Protein Science (1999), *8*:2494–2500. Cambridge University Press. Printed in the USA.

NikR is a ribbon-helix-helix DNA-binding protein

PETER T. CHIVERS AND ROBERT T. SAUER

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(RECEIVED April 15, 1999; ACCEPTED August 13, 1999)

Abstract

Escherichia coli NikR, a repressor with homologs in other bacteria and archaea, was identified as a potential new member of the ribbon-helix-helix (β-α-α) family of transcription factors in profile based sequence searches and in structure prediction experiments. Biophysical and biochemical characterization of the N-terminal domain of NikR show that it has many features expected of a β-α-α protein including α-helical content, dimeric solution form, concentration dependent thermal stability, and ability to bind DNA in sequence-specific manner. Mutation of a residue predicted to be important for DNA-binding reduces operator affinity but does not affect the secondary structure or stability of the protein.

Keywords: nickel permease; protein folding; repressor; sequence homology; sequence profile; structure prediction

The ribbon-helix-helix (β-α-α) family of DNA-binding proteins includes four prokaryotic repressors of known structure—Arc, CopG, MetJ, and Mnt—as well as the TraY protein family and relatives of CopG and MetJ (Bowie & Sauer, 1990; Burgering et al., 1994; Phillips, 1994; Raumann et al., 1994b; Gomis-Rüth et al., 1998). The ribbon-helix-helix portions of these proteins are 45–50 residues in length and form dimers in which the subunits intertwine (Fig. 1A). Each monomer contains a β-strand and two α-helices, with the strands pairing to form an antiparallel β-sheet in the dimer. Alignment of these sequences reveals conservation at the level of sequence pattern but relatively little sequence identity, typically about 15%. At this level of identity, it is very difficult to find new family members using standard sequence homology searches because false positives outnumber true positive protein pairs by ~10-fold (Rost, 1999).

In the cocrystal structures of Arc, CopG, and MetJ, each ribbon-helix-helix dimer binds to a DNA site of 8–10 bp, and residues from the β-sheet mediate most of the sequence specific contacts with DNA bases in the major groove (Somers & Phillips, 1992; Raumann et al., 1994b; Gomis-Rüth et al., 1998). Dimers can bind cooperatively to adjacent DNA subsites, although the regions of each protein used for these cooperative protein–protein interactions are different. MetJ and Mnt also contain C-terminal domains in addition to their N-terminal β-α-α domains. In MetJ, this domain mediates binding of the corepressor, S-adenosyl methionine (Somers & Phillips, 1992; Phillips & Phillips, 1994). In Mnt, the C-terminal region forms a coiled-coil tetramerization domain (Nooren, 1999). The folding of Arc dimers has been studied in detail (Bowie & Sauer, 1989a; Milla & Sauer, 1994; Jonsson et al., 1996). Monomers do not fold stably, as might be expected since both β-α-α subunits contribute to a single hydrophobic core (Bonvin et al., 1994; Raumann et al., 1994b). Hence, the stability of Arc is concentration dependent. Extensive mutagenesis of Arc has also been performed, allowing determination of which residues are important for stability or function and which are unessential (Vershon et al., 1986; Bowie & Sauer, 1989b; Brown et al., 1994; Milla et al., 1994; Smith & Sauer, 1995; Brown & Sauer, 1999).

Here, we report profile searches that identify the N-terminal domain of NikR repressor as a potential member of the ribbon-helix-helix family and demonstrate that the biochemical and biophysical properties of this NikR domain are consistent with a β-α-α fold. Finally, we show that the N-terminal domain of NikR binds to a DNA site near the promoter of the *nikABCDE* operon and also responds to mutations in the putative β-sheet as expected for a ribbon-helix-helix protein.

Results and discussion

Informatics

To search for potential new members of the ribbon-helix-helix family, a sequence profile method was used (Lüthy et al., 1994). A profile containing eight known family members including Arc, Mnt, CopG, MetJ, and TraY was constructed and a search of the August 1998 SWISSPROT/TrEMBL (Release 36) database was performed (see Materials and methods). The highest scoring matches, excluding profile sequences and their relatives, are listed in Table 1. Sequence alignments are shown in Figure 1B. Most of the high scoring proteins are encoded by open reading frames with no known function but one sequence—encoded by *nikR* (née *yhhG*)—was recently identified in a genetic screen for *Escherichia coli* mutants that fail to repress nickel permease activity (de Pina

Reprint requests to: Robert Sauer, 68-571, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139; e-mail: bobsauer@mit.edu.

Table 1. *High-scoring sequences from profile search*

ORF/protein	Organism	SWISSPROT ID	Score	Residues aligned[a]
Y420	*Haemophilus influenzae*	P43995	299	7–53 (99)
NikR[b]	*E. coli*	P28910	298	1–47 (133)
YbfE	*E. coli*	P75735	294	74–120 (120)
E-51	*SSV1*	P20217	254	11–51 (51)
Y767	*Methanococcus jannaschii*	Q58177	237	5–51 (135)

[a]Number in parenthesis is total length of protein.
[b]Also known as YhhG.

et al., 1999). A BLAST (tblastn) search of the unfinished genomes database revealed NikR homologs in three other bacterial species and in all five archaea for which genome sequences are available (Fig. 2). The N-terminal region of each NikR protein aligns with the sequence profile (Fig. 1) and shows the signature polar-hydrophobic patterning observed for the β-strand and α-helical regions of the other β-α-α proteins (Figs. 1, 2). Moreover, the highly conserved positions in the sequence alignment of NikR N-terminal domains generally correspond to positions in Arc repressor where single alanine mutations disrupt folding or function, and nonconserved NikR positions correspond to positions in Arc where mutations are silent (see Fig. 2). The NikR proteins also contain homologous C-terminal sequences of ~80 residues with a unique His-X_{13}-His-X_{10}-His-X-His-X_5-Cys motif that probably forms a nickel binding site (Fig. 2, bottom).

The N-terminal domain sequences of five NikR proteins were submitted for structure prediction by the method of Fischer and Eisenberg (1996) (see Materials and methods). The Arc structure scored highest against four of the five NikR N-terminal domain sequences and was the second highest hit for the remaining sequence, although the statistical significance of these scores was borderline (Table 2). Nevertheless, both profile searches and structure prediction suggest that NikR is a member of the ribbon-helix-helix protein family.

Biochemical and biophysical characterization

To test experimentally whether NikR has properties expected for a ribbon-helix-helix protein, we cloned, expressed, and purified the NikR protein from *E. coli*. If the N-terminal region of NikR does adopt the β-α-α fold, then it would be expected to fold independently as a dimer and to have substantial α-helical secondary structure. Digestion of full-length NikR with elastase produced a stable N-terminal fragment with a mass (6,632 Da) close to that expected

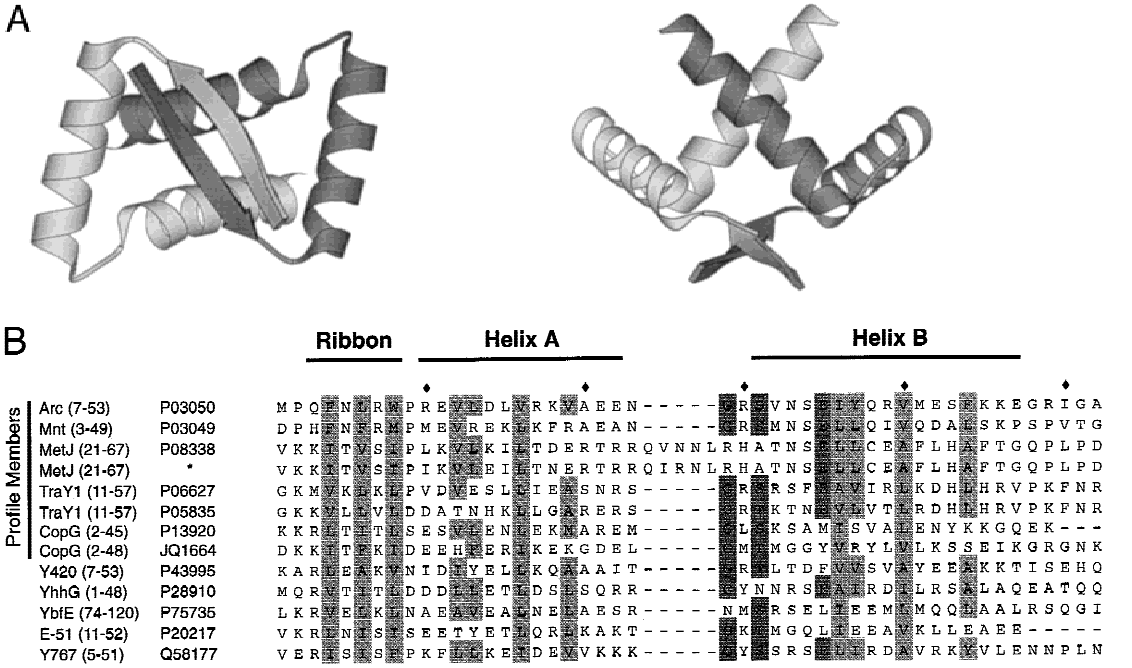

Fig. 1. Ribbon-helix-helix fold and sequence alignments. **A:** Two views of the Arc repressor dimer (Raumann et al., 1994b) generated using MOLSCRIPT (Kraulis, 1991). **B:** Alignment of known (top eight) and putative (bottom five) ribbon-helix-helix proteins. Conserved hydrophobic residues are shaded in light gray. Other conserved residues are shaded in dark gray. Every tenth residue is indicated by (♦). *, MetJ sequence from *Actinobacillus actinomycetemcomitans* from unfinished genomes database (see Materials and methods).

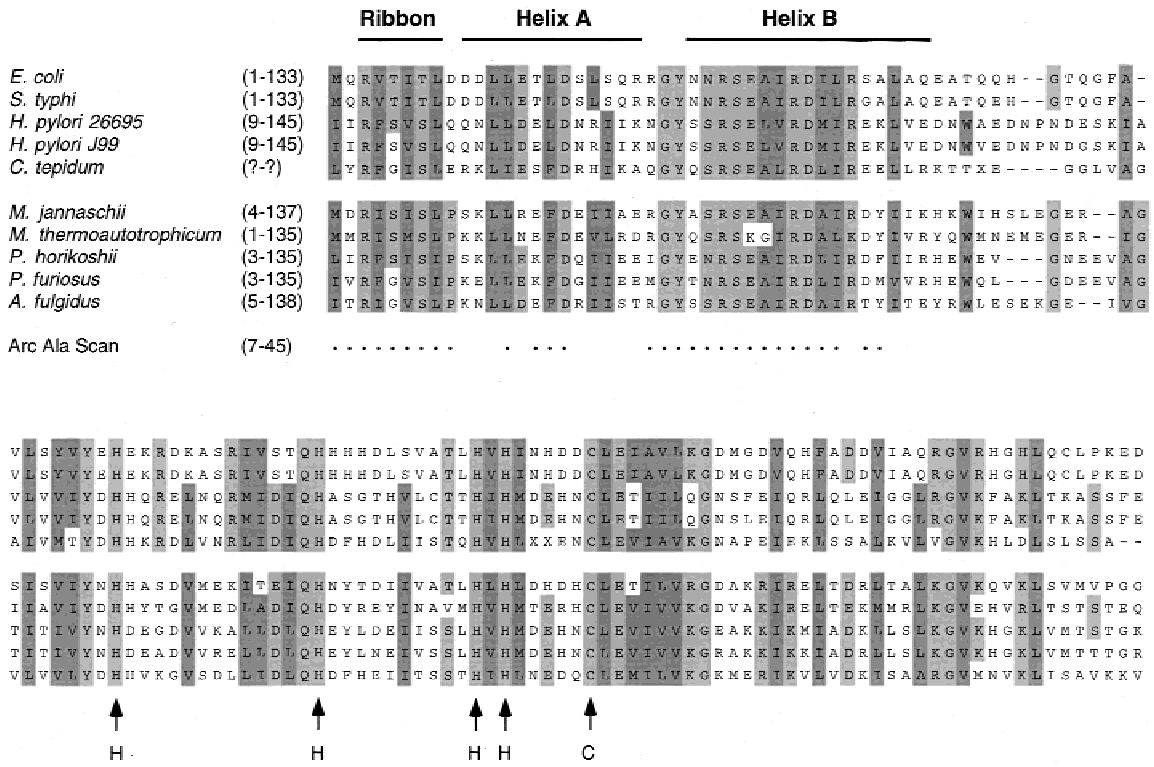

Fig. 2. Alignment of NikR family identified in BLAST (tblastn) searches (Altschul et al., 1997). Homologous positions in Arc where alanine substitutions disrupt folding of function (Brown et al., 1994; Milla et al., 1994) are marked by (●) underneath the N-terminal domain portion of the sequence. Putative nickel-binding residues in the C-terminal region are identified by (↑). Every tenth residue is indicated by (♦).

for a fragment extending from residues Met1 to Phe53. The corresponding C-terminal domain fragment of NikR (residues 54–138) also appeared to be produced by limited proteolysis but was observed only at early time points. Residues 49 to 55 of the *E. coli* NikR sequence probably form a flexible interdomain linker since this region is relatively glycine-rich and contains gaps in the alignment (Fig. 2). Therefore, a slightly shorter fragment consisting of *E. coli* NikR residues 1 to 48 with a C-terminal 6His-tag was cloned, expressed, and purified.

Table 2. *Structure prediction*[a] *results for NikR N-terminal domains*

NikR	Z-score[b]
E. coli (1–48)[c]	Arc 3.13 (1), —
M. jannaschii (4–51)	Arc 3.63 (2), Mnt 2.64 (8)
Pyrococcus horikoshii (3–50)	Arc 4.71 (1), Mnt 3.55 (2)
Helicobacter pylori (10–57)	Arc 3.97 (1), Mnt 2.38 (9)
Archaeoglobus fulgidus (5–52)	Arc 3.85 (1), —

[a] Structure prediction used the method of Fischer and Eisenberg (1996) as described in the text.

[b] Z-scores for 1myk [Arc PL8 (Schildbach et al., 1995)] and 1mnt, (Burgering et al., 1994). Mnt scores are not listed if they fell below the top 15 hits. The numbers in parenthesis are the ranking of the score.

[c] Numbers in parenthesis indicate N-terminal domain residues used for structure prediction.

The 1–48 N-terminal fragment of *E. coli* NikR had a circular dichroism (CD) spectrum characteristic of a protein with 56% α-helix (Fig. 3A). A value of 55% would be expected based on the proposed β-α-α secondary structure. In equilibrium analytical ultracentrifugation, the N-terminal domain sedimented with a mean molecular weight of 11.5 kDa, close to the value expected for a dimer of 6.4 kDa subunits (Fig. 4). The secondary structure of the NikR N-terminal domain was lost in a cooperative fashion in both thermal and GuHCl denaturation (Fig. 3B,C), as expected for a stably folded native protein. Moreover, as expected for a reaction in which native dimers and denatured monomers are in equilibrium, the thermal stability of N-terminal domain was concentration dependent (Fig. 3B). Fitting of the GuHCl denaturation data gave an equilibrium constant of 9.9×10^{-11} M for dissociation of the N-terminal domain dimer to monomers (Table 3). These results support the model that the N-terminal domain of *E. coli* NikR adopts a homodimeric β-α-α fold.

DNA binding of wild-type and mutant domains

NikR's function as a genetic repressor suggests that it is likely to be a sequence specific DNA-binding protein (de Pina et al., 1999) and the identification of the N-terminal region as having a ribbon-helix-helix fold predicts that this domain should mediate DNA recognition. To test this model, DNase footprinting was used to assay binding of the N-terminal domain of *E. coli* NikR to a DNA fragment containing the *nikABCDE* promoter region. As shown in Figure 5, the wild-type N-terminal domain protects many of the

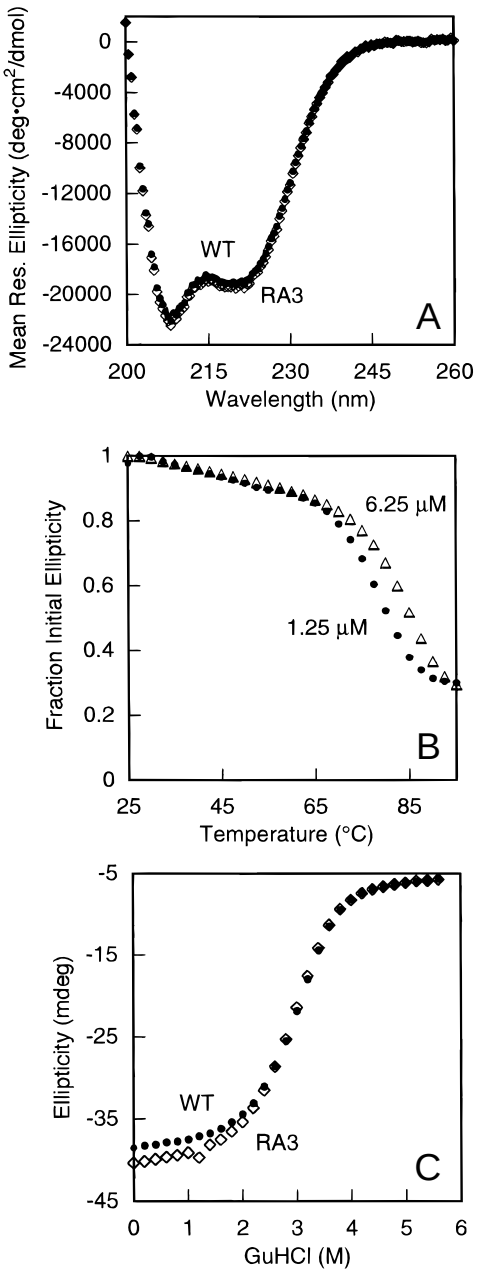

Fig. 3. Biophysical properties of NikR N-terminal domain. **A:** Far-UV CD spectrum; filled circles, 1.25 μM WT N-terminal domain; open diamonds, 1.25 μM RA3 N-terminal domain. **B:** Thermal stability; filled circles, 1.25 μM WT N-terminal domain; open triangles, 6.25 μM WT N-terminal domain. **C:** GuHCl stability; filled circles, 6.25 μM WT N-terminal domain; open diamonds, 6.25 μM RA3 N-terminal domain. Conditions for each experiment are described in Materials and methods.

bases in a region extending from bases −31 to +2 and also mediates enhancements of some bases outside of this region. These results confirm that the N-terminal domain of NikR is responsible for DNA recognition. In preliminary experiments, we have also found that a subset of mutations within the DNA region protected by the N-terminal domain reduce the affinity of NikR for the operator (data not shown).

Based on structural and mutagenic studies in Arc (Brown et al., 1994), Mnt (Knight & Sauer, 1989), and MetJ (He et al., 1992),

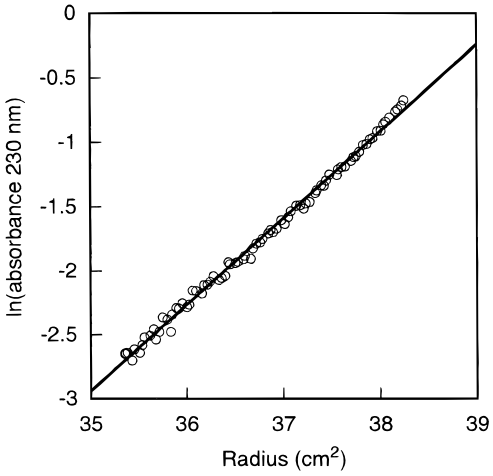

Fig. 4. Analytical ultracentrifugation of 6.25 μM NikR N-terminal domain, 31,000 rpm, 25 °C. The fitted line corresponds to a molecular weight of 11,550 Da. The expected molecular weight of a dimer of the N-terminal domain 1–48 fragment with the -LEHHHHHH tail is 13,200 Da.

another prediction of the β-α-α model for the N-terminal domain of NikR is that a mutation at the first surface position of the presumed β-sheet should interfere with DNA-binding but not affect protein stability. To test this we constructed and purified the Arg3 → Ala (RA3) mutant of the NikR N-terminal domain and assayed protein stability and DNA protection. As shown in Figures 3 and 5, when compared with wild-type protein, the RA3 mutation does not substantially alter the CD spectrum or stability of the mutant protein but does eliminate DNA binding.

The β-α-α superfamily

The prevalence of different superfamilies of transcription factors is one indication of their relative "evolutionary success" (Pabo & Sauer, 1992). Helix-turn-helix proteins probably constitute the dominant superfamily of bacterial repressor and activator proteins. By contrast, the ribbon-helix-helix group of DNA-binding proteins has been rather sparsely populated (Raumann et al., 1994a). However, the addition of the CopG and NikR families to the β-α-α class has more than doubled the number of known superfamily members in the past year alone. Many of the nonNikR sequences listed in Table 1 may also prove to be superfamily members, and there may be more examples to discover. NikR homologs are present in archaea as well as eubacteria (Fig. 2). At present, there are no validated ribbon-helix-helix proteins in eukaryotes, although a family of plant regulatory proteins does share some of the sequence features of this family (Kim et al., 1997).

Role of NikR

The *nikABCDE* operon of *E. coli* encodes proteins that appear to assemble to form an ABC-type periplasmic transporter for nickel (Navarro et al., 1993; Tam & Saier, 1993). The *nikR* gene is placed at the end of this operon but is expressed from its own promoter (de Pina et al., 1999). Genetic evidence indicates that the primary function of NikR is to repress transcription of the *nikABCDE* genes when nickel is present at concentrations of 250 μM or more

Table 3. *Stabilities of wild-type and RA3 NikR N-terminal domain dimers*[a]

Protein	T_m (°C)	C_m (M)	m (kcal/mol·M)	ΔG_u (kcal/mol)	K_u (M)
Wild-type	76.9	2.97	2.20	13.59	9.93×10^{-11}
RA3	78.9	2.90	2.11	13.20	1.92×10^{-10}

[a]Protein solutions of 1.26 and 6.25 μM were used for thermal and chemical denaturation, respectively.

(Navarro et al., 1993; de Pina et al., 1999). *E. coli* requires nickel as a cofactor for hydrogenase (Hausinger, 1987), an essential enzyme for anaerobic growth, but excessive uptake by the cell is probably toxic. It seems likely that NikR senses high concentrations of nickel through interactions of this metal with its C-terminal domain. Indeed, a fragment corresponding to the C-terminal domain of *E. coli* NikR stably folds in solution, is a monomer, and binds nickel (P.T. Chivers & R.T. Sauer, unpubl. results).

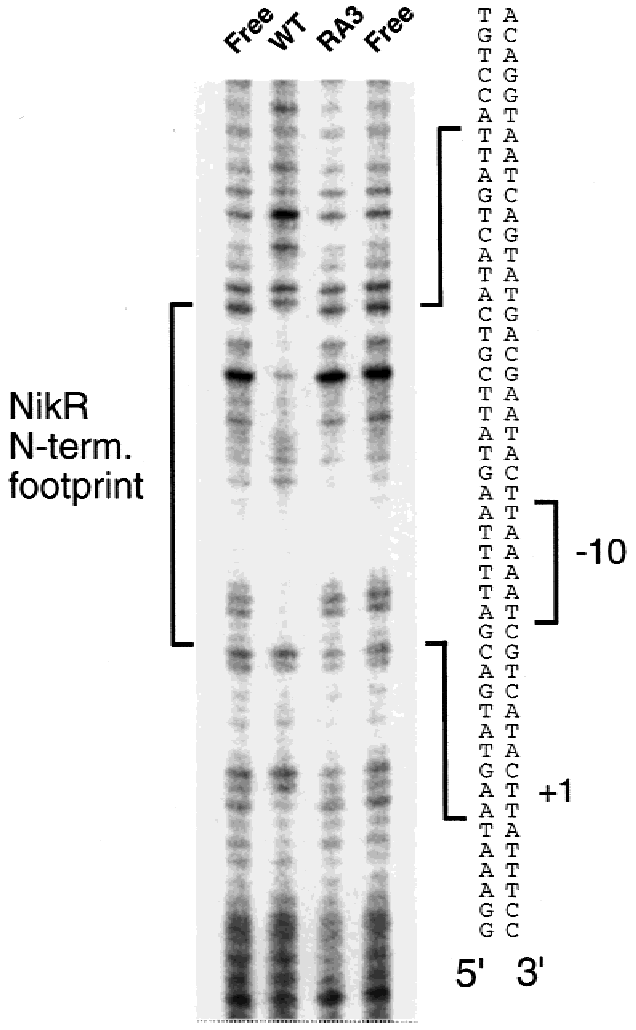

Fig. 5. Dnase I footprinting of the *nik*ABCDE promoter region by 6.25 μM of the wild-type or RA3 N-terminal domain of NikR. The region protected by the wild-type N-terminal domain is outlined, and the corresponding sequence is shown to the right side (the left strand was labeled with γ^{32}P-ATP at the 5′ end). The putative −10 region (Navarro et al., 1993) and start point of transcription are also indicated.

Validating a β-α-α fold

The *β-α-α* motif is only 45–50 residues in length and contains no residues that are absolutely conserved. The pairwise identity between sequences of members of different subfamilies is generally only 10 to 15%, with only a few cases in which identities are >20% (Table 4). Consequently, it is easy both to miss new family members and to predict family members wrongly. For these reasons, it is important to use multiple methods to confirm that a suspected relative is really a bona fide member of the family. In the case of the N-terminal domain of NikR, there are a number of indicators in addition to sequence similarity that support its assignment to the ribbon-helix-helix family: the N-terminal domain of NikR matches best to Arc, a *β-α-α* protein, in structure prediction experiments; this domain is stable to limited proteolysis and has a CD spectrum expected for a protein with the same *α*-helical content as Arc; the native N-terminal domain, like Arc, is dimeric and shows concentration dependent stability; and mutation of a predicted DNA-binding residue reduces DNA binding but does not affect protein secondary structure or stability. This combination of properties suggests strongly that NikR has a *β-α-α* fold. We also note that the pattern of sequence conservation among different NikR subfamily members matches closely to the patterns of structurally and functionally important residues that have been established by mutagenesis of Arc repressor. Finally, NikR is a transcriptional repressor and DNA-binding protein and thus shares common functions with other members of the *β-α-α* superfamily. It will obviously be important to obtain direct structural informa-

Table 4. *Percentage pairwise sequence identity between ribbon-helix-helix subfamily members*[a]

		Arc	Mnt	MetJ	TraY	CopG	NikR
Arc	P03050[b]	—	**37**	11	17	17	13
Mnt	P03049			17	15	13	11
MetJ	P08338				15	11	17
TraY	P06627					15	19
CopG	P13920						**22**
NikR	P28910						—

[a]Pairwise identity of the 46 residue alignment in Figure 1. Pairwise identity scores greater than 20% are shown in bold.
[b]SWISSPROT ID number.

tion for the N-terminal domain of NikR. In this regard, crystals of this domain have been grown that diffract to ≈3.5 Å, which with some improvement should be suitable for crystallographic analysis.

Materials and methods

Searches and structure prediction

Database profile searches (Lüthy et al., 1994) were performed using Version 2.0 of PFTOOLS (obtained from http://www.isrec.isb-sib.ch/profile/profile.html) on an SGI INDIGO 2 workstation. The PFW and PFMAKE programs were used to construct the profile matrix using the β-α-α portions of the top eight Arc, Mnt, MetJ, CopG, and TraY sequences shown in Figure 1. Database searches were performed on the August 1998 update (Release 36) of the SWISSPROT/TrEMBL database (Bairoch & Apweiler, 1998) using the PFSEARCH program. A threshold score of 100 was used. NikR homologs were identified by BLAST searches (Altschul et al., 1997) of the unfinished genomes database (http://www.ncbi.nlm.nih.gov/BLAST/unfinishedgenome.html). Searches were performed using the tblastn option and the following settings: expect=10; filter=default; descriptions=100; and alignments=100.

Structure prediction employing the method of Fischer and Eisenberg was carried out using the DOE-UCLA website service (http://www.doe-mbi.ucla.edu/people/frsvr/frsvr.html). The N-terminal domain sequences of five NikR proteins were submitted for structure prediction with the following options: H3P2=no; TOPITS=no; HTML=yes; and PROFILESEARCH=no. The structural database used in this search contained structures for Mnt, MetJ, and two variants of Arc repressor, PL8 and MYL-Arc.

Cloning and protein expression

The *nikR* gene was cloned by PCR from genomic DNA of *E. coli* strain MC1061 using oligonucleotide primers PC100 (5′-G TTA ACG CAT ATG CAA CGA GTC ACC ATC ACG-3′) and PC101 (5′-ATA ATA CTC GAG ATC TTC CTT CGG CAA GCA CTG-3′). The PCR fragment was digested with *Nde*I and *Xho*I and ligated into pET-22b (Novagen, Madison, Wisconsin), which had been cut with the same enzymes, to generate pNIK100. Protein expressed from this construct has a C-terminal extension (-LEHHHHHH) to aid in protein purification. DNA encoding the NikR N-terminal domain (residues 1–48) was obtained by PCR of the full-length gene using primers PC100 and PC105 (5′-ACC TTG CTC GAG GTG CTG CTG GGT GGC C-3′) and was cloned into pET-22b to generate pNIK101. A plasmid encoding the RA3 mutant of the N-terminal domain (pNIK101 RA3) was generated in an analogous fashion using primers PC101 and PC129 (5′-A GAT ATA CAT ATG CAA **GCG** GTC ACC ATC ACG CTT GAT GAC-3′). Constructs were verified by dideoxynucleotide sequencing using Sequenase Version 2.0 DNA Sequencing Kit (United States Biochemical, Cleveland, Ohio). Oligonucleotides were synthesized by Integrated DNA Technologies, Inc. (Coralville, Iowa).

For protein purification, *E. coli* strain BB101 (*F′ lacI^q lac$^+$ pro$^+$/ ara Δ(lac-pro) nal1 argEam rif thi-1 slyD λDE3*) transformed with the appropriate plasmid was grown in 0.5 L of LB broth to an OD_{600} of 1, and expression was induced by addition of IPTG to a final concentration of 0.5 mM. Cells were harvested 3 h after induction. The cell pellet was resuspended in 50 mL of 100 mM potassium phosphate, 10 mM Tris (pH 8.0), and 6 M GuHCl, and lysis occurred during gentle shaking for 30 min at 37 °C. The lysate was centrifuged, and the supernatant was loaded on a 3 mL column of Ni–NTA resin (Qiagen Inc., Valencia, California) that had been pre-equilibrated with five volumes of lysate buffer. After loading, the column was washed with 25 volumes of lysis buffer and eluted with three volumes of 6 M GuHCl, 0.2 M acetic acid. The eluate was dialyzed overnight against 6 L of 50 mM Tris (pH 7.6), 100 mM sodium chloride. Only about 10% of the full length NikR protein refolded to a soluble form during dialysis. The remainder of the protein was insoluble and was removed by centrifugation for 30 min at 12,000 × *g*. In contrast, more than 95% of NikR N-terminal domain was soluble after dialysis. Protein concentrations were determined using extinction coefficients of ϵ_{280} = 3,960 M^{-1} cm^{-1} for the full-length protein at pH 7.5 and ϵ_{295} = 2,600 M^{-1} cm^{-1} for the N-terminal domain of NikR at pH 12. Purified proteins were more than 95% pure as determined by 15% polyacrylamide tricine/SDS gel electrophoresis (Schagger & von Jagow, 1987).

Protease digestion

Purified NikR (100 μM) in 300 μL of 50 mM potassium phosphate (pH 7.6) was digested with 0.2 μM porcine pancreatic elastase type IV (Sigma, St. Louis, Missouri). Aliquots (20 μL) were taken at intervals from 0 to 30 min, quenched by addition to a tube containing 1 μL of 0.1 M phenylmethylsulfonyl fluoride (PMSF) and loading dyes, and immediately frozen in liquid nitrogen. Frozen samples were boiled for 5 min and analyzed by 15% polyacrylamide tricine/sodium dodecylsulfate (SDS) gels (Schagger & von Jagow, 1987). After 30 min, the remainder of the reaction was frozen in liquid nitrogen and later analyzed by MALDI-TOF mass spectroscopy using a PerSeptive Biosystems Voyager DE-STR instrument. Masses were correlated with peptide sequences using PAWS 8.1.1 (ProteoMetrics).

Circular dichroism

CD spectra were taken using an AVIV model 60CD spectropolarimeter (Aviv Associates, Lakewood, New Jersey) and a 1 cm pathlength cuvette. Scans were performed with 3.13 μM of protein in S buffer [50 mM sodium phosphate (pH 7.6)] and a 5 s averaging time at each wavelength. Fractional helicity was calculated assuming a mean residue ellipticity of 34,000 $deg \cdot cm^2 \cdot dmol^{-1}$ for a 100% helical protein. Thermal melts were performed with 1.25 μM protein in S buffer using a 1 min equilibration time and a 30 s averaging time for each 1 °C temperature increment. GuHCl melts were performed with a stirred 2 mL solution of 6.25 μM protein in S buffer from which aliquots were removed and replaced with an equal volume of 6.25 μM protein in S buffer plus 6 M GuHCl. Equilibration occurred rapidly at each concentration point as judged by the kinetics of the CD ellipticity change and signals were then averaged for 1 min. Data from thermal and GuHCl denaturation were fit to a two-state equilibrium model in which native dimers and denatured monomers are the only populated species, using nonlinear least-squares fitting and previously published equations (Bowie & Sauer, 1989a).

Analytical ultracentrifugation

A protein sample containing 6.25 μM NikR N-terminal domain in S buffer were centrifuged at 31,000 rpm in a Beckman Optima

XLA analytical ultracentrifuge. After 15 h of equilibration, absorbance data at 230 nm were recorded at 2 h intervals for 6 h and then at 1 h intervals for an additional 3 h. Data were fit as described previously (Schildbach et al., 1998), with a partial specific volume of 0.718 $cm^3 \cdot g^{-1}$ and a solution density of 1.005 $g \cdot cm^{-3}$.

DNA binding

DNase I footprinting was performed as described previously (Vershon et al., 1987; Smith & Sauer, 1995) with some modifications. Binding reactions were performed with B buffer [10 mM MES (pH 6.4), 25 mM potassium chloride, 3 mM magnesium chloride, 1.5 mM calcium chloride, and 1 mM EDTA]. Binding reactions were incubated on ice for 30 min. DNase I (Worthington Biochemical Corporation, Freehold, New Jersey) was then added to a final concentration of 4.5 μg/mL. Samples were not extracted with phenol/chloroform but precipitated with two volumes of 100% ethanol immediately after quenching. The 500 bp DNA fragment used for footprinting was from the *nikABCDE* promoter region and was generated by PCR with primers PC118 (5′-CTA TGG CCG GCC GGG CAA ACC TGC ATT TGC GCC GG-3′) and PC120 (5′-AAT CAT TGT CGA CAG CAT GGT AAC CCC AAT GGA TTA AAA-3′). The PC120 primer was 5′-end-labeled with γ^{32}P-ATP before amplification to allow detection of the digestion products. Samples from footprinting experiments were run on 8% polyacrylamide gels containing 8.3 M urea and 1 × TBE. Gels were exposed to a phosphor screen overnight, and traces of the footprints were obtained using a phosphorimager and the Image Quant 5.0 program (Molecular Dynamics, Sunnyvale, California).

Acknowledgments

We thank Catherine K. Smith, Nathan Walsh, Brian Wang, and members of the Sauer lab for technical assistance, advice, and valuable discussions. PTC was supported by an NSERC (Canada) Postdoctoral Fellowship. This work was supported by NIH grants AI-15706 and AI-16892.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res 25*:3389–3402.

Bairoch A, Apweiler R. 1998. The SWISS–PROT protein sequence data bank and its supplement TrEMBL in 1998. *Nucleic Acids Res 26*:38–42.

Bonvin AM, Vis H, Breg JN, Burgering MJ, Boelens R, Kaptein R. 1994. Nuclear magnetic resonance solution structure of the Arc repressor using relaxation matrix calculations. *J Mol Biol 236*:328–341.

Bowie JU, Sauer RT. 1989a. Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry 28*:7139–7143.

Bowie JU, Sauer RT. 1989b. Identifying determinants of folding and activity for a protein of unknown structure. *Proc Natl Acad Sci USA 86*:2152–2156.

Bowie JU, Sauer RT. 1990. TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J Mol Biol 211*:5–6.

Brown BM, Milla ME, Smith TL, Sauer RT. 1994. Scanning mutagenesis of the Arc repressor as a functional probe of operator recognition. *Nature Struct Biol 1*:164–168.

Brown BM, Sauer RT. 1999. Tolerance of Arc repressor to multiple-alanine substitutions. *Proc Natl Acad Sci USA 96*:1983–1988.

Burgering MJM, Boelens R, Gilbert DE, Breg JN, Knight KL, Sauer RT, Kaptein R. 1994. Solution structure of dimeric Mnt repressor (1–76). *Biochemistry 33*:15036–15045.

de Pina K, Desjardin V, Mandrand-Berthelot M-A, Giordano G, Wu L-F. 1999. Isolation and characterization of the *nikR* gene encoding a nickel-responsive regulator in *Escherichia coli*. *J Bacteriol 181*:670–674.

Fischer D, Eisenberg D. 1996. Protein fold recognition using sequence-derived predictions. *Protein Sci 5*:947–955.

Gomis-Rüth FX, Solà M, Acebo P, Pàrraga A, Guasch A, Eritja R, González A, Espinosa M, del Solar G, Coll M. 1998. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J 17*:7404–7415.

Hausinger RP. 1987. Nickel utilization by microorganisms. *Microbiol Rev 51*:22–42.

He Y-Y, McNally T, Manfield I, Navratil O, Old I, Phillips SEV, Saint-Girons I, Stockley PG. 1992. Probing *met* repressor-operator recognition in solution. *Nature 359*:431–433.

Jonsson T, Waldburger CD, Sauer RT. 1996. Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates. *Biochemistry 35*:4795–4802.

Kim J, Harter K, Theologis A. 1997. Protein–protein interactions among the Aux/IAA proteins. *Proc Natl Acad Sci USA 94*:11786–11791.

Knight KL, Sauer RT. 1989. Identification of functionally important residues in the DNA binding region of the Mnt repressor. *J Biol Chem 264*:13706–13710.

Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Lüthy R, Xenarios I, Bucher P. 1994. Improving the sensitivity of the sequence profile method. *Protein Sci 3*:139–146.

Milla ME, Brown BM, Sauer RT. 1994. Protein stability effects of a complete set of alanine substitutions in Arc repressor. *Nature Struct Biol 1*:518–523.

Milla ME, Sauer RT. 1994. P22 Arc repressor: Folding kinetics of a single-domain, dimeric protein. *Biochemistry 33*:1125–1133.

Navarro C, Wu L-F, Mandrand-Berthelot M-A. 1993. The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Mol Microbiol 9*:1181–1191.

Nooren IMA. 1999. Structure and dynamics of oligomeric repressor proteins [PhD dissertation]. Utrecht University, Utrecht, The Netherlands.

Pabo CO, Sauer RT. 1992. Transcription factors: Structural families and principles of DNA recognition. *Ann Rev Biochem 61*:1053–1095.

Phillips K, Phillips SEV. 1994. Electrostatic activation of *Escherichia coli* methionine repressor. *Structure 2*:309–316.

Phillips SE. 1994. The beta-ribbon DNA recognition motif. *Annu Rev Biophys Biomol Struct 23*:671–701.

Raumann BE, Brown BM, Sauer RT. 1994a. Major groove DNA recognition by β-sheets: The ribbon-helix-helix family of gene regulatory proteins. *Curr Opin Struct Biol 4*:36–43.

Raumann BE, Rould MA, Pabo CO, Sauer RT. 1994b. DNA recognition by β-sheets in the Arc repressor-operator crystal structure. *Nature 367*:754–757.

Rost B. 1999. Twilight zone of protein sequence alignments. *Protein Eng 12*:85–94.

Schagger H, von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem 166*:368–379.

Schildbach JF, Robinson CR, Sauer RT. 1998. Biophysical characterization of the TraY protein of *Escherichia coli* F Factor. *J Biol Chem 273*:1329–1333.

Smith TL, Sauer RT. 1995. P22 Arc repressor: Role of cooperativity in repression and binding to operators with altered half-site spacing. *J Mol Biol 249*:729–742.

Somers WS, Phillips SEV. 1992. Crystal structure of the *met* repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. *Nature 359*:387–393.

Tam R, Saier JMH. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev 57*:320–346.

Vershon AK, Bowie JU, Karplus TM, Sauer RT. 1986. Isolation and analysis of Arc repressor mutants: Evidence for an unusual mechanism of DNA binding. *Proteins Struct Funct Genet 1*:302–311.

Vershon AK, Liao S-M, McClure WR, Sauer RT. 1987. Bacteriophage P22 Mnt repressor. DNA binding and effects on transcription in vitro. *J Mol Biol 195*:311–322.