# PHILOSOPHICAL TRANSACTIONS OF — THE ROYAL SOCIETY A

## MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

# New Supramolecular Architectures Using Hydrogen Bonding

Steven C. Zimmerman and Thomas J. Murray

**Email alerting service**      Receive free email alerts when new articles cite this article - sign up in
the box at the top right-hand corner of the article or click   **here**

# New supramolecular architectures using hydrogen bonding

By Steven C. Zimmerman and Thomas J. Murray

*Department of Chemistry, University of Illinois Urbana, Illinois 61801, U.S.A.*

Several new multiply hydrogen bonded complexes have been studied to determine their strength and the specificity with which they form. While many factors contribute to the stability of multiply hydrogen bonded complexes, it appears that the arrangement of the hydrogen bond donor and acceptor groups is a particularly good predictor of binding strength. The results are consistent with W. L. Jorgensen's secondary electrostatic hypothesis. The heterocyclic recognition units that have been synthesized may serve as the basis for constructing new synthetic hosts or new self-assembling systems.

## 1. Introduction

Hydrogen bonding plays a key role in almost all biological recognition and assembly processes. Not surprisingly, hydrogen bonding interactions have played a prominent role in recent chemical (abiotic) models of these biological processes. Our own efforts to develop new self-assembling systems (Zimmerman & Duerr 1992) and to construct new hosts for heteroaromatic guests (Zimmerman 1993) have recently converged with a common goal of synthesizing compounds corresponding to the partial structures in figure 1. These hypothetical recognition units present two or three adjacent hydrogen bond donor (D) and acceptor (A) groups in all possible arrangements, and can form the matched complexes shown. With two adjacent donor and acceptor groups there are three possible arrangements that form two types of complexes (DA·AD and AA·DD). Three adjacent hydrogen bond donor and acceptor sites can be arranged in six different ways and these can form three different complexes (ADA·DAD, AAD·DDA, and AAA·DDD).

One aim of our work on self-assembling systems is to form discrete aggregates in solution with well-defined structures. For example, we recently reported that pyrido[4,3-g]quinoline dione (**1**) forms cyclic trimer (**2**) in solution, rather than the entropically disfavoured polymeric assemblies (**3**) (Zimmerman & Duerr 1992) (figure 2). The robustness of trimer (**2**) in solution indicates that it forms with a high degree of cooperativity. The interest in the structures in figure 1 is in exerting greater control over which supramolecular assembly is formed in this and other systems, and in maximizing their stability. To accomplish this the recognition sites need to be differentiated. Furthermore, they should recognize their complement with high specificity and affinity, and with a well defined geometry.

The stabilities of a number of complexes containing multiple hydrogen bonds were measured before our work. However, many of the components were nucleotide bases, and these were unsuitable for incorporation into our self-assembling systems. Furthermore, of the many doubly hydrogen bonded complexes that were studied, none had the DD·AA arrangement. Only four triply hydrogen bonded complexes
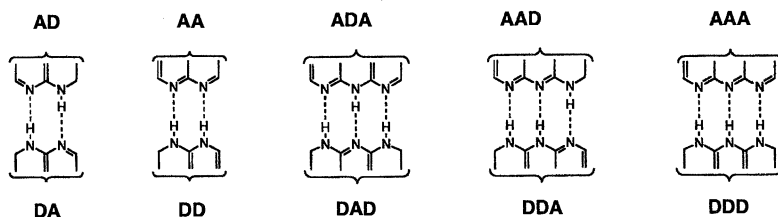
49

Figure 1. Hypothetical multiply hydrogen bonded complexes formed from recognition units containing two and three adjacent hydrogen bond donor and acceptor sites (D and A, respectively) in all possible arrangements.
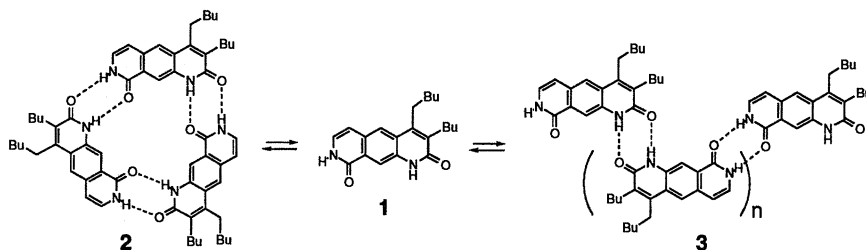


Figure 2. Cyclic and linear aggregation by pyrido[4,3-$g$]quinoline dione (1).
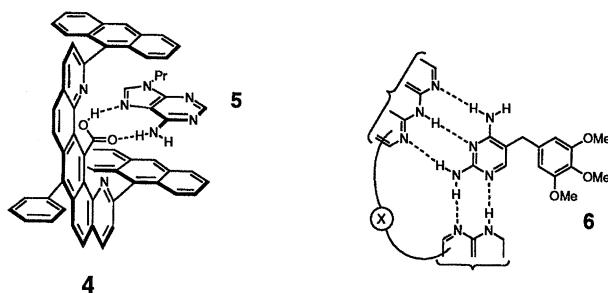


Figure 3. (a) Complex formed between host (4) and 9-propyladenine (5). (b) Complex formed between trimethoprin (6) and a hypothetical host containing an AD and ADA recognition unit.

were examined and none contained the DDD·AAA arrangement. Finally, proposals were put forward to explain the relative stability of these complexes (*vide infra*), but these were based on a small number of examples. For these reasons, we sought recognition units that: (1) were readily synthesized; (2) presented all possible arrangements of two and three hydrogen bond donor and acceptor groups, including the AAA and DDD arrays; and (3) had similar structures so that comparisons of complexation strength would be most meaningful.

The interest in these recognition units from the perspective of host–guest chemistry, arose from our efforts to develop synthetic hosts for nucleobase guests (Zimmerman 1993). We reported that (4) is an exceptionally efficient host for adenine, binding 9-propyladenine (5) in chloroform with an association constant of 120 000 $M^{-1}$ (figure 3a). The design and construction of host (4) is typical in at least one respect. Functional groups that make favourable contacts to the guest molecule of interest, in this case adenine, are attached to a more or less rigid framework that holds them in a spatial arrangement complementary to the guest. Because this is an *ad hoc* approach, each new guest will require a new host design, and a new and likely extensive synthetic effort. As a class, heteroaromatic guests may be amenable to a
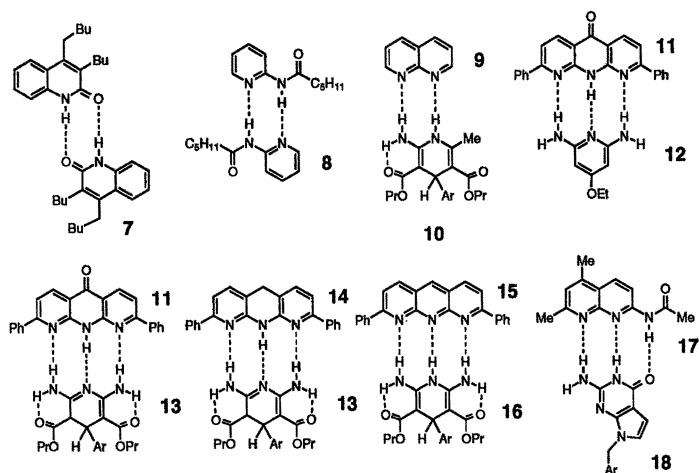
Figure 4. Doubly and triply hydrogen bonded complexes used in this study.

more general approach because they contain hydrogen bond donor and acceptor groups on their periphery in a limited number of arrangements. Indeed, the nine hypothetical structures in figure 1 are capable of recognizing any guest which contains at least one site with two or three adjacent hydrogen bond donor or acceptor groups. Furthermore, if the guest contains two or more such sites, then the appropriate structures in figure 1 might be linked to provide additional binding energy and increased specificity. The anti-malarial drug, trimethoprin (**6**), exemplifies this modular approach to host construction. The diaminopyrimidine nucleus contains hydrogen bonding functionality on three different edges, and would require a host with AD and ADA units (figure 3*b*).

Progress toward the goals outlined above is presented herein. Specifically, recognition units corresponding to all of the hypothetical structures in figure 1 were synthesized, and the stability of eight new multiply hydrogen bonded complexes were measured. The results allowed the importance of the various factors that control multiply hydrogen bonded complex stability to be evaluated.

## 2. Results

The actual compounds that were used in this study are shown in hydrogen bonded complexes with their complement in figure 4. Triply hydrogen bonded complexes studied in other laboratories are shown in figure 5. The sources of the compounds in figure 4 were reported previously (Murray & Zimmerman 1992; Zimmerman *et al.* 1993). Each of the compounds used in this study gave correct elemental analysis and had spectral data in full accord with the assigned structure. To ensure that compounds (**7**)–(**20**) were monomeric under the conditions used for binding studies, the ¹H NMR chemical shift of one or more N–H resonances was monitored as function of concentration. In each case the N–H resonance(s) showed small shifts (less than 0.1 ppm). While this negligible shift is often taken as evidence that compounds such as (**7**)–(**20**) are monomeric, the same result would obtain from a very strong aggregation. For this reason, molecular weights of compounds (**9**)–(**13**) and (**15**)–(**18**) were determined in chloroform-d using a Gonotec vapour pressure osmometer (VPO), Osmomat 070-SA from UIC, Inc. For (**10**)–(**13**) and (**15**)–(**17**), duplicate runs gave
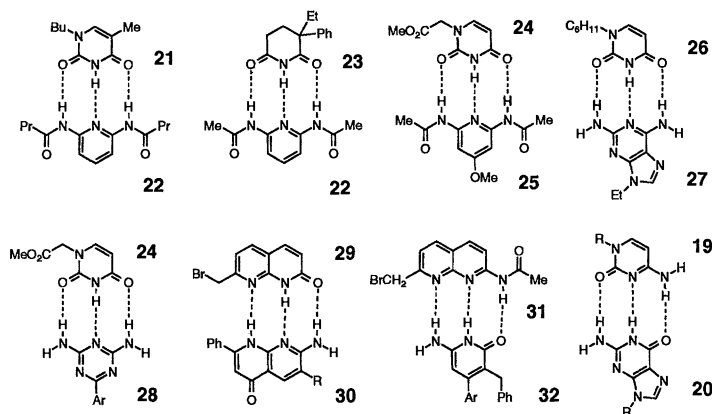
Figure 5. Triply hydrogen bonded complexes studied in other laboratories.

VPO molecular masses consistent with the compound being predominantly or entirely monomeric; the accuracy did not allow the presence of some self-association to be ruled out, but the $^1$H NMR results indicate negligible self-association. For (**18**) and (**20**) the VPO results indicated aggregation (dimer for (**18**), at least dimer for (**20**)) at concentrations just slightly above those used in the binding experiments. Thus, our previously reported association constants should be regarded as lower limits.

Meyer (1978) had reported that in dimethylformamide-d$_7$ (**16**) (Ar = 3-nitrophenyl) was entirely in the 1,4-dihydro form, while (**13**) (Ar = 2-nitrophenyl) existed in the 3,4-dihydro form exclusively. Thus, the 4-aryl substituent serves as a remote tautomeric switch that allows interconversion of DDD and DAD hydrogen bonding arrangements with a minimum change in structure. In chloroform-d, the solvent used in this study, the $^1$H-NMR spectrum of (**16**) showed it to be an approximately 67:33 mixture of 1,4-dihydro to 3,4-dihydro forms, while (**13**) remained fully in the 3,4-dihydro form. Only one diastereomer of (**13**) was seen in the $^1$H NMR, and this was tentatively assigned as the cis-isomer. It was found that 10 equivalents of (**15**) converted (**13**) (*ca.* 1 mM, CDCl$_3$) entirely into the 1,4 dihydro form. Likewise, 4 equivalents of (**11**) converted (**16**) (*ca.* 5 mM, CDCl$_3$) from a 67:33 mixture to a 44:56 mixture of 1,4-dihydro and 3,4-dihydro forms. These experiments support the assigned tautomeric structures for (**13**) and (**16**). Spectral analyses of the remaining compounds in figure 4 support the assigned tautomeric structures.

Complexation studies were performed by $^1$H NMR in chloroform-d; association constants were determined using standard methods (see table 1). The alternating hydrogen bonding motif (DAD·ADA) in (**11**)·(**12**), (**11**)·(**13**) and (**14**)·(**13**) afforded weak complexes with $K_{assoc}$ of 78 M$^{-1}$, 70 M$^{-1}$ and 65 M$^{-1}$ respectively. The DDA·AAD complexes of (**17**)·(**18**) was significantly more stable with $K_{assoc} \geqslant 10^4$ M$^{-1}$. Kyogoku (1969) had determined the strength of the GC base-pair, (**19**) (R = ribose)·(**20**) (R = ribose) as $10^4$–$10^5$ M$^{-1}$, but noted the very approximate nature of their measurements. In the current study, the association constant of the GC base-pair, (**19**) (R = Et)·(**20**) (R = ribose) was determined to be greater than $10^4$ M$^{-1}$, aggregation of (**20**) preventing a more accurate determination (*vide supra*).

The (**15**)·(**16**) complex was found to be unstable in chloroform-d. The $^1$H NMR indicated a slow reduction of (**15**) to (**14**), with concomitant oxidation of (**16**) to the corresponding pyridine. This reaction was inhibited by two equivalents of 1,8-bis(dimethylamino)naphthalene and the binding study was conducted under

these conditions. The (**15**)·(**16**) (DDD·AAA) complex was by far the tightest examined. Only approximately 15% of uncomplexed (**16**) was observed when the concentration of the 1:1 complex was about $2 \times 10^{-4}$ M. Accounting for the tautomeric equilibrium constant, the association constant can be estimated to be larger than or equal to $10^5$ M$^{-1}$.

Only three doubly hydrogen bonded complexes were investigated, but these also exhibited a wide range of stabilities. Amidopyridine (**8**) dimerized weakly with $K_{\mathrm{dimer}} = 2$ M$^{-1}$. Quinolone (**7**) self-associated with $K_{\mathrm{dimer}} = 46$ M$^{-1}$ and the association constant measured for (**9**)·(**10**) was 260 M$^{-1}$.

## 3. Discussion

What controls the strength of hydrogen bonded complexes? Many factors govern hydrogen bond strength, including geometry. It is generally believed that linear hydrogen bonds are stronger than those that are bent, although the energy required to bend a hydrogen bond is significantly less than for a covalent bond. The acidity and basicity of the hydrogen bond donor and acceptor groups is also of importance. For example, phenol complexes 4-cyanopyridine and 4-dimethylaminopyridine in carbon tetrachloride with association constants of 9 M$^{-1}$ and 550 M$^{-1}$, respectively (Hopkins *et al.* 1978). In our work a smaller dependence on acidity was found where a series of 4-substituted benzoic acids complexed 9-propyladenine in chloroform-d with a Hammett $\rho = -0.3$ (Zimmerman 1993).

Additional factors need to be considered when multiply hydrogen bonded complexes are considered. If within a group of complexes, the arrangement and nature (e.g. acidity/basicity) of the donor and acceptor sites is identical, or at least very similar, than it is possible that the relative complexation strengths will depend only on the number of hydrogen bonds within each complex. For example, Schneider (1989) has reported a linear correlation between the complexation energy and number of hydrogen bond for a series of complexes between diamidopyridines hosts and imide or barbiturate guests. This type of correlation is clearly absent in the present study. Indeed, the association constants for the triply hydrogen bonded complexes span more than three orders of magnitude, and the doubly hydrogen bonded complexes, two orders of magnitude (table 1). Furthermore, complex (**7**)·(**7**) with two hydrogen bonds is more stable than six of the complexes (e.g. (**11**)·(**12**)) that contain three hydrogen bonds.

In these cases where the arrangement of donor and acceptor groups vary between complexes, two opposing factors emerge. In the context of nucleotide base-pairing, Saenger (1984) has pointed out that when hydrogen bonding sites alternate as in the A–T base-pair, π-cooperativity allows each hydrogen bond to strengthen the other figure 6*a*. While this concept is intuitively appealing, the effect has not been quantified. The other factor that needs to be considered is the number of attractive and repulsive secondary electrostatic interactions (*vide infra*).

Jorgensen & Pranata (1990; see also Pranata *et al.* 1991), in examining four known triply hydrogen bonded complexes, noted that the alternating arrangement of hydrogen bond donor and acceptor groups in (**21**)·(**22**) and (**26**)·(**27**) (ADA·DAD) correlated with a relatively low association constant ($K_{\mathrm{assoc}} \approx 10^2$ M$^{-1}$), while the arrangement in (**31**)·(**32**), (**19**)·(**20**) (DDA·AAD) resulted in association constants of *ca.* $10^4$ M$^{-1}$. The explanation for this observation is that the ADA·DAD hydrogen bonding arrangement contains four repulsive secondary electrostatic interactions

Table 1. *Association constants for various doubly and triply hydrogen bonded complexes*

(At 298 K in chloroform-d. Method A: titration with one component at fixed concentration. Method B: dilution of 1:1 complex (or dimer).)

| complex | type | method used | $K_{assoc}$ M$^{-1}$ | $-\Delta G^0_{298}$ kcal mol$^{-1}$ | reference |
|---|---|---|---|---|---|
| 7·7 | AD·DA | B | 46 | 2.3 | this work |
| 8·8 | AD·DA | B | 2 | 0.4 | this work |
| 9·10 | AA·DD | A | 260 | 3.3 | this work |
| 11·12 | DAD·ADA | A | 78 | 2.6 | this work |
| 11·13 | DAD·ADA | A | 70 | 2.5 | this work |
| 14·13 | DAD·ADA | A | 65 | 2.5 | this work |
| 21·22 | DAD·ADA | A | 90 | 2.7 | Hamilton & Van Eugen (1987) |
| 23·22 | DAD·ADA | A | 200 | 3.1 | Scheider (1989) |
| 24·25 | DAD·ADA | A | 900 | 4.0 | Park et al. (1991) |
| 26·27 | DAD·ADA | A | 170 | 3.0 | Kyogoku et al. (1967) |
| 24·28 | DAD·ADA | A | 670 | 3.9 | Park et al. (1991) |
| 29·30 | DAD·ADA | A | 440 | 3.6 | Kelly et al. (1990) |
| 17·18 | DAA·ADD | B | $\geqslant 10^4$ | 5.4 | this work |
| 19·20 | DAA·ADD | B | $\geqslant 10^4$ | 5.5 | this work |
| 31·32 | DAA·ADD | B | $1.7 \times 10^4$ | 5.8 | Kelly et al. (1989) |
| 15·16[a] | DDD·AAA | B | $\geqslant 10^5$ | $\geqslant 7$ | this work |

[a] In the presence of 2 molar equivalents of 1,8-bis(dimethylamino)-naphthalene (see text).
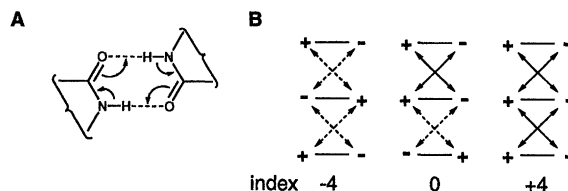
Figure 6. (a) $\pi$-Cooperativity in an A–T base-pair. (b) Jorgensen's secondary electrostatic analysis applied to triply hydrogen bonded complexes.

(figure 6b). In contrast, the two repulsive secondary electrostatic interactions in the DDA·AAD arrangement are offset by two attractive interactions. Jorgensen & Pranata (1990; see also Pranata *et al.* 1991) further pointed out that a DDD·AAA complex geometry, having all attractive secondary electrostatic interactions, would have even a higher stability.

Data collected in this study and from the literature are consistent with the Jorgensen model. Complexes with the ADA·DAD hydrogen bonding arrangement (nine in total) are weakest with $\Delta G_{298}^0$ ranging from 2.5 to 4.0 kcal mol$^{-1}$. Complexes with a DDA·AAD arrangement (three in total) have $\Delta G_{298}^0$ of $\geqslant 5.4$ kcal mol$^{-1}$, while the DDD·AAA arrangement in complex (15)·(16) is associated with a $\Delta G_{298}^0 > 7$ kcal mol$^{-1}$. Most striking is the very high stability of the (15)·(16) complex, in comparison to the weakness of complex (13)·(14) ($\Delta\Delta G^0 > 4$ kcal mol$^{-1}$), the latter containing the same number of hydrogen bonds and a very similar structure. The results with the doubly hydrogen bonded complexes are also consistent with the secondary electrostatic interaction model, but here there are far fewer examples of the two possible arrangements. Currently, we are working to obtain more accurate stability constants for the DDA·AAD and DDD·AAA complexes to quantify any advantage that either arrangement might confer.

## 4. Conclusion

Heterocyclic compounds were synthesized containing two or three adjacent hydrogen bond donor and acceptor groups in all possible arrangements. Each compound contains several pendant functional groups that allow them to be interconnected or allow the solubility properties to the modulated. The syntheses are sufficiently versatile to allow easy incorporation into more complicated supra-molecular assemblies. In this regard the AAD·DDA hydrogen bonding arrangement is particularly appealing because it is strong and affords a high degree of geometrical control absent in the symmetrical AAA·DDD and ADA·DAD hydrogen bonding arrangements.

Although many factors contribute to the strength of the complexes described herein, and our results do not confirm any one model, the highly variable stability of the doubly and triply hydrogen bonded complexes examined do correlate with the number of favourable secondary electrostatic interactions. Future efforts must be directed at quantifying the strength of the primary hydrogen bonds, and in determining the role of solvation. For example, it is possible that the AAA and DDD hydrogen bonding arrays are difficult to solvate effectively. We are currently addressing these issues and the results will be reported in due course.

# References

Hamilton, A. D. & Van Engen, D. 1987 Induced fit in synthetic receptors: nucleotide base and recognition by a 'molecular hinge'. *J. Am. chem. Soc.* **109**, 5035–5036.

Hopkins, H. P. Jr, Alexander, C. J. & Ali, S. Z. 1978 A quantitative comparison of the relative hydrogen bonding basicities and the gas-phase proton affinities of substituted pyridines. *J. phys. Chem.* **82**, 1268–1272.

Jorgensen, W. L. & Pranata, J. 1990 Importance of secondary interactions in triply hydrogen bonded complexes: guanosine-cytosine vs uracil-2,6-diaminopyridine. *J. Am. chem. Soc.* **112**, 2008–2010.

Kelly, T. R., Bridger, G. J. & Zhao, C. 1990 Bisubstrate reaction templates. Examination of the consequences of identical versus different binding sites. *J. Am. chem. Soc.* **112**, 8024–8034.

Kelly, T. R., Zhao, C. & Bridger, G. J. 1989 A bisubstrate reaction template. *J. Am. chem. Soc.* **111**, 3744–3745.

Kyogoku, Y., Lord, R. C. & Rich, A. 1967 The effect of substituents on the hydrogen bonding of adenine and uracil derivatives. *Proc. natn. Acad. Sci. U.S.A.* **57**, 250–257.

Kyogoku, Y., Lord, R. C. & Rich, A. 1969 An infrared study of the hydrogen-bonding specificity of hypoxanthine and other nucleic acid derivatives. *Biochim. Biophys. Acta* **179**, 10–17.

Meyer, H., Bossert, F. & Horstmann, H. 1978 Kondensation von aldehyden mit endiaminocarbonylverbindungen *Liebigs Ann. Chem.* 1476–1482.

Murray, T. J. & Zimmerman, S. C. 1992 New triply hydrogen bonded complexes with highly variable stabilities. *J. Am. chem. Soc.* **114**, 4010–4011.

Park, T. K., Schroeder, J. & Rebek J. Jr 1991 New molecular complements to imides. Complexation of thymine derivatives. *J. Am. chem. Soc.* **113**, 5125–5127.

Pranata, J., Wierschke, S. G. & Jorgensen, W. L. 1991 OPLS potential functions for nucleotide bases. Relative association constants of hydrogen-bonded base pairs in chloroform. *J. Am. chem. Soc.* **113**, 2810–2819.

Saenger, W. 1984 In *Principles of nucleic acid structure*, pp. 117–119. New York: Springer-Verlag.

Schneider, H.-J., Juneva, R. K. & Simova, S. 1989 Solvent and structural effects on hydrogen bonds in some amides and barbiturates. An additive scheme for the stability of corresponding host-guest complexes. *Chem. Ber.* **122**, 1211–1213.

Zimmerman, S. C. 1993 Rigid molecular tweezers as hosts for the complexation of neutral guests. In *Topics in current chemistry* (ed. E. Weber), ch. 2. Berlin: Springer-Verlag.

Zimmerman, S. C. & Duerr, B. F. 1992 Controlled molecular aggregation 1. Cyclic trimerization via hydrogen bonding *J. Org. Chem.* **57**, 2215–2217.

Zimmerman, S. C., Baloga, M. H., Duerr, B. F., Fenlon, E. E. & Murray, T. J. 1993 Multiply hydrogen bonded complexes for constructing new supramolecular assemblies *Polym. Prep. (Am. chem. Soc. Div. Polym. Chem.)* **34**, 94–95.