Detailed studies of the binding mechanism of the *Sinorhizobium meliloti* transcriptional activator ExpG to DNA

Birgit Baumgarth,[1]† Frank Wilco Bartels,[2]† Dario Anselmetti,[2] Anke Becker[1] and Robert Ros[2]

Lehrstuhl für Genetik, Fakultät für Biologie[1] and Experimentelle Biophysik, Fakultät für Physik[2], Universität Bielefeld, 33615 Bielefeld, Germany

Correspondence
Anke Becker
Anke.Becker@genetik.uni-bielefeld.de

Received 29 June 2004
Revised 5 October 2004
Accepted 7 October 2004

The exopolysaccharide galactoglucan promotes the establishment of symbiosis between the nitrogen-fixing Gram-negative soil bacterium *Sinorhizobium meliloti* 2011 and its host plant alfalfa. The transcriptional regulator ExpG activates expression of galactoglucan biosynthesis genes by direct binding to the *expA1*, *expG*/*expD1* and *expE1* promoter regions. ExpG is a member of the MarR family of regulatory proteins. Analysis of target sequences of an ExpG(His)$_6$ fusion protein in the *exp* promoter regions resulted in the identification of a binding site composed of a conserved palindromic region and two associated sequence motifs. Association and dissociation kinetics of the specific binding of ExpG(His)$_6$ to this binding site were characterized by standard biochemical methods and by single-molecule spectroscopy based on the atomic force microscope (AFM). Dynamic force spectroscopy indicated a distinct difference in the kinetics between the wild-type binding sequence and two mutated binding sites, leading to a closer understanding of the ExpG–DNA interaction.

INTRODUCTION

The Gram-negative soil bacterium *Sinorhizobium meliloti* has the ability to fix molecular nitrogen in a symbiotic interaction with plants of the genera *Medicago*, *Melilotus* and *Trigonella*. The initiation of this symbiosis is a highly specific and complex developmental process, in which both partners undergo differentiation in a concerted way. In the early stages of this interaction, flavonoids in the root exudates induce the production of Nod factors by the bacteria which, in turn, trigger a developmental programme leading to the formation of root nodules (reviewed by Long, 2001; Spaink, 2000). The bacteria infect the nodule through infection threads and colonize the cytoplasm of plant cells. Once inside the plant cell, the bacteria differentiate into nitrogen-fixing bacteroids (reviewed by Oke & Long, 1999).

Bacterial exopolysaccharides (EPSs) are important for nodule infection. *S. meliloti* is able to synthesize two acidic EPSs, succinoglycan (EPS I) and galactoglucan (EPS II). Infection of *Medicago sativa* root nodules by *S. meliloti* depends on low-molecular-mass forms of EPS I or EPS II (Glazebrook & Walker, 1989; Gonzalez *et al.*, 1996; Wang *et al.*, 1999). EPS II is composed of alternating glucose and galactose residues which are decorated by acetyl and pyruvyl groups (Her *et al.*, 1990). The biosynthesis of EPS II is directed by the 30 kb *exp* gene cluster, containing 22 genes organized in four operons (Becker *et al.*, 1997; Rüberg *et al.*, 1999).

Under standard culture conditions in a complex medium, wild-type strain *S. meliloti* 2011 synthesizes EPS I and only traces of EPS II. The biosynthesis of EPS II is increased by phosphate-limiting conditions (Zhan *et al.*, 1991) or a mutation in either of the regulatory genes *mucR* (Keller *et al.*, 1995; Zhan *et al.*, 1989) and *expR* (Glazebrook & Walker, 1989; Pellock *et al.*, 2002), which are unlinked to the *exp* gene cluster. Extra copies of the regulatory gene *expG* located in the *exp* gene cluster (Astete & Leigh, 1996; Becker *et al.*, 1997; Rüberg *et al.*, 1999) stimulate transcription of the *expA*, *expD* and *expE* operons (Rüberg *et al.*, 1999). Under phosphate-limiting conditions the enhanced transcription of these operons requires *expG*, implying that ExpG acts as a transcriptional activator of *exp* gene expression (Astete & Leigh, 1996; Rüberg *et al.*, 1999).

ExpG was grouped into the MarR family of regulatory proteins (Becker *et al.*, 1997). Like many other transcriptional regulators MarR-type regulators bind DNA through a helix–turn–helix (HTH) motif (Cohen *et al.*, 1993; Sulavik *et al.*, 1995). An assortment of biological functions, e.g. the expression of resistance to multiple antibiotics, detergents

Abbreviations: AFM, atomic force microscopy; EMSA, electrophoretic mobility shift assay; EPS, exopolysaccharide; EPS I, succinoglycan; EPS II, galactoglucan; HTH, helix–turn–helix; k_{on}, on-rate; k_{off}, off-rate; K_d, dissociation constant.

†B. B. and F. W. B. contributed equally to this work.

and oxidative stress agents, organic solvents and pathogenic factors, is controlled by members of the MarR family (Alekshun & Levy, 1999; Miller & Sulavik, 1996). Most members act as repressors and only a few as activators (Egland & Harwood, 1999; Komeda *et al.*, 1996; Oscarsson *et al.*, 1996). Recently, we showed that ExpG itself exerts positive regulation of *exp* gene expression by binding to the *expA1*, *expG*/*expD1* and *expE1* promoter regions in the *exp* gene cluster (Bartels *et al.*, 2003).

In this paper we describe three distinct DNA sequence elements of the ExpG binding sites and their contribution to the specific binding process. Association and dissociation kinetics were characterized by ensemble and single-molecule methods.

METHODS

Bacterial strains, plasmids and growth conditions. *Escherichia coli* SURE (Young & Davis, 1983), used as expression strain, was cultivated in SB (Super broth) medium (32 g tryptone l^{-1}, 20 g yeast extract l^{-1} and 5 g NaCl l^{-1}, adjusted to pH 7·6) at 37 °C. Plasmid pHisGC31 was constructed by insertion of a 589 bp *Eco*RI–*Ecl*136II fragment carrying *expG* of *S. meliloti* 2011 (Casse *et al.*, 1979) into vector pEXP2 (MBBL, Bielefeld, Germany). This fragment was amplified by PCR using the primers expG.EXP2.5 (5′-AAAAGAATTCAAACCACAGGATACTCTATCCG) and HisG.EXP2.3 (5′-AAAAGAGCTC**TCA**gtgatggtgatggtgatgGATGCCGTAGCGTGCGGC) (*Eco*RI and *Ecl*136II restriction sites are underlined, the antisense sequence of the stop codon is printed in bold face and antisense His codons are in lower case) and pARIIa as template (Becker *et al.*, 1997). Expression of the hybrid gene resulted in a fusion protein consisting of ExpG from Asn-2 to Ile-190, with a $(His)_6$ C-terminal tag and ten additional N-terminal amino acids (MAIFEMLRNS).

Proteins. Expression, of recombinant $ExpG(His)_6$ fusion protein was performed essentially as described previously (Bartels *et al.*, 2003). Purification was carried out by Ni-NTA affinity chromatography (Qiagen). Purified fusion protein was concentrated using an Ultrafree 4 centrifugal concentrator (Millipore), resuspended in buffer (250 mM NaCl, 10 mM Tris, 1 mM DTT and 50 %, v/v, glycerol) and stored at −20 °C. The concentration of purified protein was determined by using the Bio-Rad Protein Assay (Bradford, 1976).

DNA fragments. DNA fragments I, II and III (see Fig. 2a) and competitor fragments for electrophoretic mobility shift assays (EMSAs) were generated by PCR as described previously (Bartels *et al.*, 2003). Following hybridization of the oligonucleotides (Fig. 2b) and their respective antisense oligonucleotides (synthesized by Qiagen) as described by Bertram-Drogatz *et al.* (1998), the double-stranded hybridization products were inserted into pUC18 (Yanisch-Perron *et al.*, 1985). The resulting plasmids were used as templates for amplification of fragments KF-A1c, d, e, f, g and h, KF-Ge and KF-E1e by PCR (for fragment lengths, see Fig. 2b) employing primers M13uni (5′-CGCCAGGGTTTTCCCAGTCACGAC-3′) and M13rev (5′-AGCGGATAACAATTTCACACAGGA-3′). These plasmids were also used for amplification of DNA fragments KF-A1e, f, g and h for atomic force microscopy (AFM) force spectroscopy employing primers M13uni (see above) and 5′SH-labelled primer M13rev (see above). DNA fragments for AFM imaging were amplified by PCR with primer ExpG1 (5′-AAAACTCGAGAGTCGTGTCTTACCGGG-TTG-3′) and M13uni (see above) from pFPG3/41 (pUC18 carrying the 348 bp *Eco*RI–*Hin*dIII fragment that comprises the intergenic region between *expD1* and *expG*).

EMSA analysis. Cy3-labelled DNA fragments I, II and III (see Fig. 1a) in various concentrations were mixed with purified $ExpG(His)_6$ (0·013 μg μl^{-1}) in a reaction buffer containing 50 mM Tris/HCl, pH 8·0, 100 mM NaCl, 0·1 mM $MgSO_4$, 5 % glycerol, 0·05 mg sonicated herring testis DNA ml^{-1} and 0·5 mg bovine serum albumin ml^{-1} and were subjected to EMSAs (Bartels *et al.*, 2003). In competition assays, protein was added to Cy3-labelled DNA fragments in the presence of various concentrations of competitor DNA fragments. After incubation at 20 °C for 15 min, the reaction was loaded onto a 2 % non-denaturing agarose gel prepared in gel buffer (40 mM Tris base, 10 mM sodium acetate and 1 mM EDTA adjusted to pH 7·8 with acetic acid) and electrophoresed at 4 °C in gel buffer at 4·5 V cm^{-1} for 2 h. Gel images were acquired using a Typhoon 8600 Variable Mode Imager (Amersham Bioscience).

Determination of on- and off-rates. Association kinetics for the binding reactions of $ExpG(His)_6$ to fragments I, II and III (Fig. 2a) were followed by loading samples taken after various time intervals between 0 and 15 min from a standard assay on an already electrophoresing agarose gel (Fried & Crothers, 1981; Lane *et al.*, 1992). For determination of the on-rate (k_{on}) of the reaction, initial rate data were evaluated by calculating $k_{on} = d[1/(P-N)]\ \ln[N(P-PN)/P(N-PN)]/dt$ (Lane *et al.*, 1992), where P is the concentration of free protein, PN the concentration of the protein–DNA complex, N the concentration of free DNA substrate, and t the time of sampling. The kinetics of dissociation of $ExpG(His)_6$ from fragments I, II and III (Fig. 2a) were investigated by adding a 50-fold excess of unlabelled competitor fragment to a mixture containing protein–DNA complexes and determining the degree of competition after various time intervals, as described above. The off-rate (k_{off}) was determined by calculating $k_{off} = -d[\ln(PN/PN_0)]/dt$ (Lane *et al.*, 1992), where PN is the concentration of the protein–DNA complex at sampling point after the addition of competitor, PN_0 represents the concentration of the protein–DNA complex directly before the addition of competitor, and t is the time of sampling. The dissociation constant (K_d) of the protein–DNA complex was calculated as the ratio of the off-rate (k_{off}) over the on-rate (k_{on}) (Bisswanger, 1994; Lane *et al.*, 1992). Experiments to determine the on- and off-rates were repeated three times.

HPLC gel-permeation chromatography of $ExpG(His)_6$. Size distribution of $ExpG(His)_6$ was measured by gel-permeation chromatography on a TSK Gel G2000SW column (TosoHaas) with a flow rate of 0·5 ml min^{-1} (eluent: 50 mM sodium phosphate buffer, pH 7·0). Calibration was performed with an LMW Gel Filtration Calibration Kit (Amersham Biosciences). Protein absorbance was measured at 280 nm.

AFM imaging. DNA fragments for AFM imaging consisted of 158 bp of the *expG* promoter region and a non-binding sequence of 990 bp. DNA (0·2 ng μl^{-1}) and ExpG protein (0·2 ng μl^{-1}) were mixed in buffer solution (50 mM Tris, 25 mM NaCl, 4 mM $NiCl_2$, pH 8·3) and left to incubate for 5 min before being brought to a freshly cleaved mica surface (Provac AG, Balzers, Liechtenstein) which was then immediately installed under the AFM liquid cell. The concentration of Ni^{2+} counter-ions was optimized to yield a flexible immobilization of the DNA to the mica surface by electrostatic attraction (Hansma & Laney, 1996). Protein–DNA complexes were investigated at 25 °C with a commercial AFM (Multimode, Veeco Instruments) in tapping mode, using oxide-sharpened Si_3N_4 cantilevers (Veeco Instruments) at a resonance frequency of about 28 kHz. Images were taken at a scan rate of 2 Hz, while the setpoint was kept at 0·2 V. Amplitude and phase were recorded simultaneously (from the signal in trace/retrace direction) to distinguish between DNA and protein (Lysetska *et al.*, 2002).

Sample surface and AFM tip modification. For force spectroscopy measurements, sample surfaces and AFM tips were

functionalized as described previously (Bartels *et al.*, 2003). Briefly, Si_3N_4 cantilevers (Microlever; Thermomicroscopes, Sunnyvale, CA, USA) were first activated by dipping for 10 s in concentrated nitric acid and silanized in a solution of 2 % aminopropyltriethoxysilane (Sigma) in dry toluene for 2 h. After washing with toluene, the cantilevers were incubated with 1 mM *N*-hydroxysuccinimide-poly (ethylene glycol)-maleimide (Shearwater Polymers) in 0·1 M potassium phosphate buffer, pH 8·0, for 30 min at room temperature. After washing with phosphate buffer, the cantilevers were incubated overnight at 4 °C with 10 ng μl^{-1} of the respective DNA target sequence (see above) bearing a thiol label in binding buffer solution (50 mM Tris/HCl, 100 mM NaCl, 0·1 mM $NiCl_2$, pH 8·3). The cantilevers were washed with binding buffer and used for force spectroscopy experiments. Modified tips were usable for at least a week if stored at 4 °C.

Mica surfaces (Provac) were silanized with aminopropyltriethoxysilane in an exsiccator (Lyubchenko *et al.*, 1993) and incubated with 4 μM $ExpG(His)_6$ protein and 20 μM bis(sulfosuccinimidyl)suberate sodium salt (Sigma) in 0·1 M potassium phosphate buffer, pH 7·5, for 1 h at 4 °C. The sample was washed with binding buffer afterwards. Modified surfaces were stable for at least 2 days if stored at 4 °C.

Dynamic force spectroscopy. Force spectroscopy measurements were performed with a commercial AFM head (Multimode; Veeco Instruments) at 25 °C. Acquisition of the cantilever deflection force signal and the vertical movement of the piezoelectric elements was controlled by a 16 bit AD/DA card (PCI-6052E; National Instruments) and a high-voltage amplifier (600H; NanoTechTools) via a home-built software based on Labview (National Instruments). The deflection signal was low-pass filtered (<6 kHz) and box-averaged by a factor of 10, giving a typical experimental dataset of 2000 points per force–distance curve.

The spring constants of all AFM cantilevers were calibrated by the thermal fluctuation method (Hutter & Bechhoefer, 1993) with an absolute uncertainty of approximately 15 %. Spring constants of the cantilevers used ranged from 12 pN nm^{-1} to 15 pN nm^{-1}.

For loading-rate-dependent measurements, the retract velocity of the piezo was varied while keeping the approach velocity constant. The measured force–distance curves were analysed with a Matlab program (MathWorks) and corrected to display the actual molecular distances calculated from the *z* piezo extension. To obtain the loading rate, the retract velocity was then multiplied by the elasticity of the molecular system, which was determined from the slope of the corrected force–distance curves on the last 20 data points before the unbinding events.

RESULTS AND DISCUSSION

A palindromic sequence in the *exp* promoter regions is required for binding of the ExpG protein

Recently, we demonstrated the binding of ExpG to promoter regions in the *exp* gene cluster (Bartels *et al.*, 2003). To localize the DNA region recognized by ExpG more precisely, we tested the specificity of DNA binding by ExpG in competition experiments. For this purpose, an $ExpG(His)_6$ fusion protein was expressed and purified by Ni-NTA affinity chromatography. In SDS-PAGE the purified protein exhibited a major band migrating at approximately 23 kDa which was detected using an anti-His-tag antibody (data not shown). This corresponds well with the calculated molecular mass of 23·2 kDa. An apparent molecular mass of the $ExpG(His)_6$ protein of 44·4 kDa in a 50 mM sodium phosphate buffer (pH 7·0) in non-denaturing conditions was determined by gel permeation chromatography (data not shown). This major peak probably represented a dimer. In addition to this peak, three higher molecular mass forms that may include a tetramer were detected in lower concentrations. This leads to the speculation that ExpG forms dimers and binds to its target DNA at least as a dimer.

The topography of the binding site was investigated by AFM in buffer solution. The 158 bp promoter region of *expG* (Bartels *et al.*, 2003) was extended by a 990 bp non-binding sequence at one end, resulting in a 390 nm DNA fragment suitable for AFM imaging. In accordance with this experimental setup, bound proteins were observed at only one end of a given DNA fragment, confirming the binding to the promoter region. Furthermore, AFM revealed a change in DNA conformation during the process of unbinding (Fig. 1), with a different curvature of the promoter region. Such a structural change was a recurring motif and has been observed for at least eight different DNA fragment–protein complexes. Proteins binding at other DNA sites or without structural transition have not been observed. Although the unbound and bound state was not directly observed in reverse order, it can be assumed that the DNA acquires its characteristic bend during the formation of the protein–DNA complex.

A conserved 21 bp region with a palindromic sequence which may constitute the binding site of ExpG was recently found in the promoter regions of *expA1*, *expG*, *expD1* and *expE1* (Bartels *et al.*, 2003; Lloret *et al.*, 2002). In addition to this conserved sequence two further regions in the *exp* promoter fragments, box 1 and box 2, share similarities (Fig. 2b). Eight different competitor fragments (Fig. 2b) were designed to test the importance of these two boxes and the palindromic sequence for binding of $ExpG(His)_6$. The double-stranded hybridization products from 28 bp to 80 bp (see Methods and Fig. 2b) were not effective competitors in EMSA experiments. To exclude the possibility that the competitor fragments were too short for protein binding, although they may carry the specific binding-site sequence, these fragments were cloned into the pUC18 vector. Flanking sequences derived from the pUC18 vector added 102 bp to the specific sequences from the *exp* promoter regions so that the fragments measured from 130 bp to 182 bp (Fig. 2b).

Competitor fragments KF-A1e, g and h, KF-Ge and KF-E1e contained the three motifs of the *expA*, *expG*/*expD* and *expE* promoter regions, respectively. Fragments KF-A1e (Fig. 3a), KF-Ge (Fig. 3b) and KF-E1e (Fig. 3c) with the wild-type sequence were effective competitors for the binding of $ExpG(His)_6$ to DNA fragments I, II and III containing the *expA*, *expG*/*expD* and *expE* promoter regions, respectively. KF-A1d, containing only the conserved palindrome region (Fig. 2b) did not compete out binding of $ExpG(His)_6$ (Fig. 3a). This was also the case for fragment KF-A1c, which

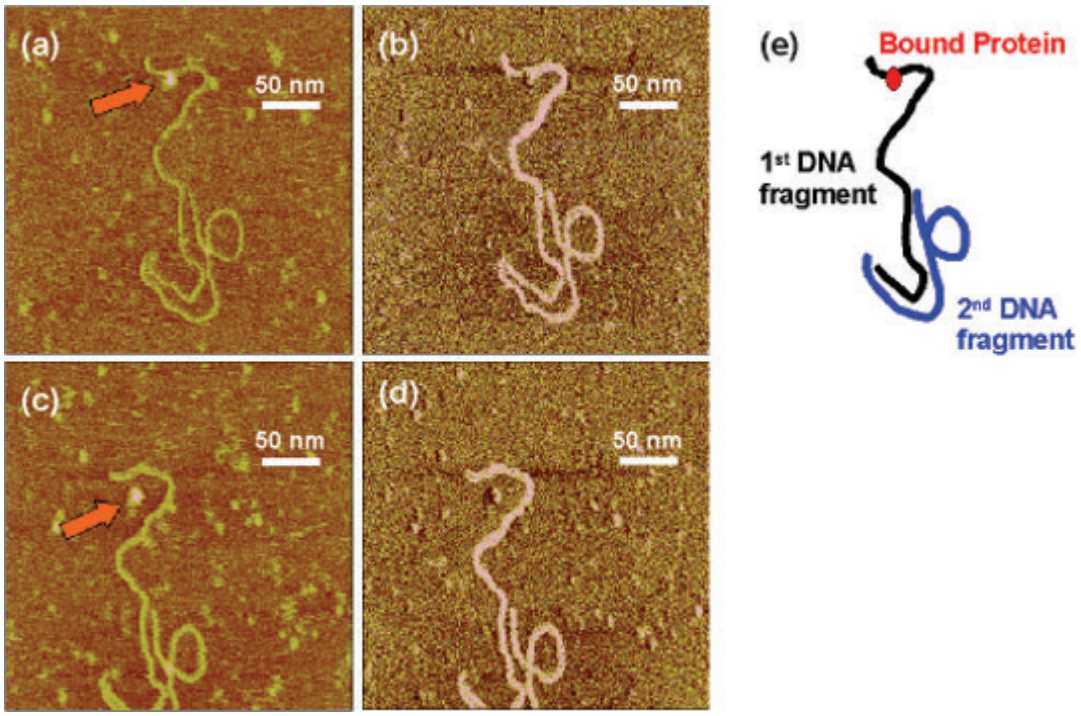

Fig. 1. AFM reveals a change in DNA conformation during unbinding of the protein–DNA complex. DNA fragments 390 nm long, containing the *expG* promoter region at one end, were imaged in buffer solution (50 mM Tris, 25 mM NaCl, 4 mM $NiCl_2$, pH 8·3) in the presence of $ExpG(His)_6$ proteins. Images on the left show the topography (*z* range 3·5 nm), while the corresponding phase information (*z* range 8·0°) is presented on the right. Height and phase signal were recorded simultaneously to distinguish between protein and DNA. The protein exhibits a negative phase shift (i.e. appears darker in the phase image) while the DNA shows a positive phase shift (i.e. appears brighter) (Lysetska *et al*., 2002). Images (a) and (b) show two DNA fragments (probably stabilizing each other by the medium of counter-ions), one of which carries a bound protein. When the protein (red arrow) breaks away (time between images: 9 min), the curvature of the DNA binding region changes (c and d). The bound state and the position of the two DNA fragments are shown in cartoon form in (e).

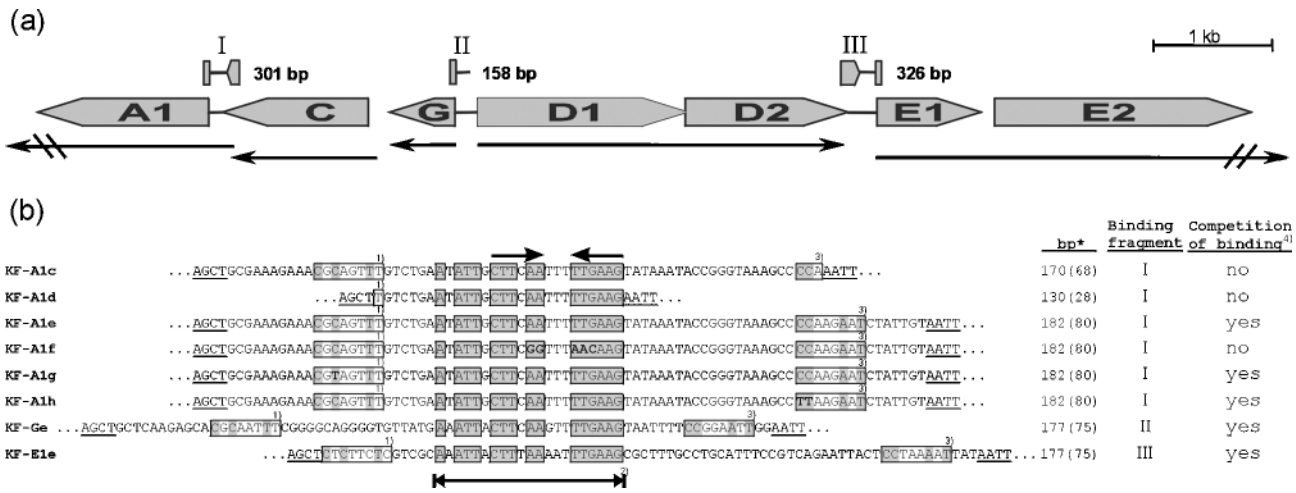

Fig. 2. (a) The *exp* gene region from *expA1* to *expE2* (10 330 nt) of the *exp* gene cluster of *S. meliloti* 2011 (Casse *et al*., 1979). Transcriptional units are marked by arrows. I, II and III are DNA fragments for EMSAs. (b) Alignment of the competitor fragments. Residues identical in the *expA1*, *expG* and *expE1* promoter regions are boxed and shaded. [1)] Box 1, [2)] conserved region, [3)] box 2. Mutations in box 1, the conserved palindrome region and box 2 are in bold. Inverted arrows indicate a palindrome found within the conserved region. AGCT, *Hin*dIII overhang; AATT, *Eco*RI overhang. *Fragment length after amplification by PCR; length of the native *S. meliloti* sequence is given in parentheses. [4)] See Fig. 3.

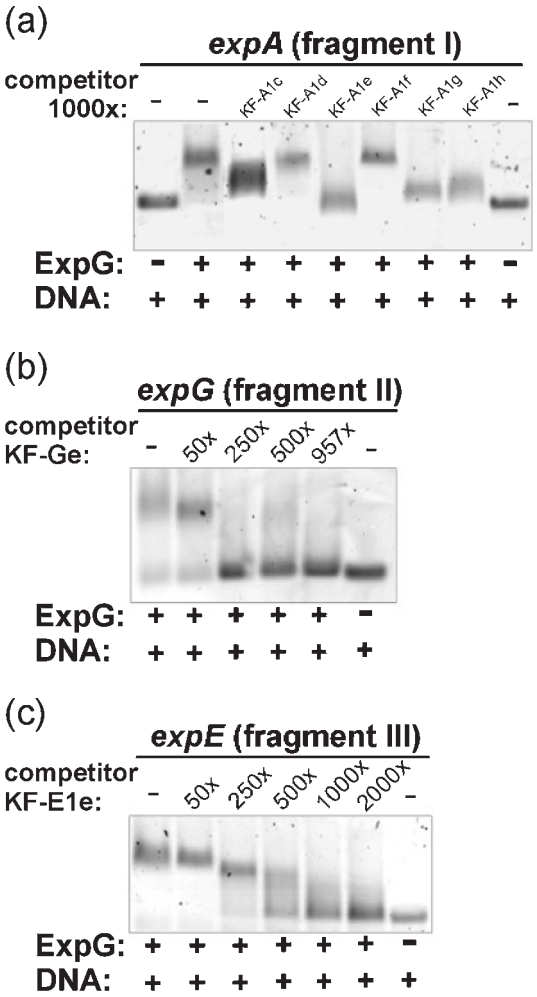

Fig. 3. EMSAs with purified ExpG(His)$_6$ fusion protein, (a) DNA fragment I, (b) DNA fragment II, (c) DNA fragment III (see Fig. 2a) and specific competitor DNA fragments (see Fig. 2b). Protein was added to all reactions to give a final protein concentration of 0·013 μg μl^{-1}. Unlabelled competitor DNA was added in increasing amounts from 50-fold to 2000-fold excess over the Cy3-labelled DNA fragments.

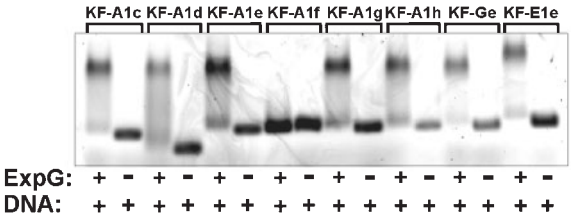

Fig. 4. EMSA with purified ExpG(His)$_6$ fusion protein and Cy3-labelled DNA fragments, which were further used as specific unlabelled competitor DNA (Figs 2b and 3a). Protein was added to all reactions to give a final protein concentration of 0·02 μg μl^{-1}. DNA was added to give a final concentration ranging from 0·05 ng μl^{-1} to 0·25 ng μl^{-1}.

included box 1, the palindrome region and 3 bp of box 2 (Figs 2b and 3a), suggesting that at least the palindrome region and box 2 are required for binding of ExpG.

Exchange of five nucleotides in the palindrome of the KF-A1e fragment resulted in fragment KF-A1f (Fig. 2b), which did not affect binding of ExpG(His)$_6$ to the wild-type sequence in the competition experiments (Fig. 3a). This indicates that the conserved palindromic sequence is essential for binding of ExpG(His)$_6$. Neither a mutation in box 1, KF-A1g (Fig. 2b), nor a mutation in box 2, KF-A1h (Fig. 2b), completely blocked binding of ExpG(His)$_6$ in the competition experiment (Fig. 3a). However, competition with fragments KF-A1c, g and h resulted in an incomplete shift (Fig. 3a), indicating a lower binding affinity of ExpG(His)$_6$ to these fragments. An EMSA using the eight specific competitor fragments as Cy3-labelled DNA fragments confirmed that the palindrome is essential for binding of ExpG(His)$_6$. ExpG(His)$_6$ completely reduced the electrophoretic mobility of each fragment except for the KF-A1f fragment, which contains the mutated palindrome (Fig. 4). Nevertheless, the failure of the fragments containing only the palindrome region (KF-A1d) or the palindrome region and box 1 (KF-A1c) in competition showed that the additional motifs box 1 and box 2 influence binding of ExpG(His)$_6$ as well. Possibly the change of 1 or 2 bp in box 1 and box 2 was not sufficient to destroy the function of these motifs.

The above results were confirmed by AFM force spectroscopy experiments. In this direct approach, binding of the ExpG(His)$_6$ protein to different DNA fragments can be observed on a single-molecule basis (Bartels *et al.*, 2003). By covalently attaching the binding partners to the AFM tip and the sample surface, respectively, unbinding forces of the protein–DNA complex can be measured during multiple approach–retract cycles of tip and surface. The unbinding forces under a single retract velocity show a nearly Gaussian distribution around the most probable unbinding force (Fig. 5). Again, mutations in the palindrome (fragment KF-A1f) led to no recognizable binding (Fig. 5b), whereas the fragments with mutations in box 1 (KF-A1g) or box 2 (KF-A1h) reached almost the same binding probability as the wild-type fragment (KF-A1e).

The function and structure of DNA-binding transcription factors of the MarR family are well investigated but little is known about their binding sites. We were able to narrow down the region required for binding of ExpG(His)$_6$ to a 63 bp region for the *expA1* promoter, a 62 bp region for the *expG* promoter, and a 72 bp region for the *expE1* promoter. We suggest that the 21 bp conserved sequence within the different *exp* promoter regions is the core region required for the binding of an ExpG dimer and the additional motifs box 1 and/or box 2 enable the complete DNA–protein interaction. A similar situation was suggested for LysR-type regulators (LTTR). A typical LTTR

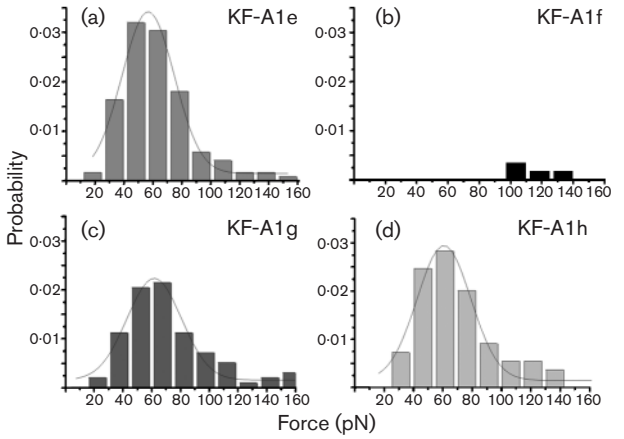

Fig. 5. Single-molecule force spectroscopy. Unbinding of the protein–DNA complex was measured in buffer solution (50 mM Tris/HCl, 100 mM NaCl, 0·1 mM $NiCl_2$, pH 8·3) for the ExpG protein and different DNA fragments: the wild-type sequence (KF-A1e) and fragments with mutations in the palindrome (KF-A1f), box 1 (KF-A1g) or box 2 (KF-A1h). Evidently, the palindrome was necessary for binding (b), while mutations in the box 1 and box 2 regions only slightly reduced the probability of binding (c–d) with respect to the wild-type sequence (a).

binds to a sequence of approximately 50–60 bp, which contains two distinct sites, a recoginition-binding site (RBS) recognized primarily by the LTTR, and an activation-binding site (ABS) (Schell, 1993). The long DNA sequence stretch containing the RBS and ABS and the size range of active LTTRs suggest that *in vivo* these regulators are multimeric and bind probably as dimers or tetramers (Henikoff *et al.*, 1988).

ExpG, which contains a HTH-MarR motif at the C-terminus (residues 70-164), is a member of the MarR family, which belongs to a supergroup of eight regulator families sharing a conserved extended sequence including the classical HTH motif (Perez-Rueda & Collado-Vides, 2001). The HTH motif is one of the most common DNA-binding motifs in proteins that control transcription

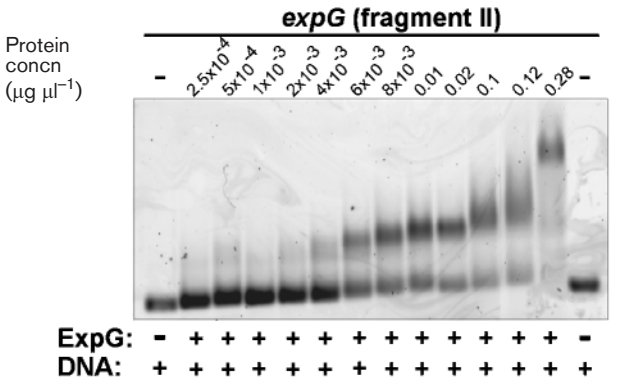

Fig. 6. EMSA with purified ExpG(His)$_6$ fusion protein of increasing concentration and DNA fragment II (see Fig. 2a). DNA was added to all reactions to give a final DNA concentration of 3·7 ng μl^{-1}.

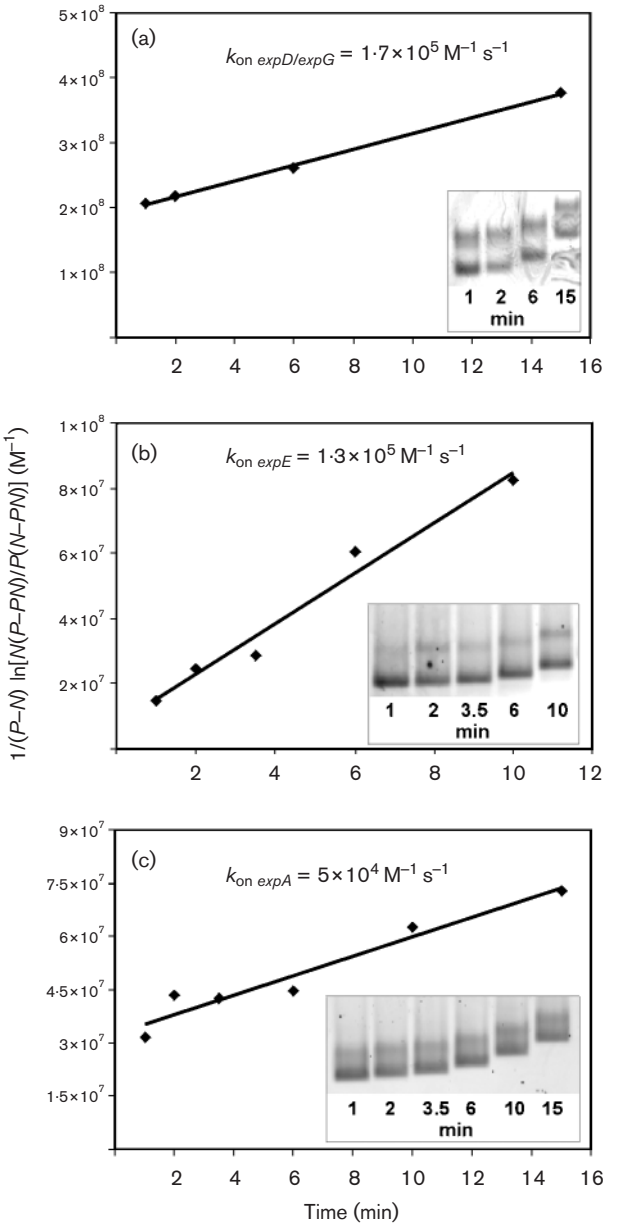

Fig. 7. On-rates (k_{on}) of the (a) ExpG–*expD*/*expG*, (b) ExpG–*expE* and (c) ExpG–*expA* protein–DNA complexes: association kinetics, represented by a plot of $1/(P-N)\ \ln[N(P-PN)/P(N-PN)]$ versus time. The ExpG protein concentration in the assay was calculated to be (a) $2{\cdot}8\times10^{-8}$ M, (b) $4{\cdot}5\times10^{-8}$ M and (c) $5{\cdot}6\times10^{-8}$ M. DNA was added to give the following final concentrations: (a) $3{\cdot}64\times10^{-8}$ M, (b) $1{\cdot}6\times10^{-8}$ M and (c) $4{\cdot}68\times10^{-8}$ M. The slope represents k_{on}. The insets show the original data obtained from an EMSA.

Table 1. On-rates (k_{on}), off-rates (k_{off}) and dissociation constants (K_d) of the ExpG(His)$_6$–DNA complexes

Promoter region	k_{on} (M^{-1} s^{-1})		k_{off} (s^{-1})	K_d (M)	
	Monomer	Dimer		Monomer	Dimer
expD/expG	$1{\cdot}7\times10^{5}$	$3{\cdot}7\times10^{5}$	$4{\cdot}3\times10^{-4}$	$2{\cdot}5\times10^{-9}$	$1{\cdot}2\times10^{-9}$
expE1	$1{\cdot}3\times10^{5}$	$5{\cdot}0\times10^{5}$	$2{\cdot}9\times10^{-4}$	$2{\cdot}2\times10^{-9}$	$5{\cdot}8\times10^{-10}$
expA1	$5{\cdot}0\times10^{4}$	$1{\cdot}0\times10^{5}$	$1{\cdot}3\times10^{-4}$	$2{\cdot}6\times10^{-9}$	$1{\cdot}3\times10^{-9}$

initiation (Sauer *et al.*, 1982). In repressor proteins the HTH binding motif is predominantly situated at the N-terminus, whereas activators mainly contain this motif at the C-terminus (Perez-Rueda & Collado-Vides, 2001). This observation is in agreement with the C-terminal position of this motif in the transcriptional activator ExpG.

Binding kinetics of the ExpG–DNA complexes

With this competition assay available, we aimed to determine the on- and off-rates of the ExpG(His)$_6$–DNA complexes of the *expD/expG*, *expE1* and *expA1* promoter fragments (see Methods). To determine the protein concentration for analysis of the binding kinetics we carried out EMSAs with increasing protein concentrations and DNA fragment II (Fig. 6), DNA fragment I and DNA fragment III (data not shown). Only at a protein concentration of 0·28 μg μl^{-1} ($1{\cdot}2\times10^{-5}$ M) was the electrophoretic mobility of the protein–DNA complex more strongly reduced in comparison to the lower protein concentrations. This may indicate the formation of a protein tetramer–DNA complex compared to a protein dimer–DNA complex that is probably formed at lower protein concentrations (Fig. 6). The protein concentration used in the EMSAs to investigate the binding kinetics of ExpG and the different *exp* promoter fragments was in the range $6{\cdot}5\times10^{-4}$ μg μl^{-1} to 0·013 μg μl^{-1} ($2{\cdot}8\times10^{-8}$ M to $5{\cdot}6\times10^{-7}$ M).

On-rates (k_{on}) measured for the binding reaction of ExpG(His)$_6$ to the different promoter fragments (Fig. 7) are summarized in Table 1. Since the results of the GPC analysis support the assumption that ExpG binds the DNA as a dimer we also calculated the on-rates for this DNA–dimer protein interaction (Table 1).

The dissociation kinetics of the ExpG(His)$_6$-mediated binding reaction were investigated by using the unlabelled fragments I, II and III (Fig. 2a) as competitors (Fig. 8). This revealed off-rates (k_{off}) (Table 1) at the lower limit of the mean off-rate $k_{off}=(1{\cdot}2\pm1{\cdot}0)\times10^{-3}$ s^{-1} for all three DNA target sequences which were previously obtained by AFM force spectroscopy experiments (Bartels *et al.*, 2003). From the on- and off-rates the different dissociation constants K_d were estimated for the complexes between the ExpG(His)$_6$ protein and the corresponding promoter fragments (Table 1). Many other transcriptional regulators, e.g. PcaU, MucR or Lrp, are characterized by quite similar dissociation constants in the nanomolar range (Bertram-Drogatz *et al.*, 1997; Calvo & Matthews, 1994; Popp *et al.*, 2002).

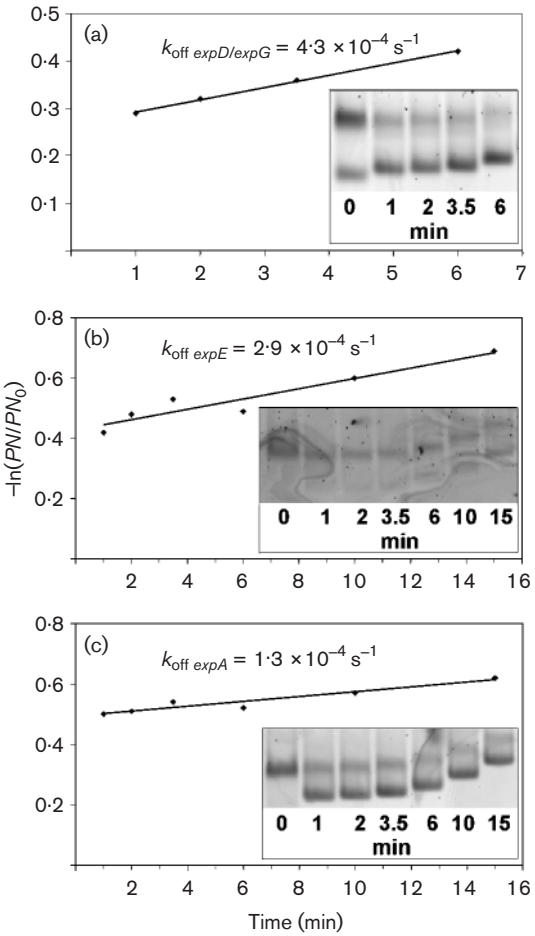

Fig. 8. Off-rates (k_{off}) of the (a) ExpG–*expD/expG*, (b) ExpG–*expE* and (c) ExpG–*expA* protein–DNA complexes: dissociation kinetics, represented by a plot of $-\ln(PN/PN_0)$ versus time. The ExpG protein concentration in the assay was calculated to be (a) $2{\cdot}8\times10^{-7}$ M, (b) and (c) $5{\cdot}6\times10^{-7}$ M. The slope represents k_{off}. DNA was added to give the following final concentrations: (a) $3{\cdot}64\times10^{-8}$ M, (b) $1{\cdot}06\times10^{-7}$ M and (c) $4{\cdot}68\times10^{-8}$ M. The insets show the original data obtained from an EMSA. The data obtained from the sample taken at 0 min (directly before the addition of competitor) was used to calculate PN_0.

We investigated the details of ExpG(His)$_6$ binding by single-molecule AFM force spectroscopy under different retract velocities. By varying the time-dependency of the external forces, i.e. the loading rate (loading rate = retract velocity × molecule elasticity), natural thermal off-rates can be measured (Bartels *et al.*, 2003; Evans & Ritchie, 1997; Merkel *et al.*, 1999; Schwesinger *et al.*, 2000; Strunz *et al.*, 1999). Dynamic force spectroscopy (Fig. 9) revealed a distinct difference in the kinetics between the wild-type sequence (KF-A1e) and its two mutated but still binding derivates (KF-A1g and KF-A1h). When the unbinding forces are plotted against the corresponding loading rates on a logarithmic scale, the experimental data can be fitted to a linear function according to the formula given by Strunz *et al.* (1999):

$$F = \frac{k_B T}{x_\beta} \ln \frac{x_\beta r}{k_B T k_{off}}$$

where F is the most probable unbinding force, $k_B T = 4{\cdot}114$ pN nm (at 298 K) is a Boltzmann factor, x_β is the distance between the minimum of the potential well of the bound state and the maximum of the energy barrier separating the bound state from the free state along the reaction coordinate, r is the loading rate, and k_{off} is the thermal off-rate under zero load. The natural thermal off-rate k_{off} can be derived by extrapolating the linear fit to the state of zero external force. We measured an off-rate $k_{off} = (4{\cdot}4 \pm 2{\cdot}5) \times 10^{-4}$ s^{-1} for the binding of ExpG(His)$_6$ to the wild-type DNA fragment, but $k_{off} = (5{\cdot}3 \pm 1{\cdot}5) \times 10^{-3}$ s^{-1} for fragment KF-A1h with a mutation in box 2 and even $k_{off} = (1{\cdot}3 \pm 0{\cdot}2) \times 10^{-2}$ s^{-1} for fragment KFA1g with a mutation in box 1. Thus, mutations in the box 1 and box 2 regions resulted in DNA sequences still capable of binding to ExpG(His)$_6$ but at the expense of a higher off-rate (i.e. a lower lifetime of the bond). We reason that the motifs box 1 and box 2 complement the 21 bp core region by fulfilling an important structural function in the binding of the ExpG(His)$_6$ protein, namely to stabilize the protein–DNA complex.

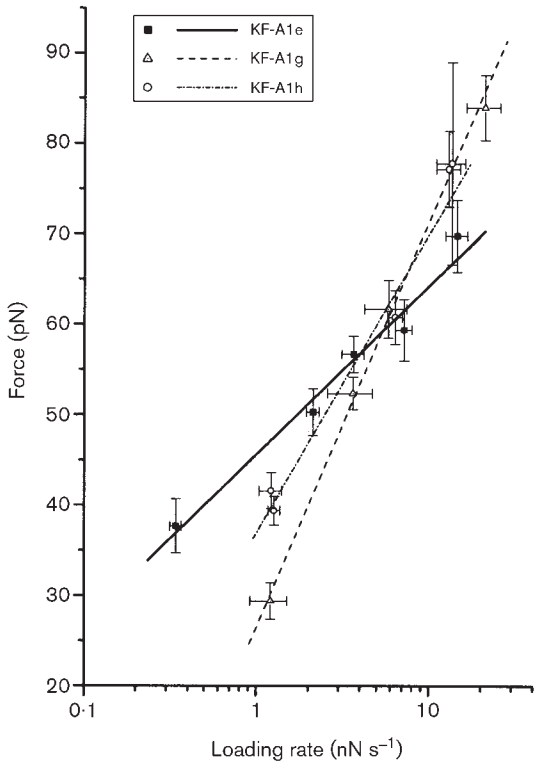

Fig. 9. Dynamic force spectroscopy. Loading rate dependent measurements (loading rate = retract velocity × molecule elasticity) are displayed for complexes formed by the ExpG(His)$_6$ protein and three DNA fragments: the wild-type sequence (KF-A1e) and mutant fragments with nucleotide changes in box 1 (KF-A1g) and box 2 (KF-A1h). The most probable unbinding force increases with the natural logarithm of the loading rate (Evans & Ritchie, 1997). By extrapolating the linear fit to the state of zero external force, the natural thermal off-rate can be derived (Strunz *et al.*, 1999). The off-rates differ by more than an order of magnitude, with k_{off}(KF-A1e) = $(4{\cdot}4 \pm 2{\cdot}5) \times 10^{-4}$ s^{-1}, k_{off}(KF-A1h) = $(5{\cdot}3 \pm 1{\cdot}5) \times 10^{-3}$ s^{-1} and k_{off}(KF-A1g) = $(1{\cdot}3 \pm 0{\cdot}2) \times 10^{-2}$ s^{-1}.

Conclusions

Our experiments demonstrate that the palindromic sequence within the *expA* promoter region is essential for binding of ExpG(His)$_6$ and suggest that the flanking sequence elements box 1 and box 2 contribute significantly to an efficient DNA–protein interaction. These findings were not only confirmed by standard ensemble methods, but also supported by data from AFM single-molecule force spectroscopy. AFM imaging explored the topography of the binding site in buffer solution conditions and suggests a change in DNA conformation upon binding of ExpG(His)$_6$. The dissociation constants K_d determined for the complex of ExpG(His)$_6$ and the corresponding promoter fragments are in good agreement with K_d values of other transcriptional regulators. The k_{off} rates obtained by the EMSA competition assay are at the lower limit of the mean off-rate for the *expA1*, *expD*/*expG* and *expE1* promoter regions previously measured by AFM force spectroscopy using a N-terminal (His)$_6$ExpG fusion protein (Bartels *et al.*, 2003). This makes it unlikely that the DNA-binding activity was severely affected in the same way by both tags. Furthermore, dynamic force spectroscopy reveals a distinct difference in the kinetics of the wild-type binding sequence and the fragments containing mutations within the box 1 and box 2 motifs. Whereas the EMSA experiments detected binding or non-binding of the DNA–protein complex, the analysis at single molecule level showed that the mutated sequences of box 1 and box 2 lead to a higher off-rate. The ExpG(His)$_6$ binding sites characterized in this study overlap with the putative PHO boxes previously predicted in the *exp* promoter regions (Rüberg *et al.*, 1999), suggesting an interference of PhoB and ExpG binding in the regulation of *exp* promoter activities.

ACKNOWLEDGEMENTS

We thank A. Pühler for helpful discussion. This work was supported by SFB 613 from the Deutsche Forschungsgemeinschaft.

REFERENCES

Alekshun, M. N. & Levy, S. B. (1999). The *mar* regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol* **7**, 410–413.

Astete, S. G. & Leigh, J. A. (1996). *mucS*, a gene involved in activation of galactoglucan (EPS II) synthesis gene expression in *Rhizobium meliloti*. *Mol Plant–Microbe Interact* **9**, 395–400.

Bartels, F. W., Baumgarth, B., Anselmetti, D., Ros, R. & Becker, A. (2003). Specific binding of the regulatory protein ExpG to promoter regions of the galactoglucan biosynthesis gene cluster of *Sinorhizobium meliloti* – a combined molecular biology and force spectroscopy investigation. *J Struct Biol* **143**, 145–152.

Becker, A., Rüberg, S., Küster, H., Roxlau, A. A., Keller, M., Ivashina, T., Cheng, H. P., Walker, G. C. & Pühler, A. (1997). The 32-kilobase *exp* gene cluster of *Rhizobium meliloti* directing the biosynthesis of galactoglucan: genetic organization and properties of the encoded gene products. *J Bacteriol* **179**, 1375–1384.

Bertram-Drogatz, P. A., Rüberg, S., Becker, A. & Pühler, A. (1997). The regulatory protein MucR binds to a short DNA region located upstream of the *mucR* coding region in *Rhizobium meliloti*. *Mol Gen Genet* **254**, 529–538.

Bertram-Drogatz, P. A., Quester, I., Becker, A. & Pühler, A. (1998). The *Sinorhizobium meliloti* MucR protein, which is essential for the production of high-molecular-weight succinoglycan exopolysaccharide, binds to short DNA regions upstream of *exoH* and *exoY*. *Mol Gen Genet* **257**, 433–441.

Bisswanger, H. (1994). *Enzymkinetik: Theorie und Methoden*. Weinheim: VCH.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.

Calvo, J. M. & Matthews, R. G. (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* **58**, 466–490.

Casse, F., Boucher, C., Hulliot, J. S., Michel, M. & Dénarié, F. (1979). Identification and characterization of large plasmids in *Rhizobium meliloti* using agarose gel electrophoresis. *J Bacteriol* **113**, 229–242.

Cohen, S. P., Hachler, H. & Levy, S. B. (1993). Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. *J Bacteriol* **175**, 1484–1492.

Egland, P. G. & Harwood, C. S. (1999). BadR, a new MarR family member, regulates anaerobic benzoate degradation by *Rhodopseudomonas palustris* in concert with AadR, an Fnr family member. *J Bacteriol* **181**, 2102–2109.

Evans, E. & Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. *Biophys J* **72**, 1541–1555.

Fried, M. & Crothers, D. M. (1981). Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res* **9**, 6505–6525.

Glazebrook, J. & Walker, G. C. (1989). A novel exopolysaccharide can function in place of the calcofluor-binding exopolysaccharide in nodulation of alfalfa by *Rhizobium meliloti*. *Cell* **56**, 661–672.

Gonzalez, J. E., Reuhs, B. L. & Walker, G. C. (1996). Low molecular weight EPS II of *Rhizobium meliloti* allows nodule invasion in *Medicago sativa*. *Proc Natl Acad Sci U S A* **93**, 8636–8641.

Hansma, H. G. & Laney, D. E. (1996). DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. *Biophys J* **70**, 1933–1939.

Henikoff, S., Haughn, G. W., Calvo, J. M. & Wallace, J. C. (1988). A large family of bacterial activator proteins. *Proc Natl Acad Sci U S A* **85**, 6602–6606.

Her, G. R., Glazebrook, J., Walker, G. C. & Reinhold, V. N. (1990). Structural studies of a novel exopolysaccharide produced by a mutant of *Rhizobium meliloti* strain Rm1021. *Carbohydr Res* **198**, 305–312.

Hutter, J. L. & Bechhoefer, J. (1993). Calibration of atomic-force microscope tips. *Rev Sci Instrum* **7**, 1868–1873.

Keller, M., Roxlau, A., Weng, W. M., Schmidt, M., Quandt, J., Niehaus, K., Jording, D., Arnold, W. & Pühler, A. (1995). Molecular analysis of the *Rhizobium meliloti mucR* gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan. *Mol Plant–Microbe Interact* **8**, 267–277.

Komeda, H., Kobayashi, M. & Shimizu, S. (1996). Characterization of the gene cluster of high-molecular-mass nitrile hydratase (H-NHase) induced by its reaction product in *Rhodococcus rhodochrous* J1. *Proc Natl Acad Sci U S A* **93**, 4267–4272.

Lane, D., Prentki, P. & Chandler, M. (1992). Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol Rev* **56**, 509–528.

Lloret, J., Martin, M., Oruezabal, R. I., Bonilla, I. & Rivilla, R. (2002). MucR and *mucS* activate *exp* genes transcription and galactoglucan production in *Sinorhizobium meliloti* EFB1. *Mol Plant–Microbe Interact* **15**, 54–59.

Long, S. R. (2001). Genes and signals in the *Rhizobium*–legume symbiosis. *Plant Physiol* **125**, 69–72.

Lysetska, M., Knoll, A., Boehringer, D., Hey, T., Krauss, G. & Krausch, G. (2002). UV light-damaged DNA and its interaction with human replication protein A: an atomic force microscopy study. *Nucleic Acids Res* **30**, 2686–2691.

Lyubchenko, Y., Shlyakhtenko, L., Harrington, R., Oden, P. & Lindsay, S. (1993). Atomic force microscopy of long DNA: imaging in air and under water. *Proc Natl Acad Sci U S A* **90**, 2137–2140.

Merkel, R., Nassoy, P., Leung, A., Ritchie, K. & Evans, E. (1999). Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature* **397**, 50–53.

Miller, P. F. & Sulavik, M. C. (1996). Overlaps and parallels in the regulation of intrinsic multiple-antibiotic resistance in *Escherichia coli*. *Mol Microbiol* **21**, 441–448.

Oke, V. & Long, S. R. (1999). Bacteroid formation in the *Rhizobium*-legume symbiosis. *Curr Opin Microbiol* **2**, 641–646.

Oscarsson, J., Mizunoe, Y., Uhlin, B. E. & Haydon, D. J. (1996). Induction of haemolytic activity in *Escherichia coli* by the *slyA* gene product. *Mol Microbiol* **20**, 191–199.

Pellock, B. J., Teplitski, M., Boinay, R. P., Bauer, W. D. & Walker, G. C. (2002). A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. *J Bacteriol* **184**, 5067–5076.

Perez-Rueda, E. & Collado-Vides, J. (2001). Common history at the origin of the position-function correlation in transcriptional regulators in archaea and bacteria. *J Mol Evol* **53**, 172–179.

Popp, R., Kohl, T., Patz, P., Trautwein, G. & Gerischer, U. (2002). Differential DNA binding of transcriptional regulator PcaU from *Acinetobacter* sp. strain ADP1. *J Bacteriol* **184**, 1988–1997.

Rüberg, S., Pühler, A. & Becker, A. (1999). Biosynthesis of the exopolysaccharide galactoglucan in *Sinorhizobium meliloti* is subject to a complex control by the phosphate-dependent regulator PhoB and the proteins ExpG and MucR. *Microbiology* **145**, 603–611.

Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M. & Pabo, C. O. (1982). Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. *Nature* **298**, 447–451.

Schell, M. A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* **47**, 597–626.

Schwesinger, F., Ros, R., Strunz, T., Anselmetti, D., Guntherodt, H. J., Honegger, A., Jermutus, L., Tiefenauer, L. & Pluckthun, A. (2000). Unbinding forces of single antibody-antigen complexes correlate with their thermal dissociation rates. *Proc Natl Acad Sci U S A* **97**, 9972–9977.

Spaink, H. P. (2000). Root nodulation and infection factors produced by rhizobial bacteria. *Annu Rev Microbiol* **54**, 257–288.

Strunz, T., Oroszlan, K., Schafer, R. & Guntherodt, H. J. (1999). Dynamic force spectroscopy of single DNA molecules. *Proc Natl Acad Sci U S A* **96**, 11277–11282.

Sulavik, M. C., Gambino, L. F. & Miller, P. F. (1995). The MarR repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. *Mol Med* **1**, 436–446.

Wang, L. X., Wang, Y., Pellock, B. & Walker, G. C. (1999). Structural characterization of the symbiotically important low-molecular-weight succinoglycan of *Sinorhizobium meliloti*. *J Bacteriol* **181**, 6788–6796.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Young, R. A. & Davis, R. W. (1983). Yeast RNA polymerase II genes: isolation with antibody probes. *Science* **222**, 778–782.

Zhan, H. J., Levery, S. B., Lee, C. C. & Leigh, J. A. (1989). A second exopolysaccharide of *Rhizobium meliloti* strain SU47 that can function in root nodule invasion. *Proc Natl Acad Sci U S A* **86**, 3055–3059.

Zhan, H. J., Lee, C. C. & Leigh, J. A. (1991). Induction of the second exopolysaccharide (EPSb) in *Rhizobium meliloti* SU47 by low phosphate concentrations. *J Bacteriol* **173**, 7391–7394.