doi:10.1016/j.jmb.2008.09.067 J. Mol. Biol. (2008) 384, 690–701

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Structural Basis of the Role of the NikA Ribbon-Helix-Helix Domain in Initiating Bacterial Conjugation

Hitoshi Yoshida[1], Nobuhisa Furuya[1], Yi-Jan Lin[2], Peter Güntert[1,3,4], Teruya Komano[1] and Masatsune Kainosho[1,5]*

[1]*Graduate School of Science, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji Tokyo 192-0397, Japan*

[2]*Faculty of Biotechnology and Center of Excellence for Environmental Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan*

[3]*Institute of Biophysical Chemistry and Center of Biomolecular Magnetic Resonance, Goethe-University Frankfurt am Main, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany*

[4]*Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany*

[5]*Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8602, Japan*

Received 13 June 2008; received in revised form 18 September 2008; accepted 23 September 2008
Available online 7 October 2008

Edited by M. F. Summers

Conjugation is a fundamental process for the rapid evolution of bacteria, enabling them, for example, to adapt to various environmental conditions or to acquire multi-drug resistance. NikA is one of the relaxosomal proteins that initiate the intercellular transfer of the R64 conjugative plasmid with the P-type origin of transfer, *oriT*. The three-dimensional structure of the N-terminal 51 residue fragment of NikA, NikA(1–51), which binds to the 17-bp repeat A sequence in R64 *oriT*, was determined by NMR to be a homodimer composed of two identical ribbon-helix-helix (RHH) domains, which are commonly found in transcriptional repressors. The structure determination of NikA(1–51) was achieved using automated NOE assignment with CYANA, without measuring filtered NOESY experiments to distinguish between the intra- and intermolecular NOEs, and without any a priori assumption on the tertiary or quaternary structure of the protein. Mutational experiments revealed that the DNA-binding region of the NikA(1–51) dimer is an anti-parallel β-sheet composed of one β-strand from each of the N-terminal ends of the two domains. Various biochemical experiments have indicated that the full length NikA(1–109) exists as a homotetramer formed through an α-helical domain at the C-terminus, and that the anti-parallel β-sheets of both dimeric domains bind to two homologous 5 bp internal repeats within repeat A. As a tetramer, the full length NikA(1–109) showed higher affinity to repeat A and bent the *oriT* duplex more strongly than NikA (1–51) did. Many RHH proteins are involved in specific DNA recognition and in protein–protein interactions. The discovery of the RHH fold in NikA suggests that NikA binds to *oriT* and interacts with the relaxase, NikB, which is unable to bind to the nick region in *oriT* without NikA.

Keywords: relaxosome; bacterial conjugation; ribbon-helix-helix fold; homodimer; NMR

**Corresponding author.* Graduate School of Science, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji Tokyo 192-0397, Japan. E-mail address: kainosho@nmr.chem.metro-u.ac.jp.

Abbreviations used: DMP, dimethyl pimelimidate; DSS, 2-dimethyl-2-silapentane-5-sulfonate sodium salt; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; PH, pleckstrin homology; RHH, ribbon-helix-helix.

Introduction

Bacterial conjugation is the main route for horizontal gene transfer in prokaryotes.[1,2] The process is encoded by diverse plasmids and conjugative transposons. Conjugation systems are remarkable in mediating transfer between a wide range of bacteria and, in some cases, from bacteria to fungal and plant cells. Conjugation allows rapid

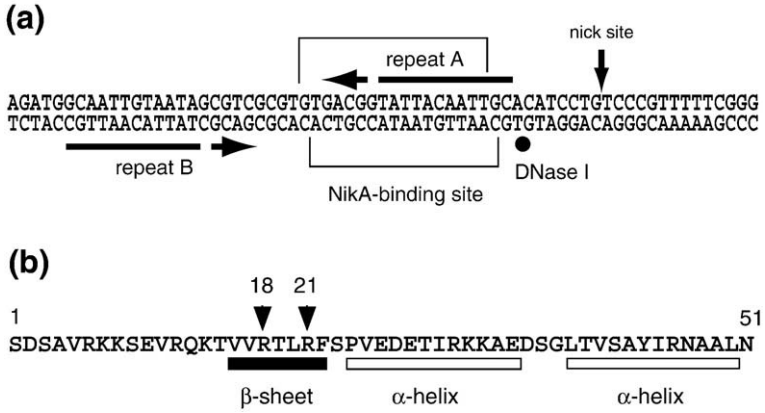

Fig. 1. (a) Sequence around the NikA-binding site within the *oriT* region of the R64 plasmid. NikA binds specifically to the repeat A sequence, indicated by an arrow, which is the first of two 17 bp inverted repeat sequences found 8 bp apart from the nick site. Repeat B, the second inverted repeat, which lacks NikA-binding ability, but is required for the termination of bacterial conjugation, differs from repeat A by a single nucleotide, marked by a gap in the arrows. Brackets mark the top and bottom strands of the sequence protected by NikA from DNase I cleavage. The position of the DNase I hypersensitive phosphodiester bond is indicated by the filled circle. The downward arrow indicates the nick site. (b) Amino acid sequence of NikA(1–51). Secondary structure elements are indicated. Two basic residues in the β-sheet, which were changed to Leu in the mutation experiments described for Fig. 5, are indicated by arrowheads.

evolution, resulting in organisms with new clinical or environmental characteristics. In particular, antibiotic resistance is related to the conjugative transfer of mobile genetic elements.

The processing of plasmid DNA during bacterial conjugation involves several steps. Upon the initiation of bacterial conjugation, a site- and strand-specific nick is introduced into the nick site in the origin of transfer (*oriT*) by an *oriT*-specific relaxase and auxiliary proteins, which form a protein–DNA complex (called the relaxosome) at the *oriT* site. The nicked single strand is then transferred from the donor to the recipient cell. Finally, religation of the transferred strand by the relaxase and replacement- and complementary-strand DNA syntheses establish double-stranded plasmid DNAs in the donor and recipient cells, respectively.

Various conjugative plasmids, such as F, R388, RP4, R100 and R64, carry their own specific *oriT* sites.[2] Each *oriT* sequence is recognized by proteins encoded on its own plasmid, consisting of a specific relaxase and auxiliary proteins. Three major groups of *oriT*-relaxase systems have been identified: P-type, Q-type and F-type.[2]

The P-type *oriT* carries the nick region sequence YATCCTG | Y (the vertical bar represents the nick site).[3-5] The relaxases encoded on the plasmids with a P-type *oriT*, R64 NikB, RP4 TraI, R751 TraI and pTF-FC2 MobA, share conserved N-terminal motifs.[5,6] It is thought that the nick region, YATCCTG | Y, is recognized by the conserved motifs of these relaxases.

The plasmids with the P-type *oriT* have two other similarities. The nick sites of R64 and RP4 are situated 8 bp away from the 17 bp and 19 bp inverted repeats, respectively, although the sequences of the inverted repeats are different.[4] Furthermore, there are 8 bp or 6 bp G + C-rich inverted repeats 6–54 bp away from the 17 bp or 19 bp inverted repeats. These two sets of inverted repeat sequences are required for the efficient termination of the DNA transfer.[7]

For the initiation of DNA transfer by R64, the binding of NikA to the first 17 bp inverted repeat, repeat A, is necessary.[3] Repeat A differs from the left inverted repeat, repeat B, by a single nucleotide (Fig. 1a). In the case of RP4, TraJ was shown to bind specifically to the first 19 bp repeat with three nucleotide mismatches with the second inverted repeat.[8] NikA shares significant homology with RP4 TraJ (30% sequence identity in a 94 amino-acid overlap).[4] Furthermore, R751 TraJ and pTF-FC2 MobB have 25% and 23% identity with the 109 residues of R64 NikA, respectively. These similarities suggest that a DNA-binding auxiliary protein, encoded on plasmids with the P-type *oriT*, binds in the vicinity of the nick site and recruits the relaxase.

The 110 codons of the R64 *nikA* gene encode a basic polypeptide. The purified full-length NikA protein consists of 109 amino acids, with the initial methionine removed after *in vivo* translation.[3] To elucidate the molecular mechanism of the initiation of DNA transfer by the R64 plasmid, the three-dimensional structure of the NikA protein was characterized by solution state NMR. The molecular mechanism of the initiation of DNA transfer of conjugative plasmids with P-type *oriT* is discussed.

Results

oriT binding by the deletion mutant NikA(1–51)

To determine the DNA-binding domain of NikA, a series of NikA deletion mutants were constructed. Gel retardation assays and DNase I footprint analysis of the deletion mutants revealed that the NikA fragment encompassing residues 1–51, NikA(1–51) (Fig. 1b), retains the ability to bind specifically to repeat A (Fig. 1a). NikA proteins with further deletions lacked the repeat A binding ability. Interestingly, the DNase I sensitive region adjacent to repeat A was not observed in the DNase I footprint analysis of NikA(1–51) (data not shown). NikA(1–51) was shown to be inactive in R64 conjugation and relaxation complex formation.

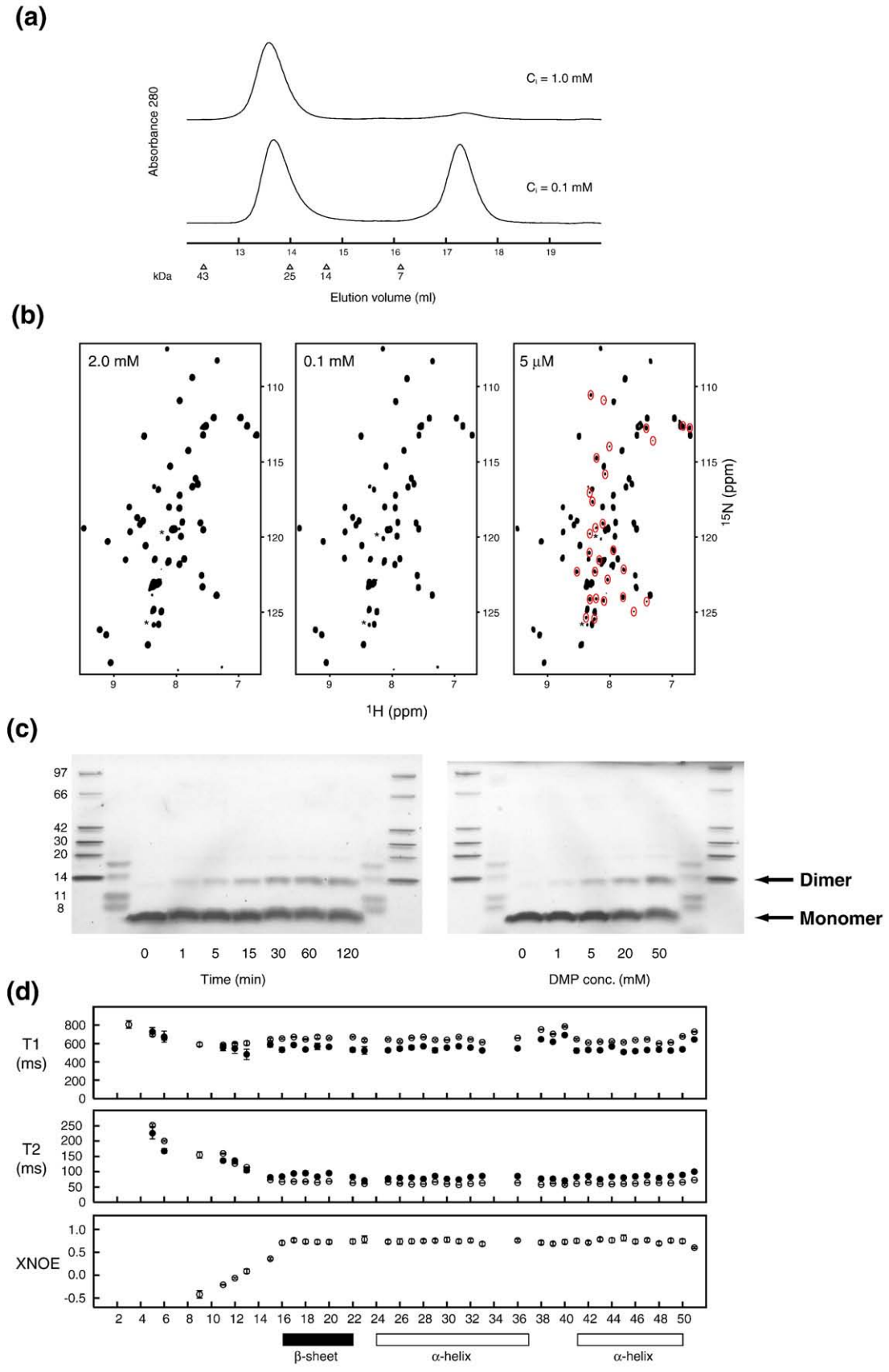

Fig. 2 (*legend on next page*)

Homodimerization of NikA(1–51)

We assessed the oligomeric state of NikA(1–51). Size-exclusion chromatography showed that NikA (1–51) exists as a single oligomer at a protein final elution concentration higher than 0.1 mM, whereas NikA(1–51) exists in a monomer–oligomer equilibrium at a lower concentration (Fig. 2a). The existence of only one set of peaks in the ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum confirmed the presence of only a single type of oligomer at a protein concentration between 0.1 mM and 2.0 mM, whereas the existence of more than one set of peaks in the ^{1}H-^{15}N HSQC spectrum at a protein concentration of 5 μM indicated a monomer–oligomer equilibrium at a lower concentration (Fig. 2b). Cross-linking experiments with dimethyl pimelimidate (DMP; SIGMA) were performed to further characterize the oligomeric state of NikA(1–51) at the optimum temperature, 37 °C, and the optimum pH, 8.5. A cross-linking time course (1–120 min) experiment, with a DMP concentration of 50 mM, indicated that 1 h is sufficient for the reaction (Fig. 2c, left-hand panel). In the presence of increasing concentrations of DMP (1–50 mM) with incubation for 1 h, greater amounts of the cross-linked dimer were detected unambiguously (Fig. 2c, right-hand panel). This cross-linking experiment demonstrated that only a single oligomeric state exists at concentrations higher than 0.1 mM, and that it is a dimer.

A search of the Protein Data Bank revealed that residues 16–51 of NikA(1–51) share relatively high sequence identity (25%) with a dimeric protein, residues 8–46 of the Arc repressor.[9,10] Interestingly, the MYL mutant (residues 8–46) of the Arc repressor has an even higher sequence identity, 28%, with residues 16–51 of NikA(1–51). In the MYL mutant, M31, Y36 and L40 contact each other via hydrophobic interactions and increase the stability of the protein.[11] Hence, it is possible that the LYA (L39, Y44 and A48) residues at the corresponding positions in NikA(1–51) contact each other in the same manner as the MYL residues of the Arc repressor mutant. Moreover, the finding that the same positions of the secondary structure elements of NikA(1–51) exhibited slow amide proton exchange with water, as well as the results of the TALOS[12] analysis of the chemical shift information, suggested that residues 16–51 of NikA(1–51) adopt a fold similar to that of the MYL mutant of the Arc repressor, which forms a homodimer. The homodimeric structure of NikA(1–51) was determined by the program CYANA, as described below, and was found to be similar to the structure of the homodimeric MYL mutant of the Arc repressor. A search for nuclear Overhauser effects (NOEs) that might identify a dimer–dimer interface and thus a tetrameric structure was unsuccessful. These results indicated that NikA(1–51) forms a homodimer at a high concentration of protein.

Although most of the experimental results suggested that NikA(1–51) exists as a homodimer, the apparent molecular mass (M_{app}) of NikA(1–51), determined by size-exclusion chromatography (Fig. 2a) and sedimentation equilibrium analysis (14 kDa at 0.15 mM and 23 kDa at 0.5 mM), was larger than that expected for a dimer. Additionally, the ^{15}N relaxation measurements, T_1, T_2, and ^{1}H-^{15}N NOE, yielded rotational correlation times, τ_m, of the α-helix and β-sheet regions of 9.1 ns and 11.8 ns at protein concentrations of 0.3 mM and 2.0 mM, respectively. These correlation times are longer than those expected from the molecular mass of this dimer. These observations seem to suggest that NikA(1–51) also forms higher-order oligomers. However, non-specific associations have reportedly changed the values of M_{app} and τ_m for several proteins, including the dynamin pleckstrin homology (PH) domain,[13] and the N-terminal domain of H-NS.[14] In these cases, it was not possible to deduce the precise molecular mass from the values of M_{app} and τm. In order to assess whether the large M_{app} and the high τ_m are the result of a non-specific association, the ^{1}H-^{15}N HSQC spectrum recorded for the 2.0 mM sample was compared with that obtained for the 0.1 mM sample. The two spectra are almost identical, with only a few cross-peaks exhibiting detectable chemical shift perturbations, and all of them were less than 0.04 ppm and 0.12 ppm in the ^{1}H and ^{15}N dimensions, respectively. The lack of concentration dependence of the chemical shift suggests that the concentration-dependent increase of τ_m is due to a non-specific association. Likewise, the relaxation parameters T_1 and T_2 in the protein core, measured in the 2.0 mM sample, exhibited a pattern similar to that for the 0.3 mM data (Fig. 2d). This indicates that the concentration dependence of the T_1 and T_2 values in the protein core is not affected by a structural

Fig. 2. Data showing the oligomeric state of NikA(1–51). (a) Size-exclusion chromatography analysis of NikA(1–51) at injected concentrations (C_i) of 1.0 mM and 0.1 mM. The samples were diluted about tenfold during chromatography. The concentrations of the eluted NikA(1–51) of C_i = 1.0 mM and 0.1 mM were diluted to approximately 0.1 mM and 0.01 mM, respectively. The elution volumes of the protein standards are indicated below the chromatogram. (b) The concentration dependence of the ^{1}H-^{15}N HSQC spectra of NikA(1–51). All spectra were collected on a Bruker DRX600 spectrometer at 298 K. The resonances that correspond to the monomer are marked with a red ellipse. The asterisks indicate impurities. (c) DMP-mediated cross-linking of NikA(1–51). In the left-hand panel, the time course of cross-linking was followed at 37 °C in the presence of 50 mM DMP. In the right-hand panel, cross-linking was performed in the presence of increasing concentrations of DMP at 37 °C for 60 min. Samples were electrophoresed in a Tris-Tricine gel. The arrows on the right indicate the different species formed. The molecular mass markers on the left are expressed in kDa. (d) Backbone ^{15}N relaxation times, T_1, T_2, and steady-state ^{1}H-^{15}N NOE values plotted as a function of the residue number. The data shown are for the 2.0 mM (open circles) and 0.3 mM (filled circles) samples. The locations of the secondary structure elements are indicated.

change, but by the change in the overall tumbling rate of the protein core. These observations revealed that non-specific association biased the M_{app} and τ_m values, and thus they could not provide a precise molecular mass. Moreover, the N-terminal flexible tail (residues 1–15) may contribute to the higher value of M_{app} determined by size-exclusion chromatography. Therefore, we concluded that NikA(1–51) forms a homodimer.

Resonance assignment

The solution structure of NikA(1–51) was solved by multi-dimensional heteronuclear NMR spectroscopy, using a protein concentration of 2.0 mM. The presence of only one set of peaks on the ^{1}H-^{15}N HSQC spectrum of NikA(1–51) indicated that the homodimer is structurally symmetric. The resonance assignments for the carbon-bound protons, the backbone amide protons and the side chain amide protons of Asn and Gln and their associated ^{13}C and ^{15}N nuclei in the structured region of residues 16–51 were complete, except for R18 γCH_2, S23 βCH_2, K32 εCH_2, K33 $\alpha/\gamma/\delta/\varepsilon CHn$, E35 γCH_2, S42 βCH_2, and R46 γCH_2. For the complete polypeptide chain of 51 residues, 90% of the aforementioned ^{1}H nuclei were assigned. The trans conformation of the single Xxx–Pro peptide bond was confirmed by the intense Xxx-H^{α} – Pro-H^{δ} sequential NOE spectroscopy (NOESY) cross-peaks,[15] and by a $\delta(^{13}C^{\beta}) - \delta(^{13}C^{\gamma})$ chemical shift difference of 3.2 ppm.[16]

Structure determination of the NikA(1–51) homodimer by CYANA

In the case of a symmetric homodimer, the assignment of NOE cross-peaks is complicated by the ambiguity between the intramolecular and intermolecular assignments.[17] In principle, the assignments of the intramolecule, intermolecule and mixed (intra- and intermolecule) NOE cross-peaks can be distinguished by two methods. In one method, the assignments are obtained by using heterodimeric samples, in which one protomer is isotopically labeled with ^{15}N and/or ^{13}C and the other is not, in combination with isotope-filtered experiments.[18] In the other method, the assignments are done by using the structure of a highly homologous protein as a template structure. The first method generally suffers from the low sensitivity of isotope-filtered experiments, and bias may be introduced by the second approach. We therefore adopted a different strategy.

For the structure determination of NikA(1–51), the assignments of the intramolecule, intermolecule and mixed (intra- and intermolecule) NOEs were obtained by automated NOE assignment with a version of the program CYANA adapted for the structure calculation of homodimeric proteins (see Materials and Methods). The NikA (1–51) structure was determined on the basis of 1672 NOE-based distance restraints, including 280 intermolecular contacts, and restraints for 16 α-helical hydrogen bonds. The structure is well defined by the NMR data, with an RMSD of 0.31 Å for the backbone atoms of residues 16–51 (Table 1). The solution structure, represented by the final 20 conformers and by a ribbon diagram, is presented in Fig. 3.

Table 1. Structural statistics for the homodimeric NikA (1–51) NMR structure

NMR distance restraints:	
Total NOE	1672
Short-range, $\lvert i-j\rvert \le 1$	776
Medium-range, $1 < \lvert i-j\rvert < 5$	430
Long-range, $\lvert i-j\rvert \ge 5$, intramolecular	186
Intermolecular	280
Restrained α-helical hydrogen bonds/monomer	16
Maximal distance restraint violation (Å)	0.15
AMBER energies:	
Total (kcal/mol) (mean±SD of 20 conformers)	−3875±138
van der Waals (kcal/mol)	−215±12
RMS deviations from idealized geometry:	
Bond lengths (Å)	0.0139±0.001
Bond angles (°)	1.74±0.05
Ramachandran plot statistics:[57]	
Residues in most favored regions (%)	85.6
Residues in additionally allowed regions (%)	13.9
Residues in generously allowed regions (%)	0.4
Residues in disallowed regions (%)	0.1
RMSD from the mean coordinates:	
N, C^{α}, C′ of residues 16–51 of both monomers (Å)	0.31±0.04
Heavy atoms of residues 16–51 of both monomers (Å)	0.74±0.06

The homodimeric structure of NikA(1–51), obtained without using any predefined intermolecular restraints, consists of an unstructured tail of residues 1–15 and a closely packed structured region of residues 16–51 with an intermolecular two-stranded antiparallel β-sheet and two α-helices. The structure overlapped well with the structure of the MYL mutant of the Arc repressor,[11] and the side chains of the LYA sequence (L39, Y44, and A48) are involved in hydrophobic interactions, like the MYL sequence of the Arc repressor mutant. The structures of NikA(1–51) are in agreement with the hydrogen bond correlations of an intermolecular antiparallel β-sheet, $V16^{HN}$-$F22^{C'}$, $R18^{HN}$-$L20^{C'}$, $L20^{HN}$-$R18^{C'}$ and $F22^{HN}$-$V16^{C'}$, which were identified in the long-

Fig. 3. The three-dimensional structure of homodimeric NikA(1–51). (a) The 20 energy-refined CYANA conformers that represent the solution structure of NikA(1–51), superposed for best fit of the backbone atoms N, C^{α} and C′ of residues 14–50. (b) Ribbon diagram of a representative structure of NikA(1–51). (c) Selected region of the two-dimensional $^{3h}J_{NC'}$ HNCO spectrum of NikA(1–51), showing the through hydrogen bond *J* correlations involving the amides of the intermolecular β-sheet. Cross-peaks marked as Res*i*/Res*j* occurred between the amide hydrogen of residue i and the carbonyl oxygen of residue *j*.

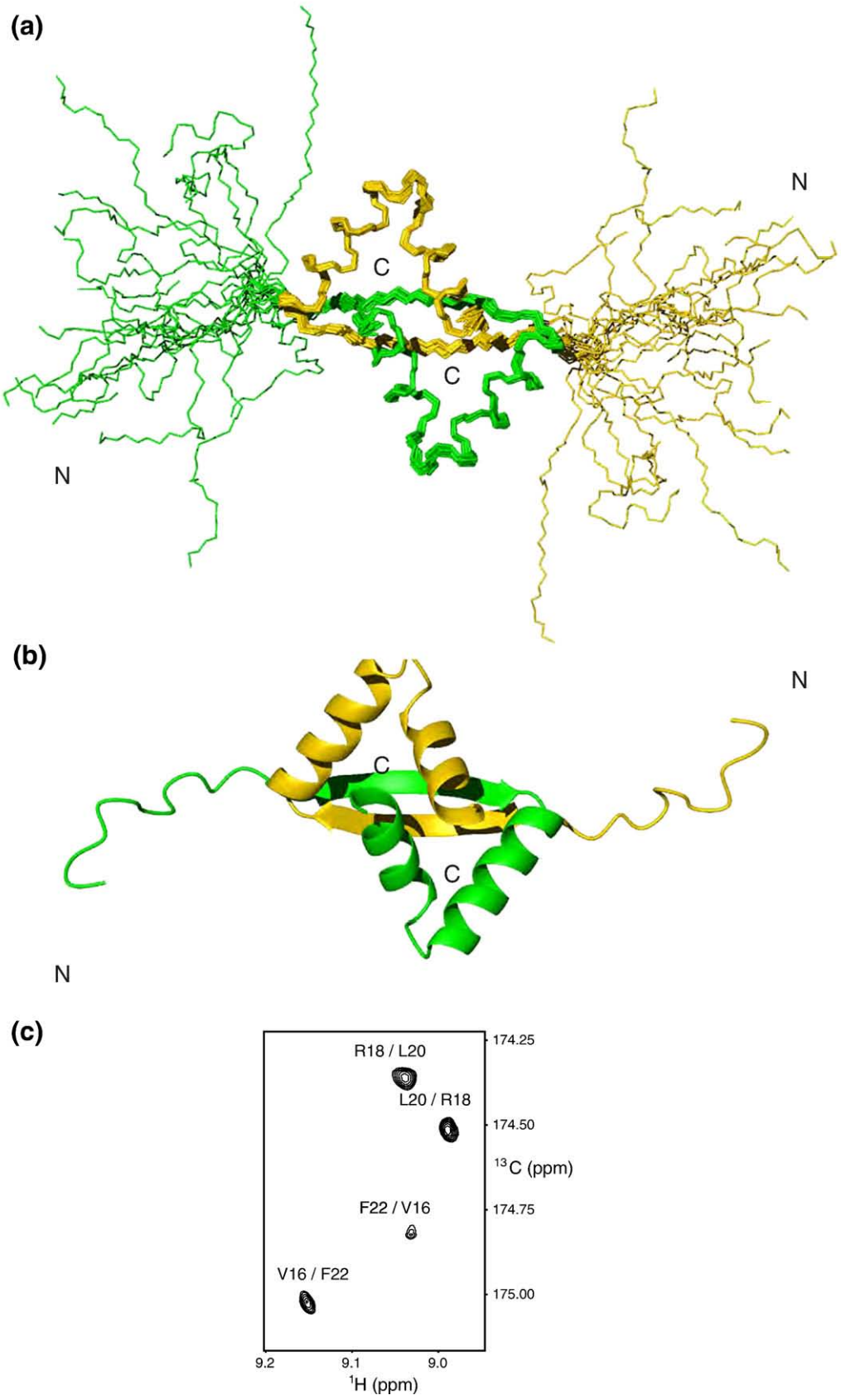

Fig. 3 (*legend on previous page*)

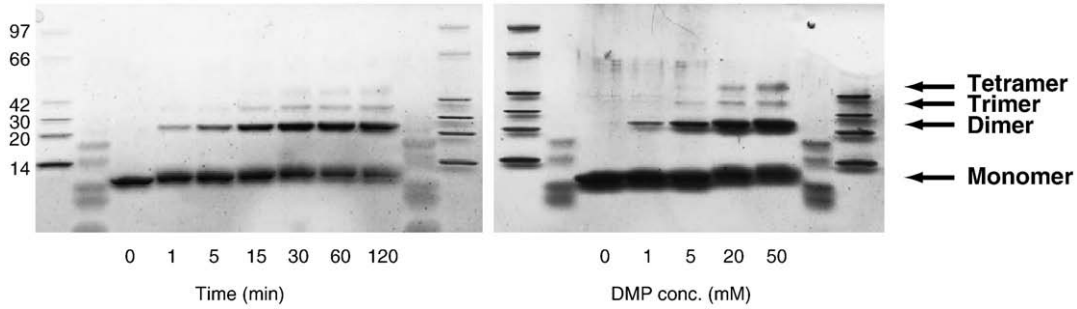

Fig. 4. DMP-mediated cross-linking of full-length NikA. In the left-hand panel, the time course of cross-linking was followed at 37 °C in the presence of 50 mM DMP. In the right-hand panel, cross-linking was performed in the presence of increasing concentrations of DMP at 37 °C for 60 min. Samples were electrophoresed in a Tris-Tricine gel. The arrows on the right indicate the different species formed, and the molecular mass markers on the left are expressed in kDa.

range $^{3h}J_{NC'}$ HNCO spectrum[19] (Fig. 3c), but were not used as input for the structure calculation.

Homotetramer formation by full-length NikA

In order to explore the oligomeric state of full-length NikA, we subjected the protein to size-exclusion chromatography. The full-length NikA protein eluted as a single peak at all concentrations tested (1.4 μM < C_i < 1.3 mM), whereas NikA(1–51) eluted as two peaks at the lower concentrations. The elution times of these single peaks were concentration-dependent, corresponding to a molecular mass of 45 kDa for C_i = 1.3 mM and C_i = 140 μM, 42 kDa for C_i = 14 μM, and 39 kDa for C_i = 1.4 μM. It is likely that full-length NikA exists in an equilibrium among various oligomers. DMP cross-linking of full-length NikA generated dimers, trimers and tetramers, indicating that it forms a tetramer (Fig. 4). The presence of a strong dimer band, as well as trimer and tetramer bands that were almost as intense as the dimer band of NikA(1–51), indicated that a new oligomeric interface exists in the C-terminal region of NikA. Secondary structure predictions suggested that two α-helices exist in the C-terminal region, involving residues 52–109. These results indicated that the C-terminal region of NikA contains an α-helix-rich domain that is responsible for homotetramer formation.

A polar amino acid in the antiparallel β-sheet in the RHH domain binds DNA

Proteins with the ribbon-helix-helix (RHH) fold usually bind to DNA with an intermolecular antiparallel β-sheet. Basic residues on this β-sheet are especially important for DNA binding. The β-sheet of NikA contains two basic residues, R18 and R21 (Fig. 1b). The positively charged side chain of R21 is oriented toward the outside of the β-sheet, whereas the side chain of R18 is oriented toward the inside of the protein. The NikA mutants R18L and R21L were constructed in order to verify that NikA utilizes its intermolecular β-sheet to bind to repeat A. The alterations of R18 and R21 to leucine were designed to maintain the hydrophobic characteristics of the arginine side chain.[20-26] A comparison of the ^{1}H-^{15}N HSQC spectra of the mutant and wild type proteins indicated that neither the structure of the β-sheet region nor that of the whole protein were affected by either the R18L or R21L mutation. The repeat A binding by the NikA mutant was measured by a gel retardation assay (Fig. 5). The absence of *oriT*-specific binding by the R21L mutant confirmed that NikA, like other RHH proteins, binds to *oriT* with its intermolecular β-sheet.

oriT bending by NikA and NikA(1-51)

The binding of NikA to repeat A bends the *oriT* DNA.[3] This NikA-induced bending of *oriT* DNA might have an important role in nicking at the *oriT* nick site. In order to assess the role of the C-terminal region of NikA in the *oriT* DNA bending, the DNA-bending activities of full-length NikA and NikA(1-51) were compared, using a permutated set of pKK524 DNA fragments in which the position of the 44-bp *oriT* core sequence was varied (Fig. 6a). The *oriT* core sequence contains repeat A and the nick site, but not repeat B. For both full-length NikA and NikA(1–51), the EcoRV-generated fragment migrated more slowly in the gel retardation assay than the Mlu- or BamHI-generated fragments. The DNA-bending assay revealed that NikA(1-51) retains the DNA-bending activity, but shows a lesser extent of bending than full-length NikA (Fig. 6). Moreover, the DNase I footprint analysis of NikA (1-51) showed no DNase I-hypersensitive phosphodiester bond, which appears when NikA binds to repeat A (data not shown). These experiments revealed that both the N-terminal RHH domains and the C-terminal region of NikA are required for the conformational change of the *oriT* DNA, whereas the N-terminal RHH domains are sufficient for specific repeat A binding.

Discussion

There are many published structures representing the RHH fold of transcriptional repressors, either alone or in complexes with their target DNA: Arc,[9,10] including the MYL mutant,[11] Mnt,[27] MetJ,[28,29] CopG,[30,31] ω,[32] ParG,[33] and NikR.[34,35] Sequence

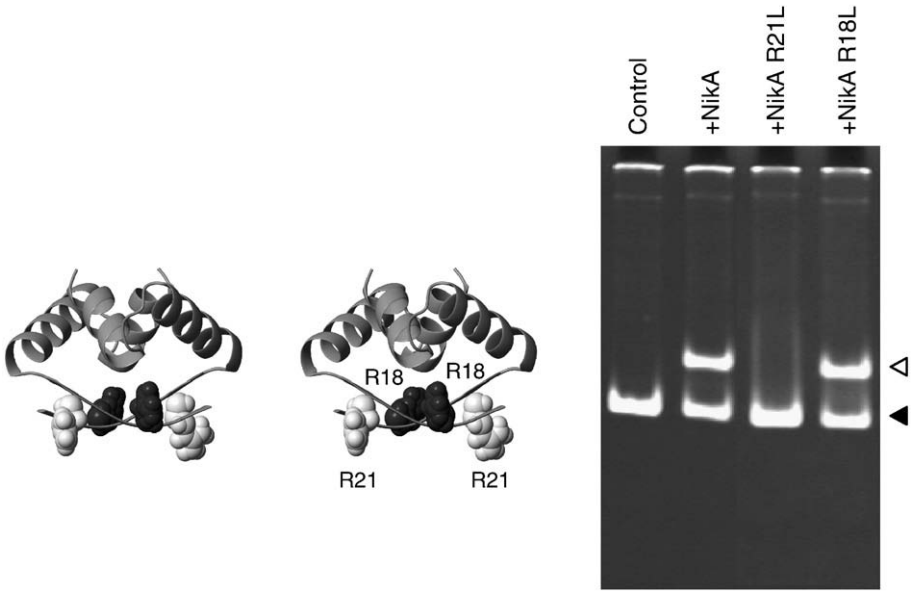

Fig. 5. Effects of the R18L and R21L mutations on the specific binding of NikA to a 162 bp DNA fragment containing the 96 bp minimal R64 *oriT* sequence. In the left-hand panel, the sidechains of R18 (black) and R21 (white) are shown in a stereo view of the ribbon representation of the RHH domain (residues 16–51). In the right-hand panel, the DNA–protein mixtures were electrophoresed in a 4% polyacrylamide gel. The open and filled triangles point to the free DNA and the NikA-bound DNA, respectively.

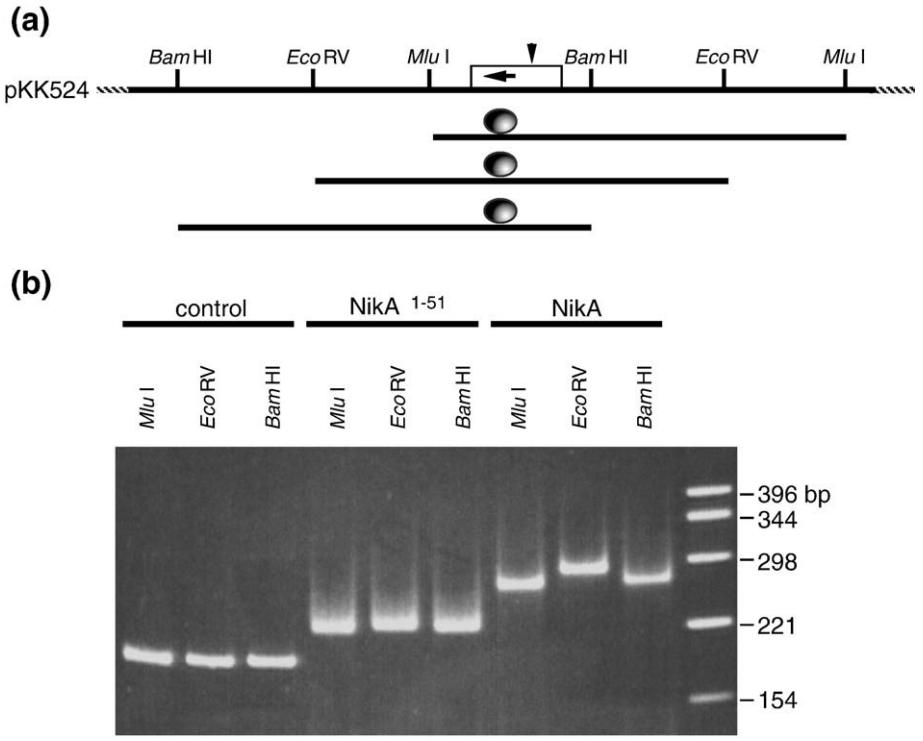

Fig. 6. DNA bending upon binding of full-length NikA or NikA(1–51). (a) DNA structure of the plasmid pKK524. pKK524 was generated by cloning the 44 bp *oriT* core sequence into the DNA-bending vector pBend2. The continuous line at the top represents a restriction map of the region in pKK524 surrounding the cloned fragment. The cloned fragment, shown as an open bar, contains the 44 bp *oriT* core sequence. The horizontal arrow within the box represents the 17 bp repeat A sequence, and the downward arrowhead indicates the nick site. The continuous lines beneath the map represent DNA fragments derived from pKK524 DNA digested with different restriction enzymes. The ovals above the lines represent NikA or NikA(1–51) protein bound to *oriT* DNA, which may induce DNA bending. (b) Gel retardation assay of the DNA fragments. DNA from pKK524 was digested with MluI, EcoRV and BamHI, and then incubated with 50 ng of full length NikA or NikA(1–51) protein at 37 °C. The mixtures were analyzed by 4% PAGE.

alignments with these RHH proteins, which function as transcriptional repressors, suggested that three DNA-binding proteins required for bacterial conjugation may have the RHH fold: TraY,[36] TrwA,[37] and TaxA.[38] Here, we determined for the first time the three-dimensional structure of the N-terminal fragment of NikA, which revealed that NikA, one of the relaxosome-constituting proteins involved in initiating the intercellular transfer of the R64 plasmid during bacterial conjugation, is an RHH protein.

Repeat A, to which NikA binds, and the nick region, YATCCTG | Y, which NikB seems to recognize, are only 1 bp apart (Fig. 1). Even a deletion or insertion of a single basepair between repeat A and the nick region is not permissible.[39] RHH proteins are involved in specific DNA recognition and in protein–protein interactions: (i) The ParG repressor interacts with DNA sequences upstream of the *parFG* genes and with the ParF partition protein, in both the absence and the presence of target DNA.[40] (ii) FitA binds to the 150 bp *fit* promoter sequence containing the translational start site and to the FitB protein.[41] (iii) The MetJ repressor binds to a number of operators with the co-repressor *S*-adenosylmethionine.[28] (iv) The plasmid stabilization protein ParD, which binds to the promoter region by the N-terminal RHH fold domain, also binds to the toxin protein ParE, probably by the C-terminal extension.[42] The finding that NikA is an RHH protein implies that NikA is involved in binding to repeat A, and is involved in binding to the relaxase, NikB, which is unable to bind to the nick region without NikA, and NikA probably recruits NikB by directly contacting NikB.

The co-crystal structures of three RHH proteins (Arc, MetJ and CopG repressor) revealed that each repressor molecule interacts directly with the bases through a pair of antiparallel β-strands inserted into the major groove of *B*-form DNA.[10,31,43] The mutation of the Arg residue on its antiparallel β-strands confirmed that NikA interacts also with the major groove of double-stranded *B*-form DNA in repeat A with the antiparallel β-strands. Most RHH proteins bind as tetramers with two dimeric DNA-binding domains to an operator sequence containing two tandem binding sites, usually arranged as an inverted repeat, which allows a mismatch of a few nucleotides. The center-to-center distance between the tandem operator subsites is typically one turn of the *B*-form DNA helix or less. The existence of this kind of inverted repeat implies that one NikA homotetramer binds to one repeat A.

The binding of NikA to repeat A causes *oriT* DNA bending.[3] The DNA-bending assay of NikA revealed that the C-terminal tetramerization domain is responsible for the DNA bending, which presumably has an important role in the initiation of the DNA transfer. This is a significant feature of NikA, because no N-terminal or C-terminal extension of other RHH proteins has been found to induce DNA bending. NikA binds to repeat A, the right repeat, whereas it does not bind to repeat B, the left inverted repeat, which differs by only one nucleotide from repeat A. This means that one of the two RHH domains of NikA recognizes this difference of a single nucleotide; therefore, one of the two RHH domains of NikA fails to recognize this difference, and has lower affinity than the other. The C-terminal tetramerization domain, which connects the two RHH fold domains, seems to be required to increase the binding affinity and the ability to bend DNA.

A structural model for the mechanism of nicking at the nick site within *oriT* by the auxiliary protein NikA and the relaxase NikB is shown in Fig. 7. Secondary structure predictions and a sequence alignment, based on the structures of NikA, R751 TraJ, RP4 TraJ and pTF-FC2 MobB, implied that the N-terminal RHH fold domain and the two C-terminal α-helices also exist in the latter three proteins. The auxiliary proteins of plasmids with

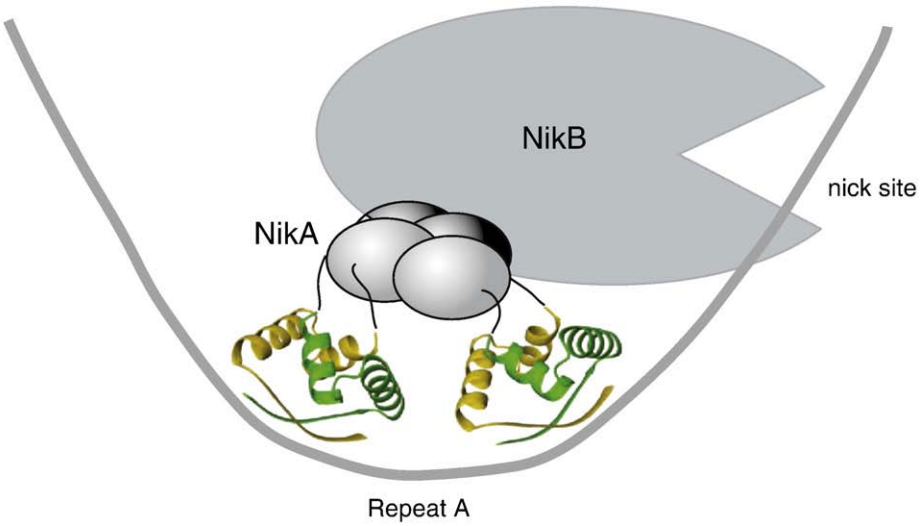

Fig. 7. Model of the action mechanism of NikA in bacterial conjugation. The two RHH fold domains of NikA are shown in green and yellow. The α-helix-rich C-terminal tetrameric region is represented by gray ellipsoids, and NikB is depicted as a large ellipsoid.

the P-type *oriT* generally adopt a homotetrameric RHH domain-containing fold with a C-terminal tetramerization domain.

Materials and Methods

Sample preparation

The NikA, NikA(1–51) and NikA(R18L/R21L) proteins were expressed in *Escherichia coli* BL21 (DE3) Star (Invitrogen) at 37 °C. Unlabeled and uniformly labeled proteins with NMR active stable isotopes were produced by growing the *E. coli* cells in M9 minimal medium. The proteins were purified by chromatography on P11-phosphocellulose (Whatman), MonoS (Amersham Biosciences) and Superdex 75 gel-filtration columns (Amersham Biosciences). NMR samples contained 2 mM NikA (1–51) dissolved in 50 mM sodium phosphate buffer (pH 6.8), containing 50 mM NaCl and 0.02% (w/v) NaN_3.

Gel retardation assay

Purified NikA/NikA(1–51) protein (0.05~0.3 μg) and DNA fragments of the plasmid pKK524[3] containing the *oriT* sequence (0.1~0.3 μg) were mixed in 25 μl of 10 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol, 7.5% (v/v) glycerol. After incubation at room temperature for 10 min, the reaction mixtures were loaded onto a 4% (w/v) polyacrylamide gel (acrylamide/bisacrylamide in a ratio of 29:1. w/w) in E buffer (40 mM Tris–acetate (pH 7.8), 1 mM EDTA) and were electrophoresed at 10 V/cm and 4 °C. Following electrophoresis, the DNA fragments in the gel were stained with ethidium bromide and visualized under long-wave UV light.

Cross-linking

Protein samples were dissolved in 100 mM triethanolamine, pH 8.5, to a final concentration of 0.3 mg/ml. Various concentrations of DMP (10, 50, 200, and 500 mM) in methanol was added to all samples (including controls) to achieve a final concentration of methanol of 10% (v/v). The reactions were incubated at 37 °C, and terminated by adding sodium phosphate buffer (pH 6.8) and 2X SDS gel sample buffer. The samples were heated to 90 °C for 5 min and analyzed by Tris-Tricine SDS-PAGE, which is commonly used to separate proteins in the mass range of 1-100 kDa.[44]

Size-exclusion chromatography

Size-exclusion chromatography was performed on a Superdex 75 HR 10/30 column equilibrated with 100 mM sodium phosphate buffer (pH 6.8), containing 200 mM NaCl. Protein samples were applied at a flow rate of 0.5 ml/min. The column was calibrated using ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), RNase A (14 kDa) and aprotinin (7 kDa) as molecular mass standards.

Analytical ultracentrifugation

Analytical ultracentrifugation experiments were carried out with a Hitachi model CP100α ultracentrifuge (Hitachi Koki Co.). Sedimentation equilibrium experiments were performed at 20 °C. A value of 0.74 ml/mg was assumed for the partial specific volume of the sample.

NMR spectroscopy

All NMR spectra were acquired at 25 °C on Bruker AV500 and DRX600 spectrometers, using triple resonance probes equipped with a pulsed field gradient coil. Standard heteronuclear multidimensional NMR experiments were performed: ^{1}H-^{15}N HSQC, HNCA, HN(CO)CA, HNCACB, and CBCA(CO)NH for the main-chain assignment, and ^{1}H-^{13}C CT-HSQC, HBHA(CO)NH, C(CCO)NH, H(CCCO)NH, HCCH-TOCSY, HCCH-COSY, and ^{15}N-edited TOCSY for the side-chain assignment.[45,46] Distance information was obtained from ^{15}N-edited NOESY and ^{13}C-edited NOESY spectra with a mixing time of 100 ms. Filtered NOESY experiments to distinguish between the intra- and intermolecular NOEs were not performed. The ^{1}H chemical shifts were referenced to the external standard 2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS). The ^{13}C and ^{15}N chemical shifts were referenced indirectly from DSS by using the frequency ratios 0.251449519 and 0.101329112 for ^{13}C/^{1}H and ^{15}N/^{1}H, respectively.[47] ^{15}N T_1, ^{15}N T_2 relaxation times and ^{1}H-^{15}N NOE data were measured as described.[48] Relaxation delays of up to 1 s for T_1 and 80 ms for T_2 measurements were used. Relaxation curves were characterized by 5 – 12 points. The long-range $^{3h}J_{NC'}$ HNCO experiment was carried out with the transfer time $2T$ set to 133 ms $\approx 2/^1J_{NC'}$ to identify unambiguously the donor and acceptor atoms involved in intermolecular hydrogen bonds in NikA(1–51).[19] All NMR data were processed using NMRPipe/NMRDraw,[49] and were analyzed using NMRView,[50] and Sparky†.[51]

Structure calculation

The automated NOE assignment[52] and the structure calculation[53] were performed with a modified version of the program CYANA 2.1, which takes the homodimer symmetry explicitly into account for the network-anchoring of the NOE assignments, ensures an identical conformation of the two monomers by dihedral angle difference restraints for all corresponding torsion angles, and maintains a symmetric relative orientation of the two monomers by distance difference restraints between symmetry-related intermolecular C^α–C^α distances.[17] The α-helices and the β-strand were identified from chemical shift information,[12] and intramolecular NOE assignments spanning more than five residues within one α-helix, or more than two residues within one β-strand were excluded in the first two cycles of automated NOE assignment and structure calculation with CYANA. Structure calculations were started from 100 starting conformers with random torsion angle values, and used 20,000 torsion angle dynamics steps per conformer.[53] The 20 CYANA conformers with the lowest final target function values were subjected to restrained energy minimization in explicit solvent against the AMBER force field,[54] using the program OPALp.[55] Structure figures were generated with the program MOLMOL.[56]

† http://www.cgl.ucsf.edu/home/sparky/

DNA-binding assay of NikA mutants

The positively charged residues R18 and R21 in the β-sheet region of full-length NikA were separately changed to Leu, using a QuikChange site-directed mutagenesis kit (Stratagene). The presence of the appropriate mutation was confirmed by DNA sequencing. The mutant protein was expressed and purified with the protocols used for wild type NikA. The DNA-binding activity of the mutant NikA was measured by a gel mobility-shift assay.

Data Bank accession codes

The coordinates of the 20 energy-refined CYANA conformers of NikA(1–51) have been deposited in the Protein Data Bank with accession code 2BA3. The chemical shifts of NikA(1–51) have been deposited in the BioMagResBank with accession code 15784.

Acknowledgements

We thank Dr Takuya Torizawa for technical advice about NMR spectroscopy. This study was supported financially by the National Project on Protein Structural and Functional Analyses and the Technology Development for Protein Analyses and Targeted Protein Research Program of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), the Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Agency (JST), a Grant-in-Aid for Scientific Research of the Japan Society for the Promotion of Science (JSPS), and by the Volkswagen Foundation.

References

1. Clewell, D. B. (1993). *Bacterial Conjugation.* Plenum Press, New York, NY.
2. Lanka, E. & Wilkins, B. M. (1995). DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* **64**, 141–169.
3. Furuya, N. & Komano, T. (1995). Specific binding of the NikA protein to one arm of 17-base-pair inverted repeat sequences within the oriT region of plasmid R64. *J. Bacteriol.* **177**, 46–51.
4. Furuya, N., Nisioka, T. & Komano, T. (1991). Nucleotide sequence and functions of the oriT operon in IncI1 plasmid R64. *J. Bacteriol.* **173**, 2231–2237.
5. Pansegrau, W. & Lanka, E. (1991). Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. *Nucleic Acids Res.* **19**, 3455.
6. Pansegrau, W., Schröder, W. & Lanka, E. (1994). Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J. Biol. Chem.* **269**, 2782–2789.
7. Furuya, N. & Komano, T. (2000). Initiation and termination of DNA transfer during conjugation of IncI1 plasmid R64: roles of two sets of inverted repeat sequences within oriT in termination of R64 transfer. *J. Bacteriol.* **182**, 3191–3196.
8. Ziegelin, G., Fürste, J. P. & Lanka, E. (1989). TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. *J. Biol. Chem.* **264**, 11989–11994.
9. Breg, J. N., van Opheusden, J. H., Burgering, M. J., Boelens, R. & Kaptein, R. (1990). Structure of Arc repressor in solution: evidence for a family of beta-sheet DNA-binding proteins. *Nature*, **346**, 586–589.
10. Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. (1994). DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature*, **367**, 754–757.
11. Nooren, I. M., Rietveld, A. W., Melacini, G., Sauer, R. T., Kaptein, R. & Boelens, R. (1999). The solution structure and dynamics of an Arc repressor mutant reveal premelting conformational changes related to DNA binding. *Biochemistry*, **38**, 6035–6042.
12. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
13. Fushman, D., Cahill, S. & Cowburn, D. (1997). The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of ^{15}N relaxation with monomer/dimer equilibration. *J. Mol. Biol.* **266**, 173–194.
14. Renzoni, D., Esposito, D., Pfuhl, M., Hinton, J. C., Higgins, C. F., Driscoll, P. C. & Ladbury, J. E. (2001). Structural characterization of the N-terminal oligomerization domain of the bacterial chromatin-structuring protein, H-NS. *J. Mol. Biol.* **306**, 1127–1137.
15. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids.* Wiley, New York.
16. Schubert, M., Labudde, D., Oschkinat, H. & Schmieder, P. (2002). A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. *J. Biomol. NMR*, **24**, 149–154.
17. Nilges, M. (1993). A calculation strategy for the structure determination of symmetrical dimers by ^{1}H-NMR. *Proteins: Struct. Funct. Genet.* **17**, 297–309.
18. Otting, G. & Wüthrich, K. (1990). Heteronuclear filters in two-dimensional [^{1}H,^{1}H]-NMR spectroscopy: combined use with isotope labelling for studies of macromolecular conformation and intermolecular interactions. *Q. Rev. Biophys.* **23**, 39–96.
19. Cordier, F. & Grzesiek, S. (1999). Direct observation of hydrogen bonds in proteins by interresidue $^{3h}J_{NC'}$ scalar couplings. *J. Am. Chem. Soc.* **121**, 1601–1602.
20. Cutler, R. L., Davies, A. M., Creighton, S., Warshel, A., Moore, G. R., Smith, M. & Mauk, A. G. (1989). Role of arginine-38 in regulation of the cytochrome c oxidation-reduction equilibrium. *Biochemistry*, **28**, 3188–3197.
21. Dowd, D. R. & Lloyd, R. S. (1989). Site-directed mutagenesis of the T4 endonuclease V gene: the role of arginine-3 in the target search. *Biochemistry*, **28**, 8699–8705.
22. He, J. J. & Matthews, K. S. (1990). Effect of amino acid alterations in the tryptophan-binding site of the trp repressor. *J. Biol. Chem.* **265**, 731–737.
23. Lin, L., Perryman, M. B., Friedman, D., Roberts, R. & Ma, T. S. (1994). Determination of the catalytic site of creatine kinase by site-directed mutagenesis. *Biochim. Biophys. Acta*, **1206**, 97–104.
24. Smulevich, G., Mauro, J. M., Fishel, L. A., English, A. M., Kraut, J. & Spiro, T. G. (1988). Heme pocket interactions in cytochrome c peroxidase studied by site-directed mutagenesis and resonance Raman spectroscopy. *Biochemistry*, **27**, 5477–5485.

25. Smulevich, G., Mauro, J. M., Fishel, L. A., English, A. M., Kraut, J. & Spiro, T. G. (1988). Cytochrome c peroxidase mutant active site structures probed by resonance Raman and infrared signatures of the CO adducts. *Biochemistry*, **27**, 5486–5492.
26. Smulevich, G., Miller, M. A., Kraut, J. & Spiro, T. G. (1991). Conformational change and histidine control of heme chemistry in cytochrome c peroxidase: resonance Raman evidence from Leu-52 and Gly-181 mutants of cytochrome c peroxidase. *Biochemistry*, **30**, 9546–9558.
27. Burgering, M. J., Boelens, R., Gilbert, D. E., Breg, J. N., Knight, K. L., Sauer, R. T. & Kaptein, R. (1994). Solution structure of dimeric Mnt repressor (1-76). *Biochemistry*, **33**, 15036–15045.
28. Rafferty, J. B., Somers, W. S., Saint-Girons, I. & Phillips, S. E. (1989). Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. *Nature*, **341**, 705–710.
29. Somers, W. S. & Phillips, S. E. (1992). Crystal structure of the met repressor-operator complex at 2.8 A resolution reveals DNA recognition by beta-strands. *Nature*, **359**, 387–393.
30. Gomis-Ruth, F. X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R. *et al.* (1998). The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J.* **17**, 7404–7415.
31. Gomis-Ruth, F. X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R. *et al.* (1998). The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J.* **17**, 7404–7415.
32. Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C. & Saenger, W. (2001). Crystal structure of omega transcriptional repressor encoded by Streptococcus pyogenes plasmid pSM19035 at 1.5 Å resolution. *J. Mol. Biol.* **314**, 789–796.
33. Golovanov, A. P., Barilla, D., Golovanova, M., Hayes, F. & Lian, L. Y. (2003). ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol. Microbiol.* **50**, 1141–1153.
34. Chivers, P. T. & Tahirov, T. H. (2005). Structure of *Pyrococcus horikoshii* NikR: nickel sensing and implications for the regulation of DNA recognition. *J. Mol. Biol.* **348**, 597–607.
35. Schreiter, E. R., Sintchak, M. D., Guo, Y., Chivers, P. T., Sauer, R. T. & Drennan, C. L. (2003). Crystal structure of the nickel-responsive transcription factor NikR. *Nature Struct. Biol.* **10**, 794–799.
36. Bowie, J. U. & Sauer, R. T. (1990). TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J. Mol. Biol.* **211**, 5–6.
37. Moncalian, G. & de la Cruz, F. (2004). DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim. Biophys. Acta*, **1701**, 15–23.
38. Núñez, B., Avila, P. & de la Cruz, F. (1997). Genes involved in conjugative DNA processing of plasmid R6K. *Mol. Microbiol.* **24**, 1157–1168.
39. Furuya, N. & Komano, T. (1997). Mutational analysis of the R64 *oriT* region: requirement for precise location of the NikA-binding sequence. *J. Bacteriol.* **179**, 7291–7297.
40. Barillà, D. & Hayes, F. (2003). Architecture of the ParF∗ParG protein complex involved in prokaryotic DNA segregation. *Mol. Microbiol.* **49**, 487–499.
41. Wilbur, J. S., Chivers, P. T., Mattison, K., Potter, L., Brennan, R. G. & So, M. (2005). *Neisseria gonorrhoeae* FitA interacts with FitB to bind DNA through its ribbon-helix-helix motif. *Biochemistry*, **44**, 12515–12524.
42. Johnson, E. P., Strom, A. R. & Helinski, D. R. (1996). Plasmid RK2 toxin protein ParE: purification and interaction with the ParD antitoxin protein. *J. Bacteriol.* **178**, 1420–1429.
43. Somers, W. S., Rafferty, J. B., Phillips, K., Strathdee, S., He, Y. Y., McNally, T. *et al.* (1994). The Met repressor-operator complex: DNA recognition by beta-strands. *Ann. N. Y. Acad. Sci.* **726**, 105–117.
44. Schägger, H. (2006). Tricine-SDS-PAGE. *Nature Protocols*, **1**, 16–22.
45. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R. & Zhu, G. (1994). Measurement of homo- and heteronuclear *J* couplings from quantitative *J* correlation. *Methods Enzymol.* **239**, 79–105.
46. Kay, L. E. (1997). NMR methods for the study of protein structure and dynamics. *Biochem. Cell Biol.* **75**, 1–15.
47. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E. *et al.* (1995). ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, **6**, 135–140.
48. Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972–8979.
49. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe - A multidimensional spectral processing system based on Unix pipes. *J. Biomol. NMR*, **6**, 277–293.
50. Johnson, B. A. & Blevins, R. A. (1994). NMR View - a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
51. Goddard, T. D. & Kneller, D. G. (2001). *Sparky 3.* University of California, San Francisco.
52. Herrmann, T., Güntert, P. & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227.
53. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
54. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M. *et al.* (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
55. Koradi, R., Billeter, M., Engeli, M., Güntert, P. & Wüthrich, K. (1998). Automated peak picking and peak integration in macromolecular NMR spectra using AUTOPSY. *J. Magn. Reson.* **135**, 288–297.
56. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
57. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.