Pergamon

Tetrahedron Letters, Vol. 35, No. 22, pp. 3729-3732, 1994
Elsevier Science Ltd
Printed in Great Britain
0040-4039/94 $7.00+0.00

0040-4039(94)E0672-K

Flexible Nucleobase Receptor — Effect of Self-preorganization of Artificial Receptor—

Yasuhisa Kuroda*, Juha M. Lintuluoto, and Hisanobu Ogoshi*

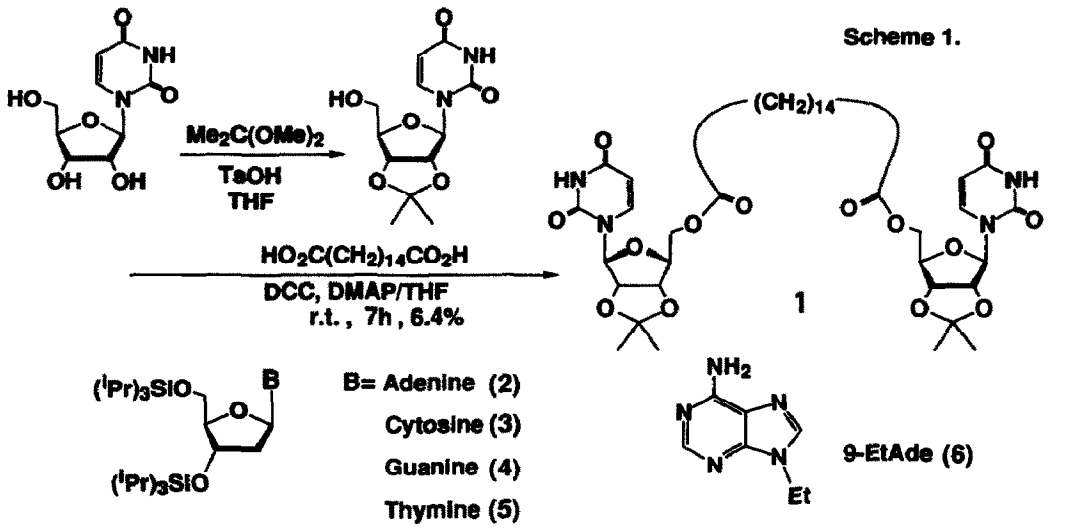
Department of Synthetic Chemistry and Biological Chemistry, Kyoto University, Sakyo-ku, Kyoto 606-01 , Japan

Abstract *: A new type flexible receptor which has two uracil moieties connected with a long alkyl chain was synthesized and its molecular recognition ability toward adenine derivatives was investigated. The present receptor shows equilibrium between two states, open and close forms. In contrast to the open form, the closed form exhibits enthalpically less favorable but entropically more favorable molecular recognition toward adenine derivatives. This observation is rationally explained by the effect of self-preorganization of the present receptor.*

Construction of artificial host molecules is one of the central themes in biomimetic chemistry and various types of molecules have been successfully realized as artificial receptors which show high molecular recognition ability.[1] During the course of these studies, the concept of "preorganization" has sometimes been playing important role,whether it has been intentional or not.[2] Thus the artificial hosts have been prepared based on relatively rigid molecules which are designed so as to minimize conformational change during complex formation. However, it is also clear that conformational (structural) change of a host molecule is inevitably necessary, if the host is required to perform further functions after the molecular recognition process such as catalyses of enzymes and signal transmission of receptors. This means rather flexible molecules seem to be suitable for mimicking of natural host molecules. In order to satisfy these incompatible requirements, rigidity for molecular recognition and flexibility for further functions, a new guideline for design of artificial receptors should be developed. In this context, natural protein molecules furnish an excellent example of the solution of the present problem, i.e., proteins in so-called active forms are well organized and keep their low entropical states, even though they basically consist of linear polypeptide chains which have extremely high motional freedom in the original states. Such conversion of natural proteins from random to organized states is sometimes spontaneous and at least partially reversible process, which allows further conformational change from its first organized form. In this sense, such "self-preorganization" of proteins is very important and strongly related to their functions.

We here report an example of self-preorganization of an artificial nucleobase receptor. Although the present system is very simple and primitive, the results clearly indicate an interesting effect of self-preorganization and is suggestive for designing and synthetizing of more refined artificial receptors.

Our receptor is prepared by condensation of 2',3'-O-isopropylidene uridine with hexadecanedioic acid as shown in Scheme 1. After purification by means of HPLC, the product showed satisfactory characteristics of ^{1}H NMR and FAB mass spectra as the expected structure of **1**.[3]



Since the compound 1 has only one kind of proton, uracil imino protons, which can form hydrogen bonding, observation of those ^1H NMR signals is very informative for estimation of the structure and/or conformation of 1.

First of all, it is noted that the uracil imino protons of 1 show strong dependence of chemical shift on temperature in halo carbon solvents. On cooling from +60 to -60°C, these protons showed the downfield shift of over 3.5 ppm and the observed temperature dependence was clearly sigmoidal within the temperature range of -90 ~ +140°C as shown in Figure 1a. This behavior was confirmed to be independent on the concentration of 1. These observations strongly indicate that 1 forms *intramolecular hydrogen bonds* and exist in the equilibrium state between two forms, 1-open and 1-closed, as shown in Figure 1b.

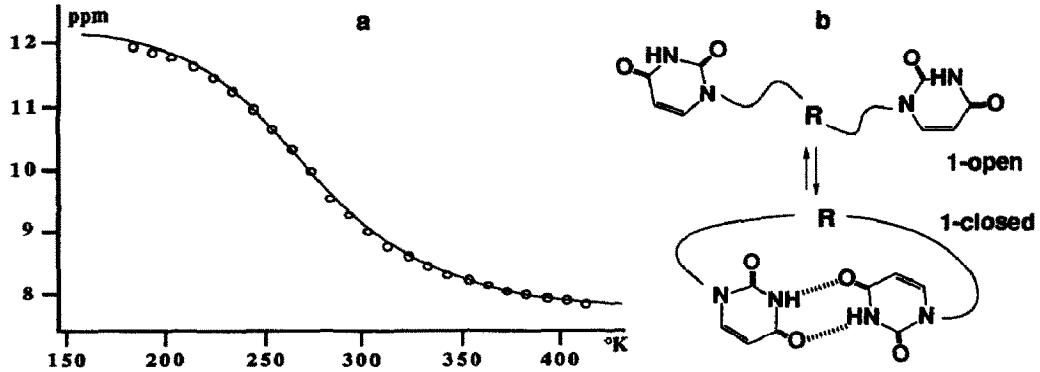


Figure 1 a) Temperature dependence of the chemical shift of uracil imino protons in 1. The data was taken in CD_2Cl_2 (183- 273°K) and $CDCl_2CDCl_2$ (283 - 413°K), respectively. The solid line is the theoretical dependence calculated from thermodynamic parameters determined in this work (see text). b) The most plausible intramolecular equilibrium between 1-open and 1-closed.

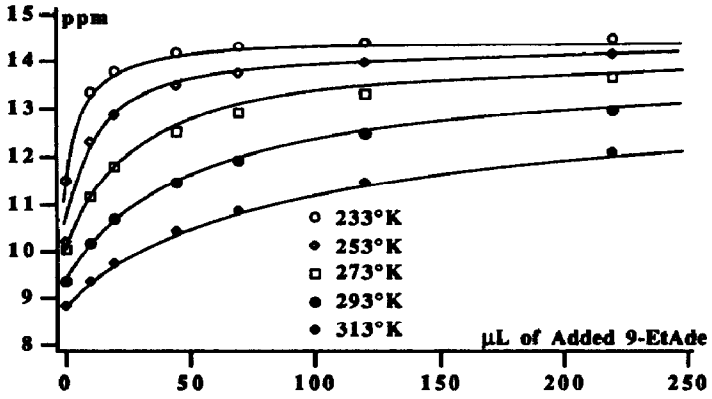
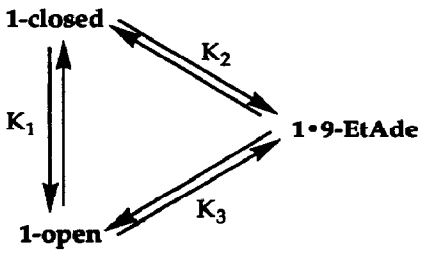


Figure 2 Plots of chemical shift of uracil imino protons in **1** vs. amount of added **9-EtAde**. The titration experiments were performed by adding the solution of **9-EtAde** (0.101M) in $CDCl_3$ to 500 μl of the $CDCl_3$ solution of **1** (0.733mM). The observed chemical shifts were plotted as a function of added volume of the **9-EtAde** solution. The solid lines were theoretical curves obtained from thermodynamic parameters (see text).

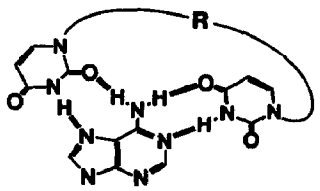
The ability of **1** to recognize nucleobases was examined by using compunds **2-5** as guest molecules (Scheme 1).[4] Titration experiments with ^1H NMR demonstrate that only **2** exhibits typical behavior of complex formation with **1** which also results in large downfield shift of imino protons[5] and other nucleosides show no signs of complex formation. Based on these results, we investigated the details of the present highly specific molecular recognition toward adenine base by using a more sterically unhindered guest, **6**. Thus, binding of **6** by **1** at various temperatures was examined by the ^1H NMR titration experiment. The results are summarized in Figure 2. Assuming a three state equilibrium as shown in Scheme 2, whole data given in Figure 1 and 2 was analyzed by a nonlinear least square method using enthalpy and entropy changes (ΔH and ΔS) for K_1 and K_2 as optimization parameters.[6] The thermodynamic parameters thus optimized are also given in Scheme 2. The theoretical curves calculated from this parameter set fit excellently with observed data which was collected under the wide range



Equilibrium constant	Definition	ΔH kJ•mol^{-1}	ΔS J•K^{-1}•mol^{-1}
K_1	$\frac{[1\text{-open}]}{[1\text{-closed}]}$	17 ± 2	62 ± 9
K_2	$\frac{[1\bullet 9\text{-EtAde}]}{[1\text{-closed}][9\text{-EtAde}]}$	-16 ± 2	-11 ± 8
K_3	$\frac{[1\bullet 9\text{-EtAde}]}{[1\text{-open}][9\text{-EtAde}]}$	-33 ± 4	-74 ± 17

Scheme 2 Equilibrium of the system containing **1** and **9-EtAde**.

of conditions of temperatures and concentrations of **6** (see Figure 1 and 2). The observed ΔH value (17 kJ•mol^{-1}) for K_1 is reasonable for the equilibrium which contains two hydrogen bonds breaking as shown in Figure 2b.[7] It should be noted that this K_1 process accompanies large entropy gain which indicates rigid characteristics of **1-closed** compared with that of **1-open**. The most interesting aspect in the present equilibrium is that the K_2 process, complex formation between **1-closed** and **6**, shows further enthalpy gain and small entropy loss. These observations



suggest that the most plausible form of the **1•9-EtAde** complex is the U•A•U type of the Hoogsteen base triplet (see left).[8] Since this type of complex contains four hydrogen bonds, the observed enthalpy gain (-16 kJ• mol^{-1}) is reasonably explained by further two hydrogen bond formation during the process from **1-closed** to **1•9-EtAde**. Similarly, small entropy loss in this process is rationalized by the fact that the most flexible alkyl chain of **1** is already fixed well in **1-closed** and structural flexibility of **1-closed** and the present triplet type complex is not expected to be so different. The observed value of entropy loss actually lies in the same range as that statistically expected for a simple bimolecular process (Rln2). Thus, **1-closed** may be considered as a "self-preorganized" form of the flexible receptor **1**. The effect of this self-preorganization is clearly demonstrated by comparison of K_2 and K_3 ; i.e., for example, the values of K_2 and K_3 at 30 °C are 170 and 96 M^{-1} respectively ,which mean **1-closed** is a better receptor for adenine base than **1-open**.[9] This result is due to the fact that the K_2 process is enthalpically unfavorable but entropically more favorable compared with the K_3 process.

References and Notes.

1) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1988**, 27, 89.; Wögtle, F. *Supramolecular Chem.*; John Wiley and Sons, Inc.: New York, 1991.; Wilcox, C. S.; Webb, T. H.; Zawacki, F. J.; Glagovich, N.; Suh, H. *ibid.* **1993**, 1, 129.; Sessler, J. M.; Furuta, H.; Král, V. *ibid.* **1993**, 1, 209.; Hirst, S. C.; Hamilton, A. D. *Tedrahedron Lett.* **1990**, 31, 2401.; Hamilton, A. D.; Little, D. *J. Chem. Soc.. Chem. Commun.* **1990**, 297.; Hamilton, A. D.; Pant, N. *ibid.* . **1988**, 765.; Chen, C.-W.; Whitlock, H. W., Jr. *J. Am. Chem. Soc.* **1978**, 100, 492.; Wilcox, C. S.; Adrian, J. C., Jr.; Webb, T. H.; Zawacki, F. J. *Jbid.* **1992**, 114, 10189.
2) Cram, D. J. *Angew.Chem.Int. Ed. Engl.* **1988**, 27, 1009
3) 1:^{1}H NMR (500 MHz, $CDCl_3$/ TMS at 10 °C), δ(ppm): 9.66 (s, 2H, 2 NH), 7.28 (d, J=7.9 Hz, 2H), 5.72 (d, J=5.8 Hz, 2H), 5.62 (s, 2H), 4.98 (m, 2H), 4.77 (m, 2H), 4.35-4.28 (m, 2H+4H), 2.30 (m, 4H), 1.56-1.34 (several peaks, 36 H + H_2O, 28H); FAB MS: m/z (%)= 841, (m^++Na), (3), 299 (2), 297 (2), 267 (2), 265 (2), 194 (5), 150 (30), 148 (12), 93 (100).
4) Ogilvie, K. K.; Thomson, E. A.; Quilliam, M. A.; Westmore, J. B. *Tetrahedron Lett.* **1974**, 33, 2865.; Williams, L. D.; Shaw, B. R. *Proc. Natl. Acad. Sci. USA* , **1987**, 84, 1779.; Williams, L. D.; Chawla, B.; Shaw, B. R. *Biopolymers* , **1987**, 26, 591.; Williams, N. G.; Williams, L. D.; Shaw, B. R. *J. Am. Chem. Soc.* **1989**, 111, 7025.
5) The apparent binding constant of **1** with **2** was K_{obs}=48 ± 2 M^{-1} in $CDCl_3$ at 10°C.
6) The extrapolated values of chemical shifts for **1-open, 1-closed** and **1•9-EtAde** are also used as optimization parameters and determined to be 7.45, 12.08 and 14.45 ppm respectively. The parameters for K_3 were obtained from the relationships of $K_3=K_2/K_1$.
7) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures* ; Springer-Verlag: Berlin Heidelberg. 1991.; Pimentel, G. C.; McClellan, A. L. *The Hydrogen Bond* ; Freeman: San Francisco. 1960.
8) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures* ; Springer-Verlag: Berlin Heidelberg. 1991, pp. 267-268.; Nguyen, T. T.; Hélène, C. *Angew. Chem. Int. Ed. Engl.* **1993**, 32, 666.
9) The values of K_2 and K_3 are the same at 5 °C and K_2 is always larger than K_3 at higher temperature.

(*Received in Japan* 24 *December* 1993; *accepted* 16 *March* 1994)