# Single-Molecule Experiments to Elucidate the Minimal Requirement for DNA Recognition by Transcription Factor Epitopes

*Katrin Wollschläger, Katharina Gaus, André Körnig, Rainer Eckel, Sven-David Wilking, Matthew McIntosh, Zsuzsanna Majer, Anke Becker, Robert Ros, Dario Anselmetti, and Norbert Sewald**

*Dedicated to the memory of Prof. Peter Welzel*

*Interactions between proteins and DNA are essential for the regulation of cellular processes in all living organisms. In this context, it is of special interest to investigate the sequence-specific molecular recognition between transcription factors and their cognate DNA sequences. As a model system, peptide and protein epitopes of the DNA-binding domain (DBD) of the transcription factor PhoB from Escherichia coli are analyzed with respect to DNA binding at the single-molecule level. Peptides representing the amphiphilic recognition helix of the PhoB DBD (amino acids 190–209) are chemically synthesized and C-terminally modified with a linker for atomic force microscopy–dynamic force spectroscopy experiments (AFM–DFS). For comparison, the entire PhoB DBD is overexpressed in E. coli and purified using an intein-mediated protein purification method. To facilitate immobilization for AFM–DFS experiments, an additional cysteine residue is ligated to the protein. Quantitative AFM–DFS analysis proves the specificity of the interaction and yields force-related properties and kinetic data, such as thermal dissociation rate constants. An alanine scan for strategic residues in both peptide and protein sequences is performed to reveal the contributions of single amino acid residues to the molecular-recognition process. Additionally, DNA binding is substantiated by electrophoretic mobility-shift experiments. Structural differences of the peptides, proteins, and DNA upon complex formation are analyzed by circular dichroism spectroscopy. This combination of techniques eventually provides a concise picture of the contribution of epitopes or single amino acids in PhoB to DNA binding.*

**Keywords:**
- AFM–DFS
- circular dichroism
- DNA recognition
- peptides
- single-molecule studies

[*] K. Wollschläger, K. Gaus, Dr. S.-D. Wilking, Prof. N. Sewald
Organic and Bioorganic Chemistry
Department of Chemistry
Bielefeld University
Universitätsstr. 25, 33615
Bielefeld (Germany)
E-mail: norbert.sewald@uni-bielefeld.de
A. Körnig, Dr. R. Eckel, Prof. D. Anselmetti
Experimental Biophysics and Applied Nanoscience
Department of Physics
Bielefeld University 33615
Bielefeld (Germany)

DOI: 10.1002/smll.200800945

Dr. M. McIntosh, Prof. A. Becker
Institute of Biology III, University of Freiburg
79104 Freiburg (Germany)

Prof. R. Ros
Department of Physics, Arizona State University
Tempe, AZ 85287-1504 (USA)

Prof. Z. Majer
Institute of Chemistry, Eötvös Loránd University
1518 Budapest 112 (Hungary)

NANO MICRO small

## 1. Introduction

Single-molecule experiments have experienced increased attention during the last decade, not only from physicists but also from chemists and biologists active in the life sciences. While biomolecular interactions in biology, biochemistry, and chemistry are usually examined with ensemble methods on the macroscopic scale, the modern biophysical methods of atomic force microscopy (AFM) and optical tweezers techniques render the interaction forces between single molecules directly observable.[1] They can be used to describe the interaction between different molecules at the single-molecule level with high sensitivity. Direct force measurements of DNA–protein complexes have been carried out at the single-molecule level with AFM for transcription factors[2,3] and with optical tweezers for investigation of restriction endonuclease binding to DNA.[4]

Dynamic force spectroscopy (DFS)[5] even allows the determination of kinetic off-rate constants by recording and suitable analysis of dissociation forces for different loading rates. The loading rate is derived from the retract velocity of the AFM cantilever and the molecular elasticity. In such experiments, not only biomolecular thermal off-rate constants $k_{off}$ and complex lifetimes $\tau$ can be determined and correlated with macroscopic values, but also inner barriers of the energy landscape may be detected. Consequently, single-molecule experiments with AFM–DFS and optical tweezers represent novel and exciting techniques for the quantitative investigation of biomolecular interactions.[6]

Molecular recognition is a basic principle used by nature to regulate interactions inside a cell. In the first place, protein expression needs to be controlled and adapted to specific biological conditions for effective operation of organisms without wasting resources. Protein biosynthesis is tightly regulated at several key steps from DNA to processed proteins. The first step of protein expression, the transcription of DNA into messenger RNA (mRNA), is controlled by transcription factors, which activate or inhibit binding of RNA polymerase to defined DNA promoter sequences. Therefore, synthetic transcription factors offer a fascinating possibility for modification of the cellular metabolism in synthetic biology. One example in the literature is a synthetic DNA-binding miniprotein mimicking bZIPprotein (GCN4).[7] A detailed understanding of the mechanisms of complex formation of transcription factors and DNA requires information on the function and binding contribution of single amino acids, ideally at the single-molecule level. Furthermore, it is of special interest to reduce the complexity of DNA-binding proteins towards peptide analogues without sacrificing binding specificity. In the class of zinc finger proteins, this is accomplished by attaching a minor groove tripyrrole binder to 31 amino acids representing the minimal DNA-binding domain (DBD) of the GAGA factor of *Drosophila melanogaster*.[8] In our studies, the transcription factor PhoB from *Escherichia coli*[9–11] was used as a model protein and the binding properties of its entire DBD were compared with those of synthetic 20-mer peptides.[12]

PhoB is a transcription activator that controls the expression of genes involved in phosphate metabolism. It is composed of two domains, an N-terminal regulatory phosphorylation domain PhoB(1–127) and a C-terminal DBD PhoB(128–229). The structure of the regulatory domain has been determined by X-ray and NMR analysis.[13] Structural information of the DBD was derived from X-ray crystal and NMR solution structures of the DBD both free and in complex with DNA.[9,14,15] All four datasets correspond to comparable tertiary structures. Furthermore, unbound DBD and DBD in complex with DNA share a high degree of structural similarity.[9,15] However, the structure of the entire protein has not yet been determined.

PhoB is part of a two-component signal transduction system. The transmembrane protein PhoR activates PhoB by phosphorylation of Asp53 in the regulatory domain,[13] and the concomitant structural change enables PhoB to bind to its cognate Pho boxes in the promoter region of the phosphate regulon *pho*. This causes an alteration of DNA structure in the complex that enables interaction of the RNA polymerase with PhoB and subsequently with DNA, thus activating the transcription.[10] The *E. coli* genome harbors nine Pho boxes containing the consensus sequence TGTCA.[10,11]

Experiments with a PhoB/PhoR-deficient strain show that the PhoB DBD without regulatory domain is also able to bind to DNA and to activate transcription. This strain was transformed with two vector constructs encoding complete PhoB and the DBD of PhoB. As phosphorylation of PhoB is not possible due to the lack of PhoR, the full, non-phosphorylated PhoB is not able to activate transcription. In contrast, the strain expressing the DBD without regulatory domain binds to DNA and transcriptional activity could be observed.[16]

Figure 1A and B show the three-dimensional structures of the DNA–protein complex (1GXP.pdb). The amino acid

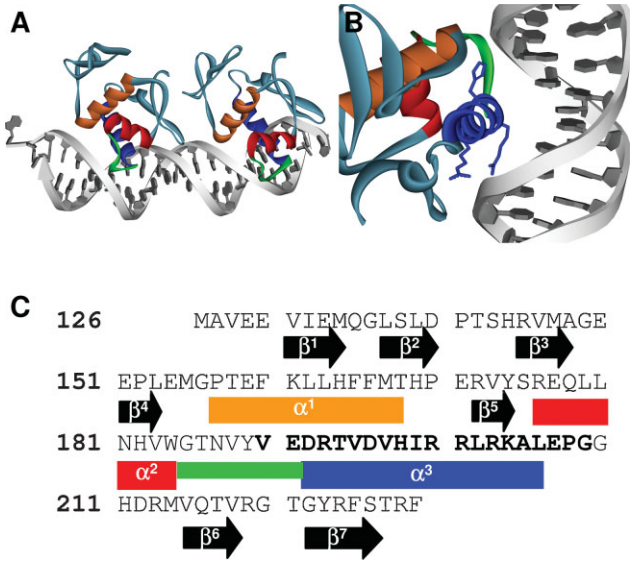



**Figure 1.** A) Crystal structure of two PhoB DBDs bound to an 18-bp double-stranded Pho box (gray); [9] helix $\alpha^1$: orange, helix $\alpha^2$: red, transactivation loop: green, helix $\alpha^3$: blue. B) Detailed view of the DNA binding helix $\alpha^3$ (blue) and the amino acids R193, H198, and R203. C) Amino acid sequence of the PhoB DBD with highlighted structural elements (arrows: $\beta$ sheets, boxes: $\alpha$ helices).

sequence of the DBD and the structural motifs are shown in Figure 1C. Like for other members of the family of winged helix-turn-helix proteins, DNA binding occurs primarily by a DNA-recognition helix (helix $\alpha^3$, residues 192–206, displayed in blue), connected to helix $\alpha^2$ (red, amino acids 176–184) by a loop (green) which replaces the turn motif normally found in the family of helix-turn-helix proteins. This loop, called the transactivation loop, is presumed to interact with the $\sigma^{70}$ subunit of the RNA polymerase leading to activation of transcription.[9,10] During complex formation, helix $\alpha^3$ specifically binds to the major groove of DNA by salt bridges, hydrogen bonds, and van der Waals contacts. Unspecific binding involves basic amino acids, such as Arg193, His198, Arg200, and Arg203, which interact by salt bridges with the negatively charged DNA backbone. Helix $\alpha^2$ is important for stabilization of the DNA binding by helix $\alpha^3$.

However, X-ray analysis gives only a static representation of the molecular system under investigation. NMR solution techniques are able to provide dynamic information on the whole ensemble of molecules and, hence, give only a macroscopic view. Therefore, the PhoB–DNA interaction was investigated at the single-molecule level by AFM–DFS Both the PhoB DBD and the recognition helix $\alpha^3$ together with a series of mutants (Ala-scan) have been examined. The proteins were obtained by intein-mediated protein purification, while synthesis of the C-terminally modified peptides succeeded by microwave-assisted solid-phase peptide synthesis. Structural analysis of the proteins, peptides, and DNA–protein complexes was carried out by circular dichroism (CD) studies. The DNA binding of the protein was additionally analyzed on the ensemble level by using the electrophoretic mobility shift assay (EMSA).

## 2. Results and Discussion

The molecular recognition between the DBD of the transcription factor PhoB and its target DNA in the corresponding *pho* regulon was analyzed at the single-molecule level. In this context, the binding properties of synthetic 20-peptides based on the primary structure of the DNA helix $\alpha^3$, PhoB (190–209), were studied to identify the contribution of single peptide modules to DNA. In addition, an alanine scan was performed, in which amino acids supposed to be important in DNA binding were replaced by alanine residues. Figure 1B shows the three amino acids R193, H198, and R203 that interact with the DNA backbone. The corresponding peptide PhoB(190–209) (**1**) and its mutants R193A (**2**), H198A (**3**), and R203A (**4**) were synthesized to investigate the contribution of each individual amino acid side chain to this DNA–peptide interaction. Likewise, the full protein, that is, the DBD of PhoB (126–229) (**5**), and the corresponding mutated analogues were examined.

### 2.1. Peptide Synthesis

The 20-peptides with a C-terminal linker suitable for attachment to a solid surface in the AFM experiments were obtained by solid-phase peptide synthesis with 2-chlorotrityl resin and 9-fluorenylmethoxycarbonyl (Fmoc) strategy. The C-terminal linker 3,6-dioxa-1,8-diaminooctane was manually coupled to the resin followed by the first amino acid glycine. Afterwards the resin was transferred to a microwave synthesizer. Subsequent to Fmoc deprotection, which was performed with 20% piperidine, coupling was performed by using 5 equivalents of 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) and 10 equivalents of *N*,*N*'-diisopropylethylamine (DIPEA). Final cleavage was achieved with 2.5% $H_2O$, 2.5% triisopropylsilane (TIS), and 95% trifluoroacetic acid (TFA). The peptide was purified by reverse-phase (RP) HPLC with TFA/water/acetonitrile elution gradients. In comparison to previously published results,[17] peptide synthesis was improved significantly by using microwave-accelerated solid-phase synthesis, especially with respect to the coupling efficiency. The purified yield amounted to 15–35% compared to 1–5% in conventional synthesis.

### 2.2. Protein Purification and Ligation

In order to obtain the DBD PhoB(126–229), intein-mediated protein purification was applied, which allows protein purification with release of a tag-free protein. Furthermore, it offers the possibility to synthesize a protein with a C-terminal thioester functionality that can be used in native chemical ligation reactions.[18,19]

The PhoB DBD was overexpressed in *E. coli* as a fusion protein with a chitin binding domain (CBD) and an intein sequence.[18,19] The efficiency of the single-step purification is presented in Figure 2. Lane 2 shows the soluble proteins after lysis and centrifugation. The band at 27 kD corresponds to the fusion protein. Due to premature self-splicing in vivo, some cleaved protein was observed in the lysate, which resulted in two bands for the released PhoB-protein (12 kD) and the CBD plus intein (25 kD). Upon loading the lysate on a chitin column (lane 3), only the proteins containing a CBD bind to the column. Unspecifically bound proteins were removed by washing (lane 4). Cleavage was induced by the addition of sodium 2-mercaptoethanesulfonate (MESNA). After overnight incubation, the protein was eluted (lanes 5–8). In the case of PhoB, this simple one-step purification method resulted in

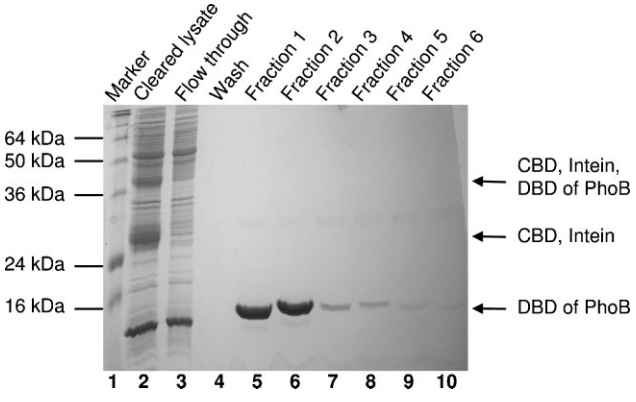



**Figure 2.** Intein-mediated protein purification of PhoB (Coomassie-stained sodium dodecyl sulfate gel, 18% acrylamide).

small NANO MICRO

**Table 1.** Proteins, peptides, and DNA fragments.

| | Construct | Calculated m/z | Measured m/z |
|---|---|---|---|
| | Peptides | | |
| **1** | Ac-PhoB(190–209)-linker | 2529.92 | 2530.27 |
| **2** | Ac-PhoB(190–209)R193A-linker | 2444.85 | 2445.44 |
| **3** | Ac-PhoB(190–209)H198A-linker | 2463.90 | 2464.57 |
| **4** | Ac-PhoB(190–209)R203A-linker | 2444.85 | 2445.38 |
| | Proteins | | |
| | PhoB(126–229) | 12155 | |
| | PhoB(127–229) | 12024 | 12024 |
| | PhoB(127–229)-MESNA | 12147 | 12149 |
| **5** | PhoB(127–229)-Cys | 12126 | 12126 |
| | PhoB(127–229)R193A | 11939 | 11940 |
| | PhoB(127–229)R193A-MESNA | 12062 | 12062 |
| **6** | PhoB(127–229)R193A-Cys | 12042 | 12043 |
| | PhoB(127–229)H198A | 11958 | 11959 |
| **7** | PhoB(127–229)H198A-Cys | 12061 | 12062 |
| | PhoB(127–229)R203A | 11939 | 11940 |
| **8** | PhoB(127–229)R203A-Cys | 12042 | 12043 |

| | DNA fragments | Number of bp | Modification |
|---|---|---|---|
| **9** | *pstS* Pho box DNA | 18 bp | 5′-Cy3-modified |
| **10** | random DNA | 18 bp | |
| **11** | *pstS* Pho box DNA | 360 bp | 5′-Thiol modified |
| **12** | KF-Ge DNA | 177 bp | 5′-Thiol modified |
| **13** | EBNA DNA | 356 bp | 5′-Thiol modified |

Linker: 3,6-dioxa-1,8-diaminooctane; MESNA: sodium 2–mercaptoethanesulfonate.

>95% pure protein with a yield of 10 mg protein per liter Luria–Bertani (LB) medium.

Intein-mediated protein splicing combined with native chemical ligation is a powerful tool to purify recombinant proteins and to ligate them to chemically synthesized peptides.[19] Here, this method was used to combine the purification process with the addition of only a single cysteine residue at the C-terminus of the PhoB proteins, which permits specific immobilization on gold surfaces. Ligation of longer peptides was also successfully performed. Ligation reactions were monitored by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The data prove the complete ligation efficiency of the protein with cysteine. Since the N-terminal methionine of the recombinant protein is followed by a small amino acid residue (Ala, see Figure 1C), the former is completely removed by the *E. coli* methionine aminopeptidase.[20,21] Therefore, only the protein PhoB(127–229) (**5**) instead of PhoB(126–229) was obtained (see Table 1). In addition to the wild-type protein (**5**), three protein point mutants (**6**–**8**) were accessible analogously to the previously analyzed peptides.[12]

## 2.3. DNA

For interaction analysis different DNA molecules were used, because EMSA and CD experiments require short DNA fragments, whereas longer DNA fragments are necessary for AFM experiments to provide enough flexibility. In nature, several Pho boxes with different sequences are found. The consensus sequence[10] is 5′-CTGTCATAWAWCTGTCACAWW-3′ (with W representing A or T) and almost all Pho boxes contain at least one TGTCA motif. The AT-rich region ATAWAW flanking the recognition motifs was found to be necessary for binding of the C-terminal $\beta$ hairpin of PhoB.[9] In EMSA and CD experiments, 18-base-pair (bp) double-stranded (ds) DNA (**9**; 5′-CTGTCATAAAACTGTCAT-3′) with two PhoB binding sites TGTCA and an 18-bp random DNA (**10**) were used. For AFM experiments, DNA with the sequence of the *pstS* Pho box was employed.[10] The 360-bp Pho box DNA (**11**) contains four binding sites in contrast to other Pho boxes, which contain two binding sites. Three of them are TGTCA motifs and one is TTACA. Two different DNA fragments were used as negative control DNA, the 177-bp KF-Ge DNA[22] (**12**) without TGTCA sequences and the 356-bp Epstein–Barr virus nuclear antigen (EBNA) DNA (**13**) containing three TGTCA motifs without the flanking regions of the Pho box DNA.

## 2.4. EMSA

EMSAs were performed to prove DNA binding in molecular ensembles. Experiments with the wild-type DBD of PhoB (**5**) and Cy3-labeled 18-bp-long Pho box DNA (**9**) clearly show DNA binding. As expected, a DNA band shift was observed upon addition of the wild-type protein (**5**; Figure 3, lanes 1 and 2). In contrast, no band shift was detected in the case of protein mutants R203A (**8**; Figure 3, lane 3), R193A (**6**), and H198A (**7**) and all peptides (data not shown) at any concentration. DNA band shifts depend on the ratio between DNA and protein (Figure 3, lanes 5–8). No band shift was observed at a ratio below 1:250 (lane 6). Also, a high excess of protein (1:5000, not shown) did not lead to a complete shift of the DNA towards higher masses. This experiment proves DNA binding of the wild-type protein PhoB(127–229) (**5**) and indicates that the binding affinities of the protein mutants R193A (**6**), H198A (**7**), and R203A (**8**) are significantly reduced or even abolished.

## 2.5. AFM–DFS

With AFM–DFS it is possible to detect and quantify dissociation forces between single binding partners in the pico-Newton range.[3,23–26] The binding partners are covalently

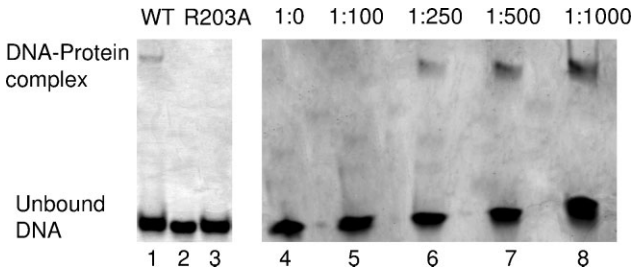



**Figure 3.** EMSA experiment of wild-type PhoB protein (**5**) and 18-bp Pho box DNA (**9**) in different ratios. Lane 1 represents a 1:250 complex of wild-type protein–DNA (**5** · **9**), lane 2 shows free DNA (**9**), and lane 3 shows a 1:250 complex between mutant protein R203A and DNA (**8** · **9**). Lane 4 shows free DNA (**9**) and lanes 5–9 represent wild-type protein–DNA (**5** · **9**) complexes at different DNA/protein ratios (10% acrylamide gel, TBE buffer).

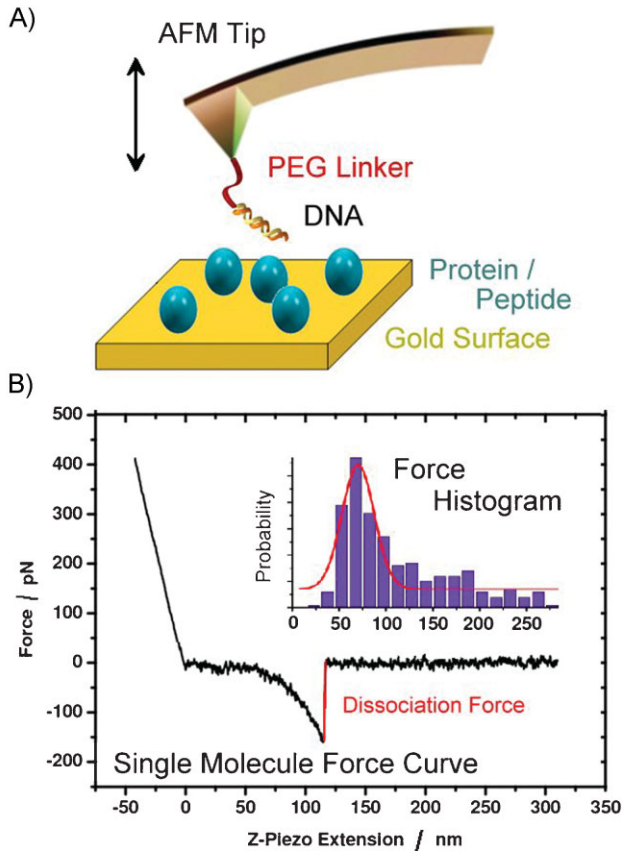



**Figure 4.** A) Schematic of the AFM–DFS setup. B) Typical AFM single-molecule force curve diagram (only retractive part shown) with force histogram (insert).

bound to either the AFM tip or the surface to investigate specific interactions, for example, between protein and DNA (Figure 4A). The dissociation of a bound DNA–protein complex can be monitored during a cycle when the tip is approached to and retracted from the surface. In AFM force curve diagrams the force acting on the AFM tip is plotted against the vertical position (piezo extension; Figure 4B), where molecular dissociation events can be identified by a sudden jump in the ''attractive'' force regime back to the curve at zero force ($F=0$).

A polymer linker, which is used to bind the PhoB-DNA to the cantilever, ensures that the binding/dissociation event occurs far enough from the surface and that the molecular binding partners have enough steric flexibility to arrange themselves in an appropriate spatial manner. Furthermore, the typical nonlinear stretching of the polymer chain, which precedes the molecular dissociation event and reflects the elasticity of the polymer chain, facilitates the detection and discrimination of a specific dissociation event from unspecific binding.[12,26,27] Unspecific tip–surface interactions occur closer to the surface and can therefore be discriminated. Since the process of molecular dissociation is of stochastic nature, many (typically hundreds) individual force curves have to be measured and combined in a force histogram (Figure 4B, insert). The most probable dissociation force $F_{max}$ (maximum) is commonly referred to as the dissociation force and can be determined by fitting an appropriate distribution to the measured force histogram.

The specificity of the binding between wild-type PhoB DBD (**5**) and the Pho box DNA (**11**) was investigated in AFM competition experiments (control experiments) with soluble Pho box DNA. Figure 5 represents a typical series of force spectroscopy experiments (force histograms). After measurement in standard reaction buffer (Figure 5A) an excess of free DNA was added, which resulted in a nearly complete loss of binding (Figure 5B) visible as the relative deterioration of the force histogram (loss of activity >85%). After washing with buffer, binding could be restored (Figure 5C). In addition, further control experiments using two different control DNA sequences (Figure 5D and E) were performed. Since KF-Ge DNA (**12**) from *Sinorhizobium meliloti* does not contain binding sites for *E. coli* PhoB, and EBNA DNA (**13**) contains TGTCA motifs without flanking regions (see above), no relevant binding to wild-type PhoB protein could be detected. From these two experiments it can be concluded that binding of the PhoB DBD occurs specifically to Pho box DNA as no

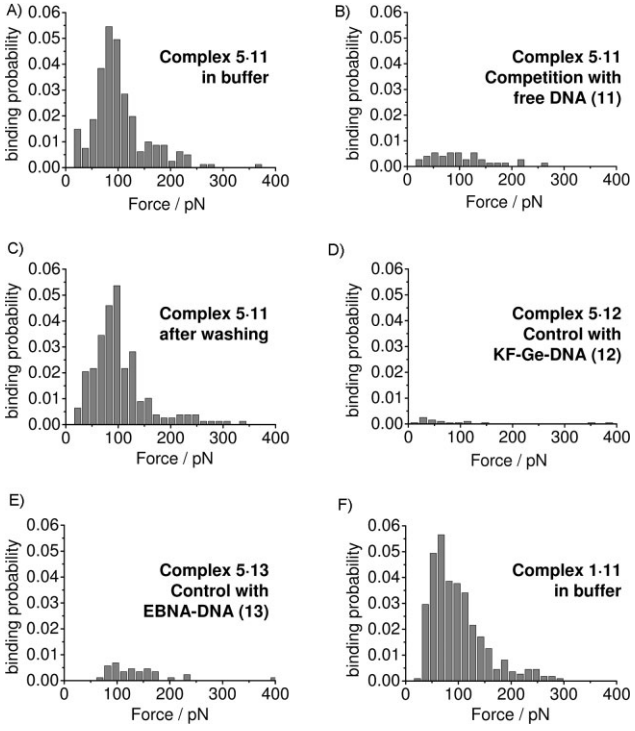



**Figure 5.** Force histograms of the molecular recognition as detected in AFM single-molecule force spectroscopy. A) Wild-type PhoB protein (**5**) and Pho box DNA (**11**) without competitor. B) Experiment (A) with an excess of free Pho box DNA (**11**) as competitor. C) Reconstitution of binding after the competition experiment by washing with buffer solution. D) Negative control experiment with KF-Ge DNA (**12**), which does not contain PhoB binding sites, and wild-type PhoB protein (**5**). E) Control experiment with EBNA-DNA (**13**) containing TGTCA motifs without flanking regions important for DNA binding of PhoB and the PhoB wild-type protein (**5**). F) Force spectroscopy experiment of wild-type peptide (**1**) and Pho box DNA (**11**) [12]. All experiments were measured in standard reaction buffer (100 mM $NaH_2PO_4$, 50 mM NaCl, pH 7.4) at 1000 nm $s^{-1}$ retract velocity.

NANO MICRO small

significant binding occurs with the DNA fragments **12** and **13**. Moreover, the results with EBNA DNA show that the TGTCA motifs alone are not sufficient for PhoB binding. The AT-rich region between the two TGTCA motifs is necessary for DNA binding. Likewise, DNA–protein binding could be observed for protein R193A (**6**) and protein H198A (**7**), but not for mutant R203A (**8**).

As previously communicated,[12] the peptides derived from PhoB helix $\alpha^3$ (PhoB(190–209); **1**) and its point mutants (**2**–**4**) were investigated by using AFM–DFS measurements. Binding events were detected for the wild-type peptide (**1**), and the peptide mutants R193A (**2**) and H198A (**3**). Peptide mutant R203A (**4**) did not bind to DNA. A typical force histogram for the wild-type peptide (**1**) is shown in Figure 5F. The specificity of the binding for all three peptides was also proven through competition experiments with an excess of DNA and peptide. Similar to the method described above, AFM competition experiments with free competitor in reaction buffer yielded a reduced binding affinity (loss of activity >90%) that could be completely restored after washing of the surface with buffer[12] (data not shown).

Notably, for both point mutants, the peptide PhoB(190–209) R203A (**4**) and protein PhoB(127–209) R203A (**8**), no specific binding events could be observed. The arginine at position 203 is conserved within the PhoB family, clearly because of its relevance for DNA binding. Recently published NMR structures in an explicit water environment[14] support our experimental result and reveal that the binding of this residue to DNA is mediated by a water molecule. Upon exchange of the arginine for alanine, this interaction is obviously deteriorated, which explains the suppression of DNA binding.

In DFS single-molecule experiments, the dissociation kinetics and lifetimes of the DNA–protein complexes can be quantitatively investigated as previously described.[12] Here, we applied this methodology to investigate the binding of different PhoB-derived mutant peptides and proteins to the Pho box DNA and compared the results with those measured for wild-type PhoB and the peptide corresponding to the binding helix $\alpha^3$.

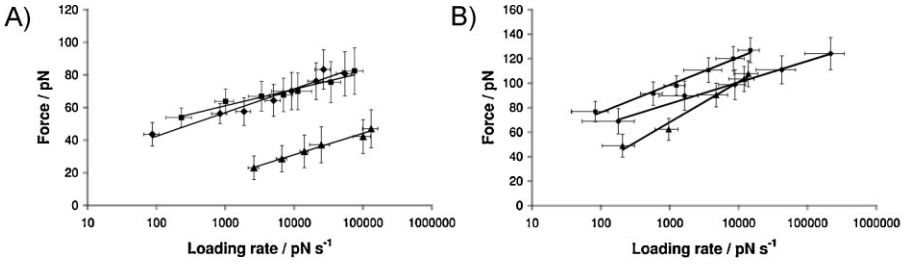
Statistical analysis of hundreds of single-molecule DNA–protein dissociation events was performed according to the standard model of force spectroscopy theory, where the most probable dissociation forces $F_{max}$ are plotted logarithmically against the loading rate $r$.[5,28] According to the standard model the measured dissociation forces depend on the loading rate, which describes how fast the force on the complex increases. The loading rate can be calculated by multiplication of the experimental retract velocity and the molecular elasticity. The latter can be deduced from fitting an elasticity model (e.g., linear) to the force curve just before dissociation for each measured dissociation event individually. Therefore, the molecular complexes have to be investigated at different loading rates, which is experimentally realized by variation of the AFM-tip retract velocity.

According to the standard model of thermally driven dissociation under external force[23] the following relationship holds:

$$F_{max} = \frac{k_B T}{x_\beta} \ln \frac{x_\beta r}{k_B T\, k_{off}} \quad (1)$$

where $F_{max}$, $k_BT$, and $r$ denote most probable rupture force, thermal energy, and loading rate, respectively. The parameter $x_\beta$ is a length parameter along the reaction coordinate, which represents the distance between the minimum of the potential well and the maximum of the energy barrier separating the bound state from the free state and is commonly termed the reaction length. $k_{off}$ is the thermal off-rate constant under zero load and can be deduced by linearly extrapolating the experimental data in the DFS plot to zero external force $F = 0$ (see, e.g., Figure 6A and B). Since $k_{off}$ is inversely related to the average lifetime $\tau$ of the complex ($k_{off} = \tau^{-1}$), this allows a direct way of evaluating the stability of a complex.

The DFS results for the interaction between the wild-type peptide (**1**) and the Pho box DNA (**11**) are shown in Figure 6A, while the results for the wild-type PhoB protein (**5**) are displayed in Figure 6B. The DNA–protein complex dissociates with $k_{off} = (0.0025 \pm 0.0021)\ s^{-1}$, which corresponds to an average lifetime of $\tau = 400\,s$. In comparison, the results obtained with the wild-type peptide PhoB(190–209) (**1**), which corresponds to the PhoB helix $\alpha^3$ alone, yielded $k_{off} = (3.1 \pm 2.1)\ s^{-1}$ and $\tau = 0.32\,s$, thus indicating that the protein complex with the entire DBD dissociates about 1000 times more slowly, most likely because the full DBD incorporates additional amino acid residues that support complex stabilization. A less pronounced effect can be observed when comparing the lifetimes of the mutants R193A and H198A on the protein and peptide level (R203A exhibited no binding). In both mutants the full




| | | $k_{off}$ [s⁻¹] peptide (A) PhoB (190-209) | $\tau$ [s] | | $k_{off}$ [s⁻¹] protein (B) PhoB (127-229) | $\tau$ [s] |
|---|---|---|---|---|---|---|
| ◆ wildtype | (1) | 3.1 ± 2.1 | 0.32 | (5) | 0.0025 ± 0.0021 | 400 |
| ■ R193A | (2) | 0.071 ± 0.053 | 14 | (6) | 0.012 ± 0.008 | 83 |
| ▲ H198A | (3) | 50 ± 21 | 0.020 | (7) | 0.76 ± 0.25 | 1.3 |
| R203A | (4) | — | — | (8) | — | — |

**Figure 6.** AFM–DFS of PhoB peptides and proteins. A) PhoB peptides 190–209. B) PhoB DBD 127–229. $k_{off}$: dissociation rate constant for the peptide/protein DNA complexes; $\tau = 1/k_{off}$: complex lifetime. The kinetic data were measured in 100 mM $Na_2HPO_4$/50 mM NaCl (pH 7.4).

protein sequences yield a distinctly longer lifetime than the mutated PhoB(190–209) peptide sequences. For protein mutant R193A (**6**) the lifetime is 83 s, and the result for peptide mutant R193A (**2**) is 14 s. For H198A the lifetimes are 1.3 s for protein mutant H198A (**7**) and 0.02 s for peptide mutant H198A (**3**). Hence, for the mutant protein H198A (**7**), where a histidine residue is replaced by an alanine, the off-rate constant is increased 300-fold compared to that of the wild-type protein (**5**). In comparison only a 16-fold increased off-rate constant is observed upon replacement of histidine by alanine in the mutant peptide H198A (**3**). This result implies that the basic histidine side chain plays an important role in DNA binding.

In contrast, the peptide and protein results for mutant R193A are puzzling. The complex of the mutant peptide R193A (**2**) and Pho box DNA (**11**) exhibits a 44-fold longer lifetime than the complex of wild-type PhoB peptide (**1**·**11**). So far no satisfactory explanation for the observation that the exchange of a basic arginine side chain by an alanine side chain improves DNA binding has been found. CD spectra do not indicate structural differences, such as enhanced $\alpha$-helical conformation (discussed below). In contrast, the DNA-binding ability of the protein mutant R193A (**6**) was reduced compared to the wild-type protein ($k_{off} = (0.012 \pm 0.008)\ s^{-1}$, $\tau = 83$ s), and exhibits a lifetime that lies between those of the wild-type protein (**5**) and the mutant H198A (**7**).

AFM competition experiments with free peptide (**3**) indicate that the binding of PhoB DNA to the entire DBD (**5**) is stronger and more specific than to the peptide (Figure 7), in agreement with the previous findings. Similar to the competition experiments with DNA (see Figure 5), the soluble wild-type peptide PhoB(190–209) (**1**) was added in a large excess (25-fold) to the buffer solution, which reduced the binding probability between DNA (**11**) and protein (**5**) to 55%, but did not result in a complete loss of binding, as observed with free DNA as competitor. Notably, the absolute binding probability measured can vary between different series of AFM–DFS experiments, since it also depends on local receptor densities that can hardly be controlled and kept constant. Therefore, a loss of binding can only be judged appropriately by quantifying the relative decrease in dissociation force distributions in the force histograms.

For molecular ensembles at thermal equilibrium, the affinity of a ligand to its receptor, represented in the case of 1:1 kinetics by the dissociation equilibrium constant $K_D = k_{off}/k_{on}$, is assumed to be mainly affected by the dissociation rate constant $k_{off}$.[25] This hypothesis would allow one to compare the thermodynamic dissociation constants $K_D$ of similar proteins or peptides by comparison of their $k_{off}$ values.

The $K_D$ for the PhoB wild-type DBD and its target DNA has been determined by Ellison et al. by using fluorescence anisotropy measurements, and amounts to $63 \times 10^{-9}$ M.[16] The authors used *E. coli* PhoB(124–229) and a dsDNA hairpin with two PhoB binding sites. Hence, assuming similar experimental conditions, the $k_{on}$ value can be approximately given as $4 \times 10^4\ M^{-1}\ s^{-1}$ based on the AFM–DFS-derived $k_{off} = 2.5 \times 10^{-3}\ s^{-1}$ for the wild-type protein (**5**; see Figure 6). Considering this $k_{on}$ as an estimate for all protein mutants, equilibrium constants $K_D = 2.0 \times 10^{-7}$ M for protein mutant R193A (**6**) and $K_D = 1.9 \times 10^{-5}$ M for protein mutant H198A (**7**) can be estimated.

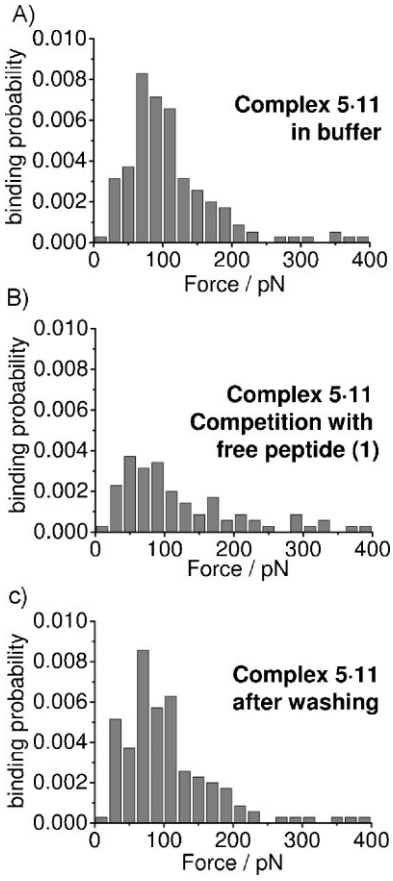



**Figure 7.** AFM competition experiments with peptide **1**. A) Force spectroscopy experiment of the wild-type protein (**5**) with Pho box DNA (**11**) in buffer without competitor. B) Competition experiment in buffer with a 25-fold excess of free wild-type peptide PhoB(190–209) (**1**) as competitor, yielding an overall decrease of binding to 55%. C) Experiment after washing and reactivation with buffer solution. All force histograms were measured at 1000 nm $s^{-1}$ retract velocity.

In summary, AFM single-molecule force spectroscopy of the interaction between DNA and protein or chemically synthesized peptides provides opportunities to quantitatively investigate and rank the affinity of protein epitopes required for DNA binding. The experiments obtained with PhoB wild-type DBD and mutants yield kinetic constants comparable to previous results with PhoB peptide epitopes and, hence, identified the amino acid residues required for DNA binding. The importance of residue Arg203 was shown by the R203A mutants of both the protein and the peptide.

## 2.6. CD Spectroscopy

CD spectra of the wild-type DBD of PhoB(127–229) (**5**) and the three PhoB(127–229) mutants (**6**–**8**) are presented in Figure 8A. The spectra show two minima at 222 and 209 nm and a maximum at about 195 nm. The percentage of $\alpha$ helices and $\beta$ sheets calculated according to Chen, Yang et al.[29,30] is in perfect agreement with the data from X-ray crystal and

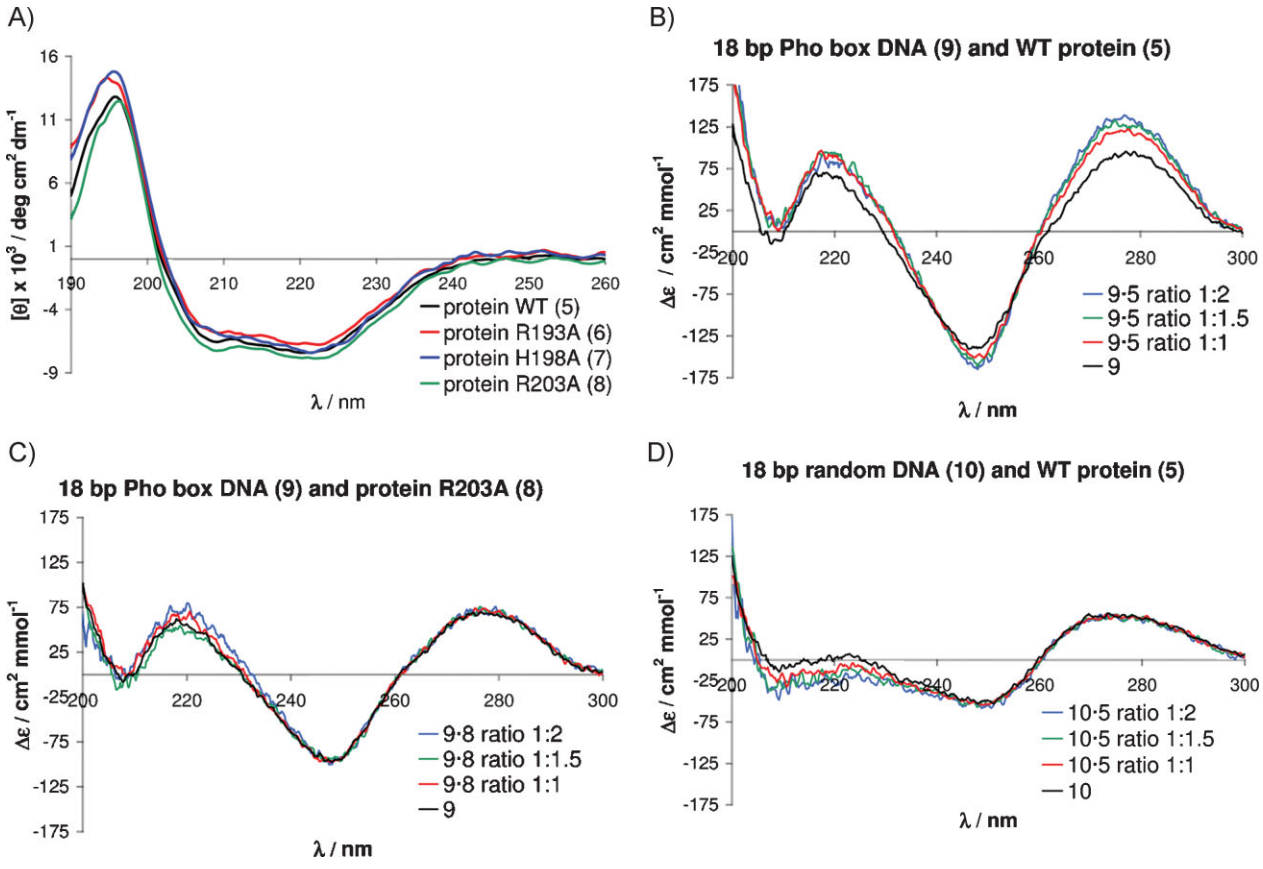



**Figure 8.** CD spectra in 10 mM $Na_2HPO_4$/5 mM NaCl (pH 7.4); all proteins are PhoB(127–229). A) CD spectra of wild-type protein and mutants. B–D) Difference spectra of DNA–protein complex minus protein at different ratios. B) 18-bp Pho box DNA (**9**) and wild-type protein (**5**). C) 18-bp Pho box DNA (**9**) and inactive mutant R203A (**8**). D) 18-bp random DNA (**10**) and wild-type protein (**5**).

NMR structure analysis. While the crystal structure of PhoB DBD contains 36% $\alpha$-helical conformation and 27% $\beta$ sheets, the values calculated from the CD spectra range from 33 to 39% for the $\alpha$ helix and 24 to 29% for the $\beta$ sheet. There are no major differences between the CD spectra of the wild-type protein and the three mutants, which substantiates the structural homology of the point mutants.

To reveal the structural changes upon binding of the proteins to DNA, CD spectra of the wild-type protein (**5**), the 18-bp DNA fragment (**9**), and the DNA–protein complex (**9**·**5**) were measured. The DNA fragment (**9**) contains two TGTCA motifs (see above). The CD spectrum displays the typical properties of B-type DNA, a positive band at 275 nm and a negative band at 248 nm with nearly equal magnitude.[31] The minimum is due to right-handed helicity and the maximum to base stacking.[32]

Differential CD spectra (expressed in molar CD $\Delta\varepsilon$), calculated by subtracting either the protein or DNA spectrum from the protein–DNA complex spectrum, reveal a change in the DNA structure (Figure 8B) but no change in protein structure upon complex formation (data not shown). The differential DNA spectra exhibit a significant increase of the CD ellipticity for both the maxima and minima. Additionally, the maximum is red-shifted from 275 to 277 nm. In line with NMR and crystal structure analysis,[9] this indicates that the DNA is being bent upon complex formation. According to previous reports,[9] the overall folding of the unbound DBD is maintained in the complex with only local differences being observed. This result is also supported by circular permutation analysis of the *pst* promoter. Probably, DNA bending facilitates the transcription activation of the RNA polymerase.[10]

Interestingly, no differences in the CD spectra of the DNA were observed when the inactive protein mutant R203A (**8**) was used (Figure 8C). This underlines the fact that the described effect is caused by the DNA binding of the wild-type protein. Furthermore, the sequence specificity of DNA binding was confirmed with an 18-bp random DNA (**10**) without the TGTCA binding motif (Figure 8D). This DNA also shows a B-type CD spectrum, but no changes in the CD spectrum can be detected upon addition of the PhoB wild-type protein (**5**).

The CD spectra of the wild-type peptide PhoB (190–209) (**1**, black) and the corresponding peptide mutants R193A (**2**, red), H198A (**3**, blue), and R203A (**4**, green) in buffer are compiled in Figure 9A. The decreased off-rate of the mutant peptide R193A (**2**) was previously hypothesized to originate from the higher helicity of the peptide.[12] However, this suggestion is not supported by CD spectroscopy. The CD spectra in phosphate buffer as solvent exhibit a predominantly random-coil structure with a minimum at 201 nm. Changing the solvent to 2,2,2-trifluoroethanol (TFE), which is known to induce an $\alpha$-helical conformation,[33] increases the $\alpha$-helical

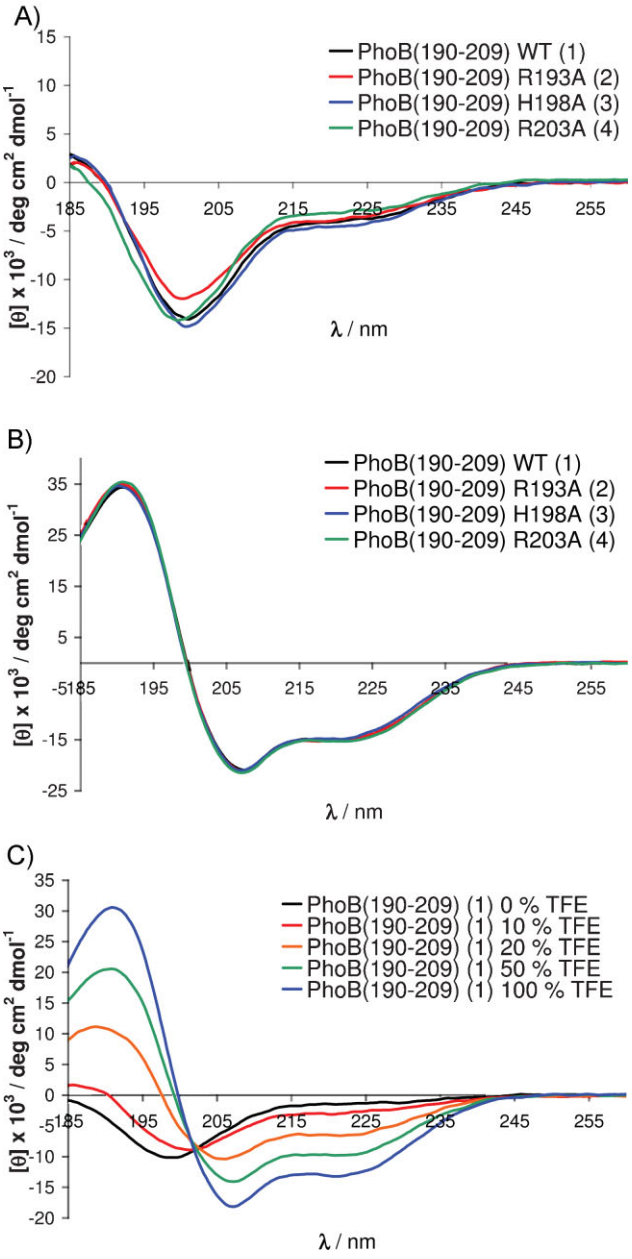



**Figure 9.** CD spectra of the peptides (0.2 mM) in A) 10 mM $Na_2HPO_4$/ 5 mM NaCl (pH 7.4), B) TFE, and C) different ratios of water and TFE.

content in the peptides (Figure 9B). A maximum at approximately 191 nm is observed and the minimum at 201 nm is shifted to 207 nm.[34] In addition, a minimum at about 220 nm starts to emerge. These spectra represent a mixture of the conformers, and the intensity of the CD bands indicates that approximately 50% of the amino acids in the sequence can adopt an $\alpha$-helical structure. The evaluation of the helical content of a peptide using CD spectra includes many inaccuracies, and therefore only allows a rough estimation of structural features. Application of several different methods[30,35,36] yielded values of 37 to 52% $\alpha$-helical residues in 100% TFE upon comparison of the molecular residual ellipticity at 222 nm.

PhoB WT peptide (190–209) (**1**) was also examined in solution at different ratios of water and TFE. The isodichroic point at 202 nm indicates that the structure of peptide can be described as a two-state equilibrium between a random coil and an $\alpha$ helix. On monitoring the different percentages of water and TFE (Figure 9C), the data show a high propensity of the peptide to adopt a helical structure. The CD shape in 50% TFE shows a high population of the helical conformation.

Differential CD spectra of the DNA–peptide complex, after subtraction of the DNA spectra, display no distinct changes in the peptide structures (data not shown). Furthermore, when subtracting the peptide spectra from the spectra of the DNA–peptide complex, the DNA spectra are not influenced by the presence of peptide (**1**), thus leading to the conclusion that the 20-mer peptides do not exert sufficient forces on the DNA to give DNA bending as observed in case of the entire PhoB DBD (see above).

## 3. Conclusions

We have demonstrated that off-rate constants can be efficiently determined by AFM–DFS for DNA–peptide and DNA–protein complexes. The single-molecule experiments also revealed subtle differences between single point mutants that could not be detected to this extent with other methods. DNA binding of the full DBD was also proven by EMSA experiments. AFM experiments proved the specific binding of the protein to its recognition sequence. Binding was not observed using control DNA without the consensus sequence. Comparison of the CD spectra of the wild-type protein or peptide with its mutants underlines the contribution of several specific amino acid residues to DNA binding. All protein mutants, and also the inactive mutant R203A, show the same secondary structure. In analogy to X-ray single-crystal and NMR structures, the proteins exhibit the expected $\alpha$-helical structure. The peptide epitopes adopt a random structure in aqueous solution, while the $\alpha$-helical content is increased in TFE. CD studies also prove the sequence-specific binding of the DBD of PhoB to Pho box DNA. A change in the structure of the DNA is observed upon binding of the protein in the case of Pho box DNA but not with random DNA.

The results presented here will be of interest for the future design of synthetic peptide and protein ligands as artificial DNA binders and transcription factors in synthetic biology.

## 4. Experimental Section

*Materials and instrumentation:* Enzymes used for DNA modification and cloning, the pTwin 2 vector, chitin beads, and *E. coli* strains were purchased from New England Biolabs (Frankfurt a.M., Germany). All primers and small DNA fragments were obtained from Operon (Köln, Germany). Chemicals and solvents were obtained from Sigma–Aldrich (Hamburg, Germany), Acros (Geel, Belgium), or Applichem (Darmstadt, Germany). All amino acids and the 2-chlorotrityl resin were purchased from Iris Biotech GmbH (Marktredwitz, Germany) or Orpegen (Heidelberg, Germany).

Analytical RP-HPLC was carried out on a Thermo Separation Products system consisting of a UV-6000 diode-array detector and a P-4000 pump, and equipped with a Phenomenex HPLC guard cartridge system (C12; 4 × 3.00 mm) and a Phenomenex Jupiter 4 μ Proteo 90A column (C12; 250 × 4.60 mm). Preparative RP-HPLC was carried out on a Thermo Separation Products system consisting of a UV-1000 detector and a P-4000 pump, and equipped with a Vydac high-performance guard column (C18) and a Phenomenex Jupiter 10 μ Proteo 90A column (C12; 250 × 21.20 mm), or on a Hitachi MERCK LaChrom system consisting of a UV/Vis L7420 detector and an L7150 pump, and equipped with a Vydac high-performance guard column (C18) and a Phenomenex Jupiter 10 μ C18 300A column (C18; 250 × 21.20 mm).

MALDI-TOF mass spectra were recorded on a Voyager DE instrument (PerSeptive Biosystems, Weiterstadt, Germany), with 2,5-dihydroxybenzoic acid (peptides) or sinapinic acid (proteins) as the matrix. Samples were desalted using ZipTips (Millipore, Schwalbach, Germany) before measurement. The mass axis was externally calibrated with calibration standard on the same target. For proteins, bovine insulin (5734.59 Da), *E. coli* thioredoxin (11674.48 Da), and horse apomyoglobin (16952.46 Da) were used; for peptides, Angiotensin 1 (1297.61 Da), ACTH (1–17) (2094.46 Da), ACTH (7–38) (3660.19 Da), and bovine insulin (5734.59 Da) were used.

*Protein expression and purification:* Genomic DNA from *E. coli* K12 was isolated using the DNeasy Tissue Kit from Qiagen (Hilden, Germany). The DBD DNA of *E. coli* PhoB (encoding for amino acids 126–229; **5**) was amplified by the polymerase chain reaction (PCR). The fragment was then inserted into a pTwin2 vector (NEB) using *Nde*I and *Sap*I restriction sites. This vector allows for tagless purification of proteins using intein-mediated protein purification by production of fusion proteins with CBD, intein, and the desired protein.[18] *E. coli* ER2566 (NEB) was used for the expression. All buffers for protein purification contained 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (20 mM), NaCl (500 mM), and ethylenediaminetetraacetic acid (EDTA; 1 mM). Additional reagents and pH values are given in brackets. Bacteria were grown at 37 °C in LB medium containing 100 μg $mL^{-1}$ ampicilin. At an OD600 = 0.6–0.7 protein production was induced by adding isopropyl-β-D-thiogalactopyranoside (0.5 mM; Applichem). After 5 h at 30 °C, the cells were harvested, resuspended in lysis buffer (pH 7.0, 20 μM phenylmethylsulfonyl fluoride, 0.1% Tween 20), and sonicated. Chitin beads (5 mL) were equilibrated with buffer (pH 7.0) and loaded with supernatant at 0.5 mL $min^{-1}$. After washing, splicing was induced with MESNA (50 mM; Sigma–Aldrich) at pH 8.5 to produce a thioester, necessary for native chemical ligation. Addition of cysteine (final concentration 1 mM) released the protein with one additional C-terminal cysteine residue. The ligation reaction was followed by MALDI-TOF MS. The ligation had to be performed immediately after protein purification, as the MESNA thioester was partially hydrolyzed after 24–48 h. For further experiments the buffer containing MESNA and cysteine was exchanged against phosphate buffer by dialysis. Protein concentrations were analyzed by using a Bradford assay (Pierce, Bonn, Germany).

Point mutations R193A (**6**), H198A (**7**), and R203A (**8**) were introduced in the PhoB DBD protein with the QuickChange site-directed mutagenesis kit (Stratagene, Amsterdam, Netherlands) and purified in analogy to the wild-type protein. The mutation has been proven by DNA sequencing and on the protein level by using MALDI-TOF MS (Table 1).

Vector construction and cloning (bold: restriction sites):

wild type: forward primer: GGTGGT**CATATG**GCGGTGGAAGAGGTGATTGAG
reverse primer: GGTGGTTG**CTCTTC**CGCAAAAGCGGGTTGAAAAACG

Site-directed mutagenesis (bold: mutated sequence):

R193A: forward primer: GTGTATGTGGAAGAC**GCC**ACGGTCGATGTCCAC
reverse primer: GTGGACATCGACCGT**GGC**GTCTTCCACATACAC

H198A: forward primer: CGCACGGTCGATGTC**GCC**ATTCGTCGCCTGCG
reverse primer: CGCAGGCGACGAAT**GGC**GACATCGACCGTGCG

R203A: forward primer: CACATTCGTCGCCTG**GCA**AAAGCACTGGAGCCC
reverse primer: GGGCTCCAGTGCTTT**TGC**CAGGCGACGAATGTG

PCR-amplification of DNA 11 and 12 for interaction studies; SH-labeled at the 5′ end for AFM studies:

M13uni: forward primer: CGCCAGGGTTTTCCCAGTCACGAC
reverse primer: AGCGGATAACAATTTCACACAGGA

*Peptide synthesis and characterization:* The C-terminally functionalized peptides were synthesized by solid-phase peptide synthesis in a Liberty microwave synthesizer (CEM, Kamp-Lintfort, Germany) using Fmoc strategy on the 0.1 mM scale. For a detailed list of synthesized peptides, see Table 1. Microwave-assisted solid-phase peptide synthesis provided yields of approximately 15% after purification.[12,17] The 2-chlorotrityl resin was preloaded with 3,6-dioxa-1,8-diaminooctane (Fluka; loading 0.5 mmol $g^{-1}$); linker (5 eq.) and DIPEA (10 eq.) were dissolved in dry dichloromethane (5 mL per gram resin). The resin was added and the mixture was stirred for 2 h under an argon atmosphere. Then Fmoc-glycine (1 eq.), TBTU (1 eq.), and DIPEA (2 eq.) in DMF (5 mL) were added to couple the first amino acid and the mixture was incubated for 2 h. This was followed by capping with acetic anhydride (2 eq.) and pyridine (2 eq.) in DMF (5 mL) for 30 min. The resin was then transferred to the microwave synthesizer. For all further reaction steps the maximum temperature was set to 75 °C. Fmoc deprotection was generally achieved by twofold treatment with 20% piperidine (7 mL) and hydroxybenzotriazole (HOBt; 0.1 M) in DMF.[37] Initial deprotection was performed for 30 s at 35 W, followed by a second deprotection step using the same method for 180 s.

Coupling reactions for all amino acids except for arginine were performed by using TBTU (5 eq., 1 mL of a 0.5 M solution), the Fmoc-protected amino acid (5 eq., 2.5 mL of a 0.2 M solution), and DIPEA (10 eq., 0.5 mL of a 2.0 M solution) for 300 s at 20 W. For arginine the coupling was performed twice. Each coupling was performed for 1500 s without microwaves followed by irradiation for 300 s at 20 W. N-terminal acetylation was performed after the last amino acid with acetic anhydride (35 eq.), DIPEA (8.75 eq.), and HOBt (7 eq., 7 mL) with 30 s at 40 W followed by 30 s without microwaves.

Final cleavage was performed manually with a solution (10 mL) of 2.5% $H_2O$, 2.5% TIS, and 95% TFA and monitored by MALDI-TOF MS. After complete deprotection the volume was reduced to 5 mL and the solution was added dropwise to chilled diethyl ether. The precipitated peptide was dissolved in a 1:1 mixture of TFA and

acetonitrile. Purification by RP-HPLC was achieved by using TFA/acetonitrile/water elution gradients. The purified peptides were analyzed by MALDI-TOF MS.

*DNA:* All DNA constructs for AFM experiments were amplified with M13 forward and reverse primer (bold typeface). For immobilization in AFM experiments, the forward primer was modified with 5′-thiol; DNA for competition experiments was not modified.

The 268-bp *pstS* Pho box,[10] which contains four PhoB binding sites (underlined), was inserted in a pUC18 plasmid via *Hind*III and *Eco*RI restriction sites. Amplification resulted in a 360-bp DNA (**11**) product:

**CGCCAGGGTTTTCCCAGTCACGAC**GTTGTAAAACGACGGCCAGTGCCAAGCTTACCGTCATCT
TCGGCTACTTTTTCTCTGTCACAGAATGAAAATTTTTCTGTCATCTCTTCGTTATTAATGTTT
GTAATTGACTGAATATCAACGCTTATTTAAATCAGACTGAAGACTTTATCTCT**CTGTCA**TAAA
**ACTGTCA**TATTCC**TTACA**TATAAC**TGTCA**CCTGTTTGTCCTATTTTGCTTCTCGTAGCCAACA
AACAATGCTTTATGAATCCTCCCAGGAGACATTATGAAAGTTATGCGTACCACCGTCGAATTC
GTAATCATGGTCATAGCTGTT**TCCTGTGTGAAATTGTTATCCGCT**

The 75-bp KF-Ge DNA[22] was inserted in a pUC18 plasmid via *Hind*III and *Eco*RI restriction sites. Amplification resulted in a 177-bp DNA (**12**) product without PhoB binding sites:
**CGCCAGGGTTTTCCCAGTCACGA**CGTTGTAAAACGACGGCCAGTGCCAAGCTGCTCAAGAGCACGCAATTTCGGGGCAGGGGTGTTATGAAATTACTTCAAGTTTTGAAGTAATTTTCCGGAATTGGAATTCGTAATCATGGTCATAGCTGTT**TCCTGTGTGAAATTGTTATCCGCT**
The 226-bp *Eco*RI–*Ava*I EBNA DNA fragment was inserted in a pK18mob plasmid, which resulted in a 356-bp DNA product (**13**) after amplification:
**CGCCAGGGTTTTCCCAGTCACGA**CGTTGTAAAACGACGGCCAGTGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCCGGGATACACTCCGCTATCGCTACGTGACTGAGCTTATCGATGATAAGC**TGTCA**AACATGAGAATTAGATCCATTTTGGCTTGAAGCCAATATGATGGATCTAGAGGATCCATTAGGATAGCATATGCTACCCAGATAGGATCCAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAA**TGTCA**TGATAATAATGGTTTCTTAGGGAATTCGTAATCA**TGTCA**TAGCTGTT**TCCTGTGTGAAATTGTTATCCGCT**

For CD spectra and EMSA experiments, 18-bp ds DNA (**9**) was used. The sequence corresponds to the consensus sequence of the *pstS* Pho box[10] (5′-CTGTCATAAAACTGTCAT-3′); 18-bp random ds DNA (**10**) (5′-CGAGGCAGCATACGGATC-3′) was used as a negative control. The corresponding forward and reverse strands were dissolved in Tris (10 mM, pH 8.5) with a final concentration of 0.1 mM. For annealing, the mixture was heated to 100 °C using a water bath and slowly cooled to room temperature. Dimerization was proven by RP-HPLC with triethylammonium acetate buffer (0.1 mM, pH 7.0) and water gradients. For EMSA experiments, the forward strand was Cy3-modified at the 5′ end.

*EMSA:* EMSA experiments were performed with 10% TBE (90 mM Tris, 90 mM boric acid, 2 mM EDTA, pH 8.0) acrylamide gels ($16 \times 16\ cm^2$). Cy3-labeled 18-bp DNA (**9**; see above) was hybridized with its complementary strand in Tris (10 mM, pH 8.5) to give a final concentration of 10 ng $\mu L^{-1}$. The proteins were dissolved in $Na_2HPO_4$ (100 mM) and NaCl (50 mM) at pH 7.4. A solution (25 μL) of DNA (25 pM) and protein in different ratios in EMSA buffer (50 mM Tris, 0.1 M NaCl, 0.1 mM $MgSO_4$, 50 μg $mL^{-1}$ herring testis sperm DNA, 0.5 mg $mL^{-1}$ bovine serum albumin, pH 8.0) was incubated for 30 min at room temperature in the dark. After adding 10% glycerol, the samples (30 μL) were loaded on the gel. The assay was run at 4 °C for 2 h at 25 mA in the dark and visualized using a Typhoon 8600 variable-mode imager (GE Healthcare, München, Germany).

*Sample surface and AFM tip modification:* For single-molecule force spectroscopy experiments, $Si_3N_4$ cantilevers (Microlever; Veeco Instruments, Santa Barbara, USA) were briefly dipped into nitric acid and then placed in a solution of 2% aminopropyltriethoxysilane (Sigma) in dry toluene (Fluka) for 2 h at room temperature. The cantilevers were washed with toluene and autoclaved water and incubated for 1 h at 25 °C with *N*-hydroxysuccinimide–polyethylene glycol–maleimide (1 mM; Nektar, Huntsville, USA) in water. After washing with phosphate buffer (100 mM $Na_2HPO_4$, 50 mM NaCl, pH 7.4), the tips were functionalized by incubating them overnight at 4 °C with the thiol-modified DNA fragment (10 ng $\mu L^{-1}$, **11**–**13**; see above) in buffer solution. After washing with buffer the cantilevers were ready for use in force spectroscopy experiments.

Mica surfaces (Provac AG, Balzers, Liechtenstein) were silanized with aminopropyltriethoxysilane in a desiccator and afterwards incubated with bis(sulfosuccinimidyl)suberate sodium salt (20 μM) and the peptides (10 μM, **1**–**4**) with the short C-terminal linker 3,6-dioxa-1,8-diaminooctane (see above) in phosphate buffer for 1 h at room temperature. Subsequently, the sample was washed with buffer solution.

The proteins **5**–**8**, C-terminally ligated with a cysteine residue (see above), were immobilized on a gold surface (Arrandee, Werther, Germany) by incubating the surface with protein (10 μM) in phosphate buffer for 2 h at 25 °C. The sample was washed with buffer solution and used for force spectroscopy experiments.

*Force spectroscopy:* DFS experiments were performed with an MFP-3D AFM instrument (Asylum Research) using the provided software based on Igor Pro 5 (Wavemetrics), and with a multimode AFM instrument (Veeco Instruments) that was equipped with a custom-made force spectroscopy control system, using a 16-bit AD/DA card (PCI 6052E, National Instruments), a high-voltage amplifier (600H, NanoTechTools), and control and data acquisition software based on LabView (National Instruments). The spring constants of the AFM cantilevers were calibrated by the thermal fluctuation method.[38] All experiments were performed at room temperature in phosphate buffer. Typically, 2000–3000 force curves were recorded. In dynamic force experiments, the retract velocity was varied while the approach velocity was kept constant. All individually measured force curves were analyzed with dedicated Matlab-based software (Math Works). To determine the molecular elasticity for calculating the experimental loading rate, all force curves were automatically and individually corrected for AFM cantilever bending upon using the *z*-piezo extension. (This procedure does not alter the measured dissociation forces.)

*CD spectroscopy:* CD spectra were recorded on a J-810 CD spectrometer (JASCO, Germany). For all experiments with buffer, $Na_2HPO_4$ (10 mM)/NaCl (5 mM) buffer of pH 7.4 was used. All spectra were recorded at a scanning rate of 50 nm $min^{-1}$, a data pitch of 0.2 nm, and three accumulations. The protein concentrations were varied between 0.2 and 0.4 mg $mL^{-1}$ (0.16–0.32 mM). Peptide spectra were recorded at a concentration of 0.2 mM. Protein and peptide spectra were recorded in a 0.2-mm quartz cell.

NANO MICRO small

Molecular residual ellipticity [$\theta$] was calculated according to Equation (2):

$$[\theta] = \theta/(10 \cdot \mathrm{N} \cdot \mathrm{c} \cdot \mathrm{l}) \quad (2)$$

where $\theta$ is the ellipticity in millidegrees, $N$ is the number of residues, $c$ is the molar concentration in mol $L^{-1}$, and $l$ is the cell path length in centimeters. For titration experiments, TFE (0.2 mM) and water solutions were mixed to reach the desired ratios.

Spectra of the DNA–protein complex, protein only, and DNA only were recorded with a 1-mm quartz cell, to analyze structural changes in the DNA–protein complex. The protein-only spectrum was subtracted from that of the DNA–protein complex to analyze structural changes upon complex formation. The initial DNA concentration was 15 μM. Protein solution or buffer was added to obtain ratios between DNA and protein from 0.5 to 2. Molar CD absorption was calculated according to Equation (3):

$$\Delta\varepsilon = \theta/(32980 \cdot \mathrm{c} \cdot \mathrm{l}) \quad (3)$$

where $\theta$ is the ellipticity in millidegrees, $c$ is the final DNA concentration in mol $L^{-1}$ after addition of the protein, and $l$ is the cell path length in centimeters.

## Acknowledgements

*Thanks to S. Braun and F. Hofmann for experimental assistance and to M. Frese for carefully reading the manuscript. This work was financially supported by DFG (SFB 613). K.G. gratefully acknowledges the financial support (PhD grant) by the International NRW Graduate School of Bioinformatics and Genome Research.*

[1] C. Bustamante, J. C. Macosko, G. J. Wuite, *Nat. Rev. Mol. Cell Biol.* **2000**, *1*, 130–136.
[2] F. W. Bartels, B. Baumgarth, D. Anselmetti, R. Ros, A. Becker, *J. Struct. Biol.* **2003**, *143*, 145–152.
[3] F. Kuhner, L. T. Costa, P. M. Bisch, S. Thalhammer, W. M. Heckl, H. E. Gaub, *Biophys. J.* **2004**, *87*, 2683–2690.
[4] G. J. Gemmen, R. Millin, D. E. Smith, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11555–11560.
[5] E. Evans, K. Ritchie, *Biophys. J.* **1997**, *72*, 1541–1555.
[6] N. Sewald, S. D. Wilking, R. Eckel, S. Albu, K. Wollschläger, K. Gaus, A. Becker, F. W. Bartels, R. Ros, D. Anselmetti, *J. Pept. Sci.* **2006**, *12*, 836–842.
[7] J. B. Blanco, M. E. Vázquez, J. Martinez-Costas, J. Castedo, J. L. Mascareñas, *Chem. Biol.* **2003**, *10*, 701–713.
[8] O. Vázquez, M. E. Vázquez, J. B. Blanco, J. Castedo, J. L. Mascareñas, *Angew. Chem.* **2007**, *119*, 7010–7014; *Angew. Chem. Int. Ed.* **2007**, *46*, 6886–6890.
[9] A. G. Blanco, M. Sola, A. G. Gomis-Rüth, M. Coll, *Structure* **2002**, *10*, 701–713.
[10] K. Makino, M. Amemura, T. Kawamoto, S. Kimura, H. Shinagawa, A. Nakata, M. Suzuki, *J. Mol. Biol.* **1996**, *259*, 15–26.
[11] B. L. Wanner, *J. Cell. Biochem.* **1993**, *51*, 47–54.
[12] R. Eckel, S. D. Wilking, A. Becker, N. Sewald, R. Ros, D. Anselmetti, *Angew. Chem.* **2005**, *117*, 3989–3993; *Angew. Chem. Int. Ed.* **2005**, *44*, 3921–3924.
[13] P. Bachhawat, G. V. T. Swapna, G. T. Montelione, A. M. Stock, *Structure* **2005**, *13*, 1353–2363.
[14] H. Okamura, H. Shinagawa, A. Nagadoi, K. Makino, Y. J. Nishimura, *J. Mol. Biol.* **2000**, *295*, 1225–1236.
[15] T. Yamane, H. Okamura, M. Ikeguchi, Y. Nishimura, A. Kidera, *Proteins: Struct, Funct, Bioinf.* **2008**, *71*, 1970–1983.
[16] D. W. Ellison, W. R. McCleary, *J. Bacteriol.* **2000**, *182*, 6592–6597.
[17] S. D. Wilking, N. Sewald, *J. Biotechnol.* **2004**, *112*, 109–114.
[18] M.-Q. Xu, T. C. Evans, *Methods* **2001**, *24*, 257–277.
[19] C. J. Noren, J. Wang, B. P. Francine, *Angew. Chem.* **2000**, *112*, 458–476; *Angew. Chem. Int. Ed.* **2000**, *39*, 450–466.
[20] P.-H. Hirel, J.-M. Schmitter, P. Dessen, G. Fayat, S. Blanquet, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8247–8251.
[21] D. D. W. Hwang, L.-F. Lui, I.-C. Kuan, L.-Y. Lin, T.-C. S. Tam, M. F. Tam, *Biochem. J.* **1999**, *338*, 335–342.
[22] B. Baumgarth, F. W. Bartels, D. Anselmetti, A. Becker, R. Ros, *Microbiology* **2005**, *151*, 259–286.
[23] U. Dammer, O. Popescu, P. Wagner, D. Anselmetti, H.-J. Güntherodt, G. N. Misevic, *Science* **1995**, *267*, 1173–1175.
[24] R. Ros, F. Schwesinger, D. Anselmetti, M. Kubon, R. Schäfer, A. Plückthun, L. Tiefenauer, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 7402–7405.
[25] F. Schwesinger, R. Ros, T. Strunz, D. Anselmetti, H.-J. Güntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9972–9977.
[26] P. Hinterdorfer, W. Baumgartner, H. Gruber, K. Schilcher, H. Schindler, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3477–3481.
[27] R. Eckel, R. Ros, B. Decker, J. Mattay, D. Anselmetti, *Angew. Chem.* **2005**, *117*, 489–492; *Angew. Chem. Int. Ed.* **2005**, *44*, 484–488.
[28] R. Merkel, P. Nassoy, A. Leung, K. Ritchie, E. Evans, *Nature* **1999**, *397*, 50–53.
[29] Y.-H. Chen, J. T. Yang, H. M. Martinez, *Biochemistry* **1972**, *11*, 4120–4131.
[30] Y.-H. Chen, J. T. Yang, K. H. Chau, *Biochemistry* **1974**, *13*, 3350–3359.
[31] V. I. Ivanov, L. E. Minchenkova, A. K. Schyolkina, A. I. Poletayev, *Biopolymers* **1973**, *12*, 89–110.
[32] P. U. Maheswari, M. Palaniandavar, *J. Inorg. Biochem.* **2004**, *98*, 219–230.
[33] P. Luo, R. L. Baldwin, *Biochemistry* **1997**, *36*, 8413–8421.
[34] M. D. Bruch, M. M. Dhingra, L. M. Gierasch, *Proteins: Struct. Funct. Genet.* **1991**, *10*, 130–139.
[35] N.-J. Greenfield, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108–4116.
[36] S. Padmanabhan, S. Marqusee, T. Ridgeway, T. M. Laue, R. L. Baldwin, *Nature* **1990**, *244*, 268–270.
[37] A. S. Palasek, Z. J. Cox, J. M. Collins, *J. Pept. Sci.* **2007**, *13*, 143–148.
[38] L. J. Hutter, J. Bechhoefer, *Rev. Sci. Instrum.* **1993**, *64*, 1868–1873.

Received: July 4, 2008