Protein Surface Recognition

DOI: 10.1002/ange.201402501

Targeted Protein Surface Sensors as a Tool for Analyzing Small Populations of Proteins in Biological Mixtures**

*Leila Motiei, Zohar Pode, Anna Koganitsky, and David Margulies**

Abstract: *Optical cross-reactive sensor arrays (the so-called chemical "noses/tongues") have recently been demonstrated as a powerful tool for high-throughput protein detecting and analysis. Nevertheless, applying this technology to biomarker detection is complicated by the difficulty of non-selective sensors to operate in biological mixtures. Herein we demonstrate a step toward circumventing this limitation by using self-assembled fluorescent receptors consisting of two distinct recognition motifs: specific and non-specific. When combined in an array, binding cooperatively between the specific and non-specific protein binders enables the system to discriminate among closely related isoform biomarkers even in the presence of serum proteins or within human urine.*

In recent years, extensive efforts have been made to develop non-reductionist approaches to disease diagnosis. Profiling the expression of multiple proteins, rather than detecting individual protein analytes, has been explored as a means of improving diagnostic accuracy and better understanding the parameters affecting disease states.[1] A promising method for obtaining multiplexed protein analysis involves the use of antibodies that can bind and detect the proteins of interest with high affinity and selectivity.[1b,c] This approach has found widespread applications in medical diagnosis; however, the need for producing an antibody for each target and for using stepwise protocols, in addition to their relatively high costs and instability, hamper high-throughput analysis.

Cross-reactive sensor arrays, inspired by the mammalian olfactory system, have recently emerged as an alternative detection method that may address these limitations.[2,3] When the "nose/tongue" approach is used, proteins can be rapidly differentiated using an array of non-specific synthetic receptors that, in combination, generate a unique optical "fingerprint" upon interacting with each protein. Unlike antibody arrays, which operate according to the "lock and key" paradigm, arrays that rely on differential sensing[4] do not require manufacturing multiple antibodies or using technically challenging procedures. As a result, such systems can straightforwardly discriminate among multiple different proteins[3] as well as profile protein mixtures in biofluids, which may indicate disease states.[3f–i]

Despite the numerous advantages of cross-reactive arrays, applications for this technology in medical diagnostics are limited by the difficulty of non-selective receptors to operate within biological mixtures. Human serum contains more than 20000 proteins, of which only about 20 proteins constitute about 99% of the serum protein mass. Thus, although such systems can effectively discriminate among combinations and concentrations of common serum proteins,[3g] detecting low-abundance disease biomarkers remains challenging. Herein, we present an integrated sensing scheme that uses both the "lock and key" and "differential sensing" strategies for discriminating among low-concentration protein biomarkers in biological mixtures. Using this approach, cross-reactive sensor arrays can be "programmed" to generate patterns that reflect the composition of specific protein groups, even in the presence of highly abundant serum proteins or within human urine.

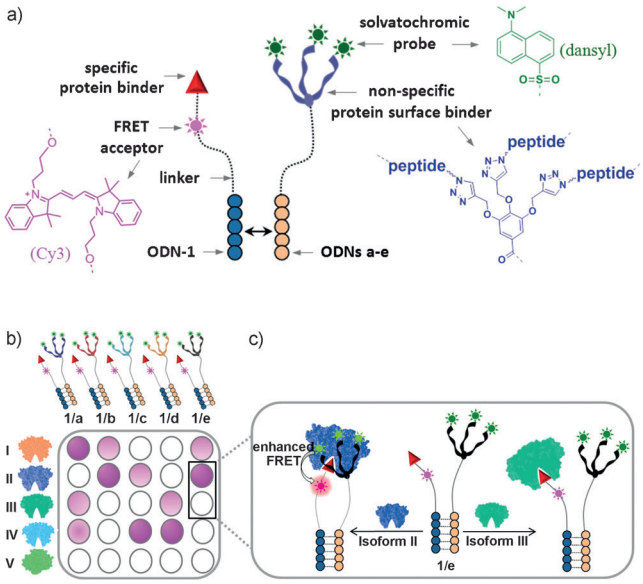

Figure 1. a) Design principles of self-assembled protein receptors consisting of two distinct recognition motifs: specific and non-specific. Two complementary oligonucleotides (ODNs) are used to form the specific (ODN-**1**) and non-specific (ODNs **a**–**e**) binding strands, which are modified with a fluorescent acceptor (Cy3) and solvatochromic fluorescent donors (dansyl), respectively. b) Schematic representation of a cross-reactive sensor array with selectivity toward a specific protein group (isoforms I–IV). Other proteins (protein V), which do not belong to this family, do not interact with the receptors in the array. c) Interaction of the receptors with isoforms having different surface characteristics (isoform II vs. III), results in different emission signals.

[*] Dr. L. Motiei, Z. Pode, A. Koganitsky, Dr. D. Margulies
Department of Organic Chemistry
The Weizmann Institute of Science, 76100 Rehovot (Israel)
E-mail: david.margulies@weizmann.ac.il

[**] This research was supported by the Israel Science Foundation (ISF), the German-Israeli Foundation (GIF), and YEDA. D.M. is the incumbent of the Judith and Martin Freedman Career Development Chair.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201402501.

Recently, we have shown that the potency and discriminating ability of pattern-generating fluorescent sensors can be largely enhanced by multivalent binding of several receptors.[5] In this study, hetero-multivalency[6] and binding cooperativity are used to create a cross-reactive sensor array with selectivity toward specific protein groups (Figure 1). As shown in Figure 1a, each fluorescent receptor in the array is self-assembled from two complementary oligonucleotides (ODNs) appended with distinct protein recognition motifs: specific or non-specific. The specific binder on ODN-**1** can be a synthetic inhibitor, a natural ligand or an aptamer that is selective toward a specific protein group, but it is also broad spectrum. In other words, this unit can strongly interact with multiple isoforms of a particular protein family. Because this binder is maintained in all different duplexes in the array (Figure 1b), it provides the receptors with high affinity and selectivity toward a small set of protein biomarkers.

The non-specific binding domain on ODNs **a**–**e** (Figure 1a), on the other hand, is a relatively weak binder that varies among the different receptors (Figure 1b). It contains a tripodal peptide group with a relatively large surface area, typical of synthetic protein surface receptors.[7] By systematically modifying its sequence, one can obtain a library of protein surface binders with distinct physicochemical properties. Generating an array of duplexes, which differ only in their protein surface binders (Figure 1b), results in an analytical device in which all the receptors can bind to a particular protein family (isoforms I–IV). However, each tripodal peptide in the array is expected to interact differently with the surfaces of distinct isoforms (Figure 1c). Distinct optical "fingerprints" for different family members can be generated by modifying the specific and non-specific binders with solvatochromic fluorescent donors (i.e., dansyls) and a fluorescent acceptor (i.e., Cy3), respectively (Figure 1a). In this way, different isoforms could be differentiated owing to changes that occur in the distance between the donor and acceptors and/or the local environment of the solvatochromic probes (Figure 1c). Notably, because binding cooperativity between the weak and strong binders can only occur with proteins that possess the specific recognition domain (Figure 1b, isoforms I–IV), other proteins that lack this site (protein V) should not affect the emission of the array.

Based on these principles, we generated DNA-based receptors with selectivity towards glutathione S-transferases (GSTs). Human GSTs are a family of enzymes that play a role in cell detoxification processes by catalyzing the conjugation of reduced glutathione (GSH) to various xenobiotics.[8] This family is subdivided into several classes, such as alpha, mu, pi, omega, sigma, theta, and zeta, which share a similar GSH binding domain, but differ in their surface characteristics (Figure S1 and Table S2 in the Supporting Information). GSTs were selected as the target proteins for this study because comparative analysis of GST isozymes in biofluids revealed a clear correlation between their expression profiles and disease states.[8b,c] In addition, a broad spectrum inhibitor (i.e., ethacrynic acid, EA) for this class of enzymes is known and could be used to generate a GST-specific binding strand (Figure 2a, ODN-**1**).

It has been shown that bivalent EA-based inhibitors, which can simultaneously bind two sites within these homodimers, exhibit significantly greater potency for GSTs.[8d] We therefore anticipated that coupling two EAs to an elongated bis-amine linker on Cy3-modified ODN-**i** (Figure 2a) would result in a specific binding strand (ODN-**1**) with a flexible bivalent inhibitor that can be accommodated within the EA binding sites of different GST isozymes. To obtain a library of protein surface binders (Figure 2b, ODNs-**a**–**e**), a complementary amino-functionalized ODN-**ii** was first modified with a trialkyne spacer **2** (see the Supporting Information for synthetic details) and the resulting strand (ODN-**iii**) was coupled to a series of dansylated and azido-modified peptides (Figure 2c and Table S1) using the click reaction.

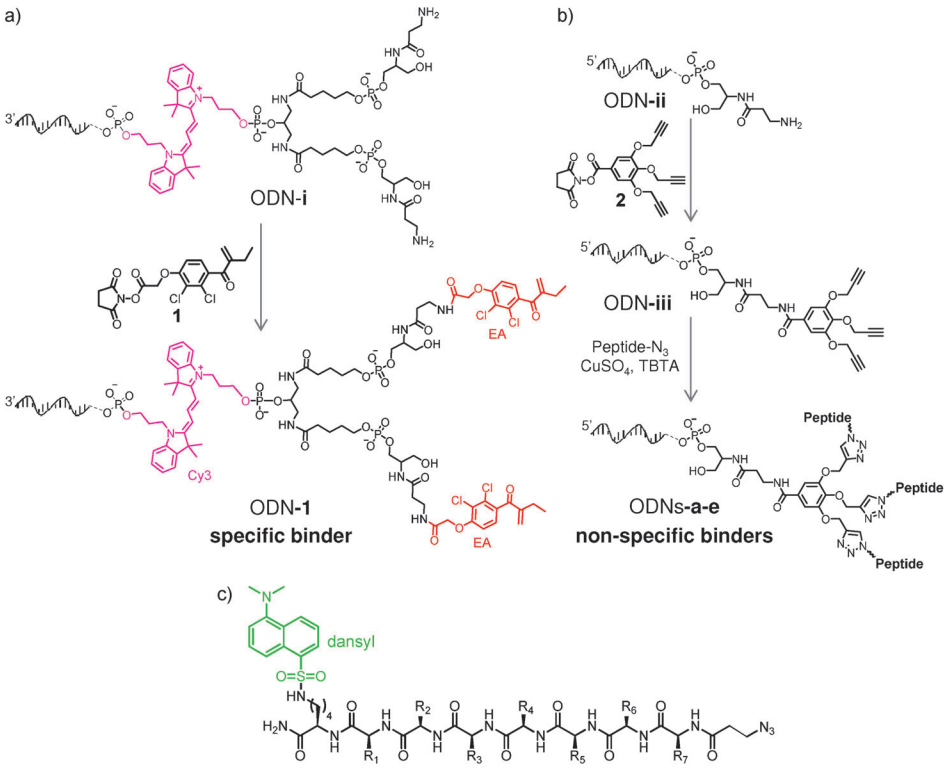

Figure 2. Synthetic steps for preparing complementary ODNs modified with a) a specific and b) non-specific GST binders. c) Representative structure of an azide-modified dansylated peptide (Peptide-N_3). The peptide sequences are listed in Table S1.

Enzymatic assays were performed to confirm broad spectrum inhibition by ODN-**1**, namely, that the specific GST binder can bind and inhibit the activity of different GST isosymes. As shown in Figure 3, the enzymatic activity of GST-A1, GST-M1, GST-A2, and GST-P1 was followed in the

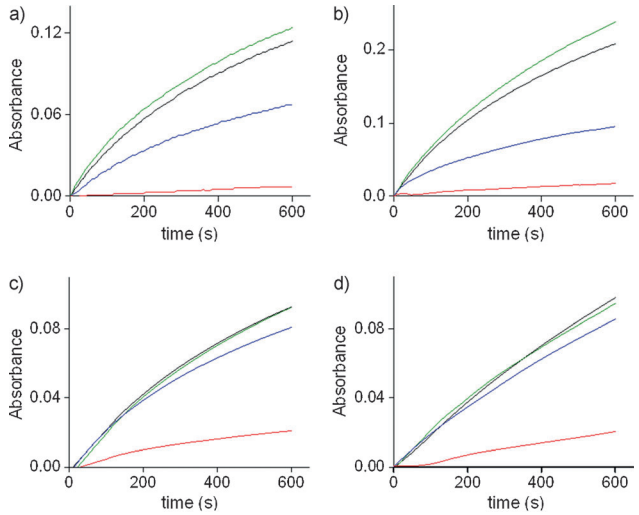

Figure 3. Enzymatic activity of a) GST-A1, b) GST-M1, c) GST-A2, and d) GST-P1 in the absence (black line) and presence of 500 nM ODN-**i** (green line), ODN-**1** (red line), or 1 μM EA (blue line). The formation of the CDNB-GSH conjugate was monitored by following the absorbance at $\lambda = 340$ nm.

absence and presence of ODN-**1** (500 nM), a known monovalent inhibitor (EA, 1 μM), and a control strand (Figure 2a, ODN-**i**), which lacks the EA groups (500 nM). The much stronger inhibition of these isozymes by ODN-**1** indicates that a bivalent, broad-spectrum GST binder with nanomolar affinities toward the four different isozymes was obtained.

One of the advantages of using DNA duplexes for scaffolding these receptors is the simplicity by which a set of water-soluble, multifunctional protein receptors can be obtained by self-assembly.[9] Accordingly, five GST receptors consisting of specific and non-specific binding strands (Figure 1b, receptors **1/a**–**1/e**) were generated simply by heating an equimolar ratio of ODN-**1** and complementary ODNs-**a**–**e** to 95 °C and slowly cooling to room temperature. The formation of these duplexes was confirmed by observing fluorescence resonance energy transfer (FRET) processes between the dansyls and Cy3 (Figure 4a). The acceptor emission was further enhanced when the different duplexes were incubated with different GSTs (Figure 4b and Figure S2), which indicates the suitability of these probes for signaling the receptor–GST interactions.

The most important principle underlying the design of this sensory system is that binding cooperativity[6] between the specific and non-specific binders would transform relatively weak tripodal peptide–GST interactions into much stronger ones (Figure 1b). To confirm the manifestation of the "chelate effect", which should endow the modified duplexes with selectivity toward GSTs, we performed three complementary experiments (Figure 5 and Figures S3 and S4). In the first, we incubated the different receptors (**1/a**–**1/e**) with different GSTs (150 nM) and compared their fluorescence response to that of duplexes, which lack a GST inhibitor (Figure 5a). In a second experiment, the different receptors were treated with common serum proteins or protein biomarkers, such as immunoglobulin A (IgA), α_1-acid glycoprotein (AGP), mixed-type haptoglobin (Hp), fibroblast growth factor-21 (FGF-21), leptin, apolipoprotein A1 (APOA1), immunoglobulin G (IgG), and transferrin (Tf) (Figure 5b). Finally, we tested the response of the receptors to more complex biological mixtures by incubating them with combinations of serum proteins (Figure S3, 1.05 μM total) or human urine (Figure S4). Remarkably, whereas duplexes **1/a**–**1/e** exhibited a strong response to nanomolar concentrations of GSTs (150 nM), only a small or negligible effect on Cy3 emission was observed for the control duplexes (Figure 5a). Similarly, the emission intensity was only slightly changed upon incubating the receptors with an excess (0.5–1 μM) of serum proteins (Figure 5b) or biological mixtures (Figures S3 and S4). Taken together, these results show that the fluorescence response in our system mainly originates from the interaction of the receptors with GSTs.

The development of targeted protein surface sensors such as **1/a**–**1/e** might lead to the realization of cross-reactive sensor arrays that can discriminate among structurally similar biomarkers in biological mixtures. It has been shown, for example, that elevated levels of GST-A1,[8e] GST-P1,[8f] GST-M1,[8b,c] or a combination of GST-A1 and GST-P1[8g,h] in biofluids indicate of distinct diseases. To demonstrate the potential use of such systems in medical diagnosis, we tested the ability of a sensor array prototype, (duplexes **1/a**–**1/e**) to discriminate among different GST isoforms and their combinations in the absence (Figure 6) and presence of other proteins (Figure 6b), as well as within human urine (Figure 7). In a typical experiment, receptors **1/a**–**1/e** (20 nM) were loaded onto a microwell plate and the changes in Cy3 emission intensities were recorded for four replicates. As shown in Figure 6a, the addition of pure GSTs (150 nM) resulted in distinct fluorescence profiles for GST-A1, GST-M1, GST-P1, GST-A2 as well as for combination of GST-A1 and GST-P1. Analyzing the patterns using linear discriminant analysis (LDA; Figure 6b, squares) resulted in markedly distinct clusters, which confirm the ability of the system to discriminate among structurally similar GST isozymes. No

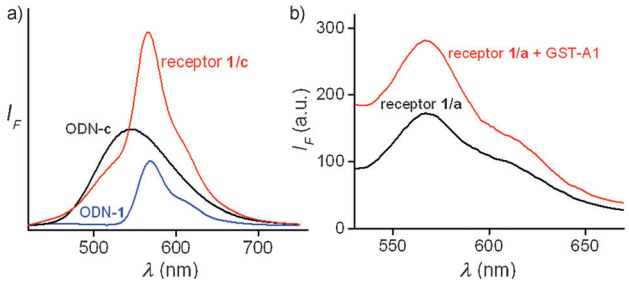

Figure 4. a) Emission spectra of ODN-**1**, ODN-**c**, and receptor **1**/**c** (1 μM). b) Representative fluorescence response of receptor **1**/**a** (20 nM) (black line) before and (red line) after the addition of GST-A1 (150 nM) in PBS buffer. Excitation: 325 nm.

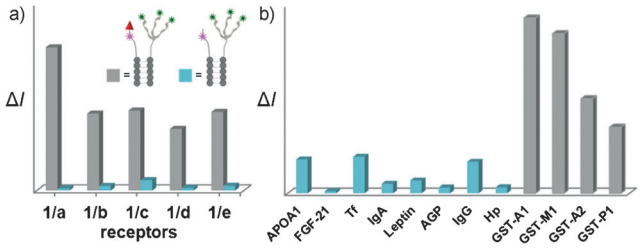

Figure 5. a) Fluorescence response of receptors **1/a**–**1/e** (gray ■) and the corresponding control duplexes (blue ■), which lack the GST inhibitor, to the addition of GST-P1 (150 nM). b) Changes in the fluorescence intensity of receptor **1/a** (20 nM) upon the addition of different GST isoforms (150 nM each) or an excess of common serum proteins: APOA1 (500 nM), FGF-21 (1 μM), Tf (500 nM), IgA (1 μM), leptin (1 μM), AGP (1 μM), IgG (500 nM), and Hp (1 μM).

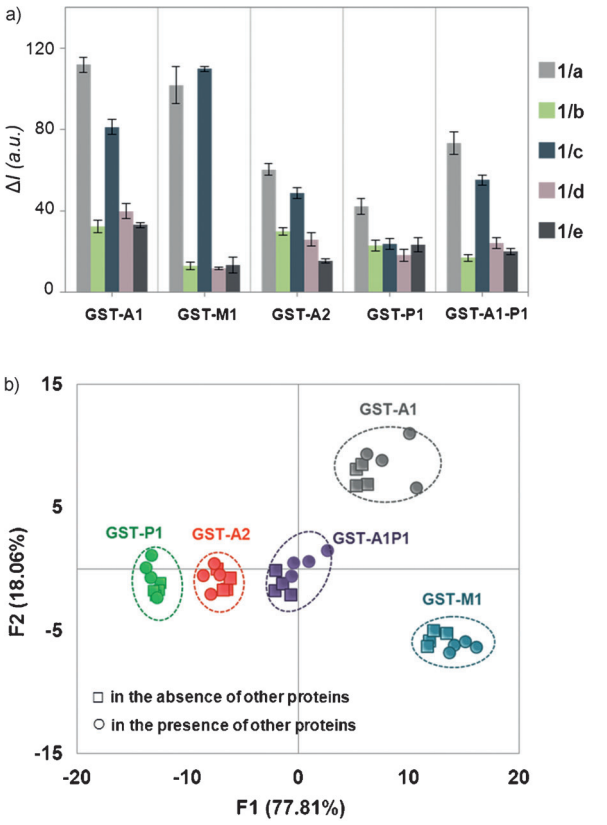

Figure 6. a) Fluorescence patterns generated by receptors **1/a**–**1/e** upon incubating with each of the four GST isoforms (150 nM) or with a combination of GST-A1 and GST-P1. b) LDA plot showing the response of the array to different GSTs (□) in PBS buffer and (○) in the presence of a 1.05 μM protein mixture containing Tf (150 nM), Hp (150 nM), IgG (150 nM), AGP (150 nM), APOA1 (150 nM), leptin (100 nM), FGF-21 (100 nM), and Tg (100 nM).

significant changes in the emission patterns were observed when treating the array with a higher concentration of enzymes (300 nM, Figure S5), indicating that under these conditions the addition of 150 nM of GSTs leads to saturation of the receptors in the array. The discrimination efficiency of the system was further demonstrated by analyzing unknown samples from the training set, with 94% accuracy.

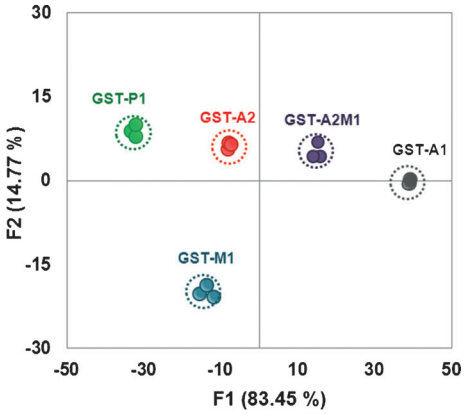

Figure 7. LDA plot showing the response of the array to different GST isozymes (500 nM) in urine.

In the next step, we spiked each receptor in the array with a mixture of eight serum proteins (with a total concentration of 1.05 μM) and fluorescence patterns were recorded under similar conditions (Figure 6b, circles). Here, a higher concentration of GSTs (300 nM) was required in order to compete with some non-specific interactions within this matrix (Figure S3) and to obtain sufficient fluorescent responses. As expected from our design (Figure 1b), the fluorescent patterns in the presence and absence of the protein mixture were very similar to each other (Figure 6b and Figure S6). Unknown GST samples containing this matrix could be identified with an accuracy of 91%. The feasibility of distinguishing among different GSTs in a more complex biological environment such as urine was also demonstrated (Figure 7 and Figure S7). Because human urine features more than 1500 proteins[10] and exhibits a slight background emission signal, higher concentrations of receptors (50 nM) and GSTs (500 nM) were required to obtain adequate signal-to-noise ratios needed for recording reproducible clusters, which enabled the analysis of unknown urine samples with 94% accuracy. Notably, our current detection limit (i.e., 500 nM) is above GST concentrations in urine, which are in the low nanomolar range.[11] However, urinary proteins can be collected in large amounts completely non-invasively and concentrated to provide adequate amounts of proteins. In addition, GSTs can be crudely enriched from this mixture with a commercially available GSH column. Thus, these results indicate the potential use of such systems in analyzing small populations of proteins in medicinally relevant samples.

In conclusion, we have demonstrated a general methodology for endowing optical cross-reactive sensor arrays[2,4,5c,12] with selectivity toward particular protein groups. The fundamental principle underlying this approach is the feasibility of creating signaling protein receptors that integrate two distinct binding motifs: specific and non-specific. The first enables the receptors to bind particular protein groups with high affinity and selectivity, whereas the second allows them to differentiate among closely related isoforms. The use of DNA-duplexes for scaffolding the receptors not only provides the system with inherent water solubility—it also facilitates device integration and modification. This should allow one

to "program" such devices to identify various other classes of protein biomarkers simply by modifying the specific binding strand. Research along this line is currently being pursued in our laboratory.

Received: February 17, 2014
Published online: May 19, 2014

Keywords: cross-reactive sensor array · fluorescent probes · multivalency · pattern-based detection · protein surface recognition

[1] a) N. L. Anderson, N. G. Anderson, *Mol. Cell. Proteomics* **2002**, *1*, 845; b) R. Gonzalez, S. Varnum, R. Zangar in *Biomarker Methods in Drug Discovery and Development* (Ed. F. Wang), Humana Press, Totowa, NJ, **2009**, pp. 273; c) S. L. Servoss, R. Gonzalez, S. Varnum, R. C. Zangar, *Methods Mol. Biol.* **2009**, *520*, 143.

[2] For recent reviews, see: a) D. Margulies, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **2010**, *14*, 705; b) O. R. Miranda, B. Creran, V. M. Rotello, *Curr. Opin. Chem. Biol.* **2010**, *14*, 728.

[3] a) H. Zhou, L. Baldini, J. Hong, A. J. Wilson, A. D. Hamilton, *J. Am. Chem. Soc.* **2006**, *128*, 2421; b) L. Baldini, A. J. Wilson, J. Hong, A. D. Hamilton, *J. Am. Chem. Soc.* **2004**, *126*, 5656; c) D. Margulies, A. D. Hamilton, *Angew. Chem.* **2009**, *121*, 1803; *Angew. Chem. Int. Ed.* **2009**, *48*, 1771; d) D. Margulies, A. D. Hamilton, *J. Am. Chem. Soc.* **2009**, *131*, 9142; e) 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; f) C.-C. You, O. R. Miranda, B. Gider, P. S. Ghosh, I.-B. Kim, B. Erdogan, S. Krovi, S. A. Krovi, U. H. F. Bunz, V. M. Rotello, *Nat. Nanotechnol.* **2007**, *2*, 318; g) M. De, S. Rana, H. Akpinar, O. R. Miranda, R. R. Arvizo, U. H. F. Bunz, V. M. Rotello, *Nat. Chem.* **2009**, *1*, 461; h) P. Wu, L.-N. Miao, H.-F. Wang, X.-G. Shao, X.-P. Yan, *Angew. Chem.* **2011**, *123*, 8268; *Angew. Chem. Int. Ed.* **2011**, *50*, 8118; i) O. R. Miranda, H.-T. Chen, C.-C. You, D. E. Mortenson, X.-C. Yang, U. H. F. Bunz, V. M. Rotello, *J. Am. Chem. Soc.* **2010**, *132*, 5285; j) S. Kolusheva, R. Zadmard, T. Schrader, R. Jelinek, *J. Am. Chem. Soc.* **2006**, *128*, 13592; k) A. T. Wright, M. J. Griffin, Z. Zhong, S. C. McCleskey, E. V. Anslyn, J. T. McDevitt, *Angew. Chem.* **2005**, *117*, 6533; *Angew. Chem. Int. Ed.* **2005**, *44*, 6375; l) D. Zamora-Olivares, T. S. Kaoud, K. N. Dalby, E. V. Anslyn, *J. Am. Chem. Soc.* **2013**, *135*, 14814; m) H. Kong, Y. Lu, H. Wang, F. Wen, S. Zhang, X. Zhang, *Anal. Chem.* **2012**, *84*, 4258.

[4] E. V. Anslyn, *J. Org. Chem.* **2007**, *72*, 687.

[5] a) B. Rout, L. Unger, G. Armony, M. A. Iron, D. Margulies, *Angew. Chem.* **2012**, *124*, 12645; *Angew. Chem. Int. Ed.* **2012**, *51*, 12477; b) B. Rout, P. Milko, M. A. Iron, L. Motiei, D. Margulies, *J. Am. Chem. Soc.* **2013**, *135*, 15330; c) B. Rout, L. Motiei, D. Margulies, *Synlett* **2014**, *25*, 1050.

[6] a) M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem.* **1998**, *110*, 2908; *Angew. Chem. Int. Ed.* **1998**, *37*, 2754; b) C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Koksch, J. Dernedde, C. Graf, E.-W. Knapp, R. Haag, *Angew. Chem.* **2012**, *124*, 10622; *Angew. Chem. Int. Ed.* **2012**, *51*, 10472.

[7] a) H. Yin, A. D. Hamilton, *Angew. Chem.* **2005**, *117*, 4200; *Angew. Chem. Int. Ed.* **2005**, *44*, 4130; b) S. Fletcher, A. D. Hamilton, *J. R. Soc. Interface* **2006**, *3*, 215; c) Q.-Q. Jiang, L. Bartsch, W. Sicking, P. R. Wich, D. Heider, D. Hoffmann, C. Schmuck, *Org. Biomol. Chem.* **2013**, *11*, 1631; d) A. J. Wilson, *Chem. Soc. Rev.* **2009**, *38*, 3289; e) P. R. Wich, C. Schmuck, *Angew. Chem.* **2010**, *122*, 4207; *Angew. Chem. Int. Ed.* **2010**, *49*, 4113.

[8] a) B. Wu, D. Dong, *Trends Pharmacol. Sci.* **2012**, *33*, 656; b) C. C. McIlwain, D. M. Townsend, K. D. Tew, *Oncogene* **2006**, *25*, 1639; c) H.-W. Lo, F. Ali-Osman, *Curr. Opin. Pharmacol.* **2007**, *7*, 367; d) S. S. Mahajan, L. Hou, C. Doneanu, R. Paranji, D. Maeda, J. Zebala, W. M. Atkins, *J. Am. Chem. Soc.* **2006**, *128*, 8615; e) L. K. Dajani, E. Paus, D. J. Warren, *Clin. Chem.* **2001**, *47*, 867; f) S. Tsuchida, Y. Sekine, R. Shineha, T. Nishihira, K. Sato, *Cancer Res.* **1989**, *49*, 5225; g) C. Jochum, M. Beste, J.-P. Sowa, M. Farahani, V. Penndorf, S. Nadalin, F. Saner, A. Canbay, G. Gerken, *Eur. J. Med. Res.* **2011**, *16*, 34; h) A. J. W. Branten, T. P. J. Mulder, W. H. M. Peters, K. J. M. Assmann, J. F. M. Wetzels, *Nephron* **2000**, *85*, 120.

[9] a) C. Battle, X. Chu, J. Jayawickramarajah, *Supramol. Chem.* **2013**, *25*, 848; b) F. Abendroth, A. Bujotzek, M. Shan, R. Haag, M. Weber, O. Seitz, *Angew. Chem.* **2011**, *123*, 8751; *Angew. Chem. Int. Ed.* **2011**, *50*, 8592; c) S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, *Nat. Biotechnol.* **2004**, *22*, 568; d) D. C. Harris, X. Chu, J. Jayawickramarajah, *J. Am. Chem. Soc.* **2008**, *130*, 14950; e) K. I. Sprinz, D. M. Tagore, A. D. Hamilton, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3908; f) B. A. Rosenzweig, N. T. Ross, D. M. Tagore, J. Jayawickramarajah, I. Saraogi, A. D. Hamilton, *J. Am. Chem. Soc.* **2009**, *131*, 5020; g) F. Diezmann, O. Seitz, *Chem. Soc. Rev.* **2011**, *40*, 5789.

[10] J. Adachi, C. Kumar, Y. Zhang, J. Olsen, M. Mann, *Genome Biol.* **2006**, *7*, R80.

[11] C. M. Walshe, F. Odejayi, S. Ng, B. Marsh, *Crit. Care Resusc.* **2009**, *11*, 204.

[12] a) K. Severin, *Curr. Opin. Chem. Biol.* **2010**, *14*, 737; b) J. P. Anzenbacher, P. Lubal, P. Bucek, M. A. Palacios, M. E. Kozelkova, *Chem. Soc. Rev.* **2010**, *39*, 3954; c) J. R. Askim, M. Mahmoudi, K. S. Suslick, *Chem. Soc. Rev.* **2013**, *42*, 8649.

