Proteins

DOI: 10.1002/anie.201307869

Tailoring Chimeric Ligands for Studying and Biasing ErbB Receptor Family Interactions**

*Andrew T. Krueger, Carsten Kroll, Edgar Sanchez, Linda G. Griffith, and Barbara Imperiali**

Abstract: *Described is the development and application of a versatile semisynthetic strategy, based on a combination of sortase-mediated coupling and tetrazine ligation chemistry, which can be exploited for the efficient incorporation of tunable functionality into chimeric recombinant proteins. To demonstrate the scope of the method, the assembly of a set of bivalent ligands, which integrate members of the epidermal growth factor (EGF) ligand family, is described. By using a series of bivalent EGFs with variable intraligand spacing, the differences in structure were correlated with the ability to bias signaling in the ErbB receptor family in a cell motility assay. Biasing away from EGFR-HER2 dimerization with a bivalent EGF was observed to reduce cell motility in an intraligand distance-dependent fashion, thus demonstrating the utility of the approach for acutely perturbing receptor-mediated cell signaling pathways.*

The ErbB/HER/epidermal growth factor receptor (EGFR) family comprises four closely related receptor tyrosine kinases (RTK) which are known to regulate diverse organismal phenotypes including cell migration, proliferation, and differentiation.[1] Evidence supports activation by receptor-specific protein–ligand binding which drives and stabilizes receptor dimerization or oligomerization. One member, HER2, has no known ligands but exists in a conformation that is poised for heterodimerization and subsequent signal transduction. Mutation, dysregulation, and overexpression of the EGFR family of receptors and their ligands are prevalent in many epithelial cancers, and specific receptor dimers, particularly heterodimers with HER2, may propagate malignant signals.[2] Understanding the complex mechanisms and interplay of these signaling processes and receptor dynamics is crucial for being able to open up new paradigms for targeting the EGFR signaling pathway. As such, the development of novel and diverse sets of tools designed to systematically manipulate receptor activation through nongenetic means should advance our knowledge towards new therapeutic approaches.

One promising strategy involves the tethering of two EGF family ligands to create bivalent ligands.[3] Preliminary studies with cell lines treated with bivalent ligands for EGFR family members exhibit phenotypic and signaling behaviors consistent with a biasing of receptor dimerization or oligomerization away from those involving HER2, as compared to treatment with monovalent ligands (as illustrated in Figure 1).

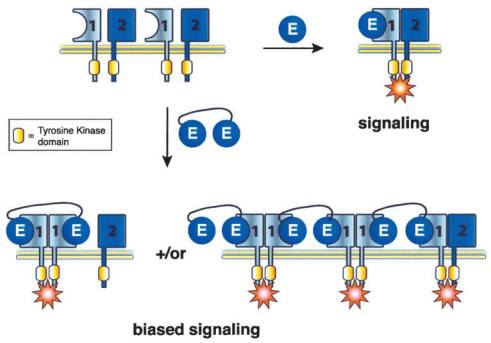

Figure 1. Illustration of biased signaling in ErbB receptors with bivalent ligands (E = epidermal growth factor).

Early prototypes of these bivalent ligands were created by recombinantly expressing individual ligands [either neuregulin-1β (NRG) or epidermal growth factor (EGF)] fused to one partner of a coiled-coil monomer motif and joining the desired ligands noncovalently (later a bivalent single-chain construct was expressed)[4] with a fixed amino-acid linker spacing.

The bivalent ligand strategy has opened the door for progress in elucidating the complex systems biology of the EGF receptor family in cancer therapeutics, regenerative medicine, as well as other physiological processes.[5] Therefore, the ability to efficiently produce a variety of bivalent ligands with tailored functionality to interrogate the EGFR signaling network, and address fundamental biophysical questions regarding receptor dynamics and trafficking, could potentially provide guidance towards, and lead to new models for therapeutic intervention.

To this end we have designed and established a general semisynthetic strategy for advancing the scaffold of bivalent protein ligands as a practical and diversely functionalized tool

[*] Dr. A. T. Krueger, Dr. C. Kroll, Prof. B. Imperiali
Departments of Biology and Chemistry
Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA (USA)
E-mail: imper@mit.edu
E. Sanchez, Prof. L. G. Griffith
Department of Biological Engineering
Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA (USA)

[**] We acknowledge Prof. David Liu, Harvard University (Sortase mutant plasmids), Sherlyn Jaimes (assistance with motility assay), the Integrated Cancer Biology Program at MIT (U54-CA112967), R01DE019523-13 (L.G.G.), CSBI Merck-MIT postdoctoral fellowship (A.T.K.), the Swiss National Science Foundation Postdoctoral Fellowship (C.K), and MIT NIEHS Toxicology Training Grant (E.S.) 5-T32-ES007020 for support.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201307869.

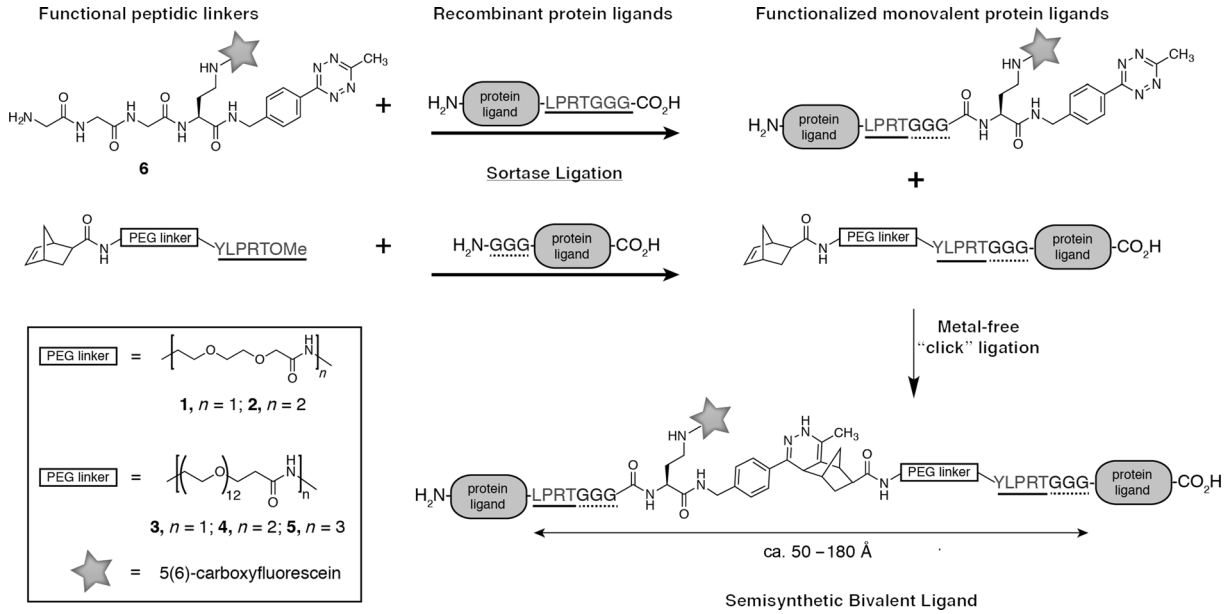

Figure 2. Semisynthetic bivalent linkage strategy assembles two protein ligands with tunable functionality. Examples of protein ligands examined in this study include epidermal growth factor (EGF) and neuregulin (NRG).

by taking a chemical approach in linking (in theory any) two recombinant protein ligands by a synthetic linker. This strategy allows distance tunability, facile label incorporation, a covalent linkage, and a single expression strategy capable of producing multiple bivalent tools with customized linkages tuned to experimental objectives. The approach (Figure 2) relies on efficient expression of two proteins modified with short N- or C-terminal amino-acid sequences for sortase-mediated ligation.[6] The ligation modifies the protein terminus with separate chemically synthesized peptidic components, including either a norbornene or tetrazine moiety, which undergo a metal-free bioorthogonal ligation[7] to link the proteins of interest (e.g. growth factors) through their termini in a C-to-N fashion. We demonstrate the utility of this method to synthesize an array of bivalent ligands based on EGF and NRG. Furthermore, the set of bivalent EGF tools is applied in a phenotypic single-cell migration assay, thus providing a structure–function relationship for effective reduction of cell migration as a function of linker length.

Our strategy initially focused on optimization of recombinant growth factor expression using the 53-amino-acid EGF ligand as a prototype. Ultimately, we could obtain efficient expression and purification of EGF in *E. coli* and reagent quantities (ca. 30–50 $mg\,L^{-1}$) of EGF could consistently be generated (see the Supporting Information for gene sequences and expression methods).[8,9] Mutagenesis allowed incorporation of either N-terminal GGG- or C-terminal -LPRTGGG sortase recognition sequences and gave similar expression yields. Similar yields with neuregulin-1β were also obtained.

Peptidic linker components for incorporation of chemical diversity into the bivalent linkage were designed to systematically vary the linker length within the bivalent ligand to probe its effect on biasing receptor interactions. To this end we synthesized a set of norbornene-modified peptides (**1**–**5**) for N-terminal ligand modification (see the Supporting Information for details). The variation of linker length was achieved during peptide synthesis by commercially available Fmoc-protected PEG_n amino acid building blocks, thus ultimately providing intraligand distances ranging from 50–180 Å. For C-terminal ligand modification, a 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzylamine-functionalized 5(6)-carboxyfluorescein-labeled triglycine peptide (**6**) was synthesized. The 6-methyl variant of the tetrazine was selected because of increased stability towards aqueous environments and peptide synthesis conditions.[7b]

By using an evolved triple mutant of the sortase enzyme[10] and expressed GGG-EGF and EGF-LPRTGGG constructs, EGFs were modified with either the tetrazine-containing peptide **6**, or peptides **1**–**5** to produce a set of modular monovalent EGFs containing bioorthogonal handles. Optimized reaction conditions allowed full conversion (as monitored by LC/MS) after 30 minutes on up to a milligram scale of protein using less than 3 equivalents of peptide. Purification involved a simple two-step process of filtering through NiNTA resin to remove the His_6-tagged sortase enzyme and subsequent size-exclusion chromatography. Neuregulin-1β was modified by sortase-mediated ligation with the same modular components. The tetrazine/norbornene PEG_n-modified EGF and NRG products were verified by LC/MS (see Tables S1 and S2, and Figure S1 in the Supporting Information) and showed biological activity comparable to wild-type EGF/NRG as confirmed by downstream kinase activation of ERK 1/2 and AKT (see Figure S2).

With the tetrazine- and norbornene-handle-modified monovalent pieces in hand, simple incubation of the appropriate components in aqueous buffer afforded bivalent ligands. The various semisynthetic bivalent EGF (EE) and NRG (NN) ligands were synthesized by combining the modified monovalent pieces in a 1.2:1 (tetrazine:norbornene

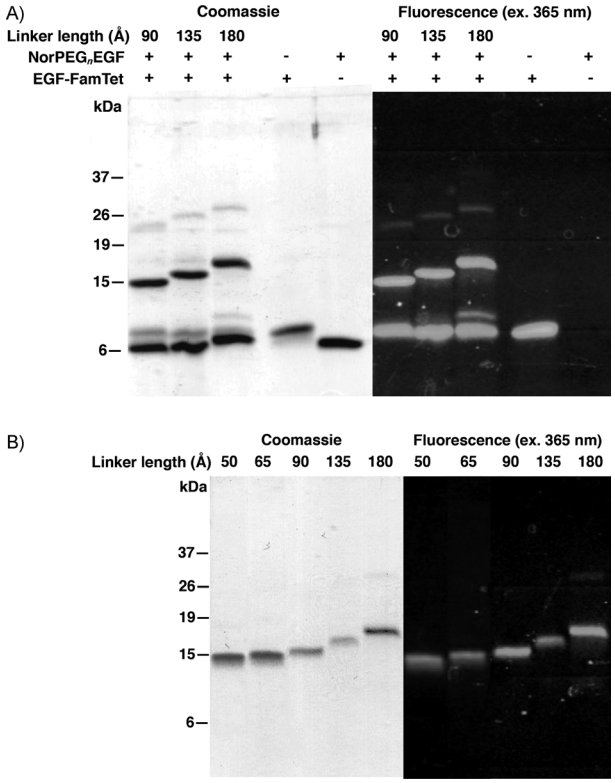

Figure 3. Representative SDS-PAGE analyses for the click ligation reaction progress (A) and final purified bivalent ligand set (B; Bivalent EGFs (EEs) shown. For bivalent NRG (NN) see the Supporting Information).

component) ratio (ca. 100 μM) and incubating at room temperature for 24 hours, after which the ligation progress appeared to plateau as observed by SDS-PAGE (Figure 3A).[11] Purification by size-exclusion chromatography yielded a set functionalized bivalent ligands with variable tether lengths and purities estimated (by LC/MS profiles) at greater than 95% (see Figure 3B and Figures S3 and S4).

To validate the new chimeric constructs and demonstrate the utility of the strategy, we implemented the set of semisynthetic bivalent EGF ligands in single-cell motility assays in hTERT mesenchymal stem cells (hTMSCs). Previous studies showed that a coiled-coil bivalent ligand induced EGFR clustering/oligomerization on the cell surface, thus resulting in biasing of EGFR-mediated signaling in hTMSCs and suggesting that a role in cell motility may also be biased upon treatment with a bivalent ligand.[11] This particular cell line is known to express EGFR, HER2, and HER3 with no quantifiable HER4. It is hypothesized that an important signaling pathway in hTMSC migration occurs through the initiation of EGFR-HER2 heterodimerization from binding of EGF to EGFR (HER2 has no associated ligand).[12] As such, biasing EGFR away from an interaction with HER2 toward EGFR homodimers should lead to a decrease in cell motility. Moreover, depending on the EE intraligand distance, this motility phenotype may be attenuated to varying degrees, thus providing a structure–function relationship. Ultimately, such a relationship may reveal insight into signaling pathways and guide potential strategies for intervention.

Treatment of hTMSCs with various concentrations of either of the bivalent ligands (50–180 Å) or their tether-modified monovalent counterparts resulted in diverse motility behaviors as quantified by tracking the paths of 50 individual cells for each set of conditions (see Figures S5–S7) and analyzing these tracks to obtain quantifiable metrics of cell motility: total path length versus net displacement from the initial position (Figure 4A and Figure S8A) and the

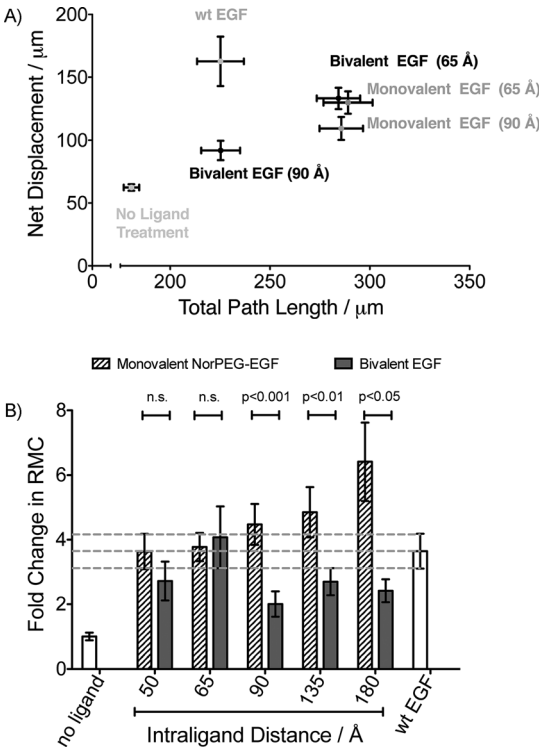

Figure 4. hTMSC migration response to treatment with 100 nM semisynthetic bivalent EGF ligand as compared to the respective PEG$_n$-modified monovalent EGF. A) Mean and S.E. of the net displacement and total path length of hTMSCs upon treatment with wild type (wt), monovalent, or bivalent (65 Å and 90 Å) EGFs. B) Fold change in the random motility coefficient (RMC/μm^{2}h^{-1}) in response to various ligand treatments. See the Supporting Information for statistical analyses.

random motility coefficient, which captures contributions from cell speed and persistence (Figure 4B and Figures S8B and S9). At 100 nM, a concentration expected to result in receptor saturation (unmodified EGF $K_D \approx 1$ nM), bivalent ligands with intraligand spacing ranging from 90–180 Å showed a marked decrease in cell motility relative to unmodified monovalent EGF or the monovalent tether-modified variants, with 90 Å-spaced bivalent exhibiting the largest reduction (Figure 4B). Interestingly, for the shorter EE ligands (50 Å, 65 Å) at 100 nM the stimulated migration phenotype was comparable to the monovalent counterparts and to unmodified EGF. Similar trends were observed at a ligand concentration comparable to K_D (1 nM), where the 90 Å spacing resulted the greatest inhibition of motility compared to the monovalent counterpart or native EGF (Figure S6B). One observation of note is that the PEG tether-

modified EGF monovalent ligands stimulated motility parameters above the levels for native unmodified EGF.

Crystallographic information from a structure of two bound EGFs in a EGFR homodimer shows an estimated distance of about 90–100 Å between the two monomeric ligands (C-to-N terminus, Figure 5).[13] As such, our ligand set straddles the minimum required distance (linker distances

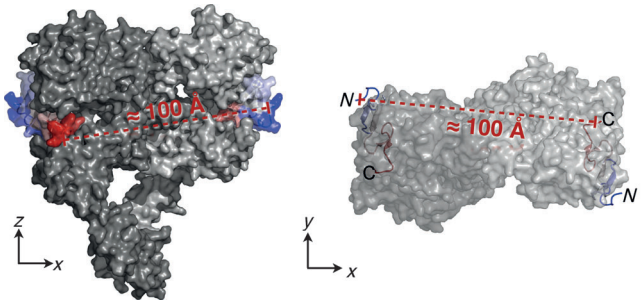

Figure 5. Frontal (left) and top (right) views of crystal structure representation of a homodimer of two extracellular EGFR domains with two bound EGFs, thus indicating the approximate distance between C and N termini of EGF ligands. (PDB ID: 3NJP).[13]

estimated from published or modeling data for peptide and PEG motifs)[14] for one bivalent ligand to sequester a homodimer. Previous bivalent ligands had an intraligand distance of about 200 Å, and while capable of exhibiting a biasing effect, our study shows the benefit of screening a variety of distances to maximize a signal bias. The strongest migration reduction at 90 Å does not seem too surprising considering the length of the PEG chain, thus allowing this bivalent ligand to span the EGFR homodimer. This observation could factor into the relatively slight increase in cell migration with longer linker lengths, thus corresponding with the increase in entropy from less preorganization.[15] The restoration of cell motility at linker lengths below 90 Å may suggest the inability of the bivalent ligand to wrap around an EGFR homodimer and effectively bias away from an EGFR-HER2 heterodimer. However, it could also be possible to bias receptor interactions by concatenation (as illustrated in Figure 1), which would not necessarily require an intraligand distance that spans the homodimer active sites, but still may be sterically hindered from occurring. Our migration data suggests that with the shorter EE ligands the concatenation is not favored. Nevertheless, the ease with which these ligands can be produced will allow future biophysical and biochemical assays to understand EGFR dynamics and phenotypic effects.

While the data support the concept that the bivalent ligands are biasing dimerization or oligomerization toward EGFR homodimers and away from EGFR-HER2 heterodimers (see Figure S10), it is also possible that the ligands are perturbing the biophysical association of the juxtamembrane domain involved in receptor activation to cause differential signaling in a manner similar to that reported for differences in TGFα and EGF.[16] The variable stimulatory response of monovalent tether-modified EGF could similarly be altering EGFR biophysical states during activation. Overall, the observation that the tether-modified monovalent EGFs stimulate cell migration beyond that of wild-type EGF is curious. It is possible the PEG tethers may sterically impede binding, and previous studies of lower-affinity, weaker-binding EGF ligands having increased migration-promoting abilities are known.[17] Quantitative binding data will be useful in confirming these observations.

Considering receptor dimerization or oligomerization is widely recognized among various cell surface receptor mediated signaling pathways, the use of the described strategy for receptor ligands could also see utility in interrogating other signaling processes.[12c,18,19] We are currently applying a similar set of bivalent neuregulins to downstream signaling, proliferation, and phenotypic assays to elucidate its relevance in cancerous cells.

Although examples of semisynthetic bivalent ligand constructs in the size range defined in this study are limited, other approaches for synthesizing bivalent assemblies have included application of solid-phase peptide synthesis (smaller peptidic ligands),[20a] site-selective incorporation of cysteines/aldehydes into antibodies/proteins for subsequent ligation and linking by copper-catalyzed azide–alkyne click chemistry or metal-free click reactions,[21,22] as well as copper-catalyzed azide–alkyne cycloaddition chemistries on peptides[23] or on proteins containing incorporated unnatural amino acids.[24] We chose to employ sortase-mediated ligation because of robust conversion and small amino acid recognition sequences, thus allowing efficient near-native protein expression and ultimately increasing the final semisynthetic bivalent ligand yields (3–6 mg L^{-1} vs. multi-microgram/10 L for the coiled-coil and single-chain constructs). Sortase-mediated ligation has shown great use in linking proteins to proteins, peptides, or other chemical entities. Recently, Ploegh and co-workers described utilizing sortase ligation and metal-free click chemistry to make N-to-N fusions of ubiquitin vinylmethylester (UbVME) and C-to-C fusions of VHH, a variable region of camelid antibodies.[25]

While in this study linkers included variable PEG spacers and a fluorescent label, one could envision incorporation of additional bioorthogonal handles for either more diverse functionality, recognition or specific delivery, or further variation of the rigidity of the spacer utilizing polyproline motifs.[20] Slight modulations in flexibility of the tether have been shown to play a remarkable role in stabilizing dimeric receptor interactions in other systems.[26]

In conclusion, we have described a robust, versatile method for incorporating tunable functionality into chimeric proteins and applied this method to bivalent growth factor ligands for studying and biasing EGFR signaling in hTMSCs using cell motility as a phenotypic readout. Although numerous studies of structure–function relationships for small molecules and their therapeutic targets are known and produce important insight, outside of mutagenesis studies, fewer examples exist with respect to larger protein ligands and their targets. Such studies, as well as methods for undertaking them will become invaluable in moving forward our knowledge, and perhaps establish new thought paradigms toward exploring the semisynthetic therapeutic space.

Received: September 6, 2013
Revised: December 20, 2013
Published online: January 31, 2014

Keywords: biological activity · ligand design · protein modifications · receptors · structure–activity relationships

[1] a) M. A. Olayioye, R. M. Neve, A. H. Lane, N. E. Hynes, *EMBO J.* **2000**, *19*, 3159 – 3167; b) M. A. Lemmon, *Exp. Cell Res.* **2009**, *315*, 638 – 648; c) B. Linggi, G. Carpenter, *Trends Cell Biol.* **2006**, *16*, 649 – 656; d) D. J. Riese II, T. M. van Raaij, G. D. Plowman, G. C. Andrews, D. F. Stern, *Mol. Cell Biol.* **1995**, *15*, 5770 – 5776.

[2] a) D. N. Amin, N. Sergina, D. Ahuja, M. McMahon, J. A. Blair, D. Wang, B. Hann, K. M. Koch, K. M. Shokat, M. M. Moasser, *Sci. Transl. Med.* **2010**, *2*, 16ra7; b) J. Baselga, S. M. Swain, *Nat. Rev. Cancer* **2009**, *9*, 463 – 475; c) F. Ciardiello, G. Tortora, *N. Engl. J. Med.* **2008**, *358*, 1160 – 1174.

[3] S. M. Jay, E. Kurtagic, L. M. Alvarez, S. de Picciotto, E. Sanchez, J. F. Hawkins, R. N. Prince, Y. Guerrero, C. L. Treasure, R. T. Lee, L. G. Griffith, *J. Biol. Chem.* **2011**, *286*, 27729 – 27740.

[4] S. M. Jay, A. C. Murthy, J. F. Hawkins, J. R. Wortzel, M. L. Steinhauser, L. M. Alvarez, J. Gannon, C. A. MacRae, L. G. Griffith, R. T. Lee, *Circulation* **2013**, *128*, 152 – 161.

[5] H. S. Wiley, S. Y. Shvartsman, D. A. Lauffenburger, *Trends Cell Biol.* **2003**, *13*, 43 – 50.

[6] a) T. Proft, *Biotechnol. Lett.* **2010**, *32*, 1 – 10; b) M. W.-L. Popp, J. M. Antos, H. L. Ploegh, *Curr. Protein Pept. Sci.* **2009**, *56*, 15.3.1-15.3.9..

[7] a) N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297 – 2299; b) M. R. Karver, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2011**, *22*, 2263 – 2270.

[8] a) J. G. Marblestone, S. C. Edavettal, Y. Lim, P. Lim, X. Zuo, T. R. Butt, *Protein Sci.* **2006**, *15*, 182 – 189; b) T. R. Butt, S. C. Edavettal, J. P. Hall, M. R. Mattern, *Protein Expression Purif.* **2005**, *43*, 1 – 9.

[9] J. Y. Lee, C. S. Yoon, I. Y. Chung, Y. S. Lee, E. K. Lee, *Biotechnol. Appl. Biochem.* **2000**, *31*, 245 – 248.

[10] I. Chen, B. M. Dorr, D. R. Liu, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11399 – 11404.

[11] A minor, third band was observed on SDS-PAGE, and attributed to an aggregate. Analysis revealed the aggregate is not dynamic, thus only appearing during the ligation reaction (Figure 3a). While the ligation with the slower reacting 6-methyltetrazine variant required overnight to reach about 70% completion, it proceeded equally well at 4 °C. Lack of conversion may be due to partial decomposition of the tetrazine. Other ring-strained partners are being explored to increase reaction kinetics.

[12] a) N. Kumar, M. H. Zaman, H.-D. Kim, D. A. Lauffenburger, *Biophys. J.* **2006**, *91*, L32 – L34; b) S. Wu, A. Wells, L. G. Griffith, D. A. Lauffenburger, *Biomaterials* **2011**, *32*, 7524 – 7531; c) M. Rodrigues, L. G. Griffith, A. Wells, *Stem Cell Res. Ther.* **2010**, *1*, 32 – 44.

[13] C. Lu, L.-Z. Mi, M. J. Grey, J. Zhu, E. Graef, S. Yokoyama, T. A. Springer, *Mol. Cell. Biol.* **2010**, *30*, 5432 – 5443.

[14] a) C. Jeppesen, J. Y. Wong, T. L. Kuhl, J. N. Israelachvili, N. Mullah, S. Zalipsky, C. M. Marques, *Science* **2001**, *293*, 465 – 468; b) F. Inagaki, K. Ogura, H. Kumeta, Y. Kobashigawa, *J. Biomol. NMR* **2009**, *43*, 145 – 150; c) COSMOS 3D structure predictor (http://cosmos.igb.uci.edu/).

[15] M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem.* **1998**, *110*, 2908 – 2953; *Angew. Chem. Int. Ed.* **1998**, *37*, 2754 – 2794.

[16] R. A. Scheck, M. A. Lowder, J. S. Appelbaum, A. Schepartz, *ACS Chem. Biol.* **2012**, *7*, 1367 – 1376.

[17] a) M. Wingens, S. Jacobs-Oomen, S. P. van der Woning, C. Stortelers, E. J. J. van Zoelen, *Biochemistry* **2006**, *45*, 4703 – 4710; b) K. Nandagopal, D. M. Popp, S. K. Niyogi, *J. Cell. Biochem.* **2001**, *83*, 326 – 341; c) C. C. Reddy, S. K. Niyogi, A. Wells, H. S. Wiley, D. A. Lauffenburger, *Nat. Biotechnol.* **1996**, *14*, 1696 – 1699; d) A. K. Iyer, K. T. Tran, L. Griffith, A. Wells, *J. Cell. Physiol.* **2008**, *214*, 504 – 512.

[18] A. Citri, Y. Yarden, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 505 – 516.

[19] J. A. Engelman, K. Zejnullahu, T. Mitsudomi, Y. Song, C. Hyland, J. O. Park, N. Lindeman, C. M. Gale, X. Zhao, J. Christensen, T. Kosaka, A. J. Holmes, A. M. Rogers, F. Cappuzzo, T. Mok, C. Lee, B. E. Johnson, L. C. Cantley, P. A. Jänne, *Science* **2007**, *316*, 1039 – 1043.

[20] a) T. Tanaka, W. Nomura, T. Narumi, A. Masuda, H. Tamamura, *J. Am. Chem. Soc.* **2010**, *132*, 15899 – 15901; b) Y. Nagel, M. Kuemin, H. Wennemers, *Chimia* **2011**, *65*, 264 – 267.

[21] J. E. Hudak, R. M. Barfield, G. W. de Hart, P. Grob, E. Nogales, C. R. Bertozzi, D. Rabuka, *Angew. Chem.* **2012**, *124*, 4237 – 4241; *Angew. Chem. Int. Ed.* **2012**, *51*, 4161 – 4165.

[22] J. G. Schellinger, A. Kudupudi, A. Natarajan, W. Du, S. J. DeNardo, J. Gervay-Hague, *Org. Biomol. Chem.* **2012**, *10*, 1521 – 1526.

[23] M. Sainlos, W. S. Iskenderian-Epps, N. B. Olivier, D. Choquet, B. Imperiali, *J. Am. Chem. Soc.* **2013**, *135*, 4580 – 4583.

[24] B. C. Bundy, J. R. Swartz, *Bioconjugate Chem.* **2010**, *21*, 255 – 263.

[25] M. D. Witte, J. J. Cragnolini, S. K. Dougan, N. C. Yoder, M. W. Popp, H. L. Ploegh, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11993 – 11998.

[26] M. M. Martino, M. Mochizuki, D. A. Rothenfluh, S. A. Rempel, J. A. Hubbell, T. H. Baker, *Biomaterials* **2009**, *30*, 1089 – 1097.

