Sensors

Protein Recognition by an Ensemble of Fluorescent DNA G-Quadruplexes**

*David Margulies and Andrew D. Hamilton**

The development of combinatorial sensory arrays inspired by the mammalian olfactory system is an exciting and promising direction in the field of supramolecular analytical chemistry.[1] Recent developments in pattern-based detection arrays which can identify a range of proteins,[2–8] be tuned to sense specific protein families,[9] and reach low-detection thresholds,[6,7] demonstrate the potential of this technology for applications in medical diagnosis, pathogen detection, and proteomics.

Nevertheless, there are several limitations accompanying the use of arrays in comparison to other homogenous, protein detection systems.[10] Although arrays can now detect proteins at nanomolar concentrations,[6,7] the need to channel each protein sample to different spatially separated receptors results in a high consumption of samples, which may complicate their use for identifying rare and low concentration proteins. Additionally, accurate and high-throughput array detection requires the development of efficient manufacturing protocols, as well as mechanisms that ensure rapid and equal distribution of analytes. Finally, unlike some solution phase protein sensors,[10] non-homogeneous systems are less efficient in monitoring real time events.
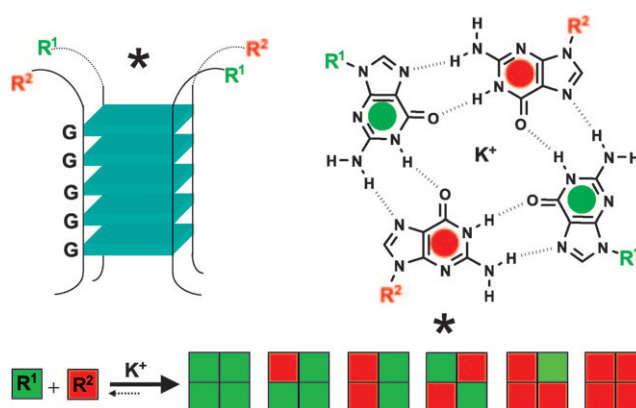
For these reasons, we aimed to develop a simple and efficient methodology for combining a range of signal-emitting protein receptors in a homogeneous solution. Anslyn and co-workers,[11] and Buryak and Severin,[12,13] have reported the solution phase recognition of biomolecules by using indicator displacement assays.[11–14] Extending these foundations to protein detection, while maintaining a high sensitivity of fluorescent signaling, would be an important step toward the realization of high-throughput systems capable of detecting proteins, which can ultimately lead to understanding their function.

Herein we report the use of fluorescent DNA G-quadruplexes as a strategy for building versatile sets of self-assembled protein receptors. We demonstrate a systematic approach to controlling both the composition and emission pattern of the ensembles in a way that enables protein

[*] Dr. D. Margulies, Prof. A. D. Hamilton
    Department of Chemistry, Yale University
    P. O. Box 208107, New Haven, CT 06520-8107 (USA)
    Fax: (+1) 203-432-6144
    E-mail: andrew.hamilton@yale.edu

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200804887.

differentiation to occur in samples as small as a single microliter drop.

To realize homogeneous, combinatorial recognition systems for proteins, we envisioned a strategy that would enable a variety of water-soluble protein receptors to coexist in a single solution. These receptors would be diverse and have a substantial surface area for nonspecific interactions with different proteins. Fluorescence is the most sensitive detection mode, measurable at the level of a single molecule[15] and from extremely small volumes.[16] Finally, practical sensory arrays should be both easily assembled and modified according to a desired application. While these requirements could not be realized with our previous porphyrin-based arrays,[3,4] our recent work has shown that DNA-based scaffolds can offer an alternative strategy for building combinations of protein receptors in solution.[17–19] Of particular interest are synthetic receptors based on functionalized G-quadruplexes, which have large surface areas that are valuable for targeting protein surfaces[18] and for sensing specific guest interactions.[19] Other studies have shown that asymmetric G-quartets could be assembled from dissimilar G-rich strands,[20] indicating the possibility of generating a variety of distinguishable receptors from only a few building blocks. Scheme 1 depicts the six receptors that could be generated from two distinct G-rich oligodeoxynucleotides (ODNs), to which a protein binding fragment (R^1, R^2) is appended. Similarly, three strands should provide 21 different receptors.

Figure 1 outlines the way in which the emission pattern from a mixture of receptors in a single solution could be tuned, simply by an appropriate choice of ODNs and preparation protocols. Three distinct fluorophores: pyrene

Scheme 1. Two distinct G-rich DNA strands, to which a protein binding fragment (R^1, R^2) is appended, can self-assemble into six possible synthetic receptors.
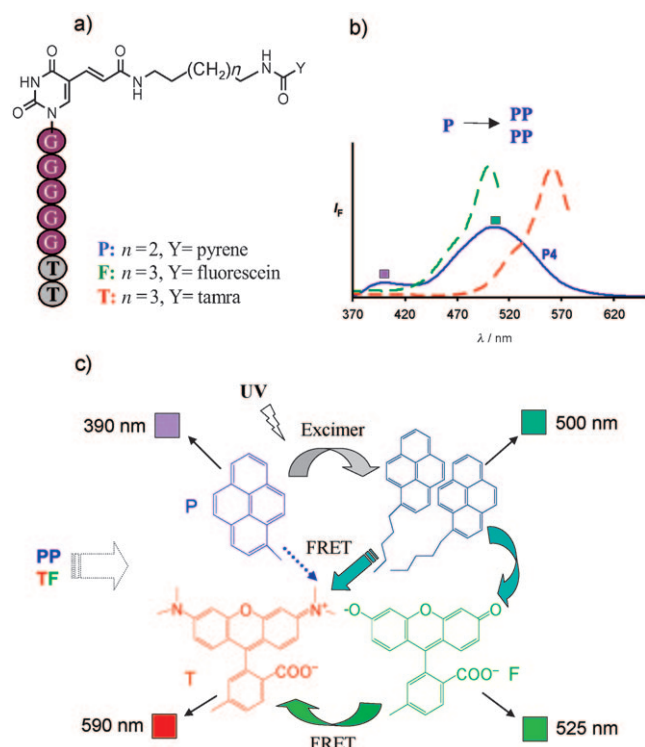
Figure 1. a) Distinguishable sets of DNA G-quadruplex ensembles can be prepared from different combinations of G-rich ODNs modified with pyrene (P), fluorescein (F), or tamra (T). b) An emission pattern of a single P_4 G-quadruplex (blue), which overlaps the excitation spectra (normalized) of fluorescein (green) and tamra (red). c) Possible optical processes that can occur among the three fluorophores.

would lead to emission patterns corresponding to pyrene (390 nm), pyrene excimer (500 nm), and fluorescein (525 nm) or tamra (590 nm) that are triggered by direct excitation or by fluorescence resonance energy transfer (FRET) processes that occur among them (Figure 1c).

Figure 2a shows how four distinct emission signatures could be generated by the choice of building blocks and by controlling the self-assembly process. In this example, the choice of ODNs, P + F versus P + T, controls the emission
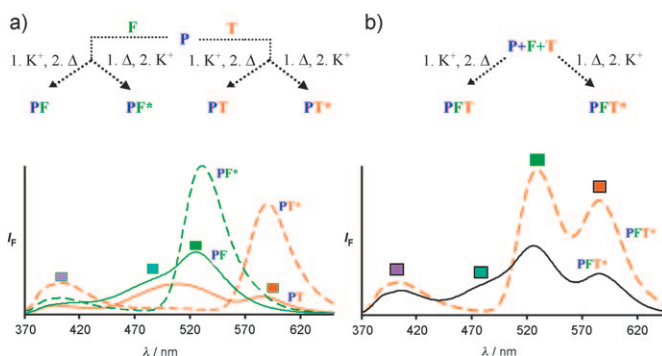

Figure 2. a) Four distinct emission patterns, corresponding to different receptor combinations, can be tuned by an appropriate choice of two ODNs and preparation methods. b) Different patterns generated by the three ODNs. Conditions: 570 nM P, 1.1 equiv F, and 1.1 equiv T, excited at 344 nm.

(P), fluorescein (F), and tamra (T) were attached to similar G-rich strands $X(G)_5TT$, where X is a modified thymine (dT) linked to the appropriate fluorophore by a diamide linker (Figure 1a). Strands composed of five guanines were chosen so that relatively stable quadruplexes would be formed, even at nanomolar concentrations. Thymine nucleotides were attached at each of the ODN termini as they have been shown to prevent quadruplex aggregation. The fluorophores were carefully chosen to have a distinct emission spectrum and spectral overlap (Figure 1b) such that optical communication and physical contact between them could lead to a range of emission patterns. Figure 1c depicts some of the optical processes that can occur, for example, in an asymmetric quadruplex composed of two pyrenes, one fluorescein, and one tamra. Alternatively, the same communication channels can take place between different quadruplexes that assemble on a protein surface. Pyrene has an emission maximum at 390 nm; however, by intramolecular π–π stacking it generates an excimer emission at 500 nm. Therefore, a four-fold symmetric G-quadruplex formed by incubating pyrene-appended ODNs with potassium ions generates fluorescence output at both 390 nm and 500 nm (Figure 1b). As this emission overlaps with the excitation spectrum of fluorescein (F) and tamra (T), a known donor-acceptor pair, it is expected that the excitation of the pyrene unit (344 nm) in G-quadruplex mixtures (formed from P, F, and T ODNs)

wavelengths, whereas environmental conditions will determine their pattern. ODNs carrying five guanines form quadruplex structures in water, however heating the solution above 95°C resulted in the breakdown of the Hoogsteen hydrogen bonding (see the Supporting Information) which enabled two distinct ODNs to assemble into asymmetric PF* and PT* quadruplex mixtures (as outlined in Scheme 1); these structures are then additionally stabilized by the addition of potassium ions. While intramolecular FRET in PF* and PT* led to a high emission intensity of fluorescein (525 nm) and tamra (590 nm), it caused a loss of excimer emission at 500 nm, which has become statistically less favored. To realize a wider spectral output, potassium was added prior to heating (as in PF and PT), in a way that maintains excimer emission by further stabilizing the symmetric species, but also reduces the emission intensity of fluorescein and tamra owing to a reduced population of the asymmetric forms. Similarly, two distinct patterns could be realized from the three ONDs (P, F, and T) assembled into different combinations of quadruplexes under the two reaction conditions (Figure 2b).

The ability to control the composition and emission of the ensembles, as well as their capacity to change their pattern in response to external stimuli, indicates that G-rich ODNs possess the right properties for building pattern recognition systems in a single solution. Considering the inherent water solubility of ODNs and the simplicity of their modification and hybridization, a large number of different binding agents can, in principle, be built by altering the number of ODNs used, their lengths,[19] and their functionalization with addi-

tional groups, such as hydrophobic[19] or polar substituents,[18] additional fluorophores, and small molecule substrates.[17]

Proteins have been shown to change the optical properties of dyes that come into contact with them.[3,4,21,22] Therefore, by the proper functionalization of quadruplexes with non-covalent recognition groups targeted towards particular protein classes,[18] we expected the binding event to result in changes in the emission pattern of the ensemble. As G-quadruplex structures carry a large negative charge on their sugar/phosphate backbone, a proof of concept could be directly demonstrated by the ability of the PFT ensemble (Figure 2b) to selectively recognize basic proteins and differentiate them.

Figure 3a shows changes that occur in the emission spectrum of the PFT ensemble upon the addition of a series of proteins (500 nM) which are different in size, charge, and composition, such that their molecular weight ranges between 2.8 kD (melittin) and 66 kD (avidin), and their isoelectric points (pI) vary from 4 (phosvitin) to 11.6 (MBP). Upon binding, the emission of the fluorophores can be affected[21,22] by direct contact with peptide side chains[23] or prosthetic groups,[3] the interruption of π–π stacking,[19] and the variation in FRET resulting from changes in the overall composition of the ensemble.[17] Figure 3b depicts signatures correspondi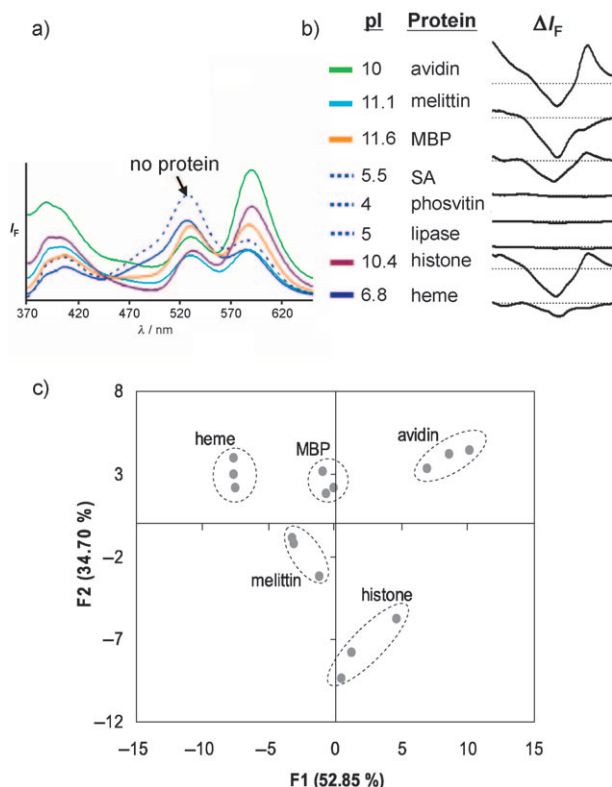ng to changes in the emission of the PFT ensemble induced by the different proteins. Except for hemoglobin (heme), whose intrinsic heme group quenches the array emission, selectivity for basic proteins was clearly observed. For example, positively charged proteins, such as MBP and melittin generated a clear pattern, whereas acidic lipase and phosvitin were invisible to the array. Even related proteins, such as avidin and streptavidin (SA) could be distinguished by the quadruplex system. A clearer picture is obtained when the change in emission is analyzed by using principal component analysis (PCA),[1] resulting in distinguishable signals for the five detectable proteins (Figure 3c).

The simplicity with which G-quadruplex ensembles can be prepared and employed indicates their potential to be used as sensitive, high-throughput protein detection systems. To demonstrate this point, we tested the compatibility of our system with a portable fluoro-spectrophotometer capable of detecting emission directly from a single microliter drop.[16] While this technology enables fluorescence detection to occur without sacrificing precious biological samples,[16] it is the G-quadruplex ensemble that allows a pattern recognition system to accommodate such volumes. Figure 4 summarizes
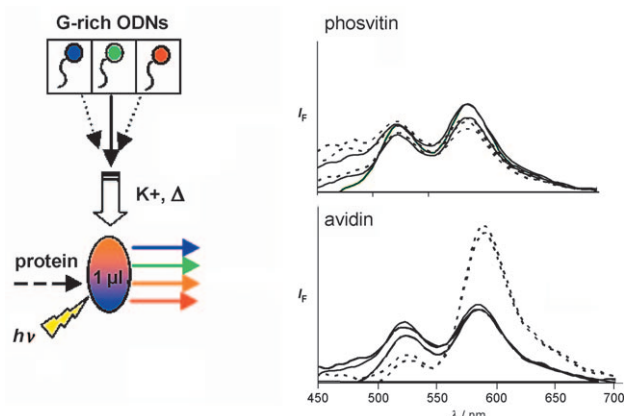

Figure 4. Left: a scheme summarizing the steps for realizing a microliter-size, pattern-based detection system for proteins. Right: emission patterns from three individual drops of the ensemble (solid lines) and drops loaded with acidic phosvitin or basic avidin (dashed lines). Conditions: 1 μL drops of 1.6 μM quadruplexes loaded with 2.8 μM proteins, excited at 365 ± 10 nm.

Figure 3. a) Emission spectra of a PFT ensemble and its response to 500 nM of different proteins. b) Monitoring the changes in emission generates unique signatures for five distinguishable proteins. SA, MBP, and heme correspond to streptavidin, myelin basic protein, and hemoglobin, respectively. pI = isoelectric point. c) PCA mapping using 57 data points corresponding to the changes in emission at 370–650 nm. Conditions: 456 nM quadruplexes, excited at 344 nm.

the steps for realizing such a device and an example for the way basic avidin can be readily distinguished from acidic phosvitin, simply by dissolving the proteins in a solution containing the minimal amounts of the ensemble and extracting a single microliter drop for analysis. The results confirm the potential of such systems to provide high throughput and selective protein detection on a very small scale.

Communications

[1] E. V. Anslyn, *J. Org. Chem.* **2007**, *72*, 687–699.
[2] B. E. Collins, E. V. Anslyn, *Chem. Eur. J.* **2007**, *13*, 4700–4708.
[3] L. Baldini, A. J. Wilson, J. Hong, A. D. Hamilton, *J. Am. Chem. Soc.* **2004**, *126*, 5656–5657.
[4] H. Zhou, L. Baldini, J. Hong, A. J. Wilson, A. D. Hamilton, *J. Am. Chem. Soc.* **2006**, *128*, 2421–2425.
[5] M. M. Reddy, T. Kodadek, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12672–12677.
[6] C.-C. You, O. R. Miranda, B. Gider, P. S. Ghosh, I.-B. Kim, B. Erdogan, S. A. Krovi, U. H. F. Bunz, V. M. Rotello, *Nat. Nanotechnol.* **2007**, *2*, 318–323.
[7] O. R. Miranda, C.-C. You, R. Phillips, I.-B. Kim, P. S. Ghosh, U. H. F. Bunz, V. M. Rotello, *J. Am. Chem. Soc.* **2007**, *129*, 9856–9857.
[8] S. Kolusheva, R. Zadmard, T. Schrader, R. Jelinek, *J. Am. Chem. Soc.* **2006**, *128*, 13592-13598.
[9] A. T. Wright, M. J. Griffin, Z. Zhong, S. C. McCleskey, E. V. Anslyn, J. T. McDevitt, *Angew. Chem.* **2005**, *117*, 6533–6536; *Angew. Chem. Int. Ed.* **2005**, *44*, 6375–6378.
[10] J. Li, Z. Tang, K. Wang, W. Tan, *Curr. Proteomics* **2004**, *1*, 315–324.
[11] S. L. Wiskur, P. N. Floriano, E. V. Anslyn, J. T. McDevitt, *Angew. Chem.* **2003**, *115*, 2116–2118; *Angew. Chem. Int. Ed.* **2003**, *42*, 2070–2072.
[12] A. Buryak, K. Severin, *Angew. Chem.* **2005**, *117*, 8149–8152; *Angew. Chem. Int. Ed.* **2005**, *44*, 7935–7938.
[13] A. Buryak, F. Zaubitzer, A. Pozdnoukhov, K. Severin, *J. Am. Chem. Soc.* **2008**, *130*, 11260–11261.
[14] M. Hafner, E. Vianini, B. Albertoni, L. Marchetti, I. Grune, C. Gloeckner, M. Famulok, *Nat. Protoc.* **2008**, *3*, 579–587.
[15] S. Weiss, *Science* **1999**, *283*, 1676–1683.
[16] D. Ash, P. Desjardins, *Genet. Eng. Biotechnol. News* **2007**, *27*, 38–39.
[17] K. I. Sprinz, D. M. Tagore, A. D. Hamilton, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3908–3911.
[18] D. M. Tagore, K. I. Sprinz, S. Fletcher, J. Jayawickramarajah, A. D. Hamilton, *Angew. Chem.* **2007**, *119*, 227–229; *Angew. Chem. Int. Ed.* **2007**, *46*, 223–225.
[19] J. Jayawickramarajah, D. M. Tagore, L. K. Tsou, A. D. Hamilton, *Angew. Chem.* **2007**, *119*, 7727–7730; *Angew. Chem. Int. Ed.* **2007**, *46*, 7583–7586.
[20] E. E. Merkina, K. R. Fox, *Biophys. J.* **2005**, *89*, 365–373.
[21] K. Schauenstein, E. Schauenstein, G. Wick, *J. Histochem. Cytochem.* **1978**, *26*, 277–283.
[22] H. J. Gruber, C. D. Hahn, G. Kada, C. K. Riener, G. S. Harms, W. Ahrer, T. G. Dax, H.-G. Knaus, *Bioconjugate Chem.* **2000**, *11*, 696–704.
[23] N. Marme, J.-P. Knemeyer, M. Sauer, J. Wolfrum, *Bioconjugate Chem.* **2003**, *14*, 1133–1139.