.

Keywords: Hoogsteen base pairing · hydrogen bonding · molecular recognition · receptors · Watson–Crick base pairing

Introduction

Structural elucidation of double-stranded DNA through X-ray analysis by Watson and Crick[1] provided the impetus for exploring simplified nucleobase complexes in the solid state throughout the 1950's and 60's, with particular emphasis placed at the time on rationalizing base-pairing structure. Early contributions by Hoogsteen,[2] Rich,[3] Sobell,[4] Frey,[5] and their co-workers explored the subtle structural differences between the complexes of substituted adenines (A) with thymine (T) and uracil (U) derivatives. Through this research was elucidated a nearly exclusive preference for Hoogsteen binding within 1:1 base-pair complexes between 9-alkyladenines (**1a**, **1b**) and 1-methylthymine (**2a**) or 1-methyluracil (**2b-d**) in the solid state (Figure 1), independent of lattice structure or crystallization conditions (vide infra).[6]



		Hoogsteen (**HG**) binding with:
2a	$R^1 = R^2 = CH_3$	**1b**[a]
2b	$R^1 = CH_3$, $R^2 = H$	**1a**[b]
2c	$R^1 = CH_3$, $R^2 = F$	**1a**[c]
2d	$R^1 = CH_3$, $R^2 = Br$	**1b**[d]
2e	$R^1 = CH_3$, $R^2 = I$	
1a	$R = CH_2CH_3$	
1b	$R = CH_3$	

Figure 1. 1:1 Complexes between 9-alkyladenines (**1a**, **1b**) and thymine (**2a**) or uracil (**2b–d**) derivatives display Hoogsteen (**HG**) binding in the solid state. [a] Refs. [2, 5]. [b] Ref. [3a]. [c] Ref. [3d]. [d] Refs. [3b,c]. The atom numbering scheme shown is used throughout the paper.

Cyclic imide derivatives **3** of Kemp's triacid[7] were introduced by Rebek and co-workers in the mid-1980's as model receptors for adenine derivatives.[8] Variants of **3** reversibly bind 9-ethyladenine (**1a**) through two-point hydrogen bonding (Figure 2).[9] Rapid interconversion on the NMR timescale is observed in solution between free **3** and bound **3** ($\mathbf{3} \cdot \mathbf{1a}$) wherein *1:1 complexation* occurs through either the Hoogsteen (**HG**) or Watson–Crick (**WC**) faces, the hydrogen-bonding motifs introduced and defined for nucleotide

[a] Prof. Dr. F. Diederich, Dr. R. K. Castellano
Laboratorium für Organische Chemie, ETH-Zentrum
Universitätstrasse 16, 8092 Zürich (Switzerland)
Fax: (+41) 1-632-1109
E-mail: diederich@org.chem.ethz.ch

[b] Prof. Dr. V. Gramlich
Laboratorium für Kristallographie, ETH-Zentrum
Sonneggstrasse 5, 8092 Zürich (Switzerland)

© WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002
0947-6539/02/0801-0118 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 1

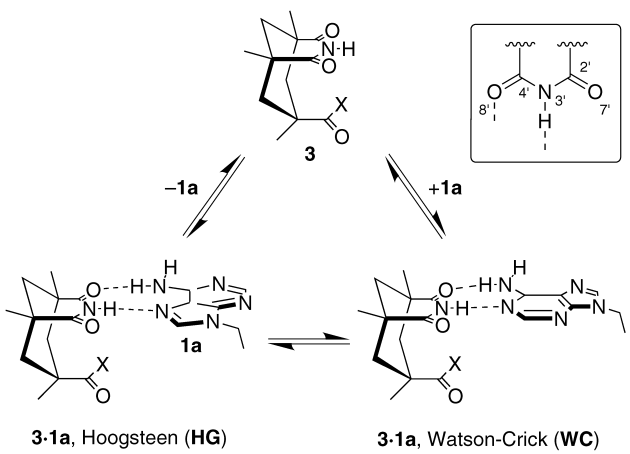
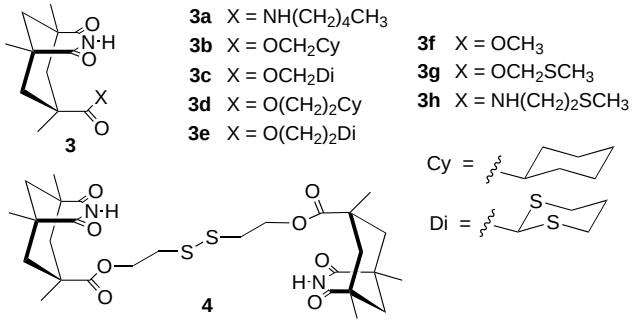


Figure 2. Rebek imides **3** can reversibly bind adenine derivatives (e.g. 9-ethyladenine, **1a**) from either the Hoogsteen (**HG**) or Watson–Crick (**WC**) face. Inset: Atom numbering scheme for **3** in **3·1a** complexes.

base pairing. NMR experiments have probed the independent contributions of hydrogen bonding and aromatic π stacking (Figure 2, in which X = phenyl, naphthyl, and anthryl derivatives) to adenine complexation and addressed the influence of π stacking on **HG** and **WC** preferences.[8, 9] Meanwhile, reduction of the cyclic imide in **3** to the corresponding lactam has provided insight into the role of secondary electrostatic interactions in hydrogen-bonded complex stability.[9i, 10] The sum of this work has established suitably substituted Rebek imides as versatile and *functional* nucleobase surrogates.

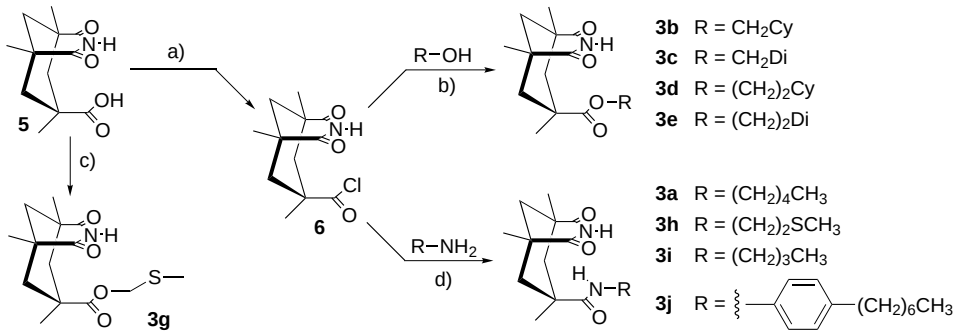
Even so, adenine complexes of the Rebek imides have yet to be studied in the solid state, for which the structures of their nucleobase congeners have been known for more than 40 years. During the course of recent investigations aimed at the study of sulfur-aromatic interactions in proteins, we have employed the Rebek imide receptor system.[11] Crystallization of imides **3a**–**h** and **4** has permitted their detailed solid-state investigation. While pentyl, cyclohexyl (Cy), and dithianyl



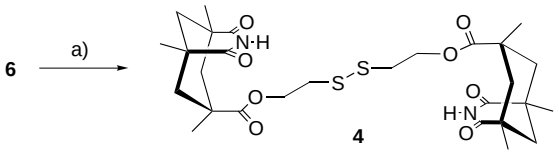
(Di) derivatives **3a**–**e** form simple hydrogen-bonded dimers in the crystalline phase, **4** forms extended ribbon-like structures, and receptors **3f**–**h**, after readily co-crystallizing with **1a**, appear as their adenine complexes. These last systems display a rich array of hydrogen-bonding motifs, which we discuss herein in terms of the known solid-state behavior and base-pairing geometries of the nucleobases. A preference for the Hoogsteen configuration within the solvated complex **3h·1a·**$CHCl_3$ inspired parallel studies in solution by ^{1}H NMR one-dimesional NOE experiments and in the gas phase by computational means. Results of these experiments demonstrate that the Hoogsteen preference persists in the solid, liquid, and gas phases for 1:1 adenine complexes of the Rebek imides.

Results and Discussion

Synthesis: Schemes 1 and 2 show the preparation of all new compounds used in our studies. Derivatives **3a**–**e**, **3h**–**j**, and



Scheme 1. Preparation of novel Rebek imide derivatives from Kemp's triacid **5**. Conditions: a) $SOCl_2$, reflux (ref. [12]). b) NEt_3, DMAP (cat.), CH_2Cl_2, RT or reflux; 35% (**3b**), 31% (**3c**), 21% (**3d**), 12% (**3e**). c) $ClCH_2SCH_3$, NEt_3, CH_3CN, RT; 24%. d) NEt_3, CH_2Cl_2, 0 °C to RT; 50% (**3a**), 77% (**3h**), 80% (**3i**), 64% (**3j**). DMAP = 4-dimethylaminopyridine.



Scheme 2. Preparation of bis-imide **4**. Conditions: a) $HO(CH_2)_2S_2(CH_2)_2OH$, NEt_3, DMAP, CH_2Cl_2, reflux; 26%.

bis-imide **4** were prepared by reaction of the appropriate alcohol or amine with known acid chloride **6**[9f, 12] as described in the literature for analogous compounds.[9] Imide **3g** was accessed by direct alkylation of Kemp's triacid **5**[7, 12] with chloromethyl methyl sulfide following the general procedure of Kim and co-workers.[13] All compounds were isolated as crystalline, white solids.

X-ray crystallography

Imides 3a–e and bis-imide 4: A search of the Cambridge Structural Database (CSD)[14] reveals four examples of Rebek

imide dimers in the solid state; however, this hydrogen-bonding motif is not discussed in the corresponding citations.[15] We describe here the solid-state structures of imide **3a** and bis-imide **4** as many of their features are representative of these two classes of compounds. The structures of **3b**–**e**, which all form hydrogen-bonded dimers, have been deposited with the Cambridge Crystallographic Data Centre (CCDC; see the Experimental Section for details).

We consider first the dimeric structure of **3a** (Figures 3a and b). The imide function within each molecule of **3a** is expectedly planar, as it appears in all of our structures—the largest deviation of the imide atoms (i.e., C(2), N(3), C(4), O(7), O(8); Figure 4) from the least-squares plane defined by these atoms is 0.02 Å. At the hydrogen-bonding interface an N(3)···O(8*)[16] bond length of 2.98(1) Å (Table 1) compares favorably with the average 2.92(3) Å observed for the four reported structures in the CSD, as does an O(8)···O(8*) distance of 3.40(1) Å (versus 3.60(9) Å). Moreover, the hydrogen-bonding angles are nearly identical. Finally, the planes of the two imide functions at the dimer junction are parallel for symmetry reasons; they are offset by only 0.09 Å. Such is also the case for imides **3b**–**e**.

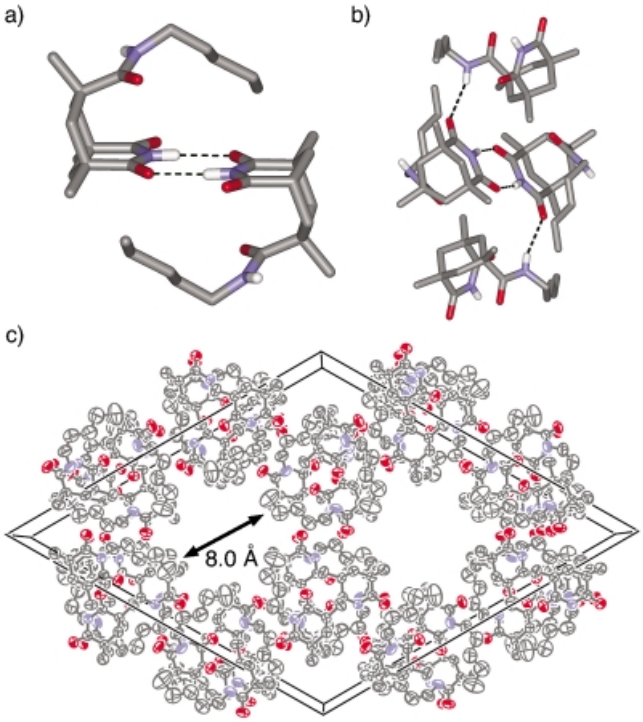


Figure 3. X-ray crystallographic analysis of **3a**. a) The **3a**·**3a** dimer. Some hydrogens have been omitted for clarity. Black dashed lines denote hydrogen bonds. b) Hydrogen bonding is also observed between the amide –NH protons and the imide carbonyl oxygens. c) An ORTEP diagram of the unit cell of **3a**, projected along the *c* axis. Hydrogens and disordered solvent molecules occupying the channels have been omitted for clarity. Atom colors: blue N, red O, white H. Thermal ellipsoids are shown at the 50% probability level.

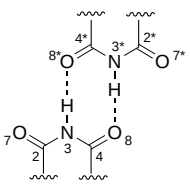


Figure 4. Atom numbering scheme used for imide dimers **3a**–**e** and bis-imide **4**.

Table 1. Hydrogen-bonding distances [Å] and angles [°] for imide **3a** and bis-imide **4** compared to known Rebek imide structures from the Cambridge Structural Database (CSD).[a]

	CSD[b]	**3a**[c]	**4**[c]
N(3)···O(8*)	2.92(3)	2.98(1)	2.91(1)
O(8)···O(8*)	3.60(9)	3.40(1)	3.49(1)
N(3)··O(8*)–C(4*)	125.5(19)	132.1(2)	126.0(2)
C(4)–N(3)···O(8*)	113.3(27)	107.7(2)	110.6(2)

[a] See Figure 4 for atom labeling. [b] Average distances calculated from the crystallographic data using Cerius² v. 4.5 (Molecular Simulations, Inc.). See ref.[15] for the CSD reference codes. Standard deviations are shown in parentheses. [c] Calculated from the crystallographic data using SHELXTL (see the Experimental Section for details). Estimated standard deviations are shown in parentheses. Symmetry operations * for the symmetry equivalent atoms: **3a**: $1.6666 - x$, $1.3333 - y$, $0.3333 - z$; **4**: $0.5 - x$, $0.5 - y$, $-z$.

A packing diagram for **3a** is shown in Figure 3c wherein the crystal displays rhombohedral ($R\bar{3}$) symmetry. The pentyl substituents are sufficient to fill the intervening space between the dimer units, but unexpectedly, lipophilic channels are formed about the lattice's trigonal axis. The distance across the channels is approximately 8 Å; possible occupants are disordered *n*-hexane solvent molecules aligned along the *c* axis in their extended chain conformation (ca. 6.5 Å end-to-end). Atoms composing the solvent molecules and Rebek imide subunits (e.g., the methyl groups) appear to be within van der Waals contact along the channel bore. The molecules of **3a** form an infinite hydrogen-bonded network throughout the lattice (Figure 3c) by means of weak contacts between the amide –NH of one imide and the carbonyl oxygen of an imide from a neighboring pair ($O(7)\cdots N_{amide} = 3.13$ Å).

A search of the CSD for Rebek-type bis(imides) does not reveal any submissions, although several alkyl-tethered cyclic bis(imides) can be found.[17] Bis-imide **4** is not unlike these examples, although it boasts more robust three-dimensional structure and added functionality within the imide component. Figure 5a shows an ORTEP diagram of the C_2-symmetric bis-imide **4**, wherein an 86.8(2)° dihedral angle is observed about the disulfide bond (CH_2-S-S-CH_2) with an S–S bond length of 2.04(1) Å (the S–S bond is at a crystallographic diad axis; $1.0 - x$, y, $0.5 - z$); these values are consistent with other disulfides in the solid state.[18]

Figure 5b shows the crystal packing diagram for **4**. Beginning at the hydrogen-bonding interface, the bonding distances and angles are nearly identical to those observed for the dimer of imide **3a** (Table 1). However, unlike within the **3a** assembly, the hydrogen-bonded imide functions in the crystal of **4** deviate by 1.02 Å from a coplanar configuration. In other words, the imide surfaces are parallel, but lie 0.51 Å off the assembly's inversion center. The packing of **4** is reminiscent of many "ditopic" receptors that self-associate by hydrogen bonding—the individual molecules of **4** form "molecular tapes"[19] and the tapes align in a parallel fashion throughout the crystal lattice.

Complexes of Rebek imides with 9-ethyladenine: In all cases, co-crystals of Rebek imides and 9-ethyladenine could be

© WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002
0947-6539/02/0801-0120 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 1

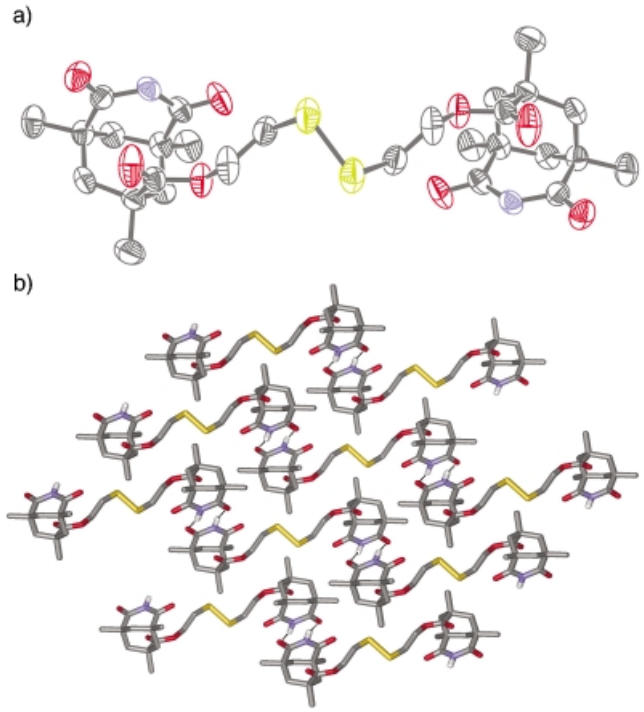


Figure 5. X-ray crystallographic analysis of bis-imide **4**. a) An ORTEP diagram of **4**. Hydrogens have been omitted for clarity. Thermal ellipsoids are shown at the 50% probability level. b) A crystal packing diagram of **4**, projected along the *c* axis. Some hydrogens have been omitted for clarity. Atom colors: yellow S, blue N, red O, white H.

obtained by slow infusion of hexane into chloroform solutions of the two components.

Complex $(\mathbf{3f})_2\cdot\mathbf{1a}$: The 2:1 complex formed between simple methyl ester **3f**[9] and **1a** (Figure 6a) bears remarkable structural similarity to the nucleobase complex featuring 1-methyl-5-iodouracil (**2e**) and **1a** (Figure 6b).[4d] In the latter, both the Hoogsteen and reversed Watson–Crick binding modes are observed. Interestingly, a search of the CSD reveals that *2:1 imide binding* to adenine in the **HG**/**WC** motif (normal or reversed pairing) is rather uncommon for isolated nucleobases, the above iodouracil complex being one of the only examples.[20] Similar to $(\mathbf{2e})_2\cdot\mathbf{1a}$, the planes defined by the two recognition surfaces within $(\mathbf{3f})_2\cdot\mathbf{1a}$ deviate from a coplanar relationship. At the **HG** face, **3f** and **1a** are twisted with respect to one another by about 10.0° (versus 9.4° in $(\mathbf{2e})_2\cdot\mathbf{1a}$), a value that increases to approximately 19.6° at the **WC** face (versus 15.7° in $(\mathbf{2e})_2\cdot\mathbf{1a}$). Such deviations presumably arise to relieve bending in the hydrogen bonds.[21]

The prima facie structural similarities between the two complexes are borne out in their hydrogen-bonding parameters (Table 2). Short N(7)···N(3′) and N(1)···N(3″) distances (ca. 2.9 Å) are observed at both faces of the purine base within $(\mathbf{3f})_2\cdot\mathbf{1a}$. The same is also true for the N(10)···O(8′) and N(10)···O(7″) values (ca. 2.9 Å). The intermolecular C(8)···O(7′) and C(2)···O(8″) distances (ca. 3.7 Å), although longer than those observed within $(\mathbf{2e})_2\cdot\mathbf{1a}$, are still within a range appropriate for favorable C–H···O interactions. The stabilizing nature of such contacts has long been debated in the context of organic crystal structures,[22] rationalized in terms of secondary electrostatic interactions within hydrogen-bonded complexes,[10] and most recently discussed in terms of base-paired nucleotide[23] and transmembrane helix structures.[24] In combination with H···O=C angles near the ideal 120° (122° for the **HG** face, 130° for the **WC** face) we take these parameters as being consistent with C–H···O hydrogen bonding in $(\mathbf{3f})_2\cdot\mathbf{1a}$.

Figure 6c shows the packing diagram of $(\mathbf{3e})_2\cdot\mathbf{1a}$ projected along the *b* axis. The crystal structure consists of layers of ternary complexes related by two-fold rotation symmetry. Stacking (ca. 3.4 Å) is observed between the Hoogsteen hydrogen-bonding face of one ternary complex and the π

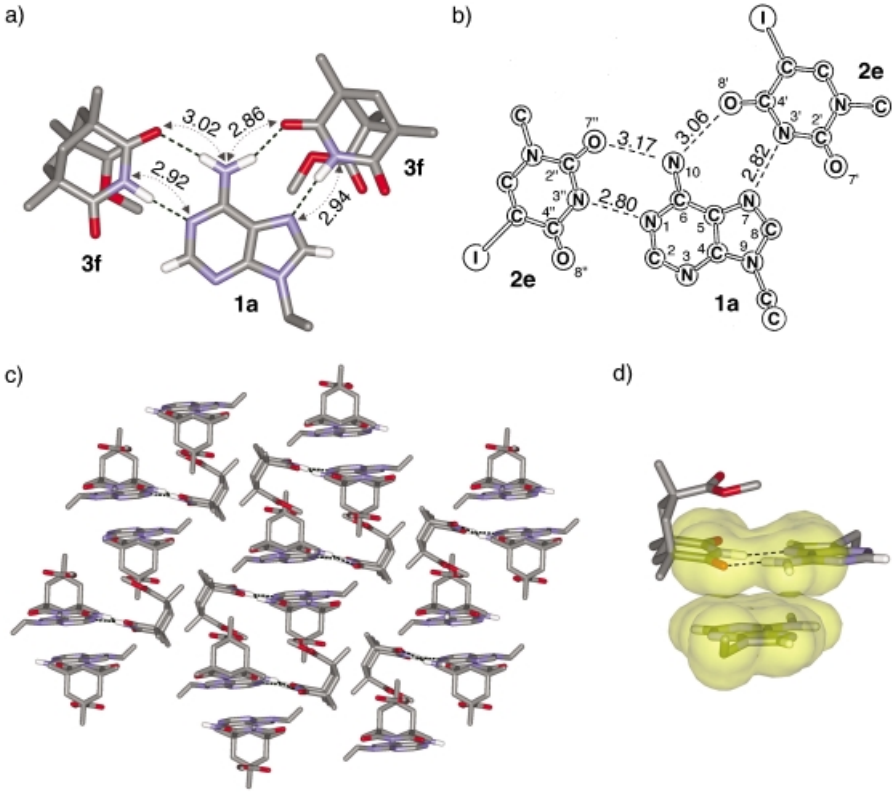


Figure 6. The hydrogen-bonding arrangements within 2:1 9-ethyladenine (**1a**) complexes with 5-iodouracil (**2e**) and **3f** are nearly identical. a) The $(\mathbf{3f})_2\cdot\mathbf{1a}$ complex. The ethyl group of **1a** is disordered in the crystal structure. Distances shown are in Å. b) The $(\mathbf{2e})_2\cdot\mathbf{1a}$ complex as described by Sobell and co-workers (ref. [4d], IMUEAD, reproduced and modified with permission of the *J. Mol. Biol.*). c) The unit cell of $(\mathbf{3f})_2\cdot\mathbf{1a}$ projected along the *b* axis. d) π Stacking between **1a** and the Hoogsteen hydrogen-bond interface. The van der Waals surfaces are shown in transparent yellow. Some hydrogens have been omitted for clarity.

 © WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002 0947-6539/02/0801-0121 $ 17.50+.50/0

Table 2. Hydrogen-bonding distances [Å][a] and angles [°][a] at the Hoogsteen and Watson–Crick faces for assemblies $(\mathbf{2e})_2\cdot\mathbf{1a}$ and $(\mathbf{3f})_2\cdot\mathbf{1a}$.[b]

Hoogsteen face	$(\mathbf{2e})_2\cdot\mathbf{1a}$[c]	$(\mathbf{3f})_2\cdot\mathbf{1a}$[d]	Watson–Crick face	$(\mathbf{2e})_2\cdot\mathbf{1a}$[c]	$(\mathbf{3f})_2\cdot\mathbf{1a}$[d]
N(7)···N(3′)	2.82	2.94(1)	N(1)···N(3″)	2.80	2.91(1)
N(7)···N(3′)-C(2′)	121.1	118.4(5)	N(1)···N(3″)-C(2″)	120.7	116.2(5)
N(7)···N(3′)-C(4′)	110.0	114.6(5)	N(1)···N(3″)-C(4″)	113.4	116.6(5)
C(5)-N(7)···N(3′)	139.5	138.1(6)	C(2)-N(1)···N(3″)	113.6	117.9(5)
C(8)-N(7)···N(3′)	114.2	118.7(5)	C(6)-N(1)···N(3″)	126.6	122.2(5)
N(10)···O(8′)	3.06	2.86(1)	N(10)···O(7″)	3.16	3.02(1)
N(10)···O(8′)-C(4′)	133.8	141.2(5)	N(10)···O(7″)-C(2″)	118.2	122.5(5)
C(6)-N(10)···O(8′)	126.1	128.7(5)	C(6)-N(10)···O(7″)	115.1	112.8(5)
C(8)···O(7′)	3.32	3.75(1)	C(2)···O(8″)	3.35	3.66(1)
C(8)···O(7′)-C(2′)	106.7	105.1(5)	C(2)···O(8″)-C(4″)	109.5	106.6(5)
N(7)-C(8)···O(7′)	84.5	78.2(5)	N(1)-C(2)···O(8″)	83.7	77.3(5)
N(9)-C(8)···O(7′)	160.8	167.1(6)	N(3)-C(2)···O(8″)	147.2	153.6(6)

[a] Calculated from the crystallographic data using Cerius² v. 4.5 for $(\mathbf{2e})_2\cdot\mathbf{1a}$ and SHELXTL for $(\mathbf{3f})_2\cdot\mathbf{1a}$. [b] See Figures 1, 2, and 6 for atom labeling. [c] Ref. [4d]. [d] Estimated standard deviations are shown in parentheses.

surface of a neighboring adenine molecule **1a**; overlap between the relevant van der Waals surfaces is shown in Figure 6d. This motif, π hydrogen-bond stacking, appears in many other hydrogen-bonded assemblies in the solid state, although its contribution to hydrogen-bond complex stability has not been well investigated.[25]

Complex $(\mathbf{3g})_2\cdot\mathbf{1a}$: X-ray analysis of the co-crystals formed from **3g** and **1a** also reveals 2:1 complex formation. Figure 7a shows the isolated ternary assembly, wherein the Hoogsteen motif is prominently displayed. The planes defined by the imide and adenine surfaces are markedly twisted (22.2(7)°) at the **HG** face, perhaps representing a compromise between the

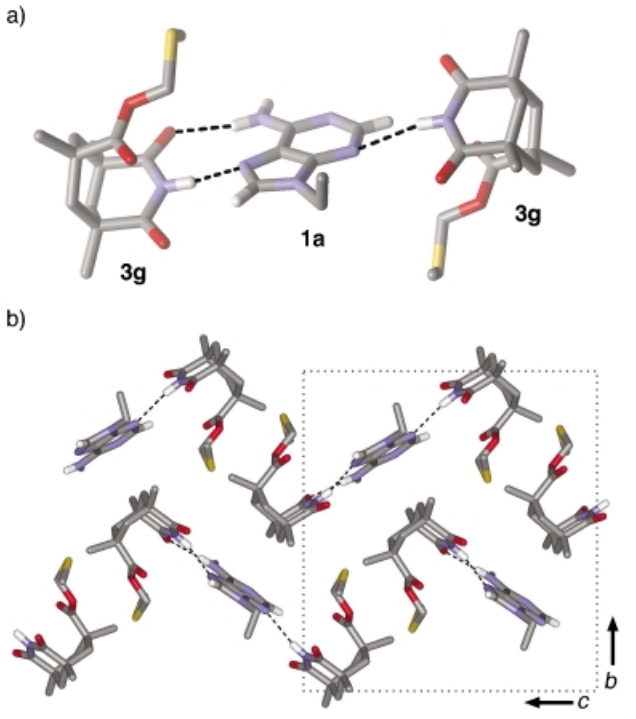


Figure 7. Single-crystal X-ray analysis reveals the adenine-binding arrangement within the $(\mathbf{3g})_2\cdot\mathbf{1a}$ complex. a) Imide hydrogen bonding to both the Hoogsteen face and to the N(3) of adenine derivative **1a** is observed. b) Packing diagram with a view along the crystallographic *a* axis; the unit cell dimensions are outlined. Some hydrogens have been omitted for clarity.

hydrogen-bonding forces operating at either end of the molecule (vide infra). For comparison of hydrogen-bonding parameters we employ the benchmark 1:1 Hoogsteen complex formed between 1-methylthymine (**2a**) and 9-methyladenine (**1b**); data were collected by Frey and co-workers using neutron diffraction techniques.[5] The N(7)···N(3′) and N(10)···O(8′) distances are nearly identical in the two complexes (ca. 3 Å), as are each of the hydrogen-bonding angles (Table 3). Similar to both $\mathbf{2a}\cdot\mathbf{1b}$ and $(\mathbf{3f})_2\cdot\mathbf{1a}$, a short C(8)···O(7′) distance (3.61(1) Å) is observed with an H···O=C angle of 115°, values again representative of C–H···O hydrogen bonding.

Table 3. Hydrogen-bonding distances [Å][a] and angles [°][a] at the Hoogsteen face for assemblies $\mathbf{2a}\cdot\mathbf{1b}$, $(\mathbf{3g})_2\cdot\mathbf{1a}$, and $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$.[b]

	$\mathbf{2a}\cdot\mathbf{1b}$[c]	$(\mathbf{3g})_2\cdot\mathbf{1a}$[d]	$\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$[d]
N(7)···N(3′)	2.93	2.90(1)	2.84(1)
N(7)···N(3′)-C(2′)	115.2	113.9(5)	113.0(3)
N(7)···N(3′)-C(4′)	118.3	118.3(5)	119.6(3)
C(5)-N(7)···N(3′)	135.7	135.5(6)	140.1(3)
C(8)-N(7)···N(3′)	119.9	119.4(5)	116.2(3)
N(10)···O(8′)	2.87	2.98(1)	3.03(1)
N(10)···O(8′)-C(4′)	137.8	134.3(5)	136.7(3)
C(6)-N(10)···O(8′)	132.1	130.2(5)	128.2(3)
C(8)···O(7′)	3.66	3.61(1)	3.38(1)
C(8)···O(7′)-C(2′)	106.2	106.6(5)	109.5(3)
N(7)-C(8)···O(7′)	77.7	77.6(5)	82.1(3)
N(9)-C(8)···O(7′)	169.1	152.2(6)	163.4(3)

[a] Calculated from the crystallographic data using Cerius² v. 4.5 for $\mathbf{2a}\cdot\mathbf{1b}$ and SHELXTL for $(\mathbf{3g})_2\cdot\mathbf{1a}$ and $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$. [b] See Figures 1 and 2 for atom labeling. [c] Ref. [5]. [d] Estimated standard deviations are shown in parentheses.

The $(\mathbf{3g})_2\cdot\mathbf{1a}$ complex features a second hydrogen-bonding motif involving N(3) of **1a** as the acceptor and the –NH group of a neighboring imide as the donor (N(3)···N(3″) = 3.07(1) Å). Pairing geometries involving this face of the adenine have recently been calculated to be *more stable* than either the Watson–Crick or Hoogsteen arrangements for *unsubstituted adenine-thymine* base pairs in the gas phase.[26] Despite these predictions, a search of the CSD[27] does not reveal *any* examples of imide –NH hydrogen bonding to the N(3) position of adenine in the context of *nucleobase complexes*,[28] although the pattern is well known within adenine complexes with barbiturate derivatives (e.g., parabanic acids, cyanuric acids, and hydantoins).[28, 29] That the N(3)···N(3″) distance of 3.07(1) Å observed in the $(\mathbf{3g})_2\cdot\mathbf{1a}$ complex is appropriate comes through comparison to the CSD structures:[27] the mean $N\cdots N(3)_{adenine}$ distance calculated from our survey is 3.05(15) Å. Finally, the deviation from planarity between **3g** and **1a** at the N(3) face is a large 27.1(7)°, a geometry that likely optimizes the N(3)···N(3″)

© WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002
0947-6539/02/0801-0122 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 1

hydrogen-bond distance while minimizing unfavorable steric repulsions between **3g** and the N(9) ethyl substituent. The twist direction at the N(3) face is opposite from that of **1g** at the **HG** face. Consequently, the ternary assemblies are pseudo-centrosymmetric. This symmetry is also evident in the packing diagram of $(\mathbf{3g})_2\cdot\mathbf{1a}$, shown in Figure 7b, wherein the ternary complexes are related in the crystal lattice by *c*-glide planes.

Complex ***3h·1a·CHCl₃***: Imide **3h** readily co-crystallizes with 9-ethyladenine (**1a**) and one molecule of chloroform to form a solvated 1:1 hydrogen-bonded complex (Figure 8). Here the Hoogsteen base-pairing configuration is *exclusively* observed. It is known that 1:1 adenine complexes with thymine and uracil derivatives prefer Hoogsteen pairing in the solid state (vide supra).[2–5, 30] A search of the CSD reveals that the Hoogsteen preference for 1:1 complexes with adenine is not limited to the nucleobases, but is also observed in flavin·adenine complexes.[31] It is the disparity between this hydrogen-bonding configuration and that what is observed in double-stranded DNA that was disconcerting to many crystallographers, and remains the impetus for ongoing solution-phase and gas-phase computational experimentation (vide infra). Identification of the Hoogsteen motif in G quartets[32] and G-DNA,[33] nucleobase triplexes,[34] and even unconstrained parallel DNA duplexes[35] more than suggests that it holds *functional* as well as structural significance. The appearance of the Hoogsteen motif within our comparatively simple 1:1 complex speaks for the generality of this binding mode, at the very least reiterating its unique prevalence within 1:1 imide·adenine complexes in the solid state.

The hydrogen-bonding angles and distances at the Hoogsteen face of **1a** are given in Table 3. The similarities between $\mathbf{2a}\cdot\mathbf{1b}$ and $(\mathbf{3g})_2\cdot\mathbf{1a}$ are transparent—the values are nearly identical. Notably, the shortest C(8)···O(7′) distance (3.38(1) Å) among the three Rebek imide complexes is observed in $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$, wherein an H···O=C angle of 126° defines the contact. Moreover, the planes formed between the imide and adenine surfaces are essentially coplanar (angle of deviation = 1.6(2)°). Whether this indicates a particularly favorable geometry for the Rebek imide relative to the nonplanar arrangement, as is observed at the **WC** face of $(\mathbf{3f})_2\cdot\mathbf{1a}$, is unclear, although it surely represents an optimization of the electrostatic contacts at the hydrogen-bonding interface.

For $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$, the crystal packing diagram (Figure 8b) shows stacking between the adenine and the hydrogen-bond surface of an adjacent complex (ca. 3.5 Å), as observed in $(\mathbf{3f})_2\cdot\mathbf{1a}$, while the individual $\mathbf{3h}\cdot\mathbf{1a}$ complexes align in an anti-parallel fashion, presumably to alleviate unfavorable dipole–dipole interactions (Figure 8a). In the context of DNA-antibiotic conjugates[36] both π stacking and dipole–dipole interactions are believed to direct the formation of the Hoogsteen configuration.[37] We do not rule out such effects as being important contributors to the Hoogsteen selectivity observed in the $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$ complex.

Solution studies

The nucleobases have enjoyed a rich history of study in solution, where much has been learned of their association behavior in a variety of media.[38] However, limitations in experimental techniques and instrumentation have made determination of relative **HG**/**WC** amounts qualitative at best.[39] More recently, Weisz and co-workers have explored the relative populations of the **HG** and **WC** configurations for A–U base pairs through low-temperature NMR studies.[38b] Their results indicate a preference (ca. 80%) for Watson–Crick bonding at 125 K in $CDClF_2/CDF_3$ solvent mixtures, but comparable, *although not quantifiable*, amounts of **HG** and **WC** complexes at ambient temperature. Upon observing Hoogsteen binding for the 1:1 $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$ complex, we were naturally interested in determining whether the **HG** preference in the solid state may reflect *intrinsic* differences in energetic stability between the competing **HG** and **WC** motifs.[40] The solution-phase analogy of the Rebek imides to the pyrimidine nucleobases well-established,[9] and their solid-state structural analogy demonstrated here, provided the momentum for these studies. On the practical front, the known 1:1 binding stoichiometry of the Rebek imides for adenines in solution combined with the NOE protocols reported by Rebek and co-workers for assessing relative **HG**/**WC** populations within these complexes by NMR spectroscopy[9c, 9d, 9f] suggested the ready solution-phase evaluation of our systems.

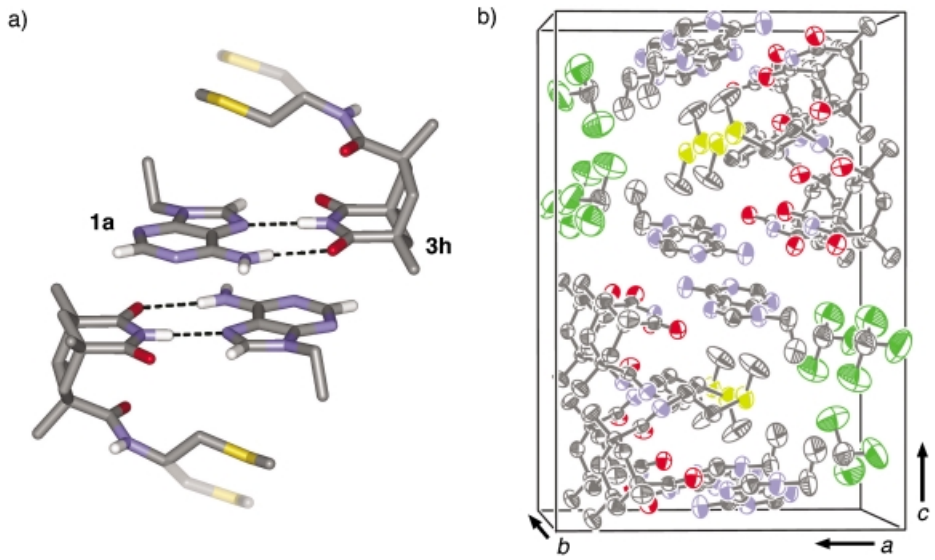


Figure 8. Single crystal X-ray analysis of $\mathbf{3h}\cdot\mathbf{1a}\cdot CHCl_3$. a) Adenine binding via the Hoogsteen face. Disorder within the sulfur-containing sidechain of **3h** gives rise to two conformers (major, 82%; minor, 18%, pale bonds). b) An ORTEP diagram of the unit cell. Thermal ellipsoids are shown at the 50% probability level. Atom colors: green Cl, yellow S, blue N, red O, white H. Some hydrogens have been omitted for clarity.

 © WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002 0947-6539/02/0801-0123 $ 17.50+.50/0

^{1}H NMR binding studies: The association constants and thermodynamic parameters describing the binding of the imides **3f–h**, and controls **3i** and **3j**, to 9-ethyladenine (**1a**) in $CDCl_3$ are given in Table 4. The K_a's were determined by nonlinear least-squares analysis of ^{1}H NMR titration data and the thermodynamic parameters (ΔH^0 and ΔS^0) were accessed by using van't Hoff plots. The average values determined for **3f–h** binding to **1a** (i.e., K_a ca. 55 M^{-1}, $-\Delta H^0$ ca. 6.7 kcal mol^{-1}, $-\Delta S^0$ ca. 14 e.u.) are in excellent agreement with the reported literature values for other Rebek imide analogues.[9g] For the assemblies **3f·1a** and **3h·1a**, 1:1 binding stoichiometry was confirmed by Job plot analysis (data not shown).

Table 4. Association constants K_a and thermodynamic parameters describing the 1:1 binding[a] of **3f–j** to **1a** in $CDCl_3$ (295 K).

Complex	K_a [M^{-1}][b]	$-\Delta G^0$ [kcal mol^{-1}]	$\Delta\delta_{sat}$[c]	$-\Delta H^0$ [kcal mol^{-1}][d]	$-\Delta S^0$ [e.u.][d]
3f·1a	62[e]	2.4	5.2	6.9	15
3g·1a	46	2.3	4.7	6.6	14
3h·1a	46[f]	2.3	4.7	6.5	14
3i·1a	40	2.2	5.0	5.8	12
3j·1a	60	2.4	5.3	6.3	13

[a] 1:1 Binding stoichiometry confirmed by Job plot analysis. [b] Determined by nonlinear least-squares curve fitting of 300 MHz ^{1}H NMR titration data. Uncertainty in K_a estimated as 15%. Values are uncorrected for the presence of the imide dimer; see ref. [9g]. [c] Downfield complexation-induced chemical shift observed for the imide –NH proton upon binding to **1a**. [d] Determined by van't Hoff analysis; e.u. = entropy units. [e] A value of 50 M^{-1} is reported in ref. [9g]. [f] *Weak* intramolecular hydrogen bonding between the sulfur atom and imide –NH is observed. The corrected K_a using the procedures outlined in ref. [9h] is ca. 50 M^{-1}.

NOE experiments: The ^{1}H NMR NOE protocol described by Rebek et al. is shown schematically in Figure 9. Irradiation of the cyclic imide –NH of **3** results in an intermolecular NOE to $H_{C(8)}$ and $H_{C(2)}$, arising from the **HG** and **WC** complexes,

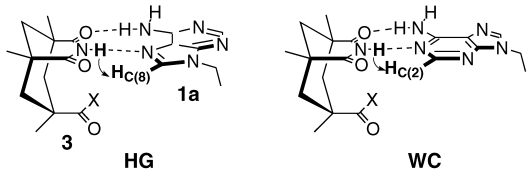


Figure 9. Graphical depictions of the NOE experiments performed on the 1:1 **3·1a** assemblies (see refs. [9c,d,f]). Irradiation of the imide –NH results in an NOE to $H_{C(8)}$ in the Hoogsteen (**HG**) complex and $H_{C(2)}$ in the Watson–Crick (**WC**) complex.

respectively. We performed these experiments on our 1:1 **3·1a** complexes and several model compounds at 300 MHz in $CDCl_3$; the results are shown in Table 5. In essentially all cases the Hoogsteen mode is favored modestly relative to its Watson–Crick counterpart. Interestingly, the **HG** to **WC** ratio is ca. 1.7:1 for the assembly involving the simple methyl ester **3f** (**3f·1a**). Presumably here the small methyl substituent does not substantially bias the **HG**/**WC** equilibrium and the latter motif is preferred for more subtle reasons. This is

Table 5. Calculated Hoogsteen (**HG**) and Watson–Crick (**WC**) ratios for 1:1 complexes based on one-dimensional steady-state NOE difference spectroscopy (irradiation of the imide –NH resonance; 300 MHz, $CDCl_3$, 25 °C).[a]

Complex	$H_{C(8)}$ [%][b]	$H_{C(2)}$ [%][b]	$H_{C(8)}/H_{C(2)}$ (**HG**/**WC**)	%**HG**
3f·1a	2.0	1.2	1.7	63
3g·1a	2.0	1.9	1.1	52
3h·1a	1.8	1.3	1.4	58
3i·1a	2.8	1.8	1.6	62
3j·1a	3.2	1.6	2.0	67
1a[c]	0.02	0.15	–	–

[a] Under the experimental conditions the imide –NH (of **3f–j**) and amino $-NH_2$ protons (of **1a**) are in fast chemical exchange and simultaneously irradiated (irradiation time = 2 s). [b] Determined by integration with the irradiated –NH peak calibrated as –100% (*absolute* intensities have significant error, ca. 50%). [c] Direct irradiation of the amino $-NH_2$ protons.

not expected to be the case for phenyl derivative **3j·1a**, investigated for comparison. Here π stacking provided by the aryl group to the surface of **1a** appears to bias the equilibrium (Figure 2) toward the **HG** configuration, for which greater π overlap exists between the purine base and aromatic ring.

Nevertheless, we recognize the limitations of the NOE technique and interpret our data both cautiously and conservatively. At best the results suggest a *modest* preference for the **HG** configuration. Systematic errors in the analysis are prevalent,[41] although several would appear to *underestimate* the **HG** contribution. For example, due to the fast chemical exchange between the imide –NH and amino $-NH_2$ protons of **1a**, all exchangeable protons are irradiated under the steady-state (irradiation time = 2 s) conditions of the one-dimensional NOE experiments. An independent measurement reveals that amino irradiation results in a greater enhancement to the $H_{C(2)}$ position than to $H_{C(8)}$ (Table 5, last entry). Similar effects would be expected upon imide binding to the N(3) position of **1a**.[42] Additionally, as the distance between the imide –NH proton and either $H_{C(8)}$ or $H_{C(2)}$ contributes to their enhancement magnitude (NOE $\propto r^{-6}$) we note that these values at the **HG** and **WC** faces are likely not identical. For example, at the **WC** face of the solid-state complex $(\mathbf{3f})_2\cdot\mathbf{1a}$ the $H_{C(2)}\cdots H_{N(3'')}$ distance is approximately 2.70 Å. The average $H_{C(8)}\cdots H_{N(3')}$ distance at the **HG** faces in the solid-state for the three complexes is longer (2.94 Å); again, a Hoogsteen underestimation results. Although we are reluctant to *directly* compare the solid-state and solution-phase structural data, the cumulative contributions of these effects may be detectable.

Computations

High-level computational work has shown the Hoogsteen mode to be the more stable binding arrangement for isolated *adenine–thymine* base pairs[43] and within A–T tetrads,[44] on the order of 0.4–1 kcal mol^{-1} and 2.3 kcal mol^{-1} more stable than the corresponding Watson–Crick configuration in the gas phase, respectively. For comparison, we have performed molecular mechanics (Amber*), semiempirical (PM3),[45] and ab initio (MP2/6–31G*) calculations on both the succinim-

© WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002
0947-6539/02/0801-0124 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 1

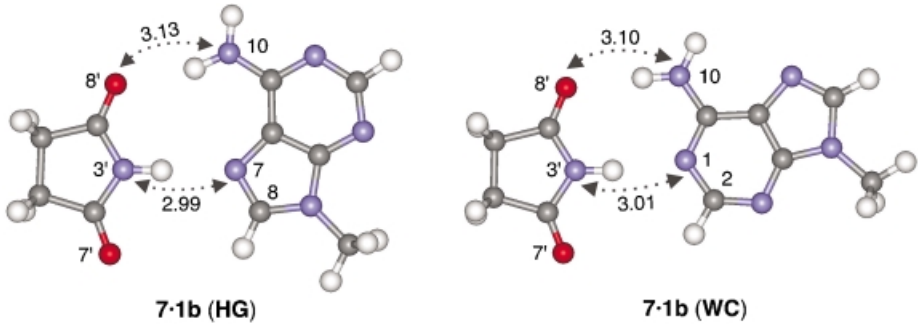


Figure 10. HF/6–31G*-optimized geometries of the Hoogsteen and Watson–Crick dimers formed from succinimide (**7**) and 9-methyladenine (**1b**). Distances shown are in Å.

ide·9-methyladenine (**7·1b**) model system (Figure 10) and the 1-methylthymine·9-methyladenine (**2a·1b**) base pair.[46] The results are presented in Table 6.

Table 6. Calculated interaction energies (ΔE)[a] for Hoogsteen (**HG**) and Watson–Crick (**WC**) geometries of **7·1b** and **2a·1b** hydrogen-bonded dimers at various levels of theory.

Level of theory	Complex	$-\Delta E_{HG}$ [kcal mol^{-1}]	$-\Delta E_{WC}$ [kcal mol^{-1}]	$-\Delta\Delta E_{(HG-WC)}$ [kcal mol^{-1}]
Amber*[b]	**7·1b**	11.20	10.84	0.36
	2a·1b	8.22	7.33	0.89
PM3[c]	**7·1b**	5.97	5.02	0.95
	2a·1b	7.59	6.67	0.92
MP2/6–31G*//	**7·1b**	11.51	10.98	0.53
HF/6–31G*[d]	**2a·1b**	12.08	11.06	1.02

[a] For **7·1b**, for example, defined as $\Delta E_{7\text{-}1b} = E_{7\text{-}1b} - (E_7 + E_{1b})$. [b] Implemented in MacroModel v. 7.0.[60] [c] Implemented in Spartan v. 5.1.3.[61] [d] Implemented in Gaussian 98.[62] Values at the MP2 level have been corrected for basis set superposition error (BSSE) using the counterpoise correction procedure of Bernardi and Boys[63] and include zero-point vibrational energy (ZPE) corrections (scaled by 0.91[64]) at the HF/6–31G* level.

At all three levels of theory the Hoogsteen motif is favored by about 0.4–1 kcal mol^{-1} for both classes of complexes.[47] Energy-minimized structures of the **HG** and **WC** configurations of the **7·1b** complex at the HF/6–31G* level of theory are shown in Figure 10. The reliability of the single-point energy determinations comes indirectly through the results obtained for the **2a·1b** base pairs. The $-\Delta E_{HG}$ value (i.e., interaction energy) for the A–T base pair of 12.1 kcal mol^{-1} compares well with the value reported by Kollman and Gould of 11.8 kcal mol^{-1} using similar computational protocols ($-\Delta\Delta E_{(HG\text{-}WC)} = 0.8$ kcal mol^{-1} in the same calculations),[43b] and, fortuitously, with the gas-phase ΔH_{298} value (13.0 kcal mol^{-1}) determined by mass spectrometry.[48] Interestingly, although the N(10)···C(8′) and N(7)···N(3′) distances are reproduced well compared to our imide X-ray structures (Figure 10), the C(8)···O(7′) distances are not (average calculated distance = 4.2 Å; average observed distance = 3.6 Å). Intrinsic differences do of course exist between solid- and gas-phase structures (as a consequence of packing), and between the structures of **7** and **2a**. However, it also seems clear that the C=O···H–C contact is not well-defined at the Hartree-Fock level in the succinimide complexes (nor well-parameterized at the molecular mechanics level[10]).[49] Nevertheless, we are optimistic that these discrepancies are neutralized upon comparing the interaction energy *differences* between the complexes.[50]

Conclusion

Investigation of Rebek imide derivatives **3a–e** and **4** in the solid state has offered considerably more than confirmation of solution-phase binding phenomena. The diverseness of functionality amenable to incorporation within the Rebek imide scaffold suggests that these platforms may themselves be versatile "supramolecular synthons" for crystal engineering.[51] In particular, molecules such as bis-imide **4**, which crystallize from solution as tape-like polymeric assemblies, offer opportunities in materials design.[19] In fact, there is still much to learn about the behavior of even simple hydrogen-bonded assemblies in the solid state. For example, one of the molecule targets posed at the 2001 Workshop on Crystal Structure Prediction[52] organized by the Cambridge Crystallographic Data Centre featured a cyclic imide, the crystal structure of which was successfully predicted by only 3 of the 18 participants.

Adenine complexes of the Rebek imides display a variety of base-pairing motifs that can be directly compared to their nucleobase congeners. This is especially true of the 2:1 (**3f**)$_2$·**1a** complex, wherein Watson–Crick and Hoogsteen binding are observed, and the 1:1 solvated complex **3h·1a**·$CHCl_3$, featuring exclusive Hoogsteen binding. In each of these cases the hydrogen-bonding distances and angles, as well as more subtle parameters, such as mean plane twist angles and imide configurational arrangement, are nearly identical to their A–U and A–T congeners, respectively. Crystallization of the 2:1 complex (**3g**)$_2$·**1a** has permitted identification of a new hydrogen-bonding motif in the solid state for nucleobase analogues; namely, one involving *pyrimidine-type* imide –NH binding to the N(3) position of adenine. The rarity of this type of hydrogen bonding in natural systems likely originates from steric hindrance provided by the bulky sugar residue attached to N(9)$_{adenine}$ in nucleosides. Single N–H···N(3)$_{adenine}$ hydrogen bonds in biological systems, however, are fairly common.[53] In the same vein, the recently solved structure of the large ribosomal subunit suggests a role for the N(3) of adenine (A2486/A2451) as a general base during peptide bond formation.[54]

Although many have recognized the preference for Hoogsteen binding in the solid state for 1:1 complexes with adenine, this observation has not stimulated extensive research to determine its origins.[38b] We are hopeful that our finding of the motif in a different class of structures, one that boasts increased functionality relative to the nucleobases, will foster new research endeavors, particularly in solution environments most relevant to natural systems. Meanwhile, we look to

expand our own efforts toward better understanding the Hoogsteen preference and its generality in the context of other systems.

Experimental Section

Materials and general methods: Compounds **1a**,[55] **3f**,[9f] **5**,[12] and **6**[12] were prepared as described in the literature. Details of the syntheses of **3b–e** can be found in the Supporting Information. Reagents and solvents were purchased from Aldrich, Fluka, or Acros and used without further purification unless otherwise specified. CH_2Cl_2, NEt_3, diisopropylamine, and *n*-butylamine were distilled from CaH_2 prior to use. THF was distilled from sodium/benzophenone prior to use. CH_3CN was distilled from K_2CO_3 and stored over 4 Å molecular sieves in an N_2 atmosphere. Flash chromatography (FC) was performed on SiO_2–60 (40–63 μm) from Fluka. Preparative thin-layer chromatography (PTLC) was performed on 2 mm Merck plates precoated with silica gel 60 F_{254}. Thin-layer chromatography (TLC) was performed on DURASIL G/UV254 glass plates from Macherey-Nagel with visualization by UV light or staining ($KMnO_4$). Melting points (m.p.) were determined on a Büchi SB-540 melting point apparatus and are uncorrected. Infrared spectra were recorded with a Perkin–Elmer FT-IR 1600 spectrophotometer. Abbreviations used are s (strong), m (medium), and w (weak). ^1H and ^{13}C NMR spectra were recorded on Varian Gemini 200 (200 MHz), Varian Gemini 300 (300 MHz), or Bruker AMX 400 (400 MHz) spectrometers as specified. Chemical shifts (δ) are given in parts per million relative to residual protonated solvent ($CHCl_3$: $\delta_H = 7.26$, $\delta_C = 77.00$). Abbreviations used are s (singlet), d (doublet), t (triplet), q (quartet), quin (quintet), br (broad), and m (multiplet). Elemental analyses were performed at the ETH-Zürich by the Mikrolaboratorium in the Laboratorium für Organische Chemie. MALDI-FTMS spectra were recorded on an Ion Spec Ultima FT-ICR mass spectrometer with DHB (2,5-dihydroxybenzoic acid) as a matrix.

Crystal growth: Crystals suitable for X-ray analysis were obtained by the slow infusion of hexane into chloroform solutions [filtered through Millipore (Bedford, MA) syringe filters] containing either the receptor **3** alone (ca. 0.01 M) or both the receptor **3** (ca. 0.01 M) and 9-ethyladenine (**1a**; ca. 0.02 M). Excess **1a** often crystallized within hours as needles,[56] whereas the **3**·**1a** complexes appeared over 3–10 days.

X-ray crystal structure determination and refinement: X-ray data were collected on either a Syntex P21 diffractometer with $Mo_{K\alpha}$ radiation (graphite monochromator, $\lambda = 0.71073$ Å) or a PICKER-STOE diffractometer with $Cu_{K\alpha}$ radiation (graphite monochromator, $\lambda = 1.54178$ Å). The structures were solved with the SHELXS program.[57] The structures were refined by full-matrix least-squares on F^2 with the SHELXL program.[58] Details are provided in Table 7. Crystallographic data (excluding structure factors) for the structures reported in this paper have been deposited with the Cambridge Crystallographic Data Centre as supplementary publication no. CCDC-167546 (**3a**), CCDC-167549 (**3b**), CCDC-167550 (**3c**), CCDC-167547 (**3d**), CCDC-167548 (**3e**), CCDC-167545 ((**3f**)$_2$·**1a**), CCDC-167542 ((**3g**)$_2$·**1a**), CCDC-167543 (**3h**·**1a**·$CHCl_3$), and CCDC-167544 (**4**). Copies of the data can be obtained free of charge on application to CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: (+44) 1223-336-033; e-mail: deposit@ccdc.cam.ac.uk).

Association constant (K_a) determination: NMR titrations were performed at 295 K on a Varian Gemini 300 spectrometer (300 MHz) in $CDCl_3$ (stored over 4 Å molecular sieves) purchased from Glaser, Basel. The receptor **3** concentration was kept constant at 0.010 M and a solution of receptor **3** (0.010 M) and **1a** (0.10 M) was added in increments by syringe to the NMR tube. The imide −NH chemical shift was recorded after each of 12 additions. Binding isotherms were created using Associate v. 1.6,[59] which provided K_a, ΔG^0, and $\Delta\delta_{sat}$. Values reported are the average of 2–3 independent runs.

Determination of thermodynamic parameters by van't Hoff analysis: Van't Hoff plots were created by monitoring the imide −NH chemical shift response to temperature given a fixed concentration of receptor **3** (0.010 M) and **1a** (0.010 M – 0.020 M). Concentrations were chosen such that 20–80 % of the saturation curve was sampled. Samples were equilibrated for 30–45 minutes initially (253–258 K), for 10–15 minutes between 258–303 K, and for 5 minutes thereafter until the maximum temperature was obtained (323 K). Values reported are the average of 2–3 independent runs.

One-dimensional steady-state NOE difference spectroscopy: Measurements were performed on a Bruker ARX 300 spectrometer (300 MHz) with an irradiation time of 2 s. $CDCl_3$ was purchased from Glaser, Basel and passed over aluminum oxide prior to use. Sample concentration: **3** 0.02 M, **1a** 0.04 M (samples were not degassed).

Computational details: Amber* calculations were performed in MacroModel v. 7.0.[60] Initial structures for semiempirical and ab initio calculations

Table 7. Crystal data and structure refinement for **3a**, **4**, (**3f**)$_2$·**1a**, (**3g**)$_2$·**1a**, and **3h**·**1a**·$CHCl_3$.

	3a	**4**	(**3f**)$_2$·**1a**	(**3g**)$_2$·**1a**	**3h**·**1a**·$CHCl_3$
formula	$C_{39}H_{56}N_4O_6$[a]	$C_{28}H_{40}N_2O_8S_2$	$C_{33}H_{47}N_7O_8$	$C_{35}H_{51}N_7O_8S_2$	$C_{23}H_{34}N_7O_3Cl_3S$
M_r	676.9	596.74	669.78	761.95	594.98
crystal system	trigonal ($\gamma = 120°$)	monoclinic	monoclinic	monoclinic	monoclinic
space group	$R\bar{3}$	$C2/c$	$C2/c$	Pc	$P2_1/c$
a [Å]	28.20(2)	21.66(3)	20.71(2)	8.070(7)	15.486(14)
b [Å]	28.20(2)	8.974(5)	12.819(9)	16.290(20)	23.042(25)
c [Å]	12.775(9)	15.643(13)	26.77(2)	14.938(27)	8.378(9)
β [°]		100.36(10)	92.78(7)	97.46(12)	98.13(8)
V [Å^3]	8796(10)	2991(6)	7099(10)	1947.1(46)	2959.5(53)
Z	9	4	8	2	4
ρ_{calcd} [g cm^{-3}]	1.15	1.325	1.253	1.300	1.335
radiation type	$Cu_{K\alpha}$	$Cu_{K\alpha}$	$Cu_{K\alpha}$	$Cu_{K\alpha}$	$Mo_{K\alpha}$
λ [Å]	1.54178	1.54178	1.54178	1.54178	0.71073
T [K]	293(2)	293(2)	293(2)	293(2)	293(2)
crystal size [mm]	0.2 × 0.1 × 0.1	0.1 × 0.1 × 0.05	0.1 × 0.08 × 0.05	0.4 × 0.2 × 0.1	0.5 × 0.3 × 0.2
independent reflections	2010	1530	3648	1741	5219
observed reflections [$I > 2\sigma(I)$]	1561	1322	2788	1691	1935
index ranges	$-23 \le h \le 242$	$-21 \le h \le 21$	$-20 \le h \le 20$	$-8 \le h \le 7$	$-18 \le h \le 18$
	$0 \le k \le 27$	$0 \le k \le 8$	$0 \le k \le 12$	$0 \le k \le 14$	$0 \le k \le 27$
	$0 \le l \le 12$	$0 \le l \le 15$	$0 \le l \le 26$	$0 \le l \le 14$	$0 \le l \le 9$
F (000)	3294	1272	2864	812	1248
goodness-of-fit (on F^2)	0.996	1.053	1.039	1.136	0.859
R_1 based on F [$I > 2\sigma(I)$]	0.0703	0.0460	0.0444	0.0499	0.0530
wR_2 based on F^2 (all data)	0.2230	0.1294	0.1267	0.1327	0.1295

[a] Formally, $C_{34}H_{56}N_4O_6$·ca. 0.6 hexane.

© WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002
0947-6539/02/0801-0126 $ 17.50+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 1

were obtained using the MMFF94 force field as implemented in Spartan v. 5.1.3.[61] Semiempirical (PM3) calculations were performed in Spartan v. 5.1.3. Prior to single-point energy calculations, all structures were fully optimized at the HF/6–31G* level of theory using the Gaussian 98 (Revision A.7) suite of programs.[62] Single-point energy calculations (MP2/6–31G*) employed the frozen-core approximation and were implemented through Gaussian 98. Interaction energies (between generic X and Y) are defined as $\Delta E_{int(XY)} = E_{XY} - (E_X + E_Y)$. Counterpoise correction calculations were made to obtain estimates of the basis set superposition error (BSSE) for the MP2/6–31G* single-point energies as described by Bernardi and Boys.[63] Vibrational frequency calculations at the HF/6–31G* confirmed that all optimized structures were true minima and provided the zero-point vibrational energies (ZPEs). The ZPE values were scaled by a factor of 0.91.[64] MacroModel v. 7.0 was run on an SGI O2 Workstation, Spartan v. 5.1.3 on an SGI Octane workstation, and Gaussian 98 through an HP V2500SCA server. Coordinates of all refined structures are provided in the Supporting Information.

Methylsulfanylmethyl 1,5,7-trimethyl-2,4-dioxo-3-azabicyclo[3.3.1]nonane-7-carboxylate (3g): Kemp's triacid **5** (0.10 g, 0.42 mmol) and dry CH_3CN (5 mL) were added to a flame-dried, 10 mL round-bottomed flask. The resulting suspension was treated with NEt_3 (0.10 mL, 0.75 mmol), yielding a clear solution, and chloromethyl methyl sulfide (0.038 mL, 0.46 mmol) was added. The mixture was stirred at RT under N_2 for 12 h. The solvent was removed in vacuo, and the residue was diluted with CH_2Cl_2 (10 mL) and washed with water (10 mL), and the organic extract was dried over Na_2SO_4. After evaporation of the solvent, the crude material was subjected to PTLC (40% EtOAc/hexane) to give the product as a white solid (0.040 g, 32%). The material was recrystallized from CH_2Cl_2/hexane (0.030 g, 24%). M.p. (CH_2Cl_2/hexane) 181.5–183.5 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 0.02 M): $\delta = 7.60$ (s, 1H), 5.09 (s, 2H), 2.72 (d, $J =$ 13.0 Hz, 2H), 2.26 (s, 3H), 1.99 (dt, $J = 13.8$, 2.1 Hz, 1H), 1.38 (d, $J =$ 13.5 Hz, 1H), 1.26 (s, 6H), 1.24 (s, 3H), 1.20 (d, $J = 14.7$ Hz, 2H); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta = 176.17$, 174.97, 69.59, 44.22, 43.77, 42.14, 39.92, 30.44, 24.16, 15.88; IR (KBr): $\tilde{\nu} = 3193$ (m), 3094 (m), 2969 (w), 2922 (w), 2867 (w), 1732 (s), 1693 (s), 1466 (m), 1325 (m), 1238 (m), 1212 (s), 1155 (s), 1090 (m), 941 cm^{-1} (m); HRMS (MALDI, DHB, $[M+Na]^+$) calcd for $C_{14}H_{21}NO_4SNa$: 322.1089; found: 322.1082; elemental analysis calcd (%) for $C_{14}H_{21}NO_4S$: C 56.17, H 7.07, N 4.68; found: C 56.19, H 7.18, N 4.71.

General procedure for reaction of 6 with primary amines: Receptors 3a and 3h–j: The acid chloride **6** (0.2–0.8 mmol) and dry CH_2Cl_2 (10–15 mL) were added to a flame-dried, round-bottomed flask. This solution was cooled to 0 °C in an ice-water bath and treated first with NEt_3 (1.5 equiv) and then with the desired amine (1.3 equiv). The mixture was warmed to RT and stirred for 12 h. The solvent was removed in vacuo and the residue was purifed by chromatography on SiO_2, either by FC or PTLC. The white, often crystalline solid obtained upon solvent evaporation was then recrystallized twice from CH_2Cl_2/hexane to remove traces of the free amine.

***N*-Pentyl 1,5,7-trimethyl-2,4-dioxo-3-azabicyclo[3.3.1]nonan-7-carboxamide (3a)**: Performed with **6** (0.4 mmol); PTLC 25% EtOAc/CH_2Cl_2 (0.060 g, 50%). M.p. 163–165 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 0.01 M): $\delta =$ 7.50 (br s, 1H), 5.47 (m, 1H), 3.14 (dt, $J = 7.4$, 5.4 Hz, 2H), 2.54 (d, $J \sim$ 14 Hz, 2H), 1.99 (dt, $J = 13.2$, 1.8 Hz, 1H), 1.50–1.39 (m, 2H), 1.37 (d, $J =$ 13.2 Hz, 1H), 1.33–1.18 (m, 4H), 1.27 (s, 6H), 1.23 (d, $J = 14.4$ Hz, 2H), 1.19 (s, 3H), 0.88 (t, $J = 7.0$ Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta =$ 176.60, 173.74, 44.51, 44.38, 42.00, 40.01, 39.90, 31.84, 28.96, 28.69, 24.34, 22.18, 13.85; IR (KBr): $\tilde{\nu} = 3047$ (m), 3215 (m), 3100 (w), 2961 (m), 2933 (m), 2856 (w), 1717 (s), 1682 (s), 1650 (s), 1531 (s), 1468 (m), 1424 (m), 1380 (m), 1239 (m), 1208 cm^{-1} (s); HRMS (MALDI, DHB, $[M+Na]^+$) calcd for $C_{17}H_{28}N_2O_3Na$: 331.1992; found: 331.1992; elemental analysis calcd (%) for $C_{17}H_{28}N_2O_3$: C 66.20, H 9.15, N 9.08; found: C 66.19, H 9.05, N 8.99.

***N*-2-(Methylsulfanyl)ethyl 1,5,7-trimethyl-2,4-dioxo-3-azabicyclo[3.3.1]nonan-7-carboxamide (3h)**: Performed with **6** (0.2 mmol); PTLC 50% EtOAc/CH_2Cl_2 (0.047 g, 77%). M.p. (CH_2Cl_2/hexane) 180–181 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 0.01 M): $\delta = 7.64$ (br s, 1H), 5.94 (m, 1H), 3.38 (q, $J = 6.0$ Hz, 2H), 2.60 (t, $J = 6.4$ Hz, 2H), 2.56 (d, $J \sim 13.5$ Hz, 2H), 2.10 (s, 3H), 2.00 (dt, $J = 13.5$, 2.1 Hz, 1H), 1.38 (d, $J = 13.5$ Hz, 1H), 1.27 (s, 6H), 1.24 (d, $J \sim 14$ Hz, 2H), 1.21 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta =$ 176.62, 173.98, 44.46, 44.35, 42.15, 40.00, 37.38, 33.15, 31.76, 24.32, 14.50; IR (KBr): $\tilde{\nu} = 3399$ (s), 3210 (m), 3100 (w), 2967 (w), 2932 (w), 2867 (w), 1717 (s), 1679 (s), 1656 (s), 1532 (m), 1244 (m), 1209 cm^{-1} (s); HRMS (MALDI, DHB, $[M+Na]^+$) calcd for $C_{15}H_{24}N_2O_3SNa$: 335.1405; found: 335.1398; elemental analysis calcd (%) for $C_{15}H_{24}N_2O_3S$: C 57.67, H 7.74, N 8.97; found: C 57.86, H 7.71, N 8.91.

***N*-Butyl 1,5,7-trimethyl-2,4-dioxo-3-azabicyclo[3.3.1]nonan-7-carboxamide (3i)**: Performed with **6** (0.2 mmol); PTLC 30% EtOAc/CH_2Cl_2 (0.046 g, 80%). M.p. (CH_2Cl_2/hexane) 163–165 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 0.01 M): $\delta = 7.49$ (s, 1H), 5.46 (m, 1H), 3.15 (dt, $J = 7.2$, 5.6 Hz, 2H), 2.54 (d, $J = 13.5$ Hz, 2H), 1.99 (dt, $J = 13.5$, 2.1 Hz, 1H), 1.45–1.29 (m, 4H), 1.37 (d, $J = 13.2$ Hz, 1H), 1.27 (s, 6H), 1.23 (d, $J = 14.4$ Hz, 2H), 1.19 (s, 3H), 0.91 (t, $J = 7.4$ Hz, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta = 176.60$, 173.76, 44.53, 44.38, 42.00, 40.01, 39.66, 31.84, 31.07, 24.34, 19.98, 13.61; IR (KBr): $\tilde{\nu} = 3338$ (s), 3200 (m), 3100 (m), 2965 (w), 2929 (w), 2864 (w), 1730 (s), 1697 (s), 1629 (s), 1547 (s), 1198 cm^{-1} (s); HRMS (MALDI-FTMS, DHB, $[M+Na]^+$) calcd for $C_{16}H_{26}N_2O_3Na$: 317.1841; found: 317.1836; elemental analysis calcd (%) for $C_{16}H_{26}N_2O_3$: C 65.28, H 8.90, N 9.52; found: C 65.46, H 9.10, N 9.32.

***N*-4-Heptylphenyl 1,5,7-trimethyl-2,4-dioxo-3-azabicyclo[3.3.1]nonan-7-carboxamide (3j)**: Performed with **6** (0.8 mmol) where the reaction solution was heated to reflux for 12 h. The crude material was purified by FC on SiO_2 (50 mL, 5% EtOAc/CH_2Cl_2) and triturated with diethyl ether to give the product as a white powder (0.20 g, 64%): M.p. (Et_2O) 192–194 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 0.02 M): $\delta = 7.55$ (s, 1H), 7.31 (d, $J = 8.4$ Hz, 2H), 7.11 (s, 1H), 7.10 (d, $J = 8.4$ Hz, 2H), 2.67 (d, $J = 14.4$ Hz, 2H), 2.55 (t, $J = 7.5$ Hz, 2H), 2.03 (d, $J = 12.9$ Hz, 1H), 1.62–1.50 (m, 2H), 1.42 (d, $J = 13.2$ Hz, 1H), 1.36–1.25 (m, 10H), 1.31 (s, 3H), 1.30 (s, 6H), 0.88 (t, $J = 6.8$ Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 176.37$, 171.80, 139.50, 134.74, 128.79, 120.84, 44.55, 44.54, 43.04, 40.19, 35.40, 31.83, 31.66, 31.52, 29.23, 29.18, 24.39, 22.68, 14.10; IR (KBr): $\tilde{\nu} = 3371$ (m), 3189 (w), 3078 (w), 2956 (w), 2926 (w), 2844 (w), 1717 (m), 1691 (s), 1530 (m), 1212 cm^{-1} (m); HRMS (MALDI, DHB, $[M+Na]^+$) calcd for $C_{25}H_{36}N_2O_3$-Na: 435.2618; found: 435.2610; elemental analysis calcd (%) for $C_{25}H_{36}N_2O_3$: C 72.78, H 8.79, N 6.79; found: C 72.69, H 8.83, N 6.68.

Dithiobis(ethane-1,2-diyl) bis(1,5,7-trimethyl-2,4-dioxo-3-azabicyclo-[3.3.1]nonan-7-carboxylate) (4): The acid chloride **6** (0.20 g, 0.78 mmol) and dry CH_2Cl_2 (15 mL) were added to a 50 mL flame-dried, round-bottomed flask. This solution was treated with NEt_3 (205 μL, 1.48 mmol), DMAP (0.23 g, 0.19 mmol), and 2,2′-dithioethanol (90% purity, 50 μL, 0.37 mmol). The mixture was stirred and heated to reflux (ca. 50 °C) under N_2 for 12 h. The solution was then diluted with CH_2Cl_2 (10 mL) and washed with HCl (1 M, 20 mL) and brine (20 mL). The organic extracts were dried over Na_2SO_4, filtered, and concentrated. The crude material was subjected to PTLC (30% EtOAc/hexane) and further recrystallized (2×) from CH_2Cl_2/hexane (0.12 g, 26%). M.p. (CH_2Cl_2/hexane) 256–258 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 0.01 M): $\delta = 8.16$ (s, 2H), 4.28 (t, $J = 7.2$ Hz, 4H), 2.90 (t, $J = 7.5$ Hz, 4H), 2.69 (d, $J = 13.2$ Hz, 4H), 2.00 (d, $J = 13.2$ Hz, 2H), 1.35 (d, $J = 12.9$ Hz, 2H), 1.25 (s, 12H), 1.23 (s, 6H), 1.17 (d, $J = 14.4$ Hz, 4H); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta = 176.67$, 175.15, 63.14, 44.35, 43.77, 42.02, 39.97, 36.57, 30.42, 24.28; IR (KBr): $\tilde{\nu} = 3213$ (m), 3083 (w), 2976 (w), 2933 (w), 1734 (s), 1695 (s), 1456 (m), 1315 (m), 1207 (s), 1159 (s), 1090 cm^{-1} (m); HRMS (MALDI, DHB, $[M+Na]^+$) calcd for $C_{28}H_{40}N_2O_8S_2Na$: 619.2124; found: 619.2126; elemental analysis calcd (%) for $C_{28}H_{40}N_2O_8S_2$: C 56.36, H 6.76, N 4.69; found: C 56.42, H 6.87, N 4.75.

Acknowledgement

R.K.C. is grateful to the US National Science Foundation for an International Scholars Research postdoctoral fellowship. We thank Prof. Bernhard Jaun for many insightful discussions concerning the NOE data and Nicolle Moonen, Dr. Maurice van Eis, and Dr. Per I. Arvidsson for advice concerning the computations.

[1] a) J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738; for a more detailed analysis of this X-ray data and DNA structure see: b) L. Pauling, R. B. Corey, *Arch. Biochem. Biophys.* **1956**, *65*, 164–181.

[2] a) K. Hoogsteen, *Acta Crystallogr.* **1959**, *12*, 822–823; b) K. Hoogsteen, *Acta Crystallogr.* **1963**, *16*, 907–916.

[3] a) F. S. Mathews, A. Rich, *J. Mol. Biol.* **1964**, *8*, 89–95; b) L. Katz, K.-I. Tomita, A. Rich, *J. Mol. Biol.* **1965**, *13*, 340–350; c) L. Katz, K.-I.

Tomita, A. Rich, *Acta Crystallogr.* **1966**, *21*, 754 – 764; d) K.-I. Tomita, L. Katz, A. Rich, *J. Mol. Biol.* **1967**, *30*, 545 – 549.

[4] a) A. E. V. Haschemeyer, H. M. Sobell, *Proc. Natl. Acad. Sci. USA* **1963**, *60*, 872 – 877; b) A. E. V. Haschemeyer, H. M. Sobell, *Acta Crystallogr.* **1965**, *18*, 525 – 532; c) T. D. Sakore, H. M. Sobell, *J. Mol. Biol.* **1969**, *43*, 77 – 87; d) T. D. Sakore, S. S. Tavale, H. M. Sobell, *J. Mol. Biol.* **1969**, *43*, 361 – 374 and references therein; e) S. S. Tavale, T. D. Sakore, H. M. Sobell, *J. Mol. Biol.* **1969**, *43*, 375 – 384; f) T. D. Sakore, H. M. Sobell, F. Mazza, G. Kartha, *J. Mol. Biol.* **1969**, *43*, 385 – 406; g) F. Mazza, H. M. Sobell, G. Kartha, *J. Mol. Biol.* **1969**, *43*, 407 – 422.

[5] M. N. Frey, T. F. Koetzle, M. S. Lehmann, W. C. Hamilton, *J. Chem. Phys.* **1973**, *59*, 915 – 924.

[6] See reference [4d] for a summary of these results.

[7] D. S. Kemp, K. S. Petrakis, *J. Org. Chem.* **1981**, *46*, 5140 – 5143.

[8] For reviews see: a) J. Rebek, Jr., *Angew. Chem.* **1990**, *102*, 261 – 272; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 245 – 255; b) J. Rebek, Jr., *Acc. Chem. Res.* **1990**, *23*, 399 – 404.

[9] a) J. Rebek, Jr., L. Marshall, R. Wolak, K. Parris, M. Killoran, B. Askew, D. Nemeth, N. Islam, *J. Am. Chem. Soc.* **1985**, *107*, 7476 – 7481; b) J. Rebek, Jr., B. Askew, C. Buhr, S. Jones, D. Nemeth, K. Williams, P. Ballester, *J. Am. Chem. Soc.* **1987**, *109*, 5033 – 5035; c) J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, A. Costero, S. Jones, K. Williams, *J. Am. Chem. Soc.* **1987**, *109*, 6866 – 6867; d) J. Rebek, Jr., K. Williams, K. Parris, P. Ballester, K. S. Jeong, *Angew. Chem.* **1987**, *99*, 1297 – 1299; *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 1244 – 1245; e) J. S. Lindsey, P. C. Kearney, R. J. Duff, P. T. Tjivikua, J. Rebek, Jr., *J. Am. Chem. Soc.* **1988**, *110*, 6575 – 6577; f) B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones, K. Parris, K. Williams, J. Rebek, Jr., *J. Am. Chem. Soc.* **1989**, *111*, 1082 – 1090; g) K. Williams, B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones, J. Rebek, Jr., *J. Am. Chem. Soc.* **1989**, *111*, 1090 – 1094; h) T. Tjivikua, G. Deslongchamps, J. Rebek, Jr., *J. Am. Chem. Soc.* **1990**, *112*, 8408 – 8414; i) K. S. Jeong, T. Tjivikua, A. Muehldorf, G. Deslongchamps, M. Famulok, J. Rebek, Jr., *J. Am. Chem. Soc.* **1991**, *113*, 201 – 209; j) V. M. Rotello, E. A. Viani, G. Deslongchamps, B. A. Murray, J. Rebek, Jr., *J. Am. Chem. Soc.* **1993**, *115*, 797 – 798; k) I. Huc, J. Rebek, Jr. *Tetrahedron Lett.* **1994**, *35*, 1035 – 1038.

[10] a) W. L. Jorgensen, D. L. Severance, *J. Am. Chem. Soc.* **1991**, *113*, 209 – 216; b) J. Pranata, S. G. Wierschke, W. L. Jorgensen, *J. Am. Chem. Soc.* **1991**, *113*, 2810 – 2819; c) W. L. Jorgensen, *Chemtracts Org. Chem.* **1991**, *4*, 91 – 119.

[11] For related work see: a) E. C. Breinlinger, C. J. Keenan, V. M. Rotello, *J. Am. Chem. Soc.* **1998**, *120*, 8606 – 8609; b) V. M. Rotello, *Heteroatom Chem.* **1998**, *9*, 605 – 606.

[12] J. G. Stack, D. P. Curran, S. V. Geib, J. Rebek, Jr., P. Ballester, *J. Am. Chem. Soc.* **1992**, *114*, 7007 – 7018.

[13] C. S. Kim, J. J. Suh, B. Y. Lee, C. S. Kim, J. W. Lee, B. C. Kim, B. H. Han (Yuhan Corporation, Ltd.), EP-0366548, **1989** [*Chem. Abstr.* **1990**, *113*, 152264].

[14] CSD System, V. 5.21, April 2001 release, as implemented through ConQuest. See: http://www.ccdc.cam.ac.uk for details. CSD reference codes are shown in parentheses (e.g. ABCDEF) throughout the paper.

[15] We have limited our search to include only unsubstituted Rebek imide derivatives (i.e., free imide –NH): a) Ref. [9d] (GADKIF, LERDIV); b) Ref. [9k] (PIVFIJ); c) A. Yanagisawa, T. Kikuchi, T. Watanabe, T. Kuribayashi, H. Yamamoto, *Synlett.* **1995**, 372 – 374 (YUFYED).

[16] Starred (*) atoms denote symmetry equivalence for the imide dimers; see Figure 4 for atom labeling.

[17] Two relevant *recent* examples include 4,4′-(Hexane-1,6-diyl)bis(piperazine-2,6-dione): a) Q. Liu, S.-W. Zhang, M.-C. Shao, *Acta Crystallogr. Sect. C* **1998**, *54*, 439 – 440 (NIFKAO); Di-(2,4-imidazolidinedione-5-ethyl)disulfide; b) D. F. Mullica, M. L. Trawick, P. W. N. Wu, E. L. Sappenfield, *J. Chem. Crystallogr.* **1998**, *28*, 761 – 765 (COQQEE).

[18] a) V. Pattabha, R. Srinivasan, *Int. J. Pept. Protein Res.* **1976**, *8*, 27 – 32; b) M. T. N. Petersen, P. H. Jonson, S. B. Petersen, *Protein Eng.* **1999**, *12*, 535 – 548.

[19] For a review see: R. E. Melendez, A. D. Hamilton, *Top. Curr. Chem.* **1998**, *198*, 97 – 129.

[20] This hydrogen-bonding motif has been observed in 9-ethyl-2,6-diaminopurine complexes with 1-methylthymine (**2a**, MTYDAP) and 1-methyl-5-iodouracil (**2e**, MIUDAP10), ref. [4f], as well as in *extended hydrogen-bonded networks* involving **1a** and barbiturate derivatives (vide infra, refs. [28, 29]).

[21] M. Takimoto, A. Takenaka, Y. Sasada, *Bull. Chem. Soc. Jpn.* **1982**, *55*, 2734 – 2738.

[22] a) G. R. Desiraju, *Acc. Chem. Res.* **1991**, *24*, 290 – 296; b) G. R. Desiraju, *Acc. Chem. Res.* **1996**, *29*, 441 – 449; c) T. Steiner, *Chem. Commun.* **1997**, 727 – 734.

[23] a) G. A. Leonard, K. McAuley-Hecht, T. Brown, W. N. Hunter, *Acta Crystallogr. Sect. D* **1995**, *51*, 136 – 139; b) E. B. Starikov, T. Steiner, *Acta Crystallogr. Sect. D* **1997**, *53*, 345 – 347; c) M. C. Wahl, M. Sundaralingam, *Trends Biochem. Sci.* **1997**, *22*, 97 – 102; d) J. Marfurt, C. Leumann, *Angew. Chem.* **1998**, *110*, 184 – 187; *Angew. Chem. Int. Ed.* **1998**, *37*, 175 – 177; e) M. Brandl, K. Lindauer, M. Meyer, J. Sühnel, *Theor. Chem. Acc.* **1999**, *101*, 103 – 113.

[24] A. Senes, I. Ubarretxena-Belandia, D. M. Engelman, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9056 – 9061.

[25] a) P. Lustenberger, E. Martinborough, T. Mordasini Denti, F. Diederich, *J. Chem. Soc. Perkin Trans. 2* **1998**, 747 – 761. For the intercalation of aromatic compounds between adjacent DNA base pairs see: b) W. Saenger, *Principles of Nucleic Acid Structure*, Springer, New York, **1984**, Chapter 16.

[26] M. Kratochvíl, J. Šponer, P. Hobza, *J. Am. Chem. Soc.* **2000**, *122*, 3495 – 3499.

[27] The CSD search was limited to –NH donors for the N(3) position (where the H···N distance is less than the sum of the van der Waals radii) of *adenine derivatives* (i.e., not including diaminopurines, isoguanines, *N*-oxides, etc.). Of the 46 structures found, 42 feature the –NH donor (e.g., adenine $-NH_2$ in 23 structures) and N(3) acceptor in the *same* molecule.

[28] For a discussion of $N-H\cdots N(3)_{adenine}$ interactions see: H.-S. Shieh, D. Voet, *Acta Crystallogr. Sect. B* **1976**, *32*, 2361 – 2367.

[29] a) H.-S. Shieh, D. Voet, *Acta Crystallogr. Sect. B* **1975**, *31*, 2192 – 2201; b) R. H. Epstein, A. V. Zeiger, C. Crocker, D. Voet, *Acta Crystallogr. Sect. B* **1976**, *32*, 2180 – 2188; c) H.-S. Shieh, D. Voet, *Acta Crystallogr. Sect. B* **1976**, *32*, 2354 – 2360; d) R. E. Cassady, S. W. Hawkinson, *Acta Crystallogr. Sect. B* **1982**, *38*, 2206 – 2209; e) A. Camerman, D. Mastropaolo, A. Hempel, N. Camerman, *Can. J. Chem.* **2000**, *78*, 1045 – 1051; for other barbiturate·adenine complexes see: f) S. H. Kim, A. Rich, *Proc. Natl. Acad. Sci. USA* **1968**, *65*, 402 – 408; g) D. Voet, A. Rich, *J. Am. Chem. Soc.* **1972**, *94*, 5888 – 5891; h) D. Voet, *J. Am. Chem. Soc.* **1972**, *94*, 8213 – 8222; i) V. R. Pedireddi, A. Ranganathan, K. N. Ganesh, *Org. Lett.* **2001**, *3*, 99 – 102.

[30] For adenine·thymine and adenine·uracil complexes displaying exclusively the Hoogsteen motif but involving multiple additional hydrogen-bonding or stacking interactions see: a) ref. [4b] (ADBURM); b) ref. [20] (BODZUP); c) S. Fujita, A. Takenaka, Y. Sasada, *Bull. Chem. Soc. Jpn.* **1983**, *56*, 2234 – 2237 (BUDVUR10); for a thiouracil derivative see: d) W. Saenger, D. Suck, *J. Mol. Biol.* **1971**, *60*, 87 – 99 (SURMAD10); relevant examples of Watson – Crick binding include the adenosine·5-bromouracil complex; e) K. Aiba, T. Hata, S. Sato, C. Tamura, *Acta Crystallogr. Sect. B* **1978**, *34*, 1259 – 1263 (ADOSBU), and the 3-(adenin-9-yl)-*N*-(2-succinimidyl)-propionamide structure, featuring hydrogen-bonded networks of a *tethered* cyclic imide and adenine residue; f) M. Takimoto, A. Takenaka, Y. Sasada, *Bull. Chem. Soc. Jpn.* **1984**, *57*, 3070 – 3073 (CUBRUM).

[31] For the reversed Hoogsteen motif see: a) D. Voet, A. Rich, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1151 – 1156 (RIBBAD); b) S. Fujii, K. Kawasaki, A. Sato, T. Fujiwara, K.-I. Tomita, *Arch. Biochem. Biophys.* **1977**, *181*, 363 – 370 (ADRBFT10); for the Hoogsteen motif see: c) T. Ishida, T. Yanagitani, Y. Nakagawa, M. Inoue, *Biochem. Biophys. Res. Comm.* **1981**, *99*, 1179 – 1184 (BACMIB); d) M. Inoue, Y. Okuda, T. Ishida, M. Nakagaki, *Arch. Biochem. Biophys.* **1983**, *227*, 52 – 70 (BACMIB10); we ignore here complexes involving the barbiturate derivatives (refs. [28, 29]), as their hydrogen-bonding interactions with adenines are much different from those involving the nucleobases or flavins, and complexes involving adenines substituted at C(8) that disfavor the Hoogsteen binding mode for steric or electronic reasons (e.g., 8-bromoadenine derivatives, refs. [4c, 29b,c]).

[32] D. Sen, W. Gilbert, *Nature* **1988**, *334*, 364 – 366.

[33] For a recent example see: C. Harrington, Y. Lan, S. A. Akman, *J. Biol. Chem.* **1997**, *272*, 24631 – 24636, and references therein.

 © WILEY-VCH Verlag GmbH, 69451 Weinheim, Germany, 2002 0947-6539/02/0801-0128 $ 17.50+.50/0

[34] For a general reference see: a) V. N. Soyfer, V. N. Potaman, *Triple-Helical Nucleic Acids*, Springer, New York, **1996**; for the importance and relative stability of the Hoogsteen component within base triples see: b) J. Tao, L. Chen, A. D. Frankel, *Biochemistry* **1997**, *36*, 3491–3495, and references therein; c) G. M. Hashem, J.-D. Wen, Q. Do, D. M. Gray, *Nucleic Acids Res.* **1999**, *27*, 3371–3379.
[35] K. Liu, H. T. Miles, J. Frazier, V. Sasisekharan, *Biochemistry* **1993**, *32*, 11802–11809.
[36] a) J. Gallego, A. R. Ortiz, F. Gago, *J. Med. Chem.* **1993**, *36*, 1548–1561; b) K. J. Addess, J. Feigon, *Nucleic Acids Res.* **1994**, *22*, 5484–5491.
[37] π Stacking has also been implicated in the Hoogsteen preference observed within 1:1 nucleobase complexes; see refs. [3c, 31a].
[38] For solution-phase studies of simple nucleobase pairs see the extensive references within: a) A. Dunger, H.-H. Limbach, K. Weisz, *Chem. Eur. J.* **1998**, *4*, 621–628; b) A. Dunger, H.-H. Limbach, K. Weisz, *J. Am. Chem. Soc.* **2000**, *122*, 10109–10114; for complementary theoretical calculations see ref. [10]; for discussion of *Hoogsteen binding* of imides to adenine derivatives in the context of synthetic receptors see, for example: c) A. Galán, E. Pueyo, A. Salmerón, J. de Mendoza, *Tetrahedron Lett.* **1991**, *32*, 1827–1830; d) O. F. Schall, G. W. Gokel, *J. Am. Chem. Soc.* **1994**, *116*, 6089–6100, and references therein; e) J. Sartorius, H.-J. Schneider, *Chem. Eur. J.* **1996**, *2*, 1446–1452, and references therein; f) Y. Kuroda, J. M. Lintuluoto, H. Ogoshi, *J. Chem. Soc. Perkin Trans. 2* **1997**, 333–339; g) M. Sirish, B. G. Maiya, *J. Porphyrins Phthalocyanines* **1998**, *2*, 327–335; h) D. G. Lonergan, G. Deslongchamps, *Tetrahedron* **1998**, *54*, 14041–14052; i) D. G. Lonergan, J. Halse, G. Deslongchamps, *Tetrahedron Lett.* **1998**, *39*, 6865–6868; j) L. Yu, H.-J. Schneider, *Eur. J. Org. Chem.* **1999**, 1619–1625; k) C.-C. Zeng, Y.-L. Tang, Q.-Y. Zheng, L.-J. Huang, B. Xin, Z.-T. Huang, *Tetrahedron Lett.* **2001**, *42*, 6179–6181.
[39] Prior to the work of Weisz and co-workers (ref. [38b]), estimates of **HG** and **WC** ratios were made by ^{1}H NMR chemical shift analysis: a) J. D. Engel, P. H. von Hippel, *Biochemistry* **1974**, *13*, 4143–4158. For the recent implementation of this strategy to study the binding of 9-butyladenine by carboxylic acids see: b) P. Rao, S. Ghosh, U. Maitra, *J. Phys. Chem. B* **1999**, *103*, 4528–4533. Worth noting is that the authors observed a distinct preference for binding to the **HG** face of the purine when aromatic carboxylic acids were employed.
[40] For salient discussions see refs. [3, 4d].
[41] For one, we have assumed a simple two-spin system, see: D. Neuhaus, M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, New York, **2000**, Chapter 5.
[42] An NOE to the $-CH_2$ group of the ethyl substituent of **1a**, however, has not been identified.
[43] a) K. I. Trollope, I. R. Gould, I. H. Hillier, *Chem. Phys. Lett.* **1993**, *209*, 113–116; b) I. R. Gould, P. A. Kollman, *J. Am. Chem. Soc.* **1994**, *116*, 2493–2499; c) F. Ryjáček, O. Engkvist, J. Vacek, M. Kratochvíl, P. Hobza, *J. Phys. Chem. A* **2001**, *105*, 1197–1202.
[44] J. Gu, J. Leszczynski, *J. Phys. Chem. A* **2000**, *104*, 1898–1904.
[45] The validity of nonempirical and semiempirical techniques has recently been evaluated for the DNA bases: P. Hobza, M. Kabeláč, J. Šponer, P. Mejzlík, J. Vondrášek, *J. Comput. Chem.* **1997**, *18*, 1136–1150.
[46] The theoretical investigation presented here is not as rigorous as those presented in ref. [43b] or the work of Hobza and co-workers (refs. [43c, 45]; for a review see: P. Hobza, J. Šponer, *Chem. Rev.* **1999**, *99*, 3247–3276). Our intention is to merely establish whether or not the gas-phase interaction energy differences between **HG** and **WC** pairs are comparable for both adenine complexes with succinimide derivatives and the nucleobases.
[47] Molecular mechanics calculations (OPLS force field) by Jorgensen and co-workers have shown the **WC** motif to be nearly isoenergetic with the **HG** motif for **7**·**1b** complexes ($\Delta\Delta E = 0.1$ kcal mol^{-1} in favor of **WC**); see ref. [10].
[48] I. K. Yanson, A. B. Teplitsky, L. F. Sukhodub, *Biopolymers* **1979**, *18*, 1149–1170; we have not, however, formally corrected ΔE to ΔH.
[49] Worth noting is that these distances are well-reproduced for the A–T base pairs.
[50] Molecular mechanics and semiempirical calculations show the N(3) binding arrangement to be >1 kcal mol^{-1} *less* stable than the Watson–Crick configuration for these complexes (i.e., with an N(9) methyl group).
[51] G. R. Desiraju, T. Steiner, *The Weak Hydrogen Bond In Structural Chemistry and Biology*, Oxford University Press, Oxford, **1999**.
[52] The workshop took place on May 10–11, 2001 in Cambridge (UK); J. D. Dunitz and W. D. S. Motherwell, personal communication.
[53] For an overview of "shallow-groove" pairings involving adenine in RNA see: N. B. Leontis, E. Westhof, *Quart. Rev. Biophys.* **1998**, *31*, 399–455.
[54] a) P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 920–930; b) D. Suárez, K. M. Merz, Jr., *J. Am. Chem. Soc.* **2001**, *123*, 7687–7690.
[55] E. Camaioni, S. Costanzi, S. Vittori, R. Volpini, K. N. Klotz, G. Cristalli, *Bioorg. Med. Chem.* **1998**, *6*, 523–533.
[56] For X-ray analysis of **1a** see ref. [29i].
[57] G. M. Sheldrick, SHELXS-86, Program for X-ray Crystal Structure Solution, Göttingen University (Germany), **1990**.
[58] G. M. Sheldrick, SHELXL-93, Program for X-ray Crystal Structure Refinement, Göttingen University (Germany), **1993**.
[59] Associate v. 1.6: B. R. Peterson, Ph. D. Thesis, University of California at Los Angeles, **1994**.
[60] MacroModel v. 7.0 (Schrödinger, 1500 SW First Ave., Suite 1180, Portland, OR, USA): F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440–467.
[61] Spartan v. 5.1.3, Wavefunction, 18401 Von Karman Ave., Suite 370, Irvine, CA, USA.
[62] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, *Gaussian 98*, Revision A.7, Gaussian, Pittsburgh PA, **1998**.
[63] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553–566.
[64] J. A. Pople, A. P. Scott, M. W. Wong, L. Radom, *Israel J. Chem.* **1993**, *33*, 345–350.

Received: July 23, 2001 [F3433]