

DNA as a Molecular Ruler: Interrogation of a Tandem SH2 Domain with Self-Assembled, Bivalent DNA–Peptide Complexes**

Hendrik Eberhard, Franziska Diezmann, and Oliver Seitz*

The recognition properties of DNA enable the sequence-programmed construction of defined DNA architectures by self-assembly.^[1] Chemically modified DNA molecules have been used to precisely control the spatial arrangement of chromophores, metals, nanoparticles, and proteins.^[2] The vast majority of studies have been concerned with the design of materials properties. Comparatively few studies have investigated the DNA-scaffolded arrangement of bioactive ligands for use in the life sciences.^[3] Among the notable examples are DNA-assembled glycoclusters^[4] and encoded self-assembled libraries.^[5] The aim was to mimic biological ligand display and/or to facilitate the drug-discovery process.^[6] Herein we present a new approach. We prepared self-assembled DNA–peptide complexes and explored whether the DNA-controlled presentation of peptides can be used to probe the structural properties of the cognate target protein. Specifically, we studied a protein-binding domain of the Syk kinase, a protein kinase that is involved in the regulation of lymphocyte activation. We demonstrate that important structural parameters, such as the preferred arrangement of protein-binding pockets and the flexibility of the connecting interdomain, can be assessed by using a set of DNA complexes. It is shown that the combination of single-stranded and double-stranded segments in self-assembled ternary complexes provides a convenient means to systematically vary the distance constraint as well as the flexibility of the ligand display.

Many signaling enzymes use two or more homologous protein-binding domains to achieve firm interactions with cognate targets. For example, the Syk and Zap-70 kinases,^[7] among others, contain tandem SH2 domains, which selectively recognize two appropriately spaced tetraphosphopeptide binding motifs of the protein substrate.^[8] In the absence of a solved structure of the protein–peptide complex, it is difficult to assess the requirements for high-affinity protein–protein binding. Chemical probing with bivalent synthetic binders has been used extensively, for example, in studies of the tandem SH2 domain of the Syk kinase.^[9,10] Since this

bivalent protein is well-characterized, we selected it as a target for spatial screening with DNA-based architectures.

The crystal structure of the Syk-tSH2 domain in a complex with a bivalent peptide shows that the two recognized phosphotyrosine residues are arranged within 34 Å of one another (Figure 1a).^[11] The asymmetric unit was found to contain six tSH2 molecules, which showed varied arrangements of the SH2 domains. Chemical probing has revealed

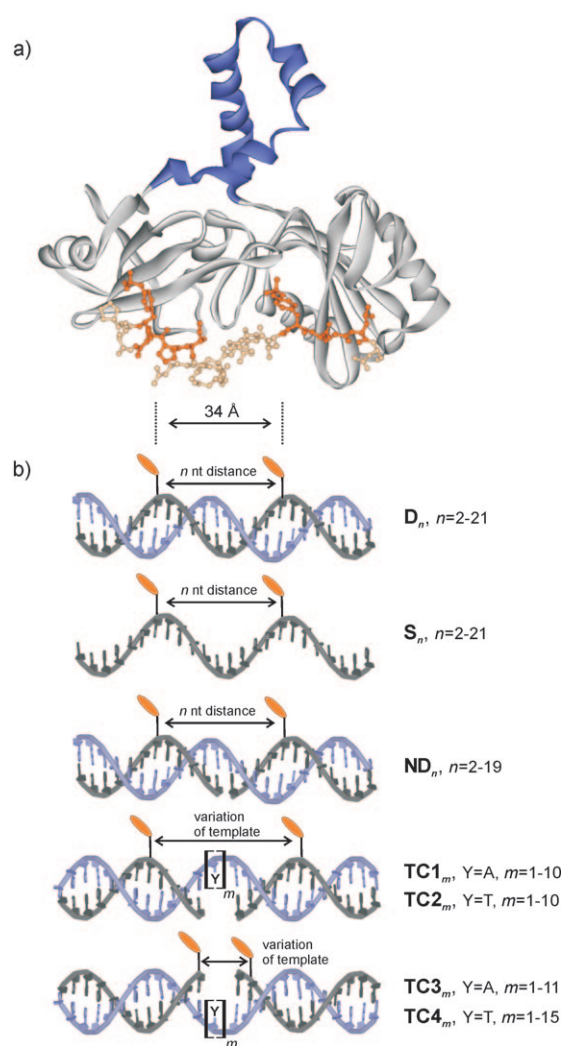


Figure 1. a) Crystal structure (PDB ID: 1A81) of the Syk-tSH2 domain in a complex with a bivalent phosphopeptide (blue, tSH2 interdomain; orange, tetraphosphopeptide binding motifs; pale orange, peptide linker). b) Bivalent presentation of phosphopeptides on single-stranded DNA (S), duplex DNA (D), nicked duplex DNA (ND), and ternary DNA (TC), which contains a single-stranded segment.

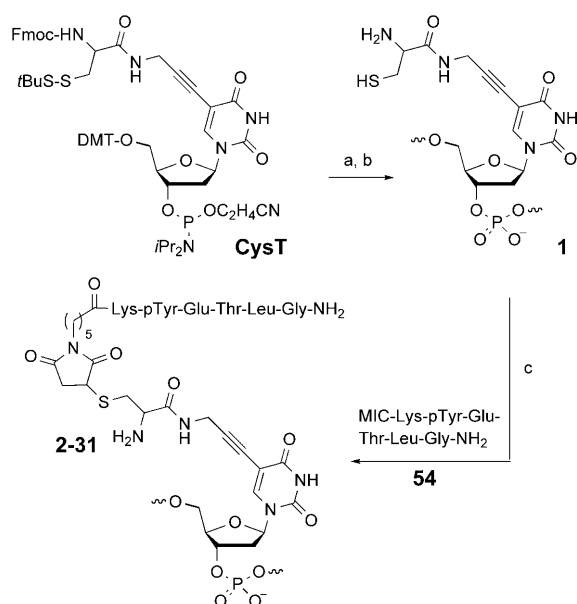
[*] H. Eberhard, F. Diezmann, Prof. Dr. O. Seitz
Department of Chemistry, Humboldt University Berlin
Brook-Taylor-Strasse 2, 12489 Berlin (Germany)
Fax: (+49) 30-2093-7266
E-mail: oliver.seitz@chemie.hu-berlin.de

[**] This research was supported by the Deutsche Forschungsgemeinschaft (SFB 765). We thank Prof. Dr. Gabriel Waksman, University College London, for providing plasmids for GST-Syk-tSH2 expression.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201007593>.

that the tSH2 domain is able to adjust to smaller distances.^[10e,12] We assumed that self-assembled DNA–peptide complexes (Figure 1b) can serve as molecular probes to provide answers to the following questions: 1) What are the tolerated distances between the two recognized binding pockets? 2) How flexible is the tSH2 interdomain in solution?

The first task was to provide robust access to covalently linked oligonucleotide–peptide conjugates. The sequence-internal cysteine modifier **CysT** was introduced during automated DNA synthesis (Scheme 1).^[13] The fully deprotected cysteine-modified oligonucleotides **1** were equipped with a peptide chain in a coupling reaction with the maleimide-functionalized phosphopeptide **54**.^[14] We then assessed the hybridization of the peptide-modified oligonucleotides **2–31** with the complementary oligonucleotides (see the Supporting Information). The peptide had little influence on the stability of the duplexes (see Figure S3 in the Supporting Information).



Scheme 1. Synthesis of oligonucleotide–peptide conjugates. a) DNA synthesis; b) 5 mM TCEP, 10 mM NaCl, 10 mM NaH₂PO₄, pH 6.5; c) 10 mM NaCl, 10 mM NaH₂PO₄, pH 6.5. DMT = dimethoxytrityl, Fmoc = 9-fluorenylmethyloxycarbonyl, MIC = maleimidocaproyl, pTyr = phosphotyrosine, TCEP = tris(2-carboxyethyl)phosphane; X denotes the peptide modifier.

We evaluated the affinity of the DNA–peptide constructs for the Syk–tSH2 protein by measuring the displacement of the FAM-labeled reference peptide FAM-KpYETGLNTRS-QETpYETLG (**55**; FAM = carboxyfluorescein) from the protein complex by means of fluorescence-polarization measurements (see the Supporting Information).^[15] To confirm the integrity of the recombinant tSH2 protein as well as bivalent binding, we tested the bivalent peptide Ac-PDpYE-PIRKGQRDLPYSGLNQRG (**56**), which spans the sequence of the native Syk substrate, TCRε TAM. The obtained IC₅₀ value of 50 nM agreed with the value reported previously.^[9,10] Neither the monovalent DNA–peptide conjugates nor the DNA scaffolds alone (see Figure S5B) exhibited significant affinity for the protein in this assay (IC₅₀ > 10 μM). The doubly peptide labeled single strands **S_n** showed largely enhanced affinity for Syk–tSH2 (IC₅₀ = 180–880 nM), as expected for bivalent binding. The distance between the peptide appendages was found to have only minor effects on protein binding (Figure 2a). The absence of large effects on binding can be attributed to the flexibility of the single strand. Such flexibility has been shown previously through fluorescence measurements of fluorophore-quencher-labeled oligonucleotides.^[16]

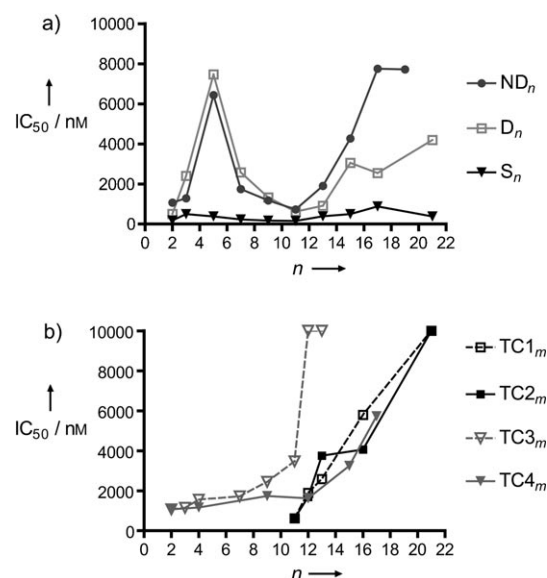


Figure 2. Distance dependence of the binding affinity of DNA–peptide conjugates for Syk–tSH2. A) single-stranded conjugates **S_n**, duplex conjugates **D_n**, and nicked-duplex conjugates **ND_n**; B) ternary complexes **TC1/TC2** and **TC3/TC4** (n = number of paired + number of unpaired spacer nucleotides $m + 1$).

For rigidification, the bivalent single strands **2–11** were hybridized with templates **2'–11'** (see Table S1 in the Supporting Information). In the resulting rigid double strands **D_{2–21}**, the peptides are anchored at a distance of 2–21 nucleotides, which corresponds to 6.8–71.4 Å on the basis of the B-helix structure (as confirmed by CD spectroscopy; see Figure S6). The plot depicted in Figure 2a exposed a maximum of affinity at a distance of 11 nucleotides (nt), which corresponds to 37 Å and is in reasonable agreement with the distance between binding sites found in the crystal structure

of Syk-tSH2 in a complex with a bivalent peptide (Figure 1).^[11] The second IC_{50} minimum at a distance of 2 nt suggests that the interdomain linker in Syk-tSH2 can arrange the SH2 domains within closer proximity. Indeed, thermodynamic analyses revealed previously that the Syk-tSH2 protein can adopt different conformations.^[12] Moreover, a recent binding study showed that the SH2 domains can adjust to distances of 12 Å.^[10e] The binding experiments also showed decreases in binding affinities for duplexes presenting the peptide ligands at distances between 3 and 9 nt and above 11 nt. The IC_{50} maximum around $n=5$ points to a possible role of the DNA helical twist. In duplex **D**₅, the linkers for peptide attachment are separated by half a helical turn and should protrude from opposite sides of the duplex. This arrangement would also explain the low binding affinity of duplexes displaying the peptides at a distance of a 1.5 helical turns (≈ 15 nt). At a distance of 21 nt, the peptides are again expected to be tethered at the same duplex site. However, the low binding affinity for duplex **D**₂₁ suggests that the tSH2 interdomain of Syk can not extend to such a degree that the two binding sites are arranged at a distance of 70 Å.

To attenuate duplex rigidity, we examined the bivalent nick duplexes **ND**₂₋₁₉. Protein-affinity measurements revealed a binding behavior that closely resembled that of the rigid duplexes **D**₂₋₂₁ (Figure 2a). The IC_{50} minima at distances of 2 and 11 nt and the IC_{50} maximum at a distance of 5 nt suggested that the nick duplexes **ND**₂₋₁₉ share structural features with the DNA double strands **D**₂₋₂₁.^[17]

We solved the orientation issue by means of the ternary complexes **TC1/TC2** and **TC3/TC4** (Figure 2b). In **TC3/TC4**, the two peptide-appendage sites were separated by one nucleotide pair and an increasing number of unpaired template nucleotides, which provide rotational degrees of freedom. The affinity for the Syk-tSH2 domain remained high when 1–9 unpaired adenine residues were present (**TC3**₁₋₉). A sharp decrease in the binding affinity was observed as the number n of paired and unpaired (m) spacer nucleotides exceeded the value of 11 (**TC3**₁₀, **TC3**₁₁). The involvement of unpaired thymidine residues in ternary complexes **TC4** led to a more relaxed distance dependence: the decrease in binding affinity occurred at $n > 13$. The ternary complexes **TC3** _{m} and **TC4** _{m} failed to show an IC_{50} maximum at a 5 nt peptide–peptide distance. In the ternary complexes **TC1** _{m} and **TC2** _{m} , the peptide-attachment sites were separated by 10 paired and 1–10 unpaired nucleotides. These complexes had high affinity for the Syk-tSH2 protein when the peptides were separated by a distance of 12–13 nt. However, the binding affinity decreased, again, as the number of spacer nucleotides increased further.

The collected results suggested that the interdomain of the Syk-tSH2 protein adjusts to the spatial arrangement of the bivalent tetraphosphopeptide motifs until a certain threshold distance is reached. The critical maximum length was obtained when the number of spacer nucleotides reached a value of $n = 12$ –13, which would correspond to a distance of approximately 41–44 Å if an increase of 3.4 Å per nucleotide is assumed. Such simple estimates seem justified for rigid double-helical architectures, but probably not for single-strand-containing ternary complexes, such as **TC1/TC2** and

TC3/TC4. We determined the distances experimentally by means of Förster resonance energy transfer (FRET) measurements. The fluorescent dyes Cy3 and Cy5 were attached to oligonucleotides **12 + 22**, **17 + 27**, and **20 + 30** instead of the peptide. We used the acceptor normalization method to calculate the FRET efficiency by measuring the ratio of acceptor emission upon donor excitation and acceptor excitation (see Figure S7).^[18] This method has been validated for distances above 8 nt between the dyes. The experimentally calculated distance between the dyes correlated well with the number of spacer nucleotides (Figure 3). This result allowed

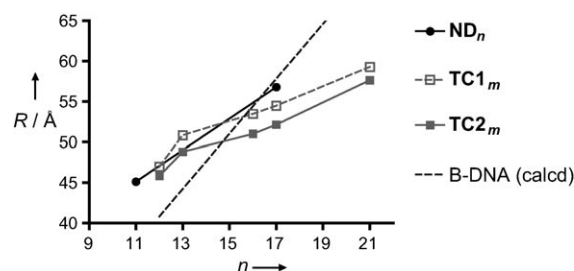


Figure 3. Correlation of the distance R determined by FRET measurements (see the Supporting Information) with the number of spacer nucleotides (n = number of paired + number of unpaired spacer nucleotides $m + 1$). The dark-gray dotted line corresponds to the theoretical 3.4 Å increase per base pair in B-DNA.

the increase in distance per nucleotide to be calculated for the ternary complexes **TC1** and **TC2**. The increase was lower (1.2 Å per nucleotide) than the increase in distance per nucleotide in a B-type duplex (3.4 Å per base pair). The difference may be explained by the higher flexibility of the single-strand regions, which enable partial bending. Though both the peptides and the dyes were attached to DNA by spacers 17 atoms long, the chemical structures and hence the flexibility of the linkers differed (alkyne-linked peptides versus alkene-linked dyes). Thus, binding and FRET experiments are comparable only to a certain extent. However, the determination of distance increase per base pair should not be affected. Importantly, the near-linear correlation between the number n of spacer nucleotides and the distance suggested that ternary complexes, such as **TC1/TC2**, can be used as probes for spatial screening. The FRET measurements suggested that the critical length, at which the interdomain of the Syk-tSH2 protein hampers bivalent binding, is reached when the two tetraphosphopeptide recognition motifs are attached to DNA at distances larger than 50 Å.

Although our study provides no direct evidence for the bivalent binding of the peptide–DNA conjugates, our findings are in line with the results of previous studies.^[8,9] Thermodynamic analyses of binders featuring rigid and semirigid linkers and crystal-structure analyses indicated a high flexibility of the protein interdomain that connects the two SH2 domains, in agreement with our results.^[10c,d,11,12] A recent study involving switchable azobenzene-containing spacers suggested that the tSH2 domain adapts to ligand–ligand distances of 12–30 Å, which is also in agreement with our measurements.^[10e] Semirigid or rigid linkers that span wider distances have not

been examined previously. Such investigations are difficult because the aggregation tendency of the commonly used rigid scaffolds complicates spatial screening when long linkers are involved.

DNA molecules confer high solubility in water regardless of length. Furthermore, the degree of flexibility of DNA-based scaffolds can be adjusted readily. Double-stranded DNA spacers are rigid rodlike structures, whereas single-stranded DNA spacers provide flexibility. The protein-binding data show that the helical twist of double-stranded DNA templates is a problem. Unpaired spacer nucleotides between double-stranded segments in ternary complexes were found to relieve the torsion constraints. To our knowledge, the display of bivalent ligands on such structures has not been studied previously. The FRET measurements suggest a clear correlation between the number of spacer nucleotides and calculated distances. Ternary complexes, such as **TC1/TC2** or **TC3/TC4**, may thus serve as semirigid molecular rulers, which can be constructed readily by self-assembly. In this study, just four covalently linked oligonucleotide-peptide conjugates were used to form 22 different ternary complexes.

In conclusion, we have demonstrated that DNA molecules are suitable scaffolds for the bivalent presentation of peptide ligands. Protein-binding experiments of the self-assembled complexes suggested important features of bivalent binding, such as a) the accessible distance between two homologous binding pockets of a protein receptor and b) the flexibility of the protein linker region connecting the two binding modules. These key parameters have important consequences for the selectivity and specificity of protein-protein interactions involved in cellular signal transduction and may thus be helpful for the evaluation of proteome interaction networks and for the design of artificial protein binders. Furthermore, cell-endogenous RNA molecules may also serve as templates for the assembly of complexes that perturb intracellular protein-protein interactions. In a fascinating scenario, the construction of artificial feedback loops may therefore be possible. Future studies will be directed toward the nucleic acid programmed construction of homo- and heterobivalent inhibitors of adaptor proteins.

Received: December 3, 2010

Revised: January 18, 2011

Published online: March 31, 2011

Keywords: DNA-peptide conjugates · multivalency · nucleic acids · protein binding · spatial screening

- [1] a) U. Feldkamp, C. M. Niemeyer, *Angew. Chem.* **2006**, *118*, 1888–1910; *Angew. Chem. Int. Ed.* **2006**, *45*, 1856–1876; b) S. K. Silverman, *Angew. Chem.* **2010**, *122*, 7336–7359; *Angew. Chem. Int. Ed.* **2010**, *49*, 7180–7201; c) P. W. K. Rothmund, *Nature* **2006**, *440*, 297–302.
- [2] a) R. Varghese, H.-A. Wagenknecht, *Chem. Commun.* **2009**, 2615–2624; b) H. Yang, K. L. Mettera, H. F. Sleiman, *Coord. Chem. Rev.* **2010**, *254*, 2403–2415; c) C. J. Ackerson, M. T. Sykes, R. D. Kornberg, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13383–13385; d) J. L. Furman, A. H. Badran, O. Ajulo, J. R. Porter, C. I. Stains, D. J. Segal, I. Ghosh, *J. Am. Chem. Soc.* **2010**, *132*, 11692–11701; e) B. Saccà, R. Meyer, M. Erkelenz, K. Kiko, A. Arndt, H. Schroeder, K. S. Rabe, C. M. Niemeyer, *Angew. Chem.* **2010**, *122*, 9568–9573; *Angew. Chem. Int. Ed.* **2010**, *49*, 9378–9383.
- [3] L. Röglin, O. Seitz, *Org. Biomol. Chem.* **2008**, *6*, 3881–3887.
- [4] a) K. Matsuura, M. Hibino, Y. Yamada, K. Kobayashi, *J. Am. Chem. Soc.* **2001**, *123*, 357–358; b) K. Matsuura, M. Hibino, T. Ikeda, Y. Yamada, K. Kobayashi, *Chem. Eur. J.* **2004**, *10*, 352–359; c) K. Gorska, K.-T. Huang, O. Chaloin, N. Winssinger, *Angew. Chem.* **2009**, *121*, 7831–7836; *Angew. Chem. Int. Ed.* **2009**, *48*, 7695–7700; d) C. Scheibe, A. Bujotzek, J. Denedde, M. Weber, O. Seitz, *Chem. Sci.* **2011**, DOI: 10.1039/C0SC00565G.
- [5] a) K. I. Sprinz, D. M. Tagore, A. D. Hamilton, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3908–3911; b) A. R. Buskirk, D. R. Liu, *Chem. Biol.* **2005**, *12*, 151–161; c) S. J. Wrenn, R. M. Weisinger, D. R. Halpin, P. B. Harbury, *J. Am. Chem. Soc.* **2007**, *129*, 13137–13143; d) R. E. Kleiner, C. E. Dumelin, G. C. Tiu, K. Sakurai, D. R. Liu, *J. Am. Chem. Soc.* **2010**, *132*, 11779–11791; e) S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, *Nat. Biotechnol.* **2004**, *22*, 568–574.
- [6] B. A. R. Williams, C. W. Diehnelt, P. Belcher, M. Greving, N. W. Woodbury, S. A. Johnston, J. C. Chaput, *J. Am. Chem. Soc.* **2009**, *131*, 17233–17241.
- [7] a) J. M. Bradshaw, *Cell. Signalling* **2010**, *22*, 1175–1184; b) S. R. Hubbard, J. H. Till, *Annu. Rev. Biochem.* **2000**, *69*, 373–398; c) M. Turner, E. Schweighoffer, F. Colucci, J. P. Di Santo, V. L. Tybulewicz, *Immunol. Today* **2000**, *21*, 148–154.
- [8] a) T. Pawson, *Curr. Opin. Struct. Biol.* **1992**, *2*, 432–437; b) Z. Songyang, S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, F. King, T. Roberts, S. Ratnoffsky, R. J. Lechleider, B. G. Neel, R. B. Birge, J. E. Fajardo, M. M. Chou, H. Hanafusa, B. Schaffhausen, L. C. Cantley, *Cell* **1993**, *72*, 767–778; c) B. A. Liu, K. Jablonowski, M. Raina, M. Arcé, T. Pawson, P. D. Nash, *Mol. Cell* **2006**, *22*, 851–868.
- [9] a) T. Chen, B. Repetto, R. Chizzonite, C. Pullar, C. Burghardt, E. Dharm, Z. Zhao, R. Carroll, P. Nunes, M. Basu, W. Danho, M. Visnick, J. Kochan, D. Waugh, A. M. Gilfillan, *J. Biol. Chem.* **1996**, *271*, 25308–25315; b) E. A. Ottinger, M. C. Botfield, S. E. Shoelson, *J. Biol. Chem.* **1998**, *273*, 729–735; c) R. A. Grucza, J. M. Bradshaw, V. Mitaxov, G. Waksman, *Biochemistry* **2000**, *39*, 10072–10081.
- [10] a) F. J. Dekker, N. J. de Mol, J. van Ameijde, M. J. E. Fischer, R. Ruijtenbeek, F. A. M. Redegeld, R. M. J. Liskamp, *ChemBioChem* **2002**, *3*, 238–242; b) J. Kuil, H. M. Branderhorst, R. J. Pieters, N. J. de Mol, R. M. J. Liskamp, *Org. Biomol. Chem.* **2009**, *7*, 4088–4094; c) F. J. Dekker, N. J. de Mol, M. J. E. Fischer, R. M. J. Liskamp, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1241–1244; d) N. J. de Mol, M. I. Catalina, F. J. Dekker, M. J. E. Fischer, A. J. R. Heck, R. M. J. Liskamp, *ChemBioChem* **2005**, *6*, 2261–2270; e) J. Kuil, L. T. M. van Wandelen, N. J. de Mol, R. M. J. Liskamp, *J. Pept. Sci.* **2009**, *15*, 685–691.
- [11] K. Fütterer, J. Wong, R. A. Grucza, A. C. Chan, G. Waksman, *J. Mol. Biol.* **1998**, *281*, 523–537.
- [12] a) R. A. Grucza, K. Fütterer, A. C. Chan, G. Waksman, *Biochemistry* **1999**, *38*, 5024–5033; b) S. Kumaran, R. A. Grucza, G. Waksman, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14828–14833; c) M. I. Catalina, M. J. E. Fischer, F. J. Dekker, R. M. J. Liskamp, A. J. R. Heck, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1039–1051.
- [13] a) F. Diezmann, H. Eberhard, O. Seitz, *Biopolymers* **2010**, *94*, 397–404; b) S. Takeda, S. Tsukiji, T. Nagamune, *Tetrahedron Lett.* **2005**, *46*, 2235–2238.
- [14] H. Eberhard, O. Seitz, *Org. Biomol. Chem.* **2008**, *6*, 1349–1355.
- [15] a) B. A. Lynch, K. A. Loicacono, C. L. Tiong, S. E. Adams, I. A. MacNeil, *Anal. Biochem.* **1997**, *247*, 77–82; b) L. Röglin, M. R. Ahmadian, O. Seitz, *Angew. Chem.* **2007**, *119*, 2759–2763; *Angew. Chem. Int. Ed.* **2007**, *46*, 2704–2707; c) L. Röglin, F. Altenbrunn, O. Seitz, *ChemBioChem* **2009**, *10*, 758–765.

- [16] a) G. Bonnet, O. Krichevsky, A. Libchaber, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8602–8606; b) N. L. Goddard, G. Bonnet, O. Krichevsky, A. Libchaber, *Phys. Rev. Lett.* **2000**, *85*, 2400–2403; c) X. Wang, W. M. Nau, *J. Am. Chem. Soc.* **2004**, *126*, 808–813.
- [17] Previous studies suggest that the nick site hardly affects the overall B-type structure; rather, the nick can serve as a hinge that may facilitate kinking; E. Protozanova, P. Yakovchuk, M. D. Frank-Kamenetskii, *J. Mol. Biol.* **2004**, *342*, 775–785.
- [18] a) R. M. Clegg, *Methods Enzymol.* **1992**, *211*, 353–388; b) R. M. Clegg, A. I. H. Murchie, A. Zechel, D. M. J. Lilley, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2994–2998; c) C. R. Sabanayagam, J. S. Eid, A. Meller, *J. Chem. Phys.* **2005**, *122*, 061103; d) M. E. Sanborn, B. K. Connolly, K. Gurunathan, M. Levitus, *J. Phys. Chem. B* **2007**, *111*, 11064–11074; e) A. Iqbal, S. Arslan, B. Okumu, T. J. Wilson, G. Giraud, D. G. Norman, T. Ha, D. M. J. Lilley, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11176–11181.
-