Effect of alkyl chain length on self-preorganization of artificial nucleobase receptors

Yasuhisa Kuroda,* Juha M. Lintuluoto and Hisanobu Ogoshi

Department of Synthetic Chemistry and Biological Chemistry, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Flexible receptors, having two uracil moieties connected through alkyl spacers of different lengths, exhibit intramolecular equilibria between two states in $CDCl_3$, CD_2Cl_2 and $(CDCl_2)_2$. The intramolecularly hydrogen bonded state (closed form) is shown to bind more strongly to adenine base by Watson–Crick and Hoogsteen base pairing than the non-hydrogen bonded state (open form) above 298 K. The closed forms of flexible receptors take advantage of their small entropy loss upon complexation (-2 to -37 J K^{-1} mol^{-1}); this is understood to be due to their self-preorganization, where closed forms are structurally similar to their complexed forms. Flexible receptors show high selectivity towards the adenine base; no binding towards other natural nucleobases is observed in ^1H NMR studies. The observed selectivity shows that a hydrogen bonding system containing a base triplet is potentially a superior molecular recognition site for nucleobases compared with duplex base pair systems.

Introduction

It has been emphasized that, for effective molecular recognition, artificial receptors should be so designed that the shapes of their binding sites are as complementary for guest molecules as possible prior to complex formation. This concept, so-called 'preorganization',[1] has been shown to be very effective in the design and construction of biomimetic host molecules and relatively rigid (preorganized) molecules have been successfully utilized as basic skeletons for artificial receptors.[2] Cram postulated that preorganization of receptors plays an important role in setting up geometric and electronic complementarity between host and guest molecules, and it can enhance binding by a factor of $>10^{12}$.[2a] On the other hand, it has also been recognized that flexibility is a very important factor in many biological processes, where molecular recognition rather acts as some kind of trigger for subsequent reactions. Such flexibility of biological molecules is found in a wide range of natural processes which include local induced fitting in enzyme active sites, global conformational change of allosteric proteins, double- and superhelix formation and rewinding of polynucleotides.[3–9]

Mimics of these biologically active molecules should perhaps be constructed based on more flexible materials which are capable of dynamic conformational change. However, flexibility in synthetic receptors is often found to be disadvantageous because of the resulting entropically unfavourable complex formation with substrate.[2a,g,10] Thus, to satisfy these two requirements, *i.e.* to have rigidity to keep complementarity during recognition, and to have flexibility to promote further chemical processes, it is necessary to accumulate chemical information on methodology for assembling and organizing flexible chemical elements that can be held in a low entropic state. For example, Rebek demonstrated an interesting example of successive binding by partially flexible receptors which could exhibit intramolecular hydrogen bonding prior to complex formation.[11] The observations suggest the attractive possibility of using self-preorganization to lower the entropy of the flexible receptor by using non-covalent weak interactions such as hydrogen bonds.

Recently, we reported a simple flexible receptor which not only binds adenine derivatives but also forms intramolecular hydrogen bonds to produce a significantly organized form (closed form) in addition to the original random form (open form).[12] Thermodynamic analyses of the complex formation system indicate that the entropy loss during adenine recognition starting from the closed form is much smaller than that from the closed one. The observation is rationalized in view of base triplet complex formation where the conformation of the receptor is expected to be similar to that of the closed form. Results clearly show that closing a recognition site with intramolecular hydrogen bonding of a flexible host molecule enhances complex formation with guest molecules at high temperatures where entropy loss during complex formation seriously retards the bimolecular process. Self-preorganization minimizes entropy loss upon complexation. Although our receptors are very primitive compared to natural receptors, this study gives reasonable mimics for thermodynamics of complex formation of natural receptors which form highly organized recognition sites in spite of their original random nature.

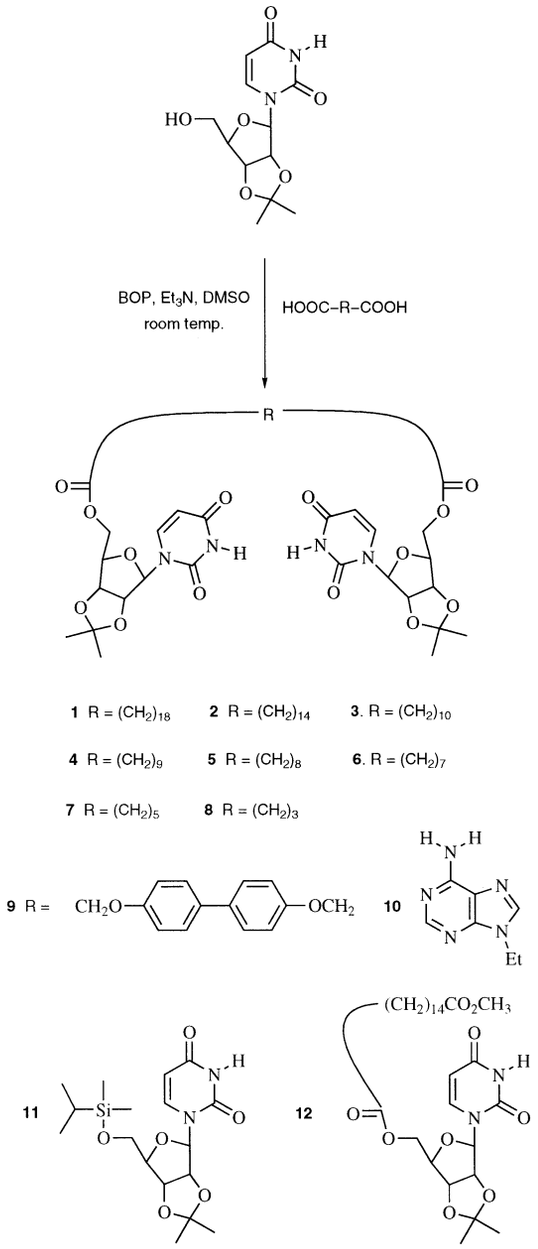
Results and discussion

Synthesis of receptors 1–9

We obtained flexible receptors by one-step reaction of 2′,3′-*O*-isopropylidene uridine with the corresponding diacid using benzotriazolyl-*N*-oxytris(dimethylamino)phosphonium hexafluorophosphate[13] as an ester coupling reagent (see Scheme 1). Most of the products are amphiphilic due to the polar uridine part and the non-polar alkyl chain and, therefore, are slightly soluble into water. This property sometimes made it difficult to purify the receptors by the usual work-up procedures and lowered their yields. Thus, final purification of all compounds was carried out by means of preparative HPLC. In contrast, esterification and purification to obtain **9** was easy due to its high solubility in organic solvents and the low lipophilic character of the biphenyl moiety.

Variable temperature ^1H NMR study—intramolecular binding of receptors

As reported previously, the present type of receptors are expected to form intramolecular hydrogen bonds between two uracil bases. Since the chemical shift of the uracil imino protons shows practically no solvent dependence on halocarbon solvents such as $CDCl_3$, CD_2Cl_2 and $(CDCl_2)_2$, it is possible to perform variable temperature NMR experiments over a wide range of temperatures using these solvents. Fig. 1 presents a typical example of the temperature dependences of the ^1H signals of imino protons of **2** and **7** at -90 to $+140$ °C, where



Scheme 1

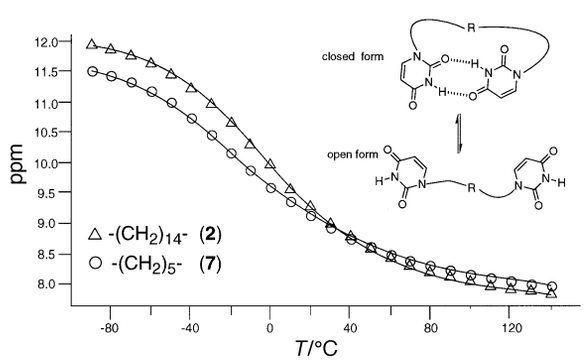


Fig. 1 Temperature dependences of ^{1}H signals of imino protons of **2** (△) and **7** (○) in CD_2Cl_2 (−90 to 20 °C), $CDCl_3$ (−60 to 60 °C) and $(CDCl_2)_2$ (0 to 140 °C). **[2]** = 7.33 and 0.73 mM. **[7]** = 4.80 and 0.48 mM. The chemical shifts observed in these solvents are identical within experimental error limits (<0.1 ppm) in each overlapped temperature range.

CD_2Cl_2, $CDCl_3$ and $(CDCl_2)_2$ were used at −90 to +20, −60 to +60 and 0 to +140 °C, respectively.[14] The data, in spite of solvent change, clearly show smooth sigmoidal temperature dependence within this temperature range which may be attributed to a single equilibrium process. Similar sigmoidal dependences were obtained for all receptors examined in this work. Furthermore, observed sigmoidal behaviour is independent of receptor concentration except for **8**. For example, the receptor **7** at a concentration of 4.80 and 0.48 mM shows the same temperature dependence within experimental errors. Saturation of chemical shifts at both termini of the sigmoidal curve and concentration-independence of chemical shifts observed for the present receptor molecules except **8** suggest that at low temperatures, intramolecular hydrogen bonds are strong and these receptors exist mainly in the cyclic form (closed form), and at high temperatures, are free from intra- and inter-molecular hydrogen bonds (open form). Chemical shifts extrapolated to both ends may correspond to those of the closed and open forms, which are determined in the next section.

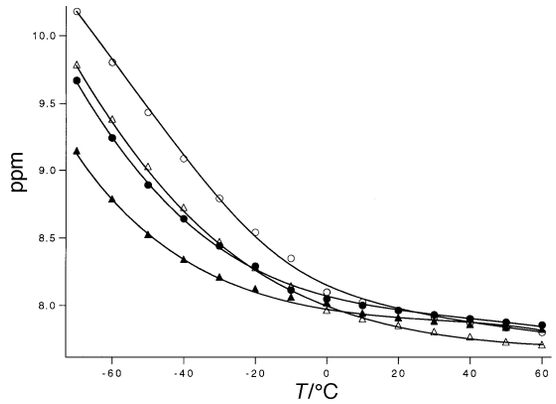


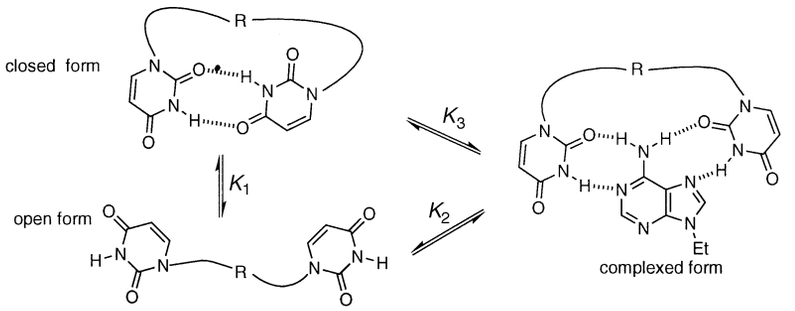
Fig. 2 Concentration and temperature dependences of ^{1}H signals of imino protons of **11** (4.45 mM, ●), (2.23 mM, ▲) and **12** (3.18 mM, ○), (1.24 mM, △) in CD_2Cl_2 (−70 to 0 °C) and $CDCl_3$ (−20 to 60 °C)

In contrast, chemical shifts of uracil imino protons of **8** showed significant concentration dependence which is possibly explained by partial dimerization to form a box-like dimer structure.[15] For example, tenfold dilution from 7.33 to 0.73 mM caused *ca.* 0.3 ppm upfield shifts for uracil imino protons below 10 °C. The observation suggests that the observed sigmoidal temperature dependence should be attributed to parallel equilibrium processes of intra- and inter-molecular hydrogen bond formation. The molecular model indicates that the methylene chain of **8** seems to be too short to construct the compact closed form similar to that of other receptors and may result in significant conformational strain in its closed form. Such strain generated in the closed form of **8** may be reflected in an unusually high upfield shift of the extrapolated chemical shift (11.17 ppm) at the low temperature end of the sigmoidal temperature-dependent curve of **8** compared to the others (11.55–12.06 ppm, see Table 1).

On the other hand, the situation is quite different for compounds **11** and **12** which mainly have only one hydrogen bonding function for bimolecular dimerization process. As shown in Fig. 2, these compounds showed strong concentration dependence of the chemical shift of the imino protons at low temperatures; twofold dilution results in *ca.* 1 ppm upfield shift of imino signals below −50 °C. Furthermore, it should be noted that the chemical shift changes observed for **11** and **12** in the experimental temperature range are much smaller than those shown in Fig. 1 and the temperature *vs.* chemical shift plot seems to be non-sigmoidal, which indicates a much weaker bimolecular complex formation of **11** and **12**.

Complexation with 9-ethyladenine

The chemical shifts of the imino protons in the present receptors are also significantly affected by addition of adenine derivatives. The observations indicate that the receptors form complexes with adenine through hydrogen bonds containing



Equilibrium constant	Definition
K_1	[open] / [closed]
K_2	[complexed] / [open][9-EtAde]
K_3	[complexed] / [closed][9-EtAde]

Scheme 2

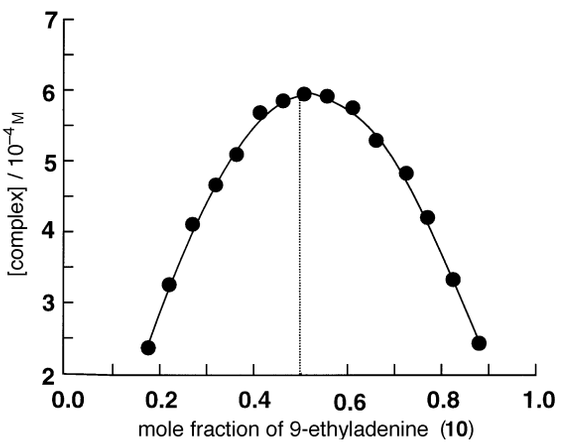


Fig. 3 Job plot of the recognition system containing **4** and **10** at 10 °C. [**4**] + [**10**] = 5.40 mM

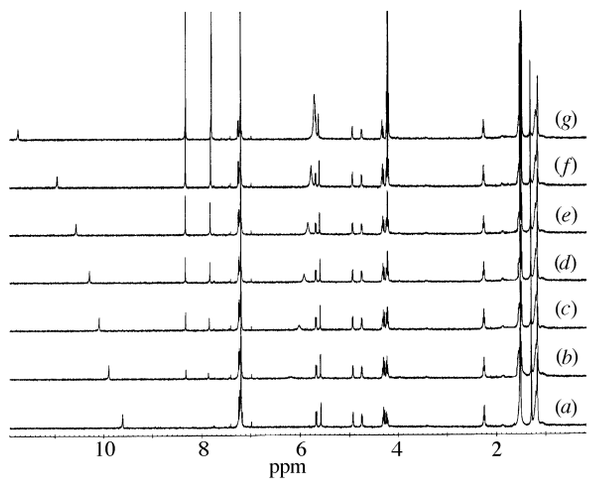


Fig. 4 500 MHz ^{1}H NMR titration spectra of **2** (0.73 mM) and **10** at 20 °C in $CDCl_3$. (*a*) 0, (*b*) 2.7, (*c*) 5.5, (*d*) 12.3, (*e*) 19.2, (*f*) 32.9 and (*g*) 60.3 equiv. of **10**.

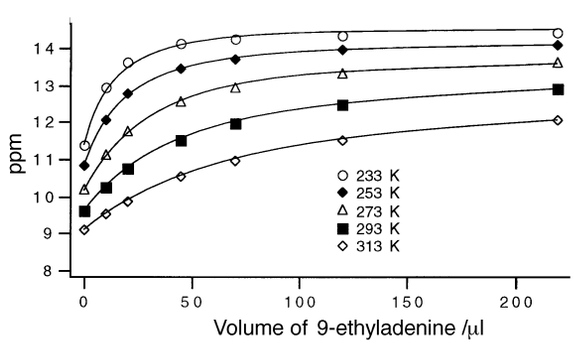


Fig. 5 Plots of chemical shifts of uracil imino protons of **4** (2.19 mM), $V_{initial}$(**4**) = 500 μl *vs.* volume of added solution of **10** (95.0 mM) at different temperatures in $CDCl_3$. The solid lines are theoretical curves obtained from optimized thermodynamic parameters.

the imino protons of uracil moieties. To clarify characteristics of this complex formation, we carried out Job plot experiments and NOE measurements. As shown in Fig. 3, the Job plot for the system of **4** and 9-ethyladenine **10** demonstrates 1:1 complex formation between the host and guest pairs. It should be noted that clear 1:1 stoichiometry is observed, even though the present artificial receptors have two binding sites of uracil moieties. Moreover, NOE experiments for **5** and **9** with 9-ethyladenine revealed simultaneous Watson–Crick and Hoogsteen binding modes. Selective irradiation of the receptor's uracil imino proton resulted in clear NOE of base protons H2 and H8 of 9-ethyladenine, and no other peaks were enhanced upon this irradiation. Thus, these observations confirm previous conclusions that the present receptors form the U·A·U type of Watson–Crick and Hoogsteen base triplet with 9-ethyladenine and the contribution of other equilibrium processes such as 1:2 complex formation is negligible under the present experimental conditions. Thus, we can identify three different states of the present artificial receptors: open form, closed form and U·A·U type complexed form. Scheme 2 represents possible equilibria and definitions for binding constants for these intramolecular binding and complexation processes with 9-ethyladenine.

Based on this entire equilibrium scheme, we can determine thermodynamic parameters for these processes using the temperature-dependent data shown in Fig. 1 and the titration data for adenine binding at different temperatures. Typical examples of ^{1}H NMR titration and plots of chemical shifts of the imino proton at five different temperatures are shown in Figs. 4 and 5, respectively.

To obtain thermodynamic parameters from these data, we employed a nonlinear curve fitting procedure which directly optimizes necessary thermodynamic parameters so as to minimize the square sum of residuals between observed chemical shifts and the theoretical values calculated for the three-state equilibrium model shown in Scheme 2 (see Experimental section for details). The method gives excellent agreement between experimental and theoretical chemical shifts under all conditions employed in this work as shown in Figs. 1 and 5. The enthalpy and entropy values thus determined are shown in Tables 1 and 2.

Since separate evaluation of intra- and inter-molecular complexation of **8** is difficult, the thermodynamic parameters for **8** were not determined in this work.

Effect of length of flexible spacer upon intramolecular binding

Although each experimental error of the estimated thermodynamic parameters is still not necessarily small enough for elucidation of detailed mechanisms, comparison of all data show clear tendencies in enthalpy and entropy change values of the intramolecular binding process towards the number of methylene groups in their spacer.

First, it should be noted that, in spite of a more than threefold

Table 1 Thermodynamic parameters for K_1 (ring opening) process. ΔH (kJ mol^{-1}), ΔS (J K^{-1} mol^{-1})

Entry	Receptor	δ_{open} [a] (ppm)	Δ_{closed} [a] (ppm)	ΔH [c]	ΔS [c]
1	**1** $(CH_2)_{18}$	7.62	12.01	19 ± 2	70 ± 8
2	**2** $(CH_2)_{14}$	7.45	12.08	17 ± 2	62 ± 9
3	**2** $(CH_2)_{10}$	7.49	12.07	18 ± 2	61 ± 5
4	**4** $(CH_2)_9$	7.44	12.21	16 ± 2	55 ± 6
5	**5** $(CH_2)_8$	7.42	12.08	16 ± 2	54 ± 6
6	**6** $(CH_2)_7$	7.40	12.16	15 ± 2	53 ± 7
7	**7** $(CH_2)_5$	7.49	11.75	15 ± 2	52 ± 9
8	**9**	7.28	10.28 [b]	13 ± 2 [d]	49 ± 7 [d]

[a] Calculated. [b] Observed in CD_2Cl_2 at −20 °C. [c] From nonlinear least-squares analysis. [d] Above −20 °C.

Table 2 Thermodynamic parameters for K_2 (*via* open form) and K_3 (*via* closed form) processes in $CDCl_3$. Guest = 9-ethyladenine. ΔH (kJ mol^{-1}), ΔS (J K^{-1} mol^{-1}), δ_{calc} (ppm)

Entry	Receptor	ΔH (K_2) [a]	ΔS (K_2) *a*	ΔH (K_3) [a]	ΔS (K_3) [a]	δ_{calc}compl.
1	**1** $(CH_2)_{18}$	−33 ± 4	−72 ± 17	−14 ± 2	−2 ± 6	14.42
2	**2** $(CH_2)_{14}$	−33 ± 4	−74 ± 17	−16 ± 2	−11 ± 8	14.45
3	**3** $(CH_2)_{10}$	−33 ± 4	−68 ± 17	−15 ± 2	−7 ± 5	14.49
4	**4** $(CH_2)_9$	−34 ± 4	−74 ± 12	−18 ± 2	−19 ± 6	14.42
5	**5** $(CH_2)_8$	−32 ± 4	−69 ± 12	−16 ± 2	−15 ± 6	14.63
6	**6** $(CH_2)_7$	−37 ± 4	−84 ± 14	−23 ± 2	−31 ± 7	14.45
7	**7** $(CH_2)_5$	−39 ± 6	−89 ± 20	−24 ± 3	−37 ± 10	14.59
8	**9**	−38 ± 4 [b]	−85 ± 17 [b]	−25 ± 2 [b]	−36 ± 8 [b]	14.12

[a] From nonlinear least-squares analysis. [b] Above −20 °C.

difference in the spacer chain length from $-[CH_2]_5-$ to $-[CH_2]_{18}-$, observed enthalpy and entropy changes for the intramolecular process of these receptors are relatively small as shown in Table 1. Similar values of chemical shifts, δ_{closed}, estimated for closed forms also indicate the structural similarity around the nucleobase moieties of closed forms of these receptors. The observation indicates that the five-methylene chain essentially satisfies the minimum structural requirement to generate the closed (preorganized) form. The receptor **8** with a three-methylene chain, however, no longer meets these criteria and contribution of intermolecular complex formation is not negligible as mentioned in the previous section. At the same time, detailed comparison of observed parameters shown in Table 1 reveals an interesting tendency that the more methylene groups a spacer contains, the larger is the enthalpy loss and the bigger is the entropy gain for ring opening process. The values obtained for the longest one **1** ($[CH_2]_{18}$, $\Delta H = 19$ kJ mol^{-1}, $\Delta S = 70$ J K^{-1} mol^{-1}) in the ring opening process suggest that this receptor can obtain the most preorganized structure (closed form) compared to its random form (open form). The length of the five-methylene spacer in the receptor **7** would be a critical length to form the most stable closed form and result in some distortion of the intramolecular hydrogen bond system of **7**. Such critical properties of the receptor **7** are supported by the flatter shape of the sigmoidal curve in Fig. 1 and its chemical shift, $\delta_{closed} = 11.55$, for the closed form which shows a relatively high upfield shift compared with those of **1**–**6**, $\delta_{closed} = 11.97 \pm 0.1$. Furthermore, the thermodynamic parameters obtained for **7** are interestingly close to the values obtained for the receptor **9** having a biphenyl spacer which possesses more rigidity in its spacer and which would have more strain in its perfect closed form than the methylene group spacers in receptors **1**–**7**.

Complexation *via* open form *vs.* closed form

There are two thermodynamically feasible processes to bind adenine, *i.e.* complex formation processes *via* open form (K_2 process) and closed form (K_3 process). The thermodynamic parameters listed in Table 2 reveal that the K_2 process generally accompanies a much larger enthalpy gain and entropy loss compared with the K_3 process. The larger enthalpy gain observed for the K_2 process may correspond to the formation of four hydrogen bonds during the equilibrium, while enthalpy change for two newly formed hydrogen bonds during the K_3 process is reflected in the smaller enthalpy gain for this equilibrium. Small entropy loss in the K_3 process is understood by the fact that alkyl chains of receptors are already fixed well in the closed form and the structural flexibility of the closed form is similar to that of the complexed form. The receptors with long alkyl chains show remarkably small entropy loss for complexation *via* the closed form and, for example, the observed value for **1** is negligible (−2 ± 6 J K^{-1} mol^{-1}). Molecular mechanism calculations [16] suggest that the smaller entropy loss for the K_3 process of long spacer receptors is actually due to small differences in structural flexibility between the closed form and the complexed form (Fig. 6). The distance between two uracil imino protons in the closed form of **2** is only *ca.* 4 Å shorter than in the complexed form when the total length of the spacer is *ca.* 23 Å for $-[CH_2]_{14}-$.

Thus, such characteristic behaviour of the closed form of which adenine binding is less entropically controlled makes the closed form a better receptor for adenine bases than the open at higher temperatures around room temperature, where the unfavourable entropic contribution for complex formation becomes more important. It is also evident from the data shown in Table 2 that observed enthalpy and entropy changes for the K_3 process are affected more by the length of the methylene spacer compared to those of the K_1 and K_2 processes. The shortest receptor **7** exhibits entropy and enthalpy values rather close to those of **9** which has the most rigid spacer of the receptors used in this work. This indicates, again, that the closed form of **7** is more unstable than that of the longer receptors and is expected to have some strain in its structure. The release of such strain during complex formation with adenine may result in the observed relatively large enthalpy gain for the K_3 process of **7**, though the energy gain is largely compensated for by the relatively large entropy loss. Larger entropy losses for the short spacer receptors show that complexation rigidifies all methylene groups in the spacer. ^{1}H NMR titration observations at −40 °C also support such rigidification of the short methylene spacer upon complexation. For example, the peak pattern for the methylene groups of **7** adjacent to the carbonyl groups in the spacer at *ca.* 2.3 ppm split strongly upon adenine addition as shown in Fig. 7; this effect is not found for longer spacer receptors (see Fig. 4). Thus, the more flexible receptors, such as **1** in this work, do not necessarily have poor recognition ability if the self-

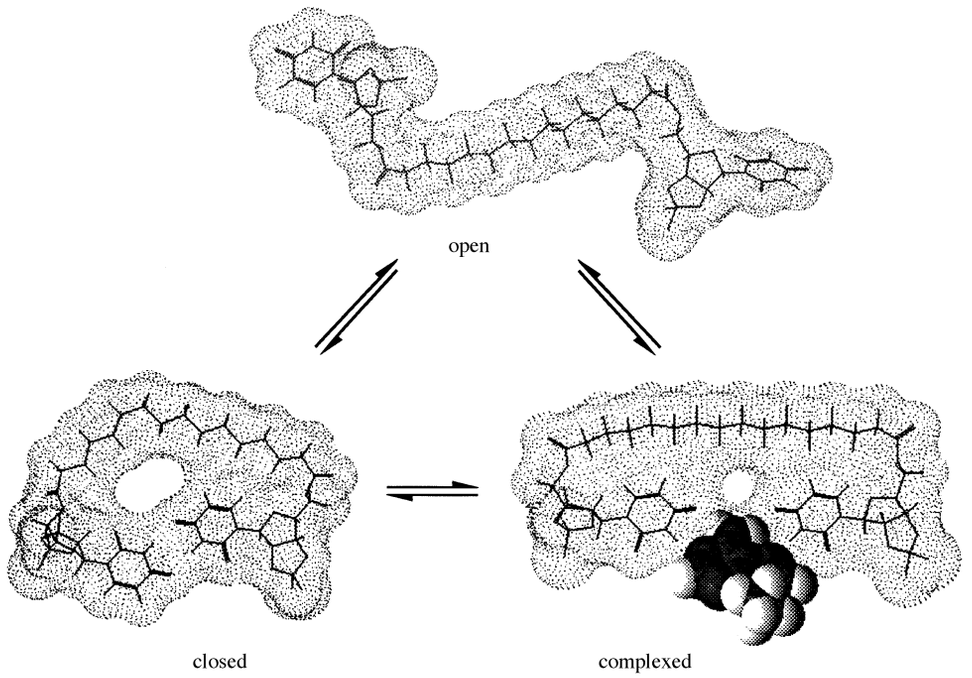


Fig. 6 Energy minimized molecular models of three states of **2**

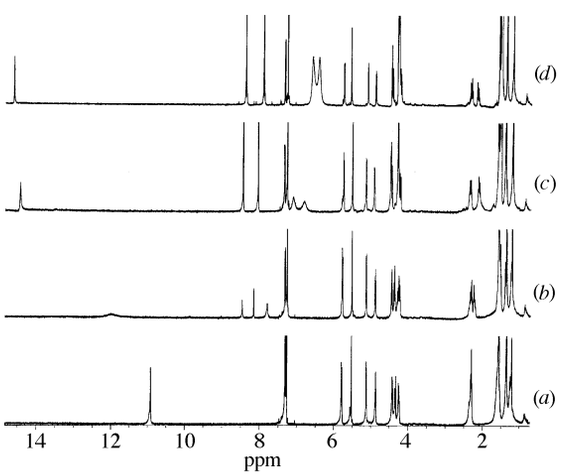
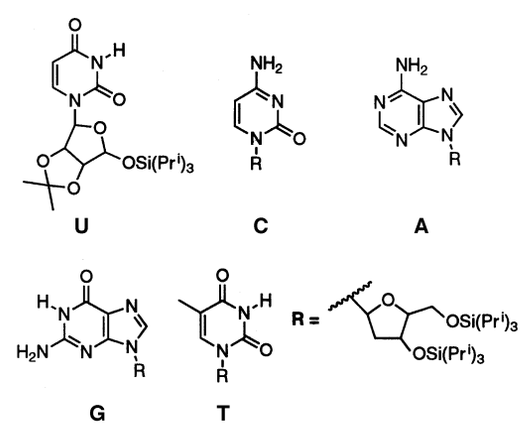


Fig. 7 500 MHz ^{1}H NMR titration spectra of **7** (4.77 mM) and **10** at −40 °C in $CDCl_3$. (*a*) 0, (*b*) 0.4, (*c*) 1.6 and (*d*) 4.2 equiv. of **10**.



Silylated nucleosides used in this work

preorganized forms, such as the closed form, can participate in the recognition process.

Selectivity of flexible receptors

Closing the binding site and performing self-preorganization, flexible receptors have higher selectivity than uridine itself towards adenine base in certain conditions.[17,18] To establish the selectivity of receptors towards other natural nucleosides, complexation between the long spacer receptor **2** and silylated adenosine **A**, thymidine **T**, guanosine **G** and cytidine **C** derivatives[19–21] was investigated. On addition of up to 23 equivalents of each nucleobase at 10 °C into the receptor solution, only adenine derivatives **10** and **A** showed clear complex formation whereas other nucleobase derivatives **T**, **G** and **C** showed practically no signs of complex formation judged by the chemical shift changes of a receptor's uracil imino protons (less than 0.05 ppm) as shown in Fig. 8.

On the other hand, the uracil base is known to be capable of forming 'wobble' type mismatched base pairs with other nucleobases.[22] For example, **U** complexed with **T** and indeed

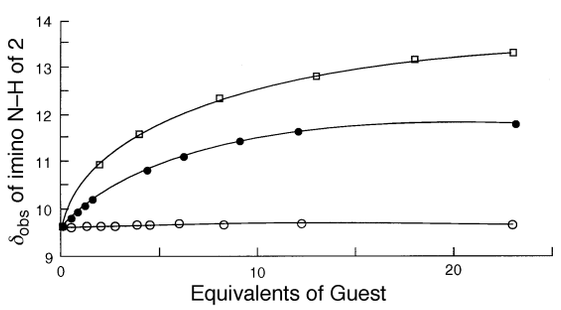


Fig. 8 Plots of chemical shifts of uracil imino protons of **2** *vs.* amounts of added **10** (□), **A** (●), **T** (○), **G** (○) and **C** (○) in $CDCl_3$ at 10 °C

uridine derivative **U**, which does not possess a closed binding site, showed weak but clear binding towards nucleobase derivatives **T**, **G** and **C**. For example, **U** complexed with **T** where $\Delta H = -9$ kJ mol^{-1}, $\Delta S = -20$ J K^{-1} mol^{-1} and $\Delta\delta_{calc}$ N–H = 2.42 ppm, which corresponds to *ca.* 4 dm^3 mol^{-1} of association constant at 10 °C. Although this type of complex formation, of course, is expected for systems containing the open form of receptors **1**–**9**, it should not be significant judging from this association constant. The observed selectivity of these flex-

ible receptors shows that the hydrogen bonding system containing base triplet is potentially a much superior recognition site for the nucleobase compared with duplex base pair systems which form a complex *via* a simple bimolecular process.

Conclusions

Flexible receptors **1**–**7**, which are also able to perform intramolecular preorganization by hydrogen bonding, can bind adenine derivatives effectively. Both their open and closed forms recognize guest molecules with similar effectiveness. However, the contents of the enthalpy and entropy terms for these recognition processes are quite different from each other. At higher temperatures, the binding process starting from the closed form becomes more favourable due to its less entropically controlled nature. Thus, the present results reveal the possibility of constructing flexible receptors using the self-preorganization process. It should be noted that effects of self-preorganization are evident, even when the simple methylene chain, which is one of the most flexible elements, is used as the spacer, though their enthalpy and entropy terms are still significantly compensatory.

Experimental

Apparatus

^{1}H NMR measurements were performed on a JEOL ALPHA 500 MHz spectrometer. Chemical shifts are relative to residual solvent peak, 7.25 ppm for $CHCl_3$, 5.97 ppm for $CHCl_2CHCl_2$ and 5.33 ppm for CH_2Cl_2. *J* values are in Hz. FAB mass spectra were recorded with a JEOL JMS-SX 102A spectrometer in 2-nitrobenzyl alcohol (NBA). HPLC was performed on a Waters 600E Multisolvent Delivery System equipped with TOSOH UV-8010 or Waters 991J Photodiode Array detector using two connected YMC-Pack ODS-AQ SH-343-5AQ 250 × 20 mm columns or a YMC-Pack ODS-AQ-313 S-5 120A ODS column, respectively.

Materials and methods

$CDCl_3$ (99.96% D, CEA), $CDCl_2CDCl_2$ (99.5% D, CEA) and CD_2Cl_2 (99.75% D, Wako Pure Chemical Industries) were passed through dry, purified basic alumina and stored overnight with 4A type molecular sieves before use. Commercially available compounds (Wako Pure Chemical Industries, Nacalai Tescue, Inc., Aldrich, Sigma) used for the syntheses were used without further purification. Solvents for syntheses were dried by known methods. NMR tubes were baked for several hours and cooled under vacuum before use.

Titrations

NMR titrations for defining thermodynamic parameters were done in a 5 mm internal diameter NMR tube at five different temperatures [−40, −20, 0, 20 and 40 °C (±0.5 °C)] by titrating receptors (1–5 mM) with 9-ethyladenine (70–100 mM) in $CDCl_3$. In each experiment, the solution of 9-ethyladenine (10–300 μl) was added directly to the receptor solution and the concentrations of the receptor and guest were corrected by the dilution factor in eqn. (1), where *V* and *v* are the volumes of the

$$[H]_T = \frac{V}{V+v}[H]_{org},\ [G]_T = \frac{v}{V+v}[G]_{org} \quad (1)$$

initial receptor solution and the added guest solution and $[H]_{org}$ and $[G]_{org}$ are their concentrations, respectively. The corrected concentrations, $[H]_T$ and $[G]_T$ were used for the following calculation of thermodynamic parameters.

Determination of thermodynamic parameters

A nonlinear least squares method was applied to determine ΔH and ΔS for all equilibria shown in Scheme 2. Optimization calculations were carried out by using all temperature dependence data as shown in Fig. 1 and titration at different temperatures as shown in Fig. 5. In order to determine the thermodynamic parameters directly from these data, the following equation: $\delta_{obs} = \alpha\delta_{open} + \beta\delta_{closed} + \gamma\delta_{complexed}$ was used to give the observed chemical shift (δ_{obs}) as a function of total concentration of the receptor ($[H]_T$), and guest ($[G]_T$) and temperature, where δ_{open}, δ_{closed} and $\delta_{complexed}$ are the chemical shifts of open, closed and complexed forms, respectively and α, β and γ are the molar fractions of each species which are given as $[H_{open}]/[H]_T$, $[H_{closed}]/[H]_T$ and $[H_{complexed}]/[H]_T$, respectively. Using eqns. (2)–(4), the values of α, β and γ are represented as in eqns. (5)–(7).

$$[H]_T = [H_{open}] + [H_{closed}] + [H_{complexed}] \quad (2)$$

$$[G]_T = [G] + [H_{complexed}],\ \alpha + \beta + \gamma = 1 \quad (3)$$

$$K_1 = [H_{open}]/[H_{closed}] \text{ and } K_3 = [H_{complexed}]/[H_{closed}][G] \quad (4)$$

$$\gamma = [\{K_3([H_T] + [G_T]) + K_1 + 1\} - (\{K_3([H_T] + [G_T]) + K_1 + 1\}^2 - 4K_3^2[H_T][G_T])^{1/2}]/(2K_3[H_T]) \quad (5)$$

$$\beta = (1 - \gamma)/(1 + K_1) \quad (6)$$

$$\alpha = \beta K_1 \quad (7)$$

Since the equilibrium constants K_1 and K_3 are given as $K_1 = \exp[-(\Delta H_1 - T\Delta S_1)/RT]$ and $K_3 = \exp[-(\Delta H_3 - T\Delta S_3)/RT]$, δ_{obs} is represented finally as the function which contains $[H]_T$, $[G]_T$ and T as independent variables and ΔH_1, ΔH_3, ΔS_1, ΔS_3, δ_{open}, δ_{closed} and $\delta_{complexed}$ as optimized parameters. The Gauss–Newton method was used as the nonlinear least-squares optimization procedure which usually gave these parameters with standard deviations of less than 8% for ΔH and 40% for ΔS. The values of ΔH_2, and ΔS_2 were obtained from the relationship $K_2 = K_3/K_1$. To the values obtained from the nonlinear least-squares optimization, 100% additional standard deviation was used for sufficient reliability.

Variable temperature ^{1}H NMR measurements

Measurements were done in a 5 mm internal diameter NMR tube in CD_2Cl_2 and $(CDCl_2)_2$ at −90 to 20 and 0 to 140 °C (±0.5 °C). respectively. Spectra were recorded at 10 °C intervals and spectra at 0, 10 and 20 °C were recorded in both solvents to check the solvent dependence of chemical shifts.

Synthesis: general procedure for flexible receptors

2′,3′-*O*-Isopropylideneuridine[23] (100 mg, 0.35 mmol), dicarboxylic acid (0.17 mmol), benzotriazolyl-*N*-oxytris(dimethylamino)phosphoniumhexafluorophosphate (BOP reagent) (155 mg, 0.17 mmol), triethylamine (71 mg, 0.70 mmol) and DMSO (dimethyl sulfoxide) (0.3 ml) were mixed overnight at room temp. DMSO was removed under reduced pressure and resulting brown residue was dissolved in chloroform (10 ml). The organic phase was washed with water (6 × 10 ml), dried over anhydrous $MgSO_4$, filtered and evaporated. Raw product was purified by HPLC (100% MeOH).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) icosanedioate 1.† Compound **1** was obtained in 18% yield, δ_H($CDCl_3$, 0 °C) 9.64 (s, 2 H), 7.25 [d (apparent), 2 H (apparent)], 5.71 (m, 2 H), 5.63 (s, 2 H), 4.98 (m, 2 H), 4.81 (m, 2 H), 4.30–4.42 (m, 6 H), 2.31 (m, 4 H), 1.22–1.61 (m, 40 H). FAB MS in NBA, *m/z* 875 ($C_{44}H_{66}N_4O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) hexadecanedioate 2.† Compound **2** was obtained in 40% yield, δ_H($CDCl_3$, 20 °C) 9.36 (s, 2 H), 7.29 [d (apparent), 2 H (apparent)], 5.73 [d (apparent),

2 H], 5.65 (s, 2 H), 4.99 (m, 2 H), 4.81 (m, 2 H), 4.28–4.37 (m, 6 H), 2.32 (t, 4 H, *J* 7.7), 1.24–1.58 (m, 36 H). FAB MS in NBA, *m/z* 841 ($C_{40}H_{58}N_4O_{14}$ + Na).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) dodecanedioate 3.† Compound **3** was obtained in 8% δ_H($CDCl_3$, 20 °C) 9.59 (s, 2 H), 7.26 [d (apparent, 2 H (apparent)], 5.73 (m, 2 H), 5.62 (s, 2 H), 5.01 (m, 2 H), 4.81 (m, 2 H), 4.25–4.42 (m, 6 H), 2.30 (m, 4 H), 1.22–1.56 (m, 28 H). FAB MS in NBA, *m/z* 763 ($C_{36}H_{50}N_4O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) undecanedioate 4. Compound **4** was obtained in 18% yield, δ_H($CDCl_3$, 0 °C) 10.20 (s 2 H), 7.29 (d, 2 H, *J* 8.2), 5.72 (d, 2 H, *J* 7.9), 5.60 (s, 2 H), 5.02 (d, 2 H, *J* 6.4), 4.82 (m, 2 H), 4.23–4.40 (m, 6 H), 2.29 (t, 4 H, *J* 7.6), 1.22–1.75 (m, 26 H). FAB MS in NBA, *m/z* 749 ($C_{35}H_{48}N_{54}O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) decanedioate 5. Compound **5** was obtained in 44% yield, δ_H($CDCl_3$, 20 °C) 9.65 (s, 2 H), 7.29 (d 2 H, *J* 8.0), 5.71 (d, 2 H, *J* 7.9), 5.59 (s, 2 H), 5.03 (d, 2 H, *J* 6.1), 4.81 (m, 2 H), 4.27–4.37 (m, 6 H), 2.29 [t (apparent), 4 H], 1.24–1.56 (m, 24 H). FAB MS in NBA, *m/z* 735 ($C_{34}H_{46}N_4O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) nonanedioate 6. Compound **6** was obtained in 27% yield, δ_H($CDCl_3$, 20 °C) 9.41 (s, 2 H), 7.27 [d (apparent), 2 H (apparent)], 5.71 (d, 2 H, *J* 7.9), 5.59 (s, 2 H), 5.04 (s, 2 H), 4.82 (s, 2 H), 4.29–4.37 (m, 6 H), 2.29 (m, 4 H), 1.27–1.56 (m, 22 H). FAB MS in NBA, *m/z* 721 ($C_{33}H_{44}N_4O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) heptanedioate 7.† Compound **7** was obtained in 21% yield, δ_H($CDCl_3$, 20 °C) 9.35 (s, 2 H), 7.28 [d (apparent), 2 H (apparent)], 5.74 (d 2 H, *J* 7.9), 5.58 (s, 2 H), 5.08 (d, 2 H, *J* 6.4), 4.87 [t (apparent), 2 H], 4.27–4.39 (m, 2 H), 2.32 (t, 4 H, *J* 7.1), 1.26–1.64 (m, 18 H). FAB MS in NAB, *m/z* 693 ($C_{31}H_{40}N_4O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) pentanedioate 8. Compound **8** was obtained in 8% yield, δ_H(CD_2Cl_2, 20 °C) 9.27 (s, br, 2 H), 7.33 (d, 2 H, *J* 7.9), 5.73 (d, 2 H, *J* 8.2), 5.64 (d, 2 H, *J* 1.8), 5.03 (m, 2 H), 5.83 (m, 2 H), 4.26–4.39 (m, 6 H), 2.37 (t, 4 H, *J* 7.3), 1.91 (m, 2 H), 1.56 (s, 6 H), 1.34 (s, 6 H). FAB MS in NBA, *m/z* 665 ($C_{29}H_{36}N_4O_{14}$).

Bis(2′,3′-*O*-isopropylideneuridine-5′-yl) *p*-phenoxyacetate 9. Compound **9** was obtained in 80% yield, δ_H($CDCl_3$, 20 °C) 9.17 (s, 2 H), 7.42 (d, 4 H, *J* 8.5), 7.20 (d, 2 H, *J* 8.0), 6.90 (d, 4 H, *J* 8.8), 5.62–5.65 (m, 4 H), 4.79 (m, 2 H), 4.68–4.71 (m, 6 H), 4.38–4.49 (m, 6 H), 1.55 (s, 6 H), 1.28 (s, 6 H). FAB MS in NBA, *m/z* 835 ($C_{40}H_{42}N_4O_{16}$).

Acknowledgements

This study is supported by the Ministry of Education, Science and Culture: Large-scale Research Project under the new Program System in Grants-in-Aid for Scientific Research.

† Solvent independent and concentration independent chemical shifts of receptors **1**, **2**, **3** and **7** are available as supplementary data from the British Library (Supp. Publ. No. 57213). For details of the Supplementary Publications Scheme see 'Instructions for Authors', *J. Chem. Soc., Perkin Trans. 2*, Issue 1, 1997.

References

1 (*a*) F. Vögtle and D. J. Cram, in *Host Guest Complex Chemistry I-II*, ed. F. Vögtle, Springer, Berlin, Heidelberg, 1981; (*b*) D. J. Cram, G. M. Lein, *J. Am. Chem. Soc.*, 1985, **107**, 3657.
2 (*a*) D. J. Cram, *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 1009; (*b*) H. Ogoshi, H. Hatakeyama, J. Kotani, A. Kawashima and Y. Kuroda, *J. Am. Chem. Soc.*, 1991, **113**, 8181; (*c*) Y. Kuroda, Y. Kato, T. Higashioji and H. Ogoshi, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 723; (*d*) S. Goswami, A. D. Hamilton and D. Van Engen, *J. Am. Chem. Soc.*, 1989, **111**, 3425; (*e*) M. M. Conn, G. Deslongchamps, J. de Mendoza and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1993, **115**, 3548; (*f*) S. C. Hirst and A. D. Hamilton, *Tetrahedron Lett.*, 1990, **31**, 2401; (*g*) S. C. Zimmerman, C. M. Vanzyl and G. S. Hamilton, *J. Am. Chem. Soc.*, 1989, **111**, 1373; (*h*) Y. Aoyama, M. Asakawa, Y. Matsui and H. Ogoshi, *J. Am. Chem. Soc.*, 1991, **113**, 6233; (*i*) V. Hedge, C.-Y. Hung, P. Madhukar, R. Cunningham, T. Höpfner and R. P. Thummel, *J. Am. Chem. Soc.*, 1993, **115**, 782; (*j*) C.-Y. Huang, L. A. Cabell and E. V. Anslyn, *J. Am. Chem. Soc.*, 1994, **116**, 2778; (*k*) B. R. Peterson and F. Diederich, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 1625; (*l*) R. Güther, M. Nieger and F. Vögtle, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 601; (*m*) M. D. Cowart, I. Sucholeiki, R. R. Bukownik and C. S. Wilcox, *J. Am. Chem. Soc.*, 1988, **110**, 6204.
3 L. W. Guddat, L. Shan, J. M. Anchin, D. S. Linthicum and A. B. Edmundson, *J. Mol. Biol.*, 1994, **236**, 247.
4 S. Nakagawa and H. Umeyama, *FEBS Lett.*, 1982, **139**, 181.
5 R. S. Spolar and M. T. Record, Jr., *Science*, 1994, **263**, 777.
6 A. Wishnia, *Biochemistry*, 1969, **8**, 5070.
7 J. Baldwin and C. Chothia, *J. Mol. Biol.*, 1979, **129**, 175.
8 B. R. Gelin, A. W.-M. Lee and M. Karplus, *J. Mol. Biol.*, 1983, **171**, 489.
9 R. G. Shulman, J. J. Hopfield and S. Ogawa, *Quart. Rev. Biophys.*, 1975, **8**, 325.
10 S. C. Zimmerman and C. M. VanZyl, *J. Am. Chem. Soc.*, 1987, **109**, 7894.
11 T. Tjivikua, G. Deslongchamps and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1990, **112**, 8408.
12 Y. Kuroda, J. M. Lintuluoto and H. Ogoshi, *Tetrahedron Lett.*, 1994, **35**, 3729.
13 B. Castro, J.-R. Dormoy, B. Dourtoglou, G. Evin, C. Selve and J.-C. Ziegler, *Synthesis*, 1976, 751.
14 Within this temperature range, the chemical shift changes for the olefinic protons H5 and H6 are negligibly small (<0.05 ppm). See Figs. 4 and 7 also.
15 D. F. Schall and G. W. Gokel, *J. Am. Chem. Soc.*, 1994, **116**, 6089.
16 About Dreiding force field, see: S. L. Mayo, B. D. Olafson and W. A. Goddard III, *J. Phys. Chem.*, 1990, **94**, 8897.
17 Receptors **1**–**7** in this work bind stronger through closed form above room temp.
18 Closed form can resist to bind water to be able to bind substrate, see: J. C. Adrian and C. S. Wilcox, *J. Am. Chem. Soc.*, 1992, **114**, 1398.
19 L. D. Williams, B. Chawla and B. R. Shaw, *Biopolym*, 1087, **26**, 591.
20 L. Williams, Dean, Shaw and B. Ramsay, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 1779.
21 N. Williams, Gupta, L. Williams, Dean and B. R. Shaw, *J. Am. Chem. Soc.*, 1989, **111**, 7205.
22 G. A. Jeffrey and W. Sanger, *Hydrogen Bonding in Biological Structures*, Springer, Berlin, Heidelberg, 1991.
23 H. P. M. Fromageot, B. E. Griffin, C. B. Reese and J. E. Sulston, *Tetrahedron*, 1967, **23**, 2315.

Paper 6104278C
Received 18*th June* 1996
Accepted 10*th September* 1996