The N-terminus is unstructured, but not dynamically disordered, in the complex between HK022 Nun protein and λ-phage *BoxB* RNA hairpin

Ashley C. Stuart[1], Max E. Gottesman*, Arthur G. Palmer III**

Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, NY 10032, USA

Received 3 July 2003; revised 23 August 2003; accepted 25 August 2003

First published online 9 September 2003

Edited by Giovanni Cesareni

Abstract The Nun protein of lambdoid phage HK022 excludes λ-phage superinfection by blocking expression of genes downstream from the λ *nut* sequences. Heteronuclear NMR studies have been performed on a Nun peptide comprising residues 1–49 bound to the *nutR BoxB* RNA. The pattern of ^{13}C chemical shifts indicates that the arginine-rich motif of Nun forms an induced α-helix, consisting of residues 23–43, when bound to *BoxB* RNA, consistent with the structure of a shorter (residues 22–44) Nun peptide/*BoxB* RNA complex [Faber, C., Schärpf, M., Becker, T., Sticht, H. and Rösch (2001) J. Biol. Chem. 276, 32064–32070]. The N-terminal extension, residues 1–22, does not show chemical shifts or nuclear Overhauser effects characteristic of stable secondary structure. Nonetheless, ^{15}N relaxation measurements indicate that this region is not completely disordered, as expected for a random coil peptide. Restriction of conformation flexibility in the N-terminal extension of Nun may be important for binding to other target molecules involved in transcription termination.

Key words: Transcription termination; Dynamics; Disordered protein; Nuclear magnetic resonance spin relaxation; Arginine-rich motif

1. Introduction

Prophage HK022 Nun protein excludes superinfection by λ-phage by subverting the λ transcription antitermination system, which is required for temporal regulation of gene expression [1–3]. Nun and N bind competitively to the same λ *nut*R and *nut*L *BoxB* RNA sequences and interact with the host *Escherichia coli* RNA polymerase and Nus factors to terminate or antiterminate RNA transcription, respectively. Notably, Nun actively terminates transcription; it does not act simply by blocking interactions between N and *BoxB* RNA [2]. Although Nun is expressed in HK022 lysogens at immunologically undetectable levels from a weak prophage promoter [4], Nun out-competes N in vivo and only N transcribed from a strong promoter on a high copy number plasmid overcomes Nun activity and allows λ gene expression [5,6].

Both N and Nun are members of the arginine-rich motif (ARM) family of RNA-binding proteins. Other well-known members include the P22 and ϕ21 phage N proteins and the HIV viral proteins Rev and Tat [7]. The sequences of N and Nun show some homology in the ARM; however, sequences of the domain C-terminal to the ARM are disparate, consistent with the divergent functions of these proteins, and Nun contains an N-terminal extension not found in the N proteins. Sequence alignments of the ARM sequences of HK022 Nun and λ N are shown in Fig. 1. Structures have been determined by nuclear magnetic resonance (NMR) spectroscopy for short peptides encompassing the ARM motif of N (residues 1–22) [8], N (residues 1–36 [9], and Nun (residues 20–44) [10] bound to *BoxB* RNA. In both N and Nun peptides, the ARM forms a bent α-helix in the major groove of the *BoxB* RNA hairpin.

Previous investigations have demonstrated that the N-terminal extension is resistant to trypsin proteolysis when Nun is bound to *BoxB* RNA, whereas residues C-terminal to the ARM are not [11]. In the present investigation, NMR spectroscopy is used to characterize the interaction between *BoxB* RNA and a Nun peptide, consisting of residues 1–49, that encompasses both the ARM and the unique N-terminal extension. Secondary ^{13}C chemical shifts [12–14] and backbone amide ^{15}N spin relaxation rate constants [15,16] are used to provide insight into structure and dynamics of the Nun peptide on ps-ns and μs-ms time scales.

2. Materials and methods

2.1. BoxB17 RNA

The BoxB17 RNA sequence is 5′-**GG**CCCUGAAAAAGGGC**C**-3′. An additional closing GC base pair (shown in boldface) was added to the 15-nucleotide *nut*R *BoxB* sequence for added stability of the hairpin. The pentaloop is underlined. BoxB17 was transcribed in vitro from a single-stranded template using T7 polymerase [11,17]. The top and template DNA strands were 5′-AGA AAT TAA TAC GAC TCA CTA TA-3′ and 5′-GGC CCT TTT TCA GGG CCT GAA TGT GTA CCT ATA GTG AGT CGT ATT AAT TTC-3′, respectively. BoxB17 was purified using 20% polyacrylamide/8 M urea denaturing gel electrophoresis.

2.2. Nun (1–49) peptide

The peptide was produced by cyanogen bromide cleavage of full-length F49M mutant HK022 Nun; after cleavage the C-terminal M49 is converted to a homoserine residue. The F49M mutant was produced using the Quik Change Site-Directed Mutagenesis kit (Stratagene) with 5′ primer GGT GTG ACA CCT GGA ATG AAT GCT ATA GAT GAC GG and 3′ primer CCG TCA TCT ATA GCA TTC ATT CCA GGT GTC ACA CC. Full-length F49M Nun (1–109) was

*Corresponding author. Fax: (1)-212-305 1741.
**Corresponding author. Fax: (1)-212-305 6949.
E-mail addresses: gottesman@cancercenter.columbia.edu (M.E. Gottesman), agp6@columbia.edu (A.G. Palmer III).

[1] Present address: Rockefeller University, 1230 York Ave, New York, NY, USA.

0014-5793/03/$22.00 © 2003 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)00979-7

expressed using a pET-21d plasmid transformed into the BL21(DE3)/pLysS *E. coli* strain (Novagen) [11]. [U-^{15}N]Nun or [U-^{13}C, U-^{15}N]-Nun was produced using M9 media [18] prepared with $^{15}NH_4Cl$ (Cambridge Isotopes) and ^{13}C-glucose (Cambridge Isotopes), as needed. Nun was purified using S-sepharose ion exchange chromatography (Pharmacia). The protein was cleaved using 0.5 M cyanogen bromide in 5% trifluoroacetic acid. The Nun (1–49) peptide fragment was purified using reverse phase HPLC using a Delta-Pak C4 column followed by Mono-S ion exchange column chromatography (Pharmacia). Purified Nun (>97%) was dialyzed into aqueous buffer (10 mM KPO_4, 100 mM KCl, 2 mM NaN_3, pH 7.0) and concentrated using a centricon-3 (Amicon).

2.3. Nun (1–49)/BoxB17 complex

Binding of Nun (1–49) to *BoxB* RNA was assessed using a gel mobility shift assay [11]. BoxB17 was 5′-end labeled using [γ-^{32}P]ATP (Amersham). A fixed concentration (0.4 nM) BoxB17 was incubated with increasing concentrations of Nun (1–49) for 20 min at 0°C. The reaction mixtures were analyzed by 7.5% native polyacrylamide gel electrophoresis at 4°C. Samples for NMR were prepared by adding an aliquot of 48.2 μl of 4.54 mM BoxB17 RNA to 930 μl of 235 μM Nun (1–49). The solution was concentrated to 200 μl, 50 μl of 99.9% D_2O (Isotec) was added, the final volume was adjusted to 0.25 ml, and the solution was transferred to an NMR microtube (Shigemi).

2.4. NMR spectroscopy

NMR spectra were recorded at a temperature of 30°C using standard experimental protocols [19]. Triple-resonance, HCCH-TOCSY and ^{13}C-separated NOESY experiments used [U-^{13}C, U-^{15}N]Nun (1–49) and unlabeled BoxB17; other experiments used [U-^{15}N]Nun (1–49) and unlabeled BoxB17. HNCA, HNCO, HNCACB, and C(CO)NH experiments were performed on a 600 MHz Bruker DMX NMR spectrometer (NMRFAM, University of Wisconsin, Madison, WI, USA). CBCA(CO)NH and HCCH-TOCSY experiments were performed on a 600 MHz Varian INOVA NMR spectrometer (University of Cincinnati). HBHA(CO)NH and ^{15}N- and ^{13}C-separated 3D NOESY experiments were performed on a 600 MHz Bruker DRX spectrometer (Columbia University). Relaxation experiments [20] were performed on a Bruker DRX-500 NMR spectrometer (Columbia University). R_2 was measured from 15 spectra with relaxation delays of 16, 32, 64 (×2), 96 (×2), 128 (×2), 160, 192 (×2), 224, 288 (×2), 352 ms. Three replicates of {^{1}H-^{15}N} NOE experiments (with and without ^{1}H saturation) were collected. All NMR spectra were processed in FELIX 97 (Molecular Simulations). Resonance assignments were obtained using the ANSIG program [21]. Relaxation rate constants were determined with the Curvefit program (www.palmer.hs.columbia.edu). Secondary chemical shifts were determined as $\Delta\delta = \delta - \delta_{rc}$, in which δ is the observed shift and δ_{rc} is the random coil shift corrected for amino acid sequence dependence [12,14].

3. Results

Fig. 2 shows a gel mobility shift assay for binding of Nun (1–49) to BoxB17. The peptide exhibits strong 1:1 binding to BoxB17 similar to that of the full-length Nun protein binding to longer 53-nucleotide *BoxB* RNA sequences [11]. The assignments of the backbone amide resonances for Nun (1–49) in complex with BoxB17 are shown in the ^{1}H-^{15}N HSQC spectrum in Fig. 3. Resonance signals were not observed for V1 and K2. Doubling of resonances are observed for residues T4, I5, Y6, and V7. This sample heterogeneity could reflect *cis-trans* isomerization of P9 or a fraction of Nun (1–49) for which the N-terminal methionine residue has not been removed.

Secondary chemical shifts for $^{13}C^{\alpha}$ in Nun (1–49) in complex with BoxB17 are shown in Fig. 4. Shifts $\Delta\delta_{C\alpha} > 1$ ppm for residues S24-K42 are consistent with formation of an α-helix in the Nun (1–49)/BoxB17 complex. Secondary $^{13}C^{\alpha}$ shifts for N43 and G44 are smaller, but still positive, suggesting that the terminal residues of the helix are frayed (secondary $^{13}C'$ shifts are 2.0 and 0.3 ppm for N43 and G44, respectively). Values of $\Delta\delta_{C\alpha} = -1.4$ ppm and $\Delta\delta_{C\beta} = 3.0$ ppm for T23 suggests that this residue is involved in a helix N-capping interaction [22]. The reduction of the secondary shift at R32 is consistent with distortion of the helix at this point. A number of medium-range $d_{\alpha N}(i, i+3)$ ^{1}H-^{1}H NOEs characteristic of α-helical structure [23] are observed, beginning with residues T23-D26 and ending with Y39-K42 (data not shown). Secondary shifts for residues K3-L22 in the unique N-terminal extension of Nun are small (<1 ppm). No characteristic medium-range NOEs were observed in this region of the peptide.

```
                5    10    15    20    25
Nun   VKKTIYVNPDSGQNRKVSDRGLTSR
λ N                         gsMDAQ

                30    35    40    45
Nun   DRRRIARWEKRIAYALKNGVTPGM
λ N   TRRRERRAEKQAQWKAAN
      5     10    15    20
```

Fig. 1. Sequence alignments of HK022 Nun and λ N peptides. The sequence of Nun (1–49) used in the present work is shown; the F49M mutation is shown in boldface; after cyanogen bromide cleavage, this residue is homoserine. Sequence numbering is indicated above the amino acid sequence. The sequence of the Nun (20–44) peptide used by Rösch and coworkers [10] for structural studies is underlined. The sequence of the λ N peptide fragment used by Greenblatt and coworkers [8] for structural studies is shown; the N-terminal GS residues shown in lowercase result from the cloning strategy employed. Sequence numbering is indicated above or below the amino acid sequences. A longer N (1–36) fragment was used by Rösch and coworkers [9].

Fig. 5 shows the backbone ^{15}N spin relaxation data for Nun (1–49) in complex with BoxB17. Values of the {^{1}H}-^{15}N *NOE* > 0.5 obtained for residues T23-K42 indicate that the structure of the backbone is well-ordered on ps-ns times scales. The slightly smaller (0.46) value for N43 indicates an increased distribution of conformations. The small value (0.16) for G44 and increasingly negative values for residues V45-G48 indicate that the peptide is highly disordered at the C-terminus. Surprisingly, values of the {^{1}H}-^{15}N *NOE* for the N-terminal extension between K3 and R20 are small in magnitude, ranging from −0.04 to 0.18, and relatively uniform; large negative values characteristic of a highly disordered N-terminus are not observed. Values of $R_2 > 9$ s^{-1} are observed for residues S24-K42. Values of R_2 are slightly smaller for T23 and N43. Residues G44-M49 have $1 < R_2 < 5$ s^{-1}, with the more C-terminal residues having smaller values. The N-terminal residues K3-R20 are relatively uniform, rang-

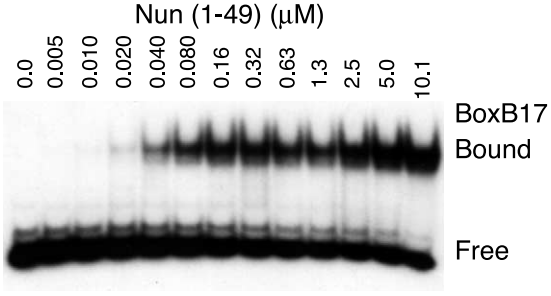

Fig. 2. Gel mobility shift of BoxB17 RNA by Nun (1–49). The concentration of BoxB17 was 0.4 nM; concentrations of Nun (1–49) are indicated above each lane of the gel. Fitting the intensities of the bands for the bound complex yields $K_D = 50 \pm 10$ nM.

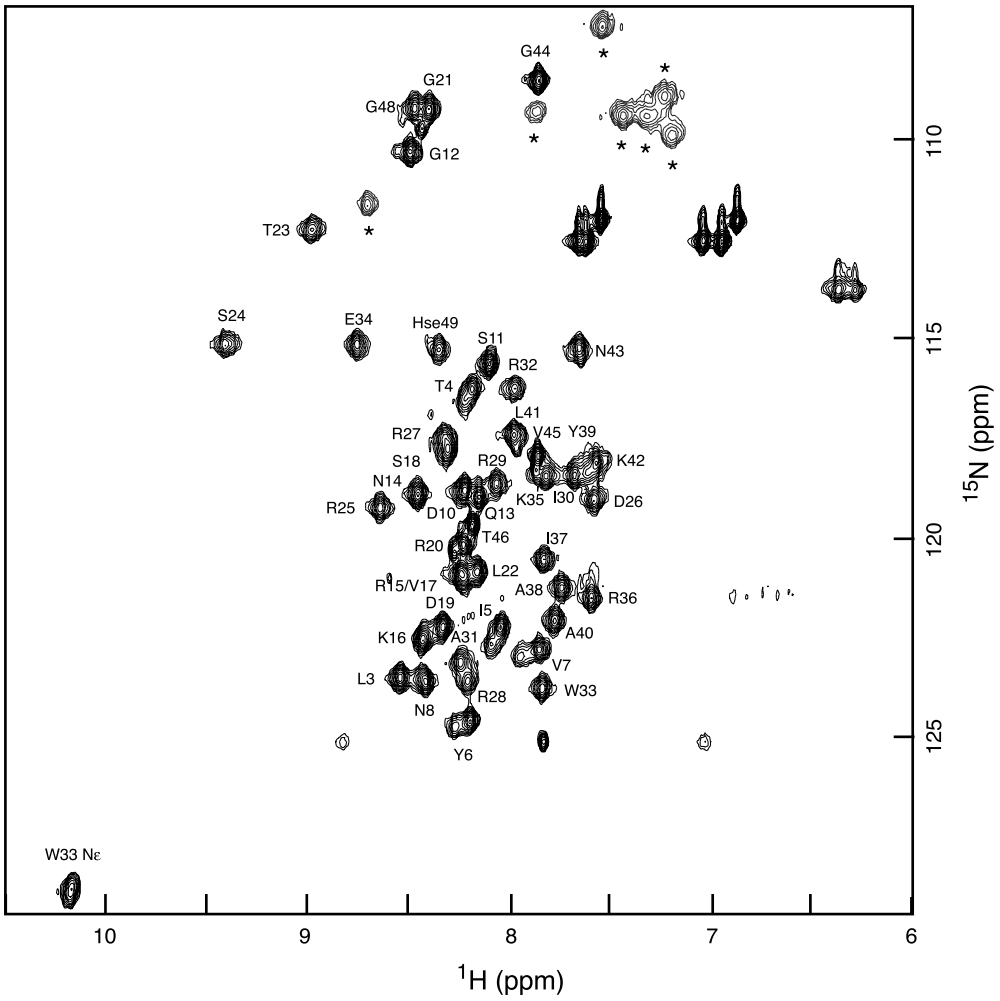

Fig. 3. ^{1}H-^{15}N HSQC NMR spectrum of Nun (1–49) complexed with BoxB17. Resonance assignments for backbone amide groups are indicated. Hse49 is the C-terminal homoserine residue resulting from cyanogen bromide cleavage. Peaks marked with a '*' are aliased resonances for arginine Nε moieties. All non-helical residues (3–22 and 44–49) have ^{1}H chemical shifts between 7.8 and 8.5 ppm; 12 of the 21 residues in the α-helix (23–43) have ^{1}H shifts outside this region.

ing from 3.5 to 5.3 s^{-1}. Within the α-helical region, R_2 values are more variable than the {^{1}H}-^{15}N *NOE*. For example, residues R29, R32, Y39, A40, L41, and K42 have R_2 between 9.0 and 10.0 s^{-1}; in contrast, residues S24, R25, D26, A31, W33, I37, and A38 have $R_2 > 11$ s^{-1}. The increased R_2 for the latter set of residues is not correlated with position within the helical structure. Thus, the increased relaxation rate for the latter set of residues appears to reflect chemical exchange line broadening resulting from dynamic processes on μs-ms time scales [16].

4. Discussion and conclusion

The HK022 Nun protein is largely disordered in the absence of other molecules. Protease protection assays have shown that the N-terminal half of the Nun protein, containing both the unique N-terminal extension and the ARM, is protected from degradation upon binding to *BoxB* RNA [11]. A previous structural study used Nun peptide fragments encompassing primarily the ARM [10]; the present study employed NMR spectroscopy to characterize structure and dynamics of a Nun (1–49) peptide that comprised the entire N-terminal extension and the ARM.

The ^{13}C chemical shifts, medium-range NOEs, and ^{15}N spin relaxation rate constants provide evidence that the longer Nun (1–49) construct employed in the present work binds to *BoxB* RNA in the same manner as the isolated ARM motif [10]. Specifically, the recognition helix is initiated by an N-capping box at residue T23, is bent near residue R32, and is terminated by a frayed residue N43. Dissociation of the complex is slow on the chemical shift time scale; thus, chemical exchange line broadening of ^{15}N residues in the ARM must reflect intramolecular conformational changes on μs-ms chemical shift time scales. Similar line broadening effects have been noted in the interface of other protein–nucleic acid complexes and have been attributed to multiple side chain conformations [24].

The function of the unique N-terminal extension of HK022 Nun protein is not known. The unusual patterns of the {^{1}H}-^{15}N *NOE* and ^{15}N R_2 for residues K3-R20 suggest that the distribution of backbone conformations is restricted on the ps-ns relaxation time scale even though no stable structure is evident on slower, μs-ms, chemical shift time scales. Residues I5-N8 have larger values of the {^{1}H}-^{15}N *NOE* and R_2 and more negative values of $\Delta\delta_{C\alpha}$ compared with other residues within the N-terminal extension. These observations suggest that transiently populated extended structures are more

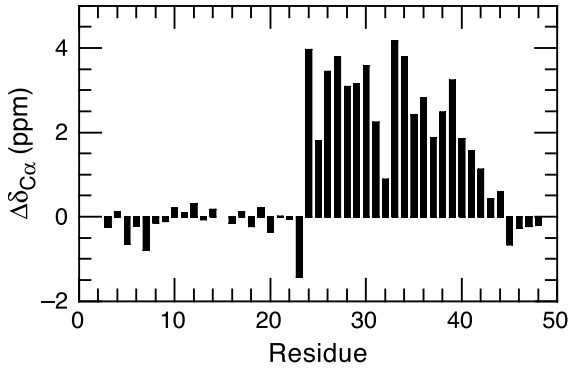

Fig. 4. Backbone $^{13}C^{\alpha}$ secondary chemical shifts, $\Delta\delta_{C\alpha}$, for Nun (1–49) complexed with BoxB17. The shift for the C-terminal homoserine residue is not shown. Values of $\Delta\delta_{C\alpha} > 1$ ppm are characteristic of α-helical conformations; values near 0 indicate the absence of regular secondary structure.

stabilized in this short peptide sequence. The PHD secondary structure prediction program [25] identifies residues R25-N43 as α-helical, consistent with the structural results, and residues K3-N8 as an extended β-strand, consistent with the conformational preferences derived from the chemical shift and relaxation data.

Recently, considerable interest has emerged in disorder–order transitions involved in functions of nucleic acid-binding proteins [26,27] and in other 'intrinsically disordered' proteins [28,29]. The present results suggest that initial binding of Nun to *BoxB* RNA restricts the conformational ensemble of the N-terminal extension. This would reduce the configurational entropic penalty associated with subsequent interactions between the N-terminal extension and other molecules involved in the transcription termination reaction. Binding of the N-terminal extension to other protein or nucleic acid molecules may involve stabilization of an extended conformation

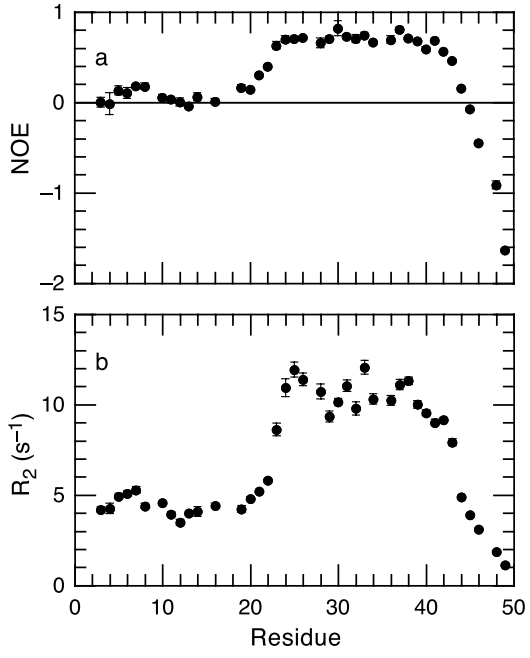

Fig. 5. Backbone (a) $\{^1H\}$-^{15}N *NOE* and (b) ^{15}N R_2 for Nun (1–49) complexed with BoxB17. Large values of $NOE > 0.6$ are characteristic of ordered conformations; values of $NOE < 0$ and of $R_2 \rightarrow 0$ indicate highly disordered conformations.

for residues I5-N8. Sequential accretion of order may be a mechanism for facilitating and regulating the assembly of macromolecular complexes, such as the termination machinery [30]. The biological implication of this work is that Nun may be capable of interactions that N is not, thereby explaining the observation that Nun strongly out-competes N in vivo [5,6].

Acknowledgements: This work was supported by NIH grant GM37219 (M.E.G.) and NSF grant MCB-9722392 (A.G.P.). We thank Mark Rance (Univ. Cincinnati) for access to a 600 MHz NMR spectrometer. Some NMR studies were performed at the National Magnetic Resonance Facility at Madison (NMRFAM) with support from the NIH (RR02301, RR02781, RR08438, GM66326), NSF (BIR-9214394, DMB-8415048) and USDA.

References

[1] Das, A. (1993) Annu. Rev. Biochem. 62, 893–930.
[2] Hung, S.C. and Gottesman, M.E. (1995) J. Mol. Biol. 247, 428–442.
[3] Nudler, E. and Gottesman, M.E. (2002) Genes Cells 7, 755–768.
[4] King, R.A., Madsen, P.L. and Weisberg, R.A. (2000) J. Bacteriol. 182, 456–462.
[5] Robert, J., Sloan, S.B., Weisberg, R.A., Gottesman, M.E., Robledo, R. and Harbrecht, D. (1987) Cell 51, 483–492.
[6] Washburn, R.S., Wong, Y. and Gottesman, M.E. (2003) J. Mol. Biol. 329, 655–662.
[7] Weiss, M.A. and Narayana, N. (1998) Biopolymers 48, 167–180.
[8] Legault, P., Li, J., Mogridge, J., Kay, L.E. and Greenblatt, J. (1998) Cell 93, 289–299.
[9] Schärpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S. and Rösch, P. (2000) Eur. J. Biochem. 267, 2397–2408.
[10] Faber, C., Schärpf, M., Becker, T., Sticht, H. and Rösch, P. (2001) J. Biol. Chem. 276, 32064–32070.
[11] Chattopadhyay, S., Hung, S.C., Stuart, A.C., Palmer, A.G., Garcia-Mena, J., Das, A. and Gottesman, M.E. (1995) Proc. Natl. Acad. Sci. USA 92, 12131–12135.
[12] Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Chung, J., Wright, P.E. and Dyson, H.J. (2001) J. Am. Chem. Soc. 123, 2970–2978.
[13] Spera, S. and Bax, A. (1991) J. Am. Chem. Soc. 113, 5490–5492.
[14] Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) J. Biomol. NMR 5, 67–81.
[15] Palmer, A.G. (2001) Annu. Rev. Biophys. Biomol. Struct. 30, 129–155.
[16] Palmer, A.G., Kroenke, C.D. and Loria, J.P. (2001) Methods Enzymol. 339, 204–238.
[17] Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Nucleic Acids Res. 15, 8783–8798.
[18] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[19] Cavanagh, J., Fairbrother, W.J., Palmer, A.G. and Skelton, N.J. (1996) Protein NMR Spectroscopy: Principles and Practice, Academic Press, San Diego, CA.
[20] Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. and Kay, L.E. (1994) Biochemistry 33, 5984–6003.
[21] Kraulis, P.J. (1989) J. Magn. Reson. 84, 627–633.
[22] Gronenborn, A.M. and Clore, G.M. (1994) J. Biomol. NMR 4, 455–458.
[23] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[24] Härd, T. (1999) Q. Rev. Biophys. 32, 57–98.
[25] Rost, B. and Sander, C. (1993) J. Mol. Biol. 232, 584–599.
[26] Bracken, C., Carr, P.A., Cavanagh, J. and Palmer, A.G. (1999) J. Mol. Biol. 285, 2133–2146.
[27] Spolar, R.S. and Record, M.T. (1994) Science 263, 777–784.
[28] Dyson, H.J. and Wright, P.E. (2002) Curr. Opin. Struct. Biol. 12, 54–60.
[29] Dunker, A.K., Brown, C.J., Lawson, D., Iakoucheva, L.M. and Obradovic, Z. (2002) Biochemistry 41, 6573–6582.
[30] Namba, K. (2001) Genes Cells 6, 1–12.