histidine, stabilizing the motif into a single folded domain[12,14]. As well as the conservation of the Zn^{2+} ligands, there are additional conserved residues between MDM2 and the rest of the RING-finger family, including a number of hydrophobic residues implicated in stabilizing the RING-finger fold (Fig. 1a; Ref. 14). Interestingly, two of the proteins shown in the alignment are implicated in the control of cell fate, namely IAP[15] and DROC3H[16]. ICPO is involved in gene regulation[14] whereas the functions of CELRO5 (a *Caenorhabditis elegans* protein)[17] and RING1 (Ref. 12) are not known. MDM2 and CELR05 show significant sequence homologies outside the RING finger, and both can be aligned (17% identity over 491 residues; Fig. 1b). To further assess the significance of the alignment, the zinc finger region of MDM2 (434–491) was used to search the OWL database (version 22.1) with the program PROSRCH[18] using PAM matrices of 100 and 250. Statistically significant matches (excluding itself) were observed only between MDM2 and IAP, DROC3H, CELR05 and ICPO. Furthermore, no matches were observed in the top 100 alignments between MDM2 and any C2H2 or C4 zinc-finger-containing protein, supporting the proposal that MDM2 contains a RING finger rather than two tandemly arranged zinc fingers as previously postulated[8–10].

The extensive RING-finger family contains a number of known oncoproteins including, for example, MEL18, BMI-1, RFP, PML and T18 (reviewed in Ref. 13). A large number of proteins in this family are also implicated in the control of cell growth, differentiation and development[13]. The presence of the RING-finger domain in MDM2 may infer a common functionality with other members of the RING-finger family. This function may involve protein–protein or protein–nucleic acid interaction, which would be distinct from the proposed inactivation of p53, but necessary for the other important biological activities of MDM2.

References

1 Eliyahu, D. *et al.* (1989) *Proc. Natl Acad. Sci. USA* 86, 8763–8767
2 Finlay, C. A., Hinds, P. W. and Levine, A. J. (1989) *Cell* 57, 1083–1093
3 Kern, S. E. *et al.* (1991) *Science* 252, 1708–1711
4 Zambetti, G. P. *et al.* (1992) *Genes Dev.* 6, 1143–1152
5 Levine, A. J. (1990) *Virology* 177, 419–426
6 Momand, J. *et al.* (1992) *Cell* 69, 1237–1245
7 Fakharzadeh, S. S., Trusko, S. P. and George, D. L. (1991) *EMBO J.* 10, 1565–1569
8 Oliner, J. D. *et al.* (1992) *Nature* 358, 80–83
9 Chen, J., Marechal. V. and Levine, A. J. (1993) *Mol. Cell. Biol.* 13, 4107–4114
10 Oliner, J. D. *et al.* (1993) *Nature* 362, 857–860
11 Freemont, P. S., Hanson, I. M. and Trowsdale, J. (1991) *Cell* 64, 483–484
12 Lovering, R. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 2112–2116
13 Freemont, P. S. (1993) *Anal. NY Acad. Sci.* 684, 174–192
14 Everett, R. D. *et al.* (1993) *J. Mol. Biol.* 234, 1038–1047
15 Crook, N. E., Clem, R. J. and Miller, K. L. (1993) *J. Virol.* 67, 2168–2174
16 Price, B. D. *et al.* (1993) *EMBO J.* 12, 2411–2418
17 Sulston, J. *et al.* (1992) *Nature* 356, 37–41
18 Collins, J. F., Coulson, A. F. W. and Lyall, A. (1988) *CABIOS* 4, 67–71
19 Barton, G. J. and Sternberg, M. J. E. (1990) *J. Mol. Biol.* 212, 389–402
20 Needleman, S. B. and Wunsch, C. D. (1970) *J. Mol. Biol.* 48, 443–453

MICHAEL N. BODDY AND PAUL S. FREEMONT

Protein Structure Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London, UK WC2A 3PX.

KATHERINE L. B. BORDEN

Laboratory of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London, UK NW7 1AA.

The modular structure of NifU proteins

The *nif* cluster contains genes responsible for nitrogen fixation in several prokaryotes[1–3]. With the exception of the essential genes coding for and regulating the expression of the nitrogenase subunits, the functions of other *nif* genes appear to be associated with cofactor biosynthesis, amino acid interconversion and regulation of ammonia storage. However, not all the biochemical activities of the *nif* gene products have been identified yet. Some of these proteins might represent specialized versions of general metabolic enzymes; thus their presence cannot be precluded from non-nitrogen fixing organisms. A recent example is the identification of the NifS family as pyridoxal-phosphate-dependent aminotransferases[4].

Here we report the identification by sequence analysis of two distinct domains in the NifU protein family (Fig. 1a). Although the biochemical function of NifU proteins remains unknown, it has been shown that NifU exists as a homodimer, containing one redox 2Fe–2S cluster per subunit[5].

The carboxy-terminal domain, known to be the only common region of NifU between nitrogen-fixing bacteria and several rhodobacterial species[6,7], is shown here to be present in: the open reading frame (ORF) YKL253 of chromosome XI of yeast[8]; a hypothetical 22 kDa protein from *Haemophilus influenzae*[9]; and a short ORF (ORF2) in *Azotobacter vinelandii*[10] (Fig. 1b). Thus, *A. vinelandii* has two NifU-like proteins, one of which contains only the carboxy-terminal domain (Fig. 1a, b).

An internal domain of NifU proteins from nitrogen-fixing bacteria, not present in the so-called NifU proteins from two *Rhodobacter* species or the NifU-like protein from yeast, shows significant sequence similarity (Fig. 1c) to two internal repeats of nitrite reductases from *Klebsiella pneumoniae* (NasB)[11], *Escherichia coli* (NirB)[12], *Emericella nidulans* (NiiA)[13] and *Neurospora crassa* (Nit-6)[14], as well as to a single carboxy-terminal domain of nitrate reductase of *K. pneumoniae* (NasA)[11], the carboxyl terminus of NifE from *Bradyrhizobium japonicum*[15] and a short ORF (gp64)[16] from *E. coli*. Interestingly, the domain is present in NifE of *B. japonicum* at its extreme carboxy-terminal end and is absent from the other NifE homologues. In *E. nidulans*, the domain is encoded by the fifth exon of the *niiA* (nitrite reductase) gene[13]. This domain contains two pairs of conserved cysteines that might participate in the formation of the 2Fe–2S cluster in NifU and the other proteins identified (Fig. 1a). This cluster might serve as an electron carrier in the nitrate and nitrite reductases.

Modular proteins, frequently found in higher eukaryotes, are rare in prokaryotes, probably because of a lack of appropriate genetic shuffling mechanisms (see Refs 17, 18 and references therein). The mobility of the NifU-like domains in both prokaryotes and lower eukaryotes therefore comes as a surprise. It will be a challenge to identify the genetic mechanism(s) that enabled the various species to acquire the two domains. In addition, knowledge of the modular structure of NifU-like proteins might contribute towards elucidating the functions of their domains.

References

1 Beynon, J. *et al.* (1987) *J. Bacteriol.* 169, 4024–4029
2 Arnold, W. *et al.* (1988) *J. Mol. Biol.* 203, 715–738
3 Jacobson, M. R. *et al.* (1989) *J. Bacteriol.* 171, 1017–1027
4 Ouzounis, C. and Sander, C. (1993) *FEBS Lett.* 322, 159–164
5 Dean, D. R. *et al.* (1993) *J. Bacteriol.* 175, 6737–6744
6 Meijer, W. G. and Tabita, F. R. (1992) *J. Bacteriol.* 174, 3855–3866
7 Masepohl, B. *et al.* (1993) *Mol. Gen. Genet.* 238, 369–382
8 Purnelle, B., Skala, J., Van Dyck, L. and Goffeau, A. (1992) *Yeast* 8, 977–986
9 Tomb, J. F., El-Hajj, H. and Smith, H. O. (1991) *Gene* 104, 1–10
10 Joerger, R. D., Jacobson, M. R. and Bishop, P. E. (1989) *J. Bacteriol.* 171, 3258–3267
11 Lin, J. T., Goldman, B. S. and Stewart, V. (1993) *J. Bacteriol.* 175, 2370–2378

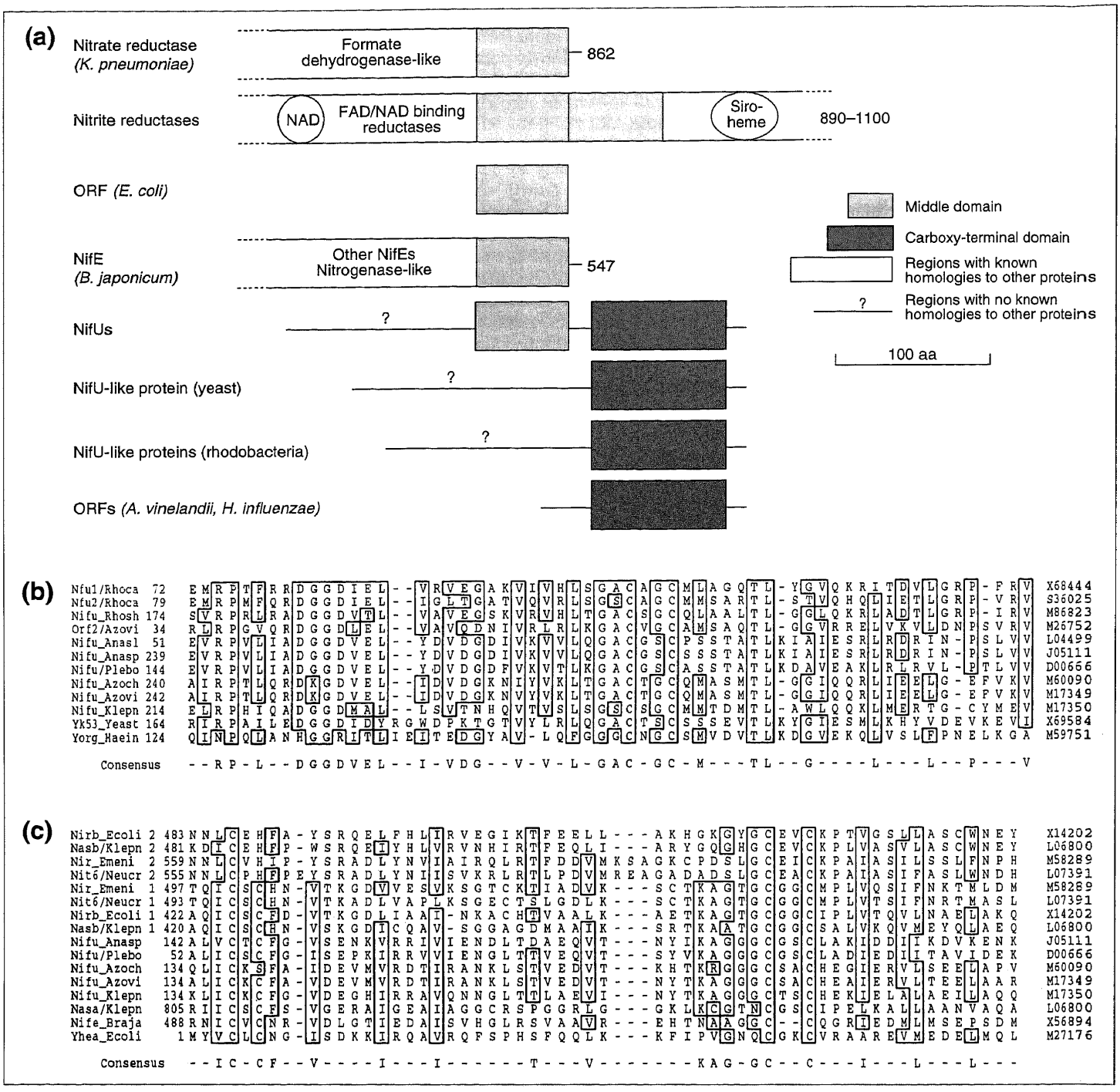

Figure 1

(a) Modular architecture of proteins that contain NifU-like domains. For large proteins (fragments only shown), the length of sequence shown is indicated. **(b)** Alignment of the approximately 65-residue carboxy-terminal domain. Proteins are indicated by their Swissprot names if available [these contain an underscore (_) in the entry code]. Nfu1/Rhoca and Nfu2/Rhoca, *R. capsulatus* Nfu1 and Nfu2; Nifu_Rhosh, *R. sphaeroides* Nifu; Orf2/Azovi, *A. vinelandii* ORF 2; Nifu_Anas1, Nifu_Anasp, Nifu/Plebo, Nifu_Azoch, Nifu_Azovi and Nifu_Klepn, NifU proteins from *Anabaena* (two species), *Plectoneme boryanum*, *A. chroococcum*, *A. vinelandii* and *K. pneumoniae*, respectively; Yk53_Yeast, yeast NifU-like protein; Yorg_Haein, *H. influenzae* hypothetical 22 kDa protein encoded by ORF G. The position of the domain within the corresponding proteins is indicated, as is the EMBL database accession number. Residues conserved in more than half of the sequences are boxed and displayed as consensus (bottom line). The newly identified domains show significant sequence similarity to known NifU proteins (amino acid identity: Yk53_Yeast, 28–36%; Yorg_Haein, 24–33%; Orf2/Azovi 35–50%). Database searches with the candidate proteins result in BLASTP P-values[19] that are below $10e \times 10^{-6}$ for at least one NifU. In order to verify these similarities, various sequence analysis methods were used including profile and pattern searches[20]. **(c)** Alignment of the approximately 60-residue middle domain. Nirb_Ecoli, *E. coli* nitrite reductase B; Nasb/Klepn, *K. pneumoniae* nitrite reductase B; Nir_Emeni, *E. nidulans* nitrite reductase; Nit6/Neucr, *N. crassa* nitrite reductase; Nasa/Klepn, *K. pneumoniae* nitrite reductase A; Nife_Braja, *B. japonicum* Nife protein; Yhea_Ecoli, *E. coli* hypothetical protein encoded by the *bfr* 3'-region. The newly identified domains show significant sequence similarity to known NifU proteins as follows: Nasa/Klepn, 19–30%; Nirb_Ecoli domain 1 (as an example of nitrite reductases from prokaryotes and fungi), 21–38%; Yhea_Ecoli, 19–26%; Nife_Braja, 22–38%. The similarities were verified as in **(b)**.

12 Peakman, T. *et al.* (1990) *Eur. J. Biochem.* 191, 315–323
13 Johnstone, I. L. *et al.* (1990) *Gene* 90, 181–192
14 Exley, G. E., Colandene, J. D. and Garrett, R. H. (1993) *J. Bacteriol.* 175, 2379–2392
15 Aguilar, O. M. *et al.* (1990) *Mol. Gen. Genet.* 224, 413–420
16 Andrews, S. C., Harrison, P. M. and Guest, J. R. (1989) *J. Bacteriol.* 171, 3940–3947
17 Bork, P. (1993) *Proteins* 17, 363–374
18 Doolittle, R. F. and Bork, P. (1993) *Sci. Am.* 269(4), 50–56
19 Altschul, S. F. *et al.* (1990) *J. Mol. Biol.* 215, 403–410
20 Bork, P. *et al.* (1992) *Protein Sci.* 1, 1677–1690

CHRISTOS OUZOUNIS, PEER BORK AND CHRIS SANDER

European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany.

The glycine-rich sequence of protein kinases: a multifunctional element

Dirk Bossemeyer

Evolution favours the use of glycine-rich loops for nucleotide binding in proteins. In the large family of protein kinases, the catalytic domain of which has one of the highest degrees of conservation among all known proteins, the structure of the nucleotide-binding site differs from classical folds. We are now beginning to understand the multiple functional roles of the glycine-rich sequence in protein kinases and some of the structural constraints leading to its conservation.

SINCE ROSSMANN *et al.* described the common nature of the dinucleotide-binding fold[1], the glycine clusters of nucleotide-binding proteins have been a continuing matter of interest for many researchers[2,3]. A cluster of glycines is a conserved feature of many nucleotide-binding proteins and has been used successfully in their identification[4]. In most cases, two related consensus-sequence patterns can be found: Gly-X-X-Gly-X-Gly-Lys-[Ser/Thr] in proteins that bind mononucleotides, or Gly-X-Gly-X-X-Gly in proteins with a preference for dinucleotides[5].

The glycine loop in nucleotide-binding proteins

The many crystal structures of nucleotide-binding proteins that are now available reveal that the architecture of the nucleotide-binding sites varies considerably between the different classes of enzymes, as outlined below*.

Dinucleotide-binding proteins. In proteins that bind dinucleotides, such as nicotinamide adenine dinucleotide (NAD), nicotinamide adenine dinucleotide phosphate (NADP) and flavine adenine dinucleotide (FAD), and where the fingerprint of a Rossmann fold is conserved, the glycine consensus typically comprises a tight turn between a β-strand and an α-helix. The β-strand is part of a βαβ-motif with parallel β-strands. The amino-terminal helix dipole orients towards the pyrophosphate moiety, providing a favourable interaction[6]. Of special significance is the second glycine in the conserved pattern. Its position marks the beginning of the helix and provides space for the approaching pyrophosphate. The first and third glycines allow a tight turn and a very compact βαβ-unit with an accessible helix at the amino terminus. An alanine in the position of the third glycine is characteristic of NADP-binding proteins, and potentially of relevance for coenzyme specificity[7,8].

Proteins with different types of dinucleotide binding, which do not involve the Rossmann fold, are aldose reductase (an α/β-barrel protein, which binds NADP in an unusual manner[9,10]), and some bacterial toxins such as diphtheria toxin and *Escherichia coli* enterotoxin (these are β-barrel proteins with a new kind of NAD-binding site[11,12]).

Mononucleotide-binding proteins. The structural relationship between the different classes of mononucleotide-binding proteins is weak; thus convergence in the evolution of the binding sites appears more likely than divergence. Several high-resolution crystal structures are available for proteins containing the classical consensus sequence Gly-X-X-Gly-X-Gly-Lys-[Ser/Thr]. The common features and the variations within this phosphate-binding loop (P-loop) have been investigated by Saraste *et al.*[3] A superposition of the P-loops of adenylate kinase[13], RecA[14], elongation factor Tu[15] and the H-*ras* oncogene product p21 (Ref. 16) revealed equivalent positions for the α- and β-phosphoryl groups of the bound mononucleotide[14]. The loop also links the first of two parallel β-strands to a following α-helix but, unlike in dinucleotide-binding proteins, the P-loop is large and wraps around the phosphate group (Fig. 1). In $p21^{H\text{-}ras}$ the amide nitrogens of the glycine loop provide a positively polarized electrostatic field for the GTP analogue. The

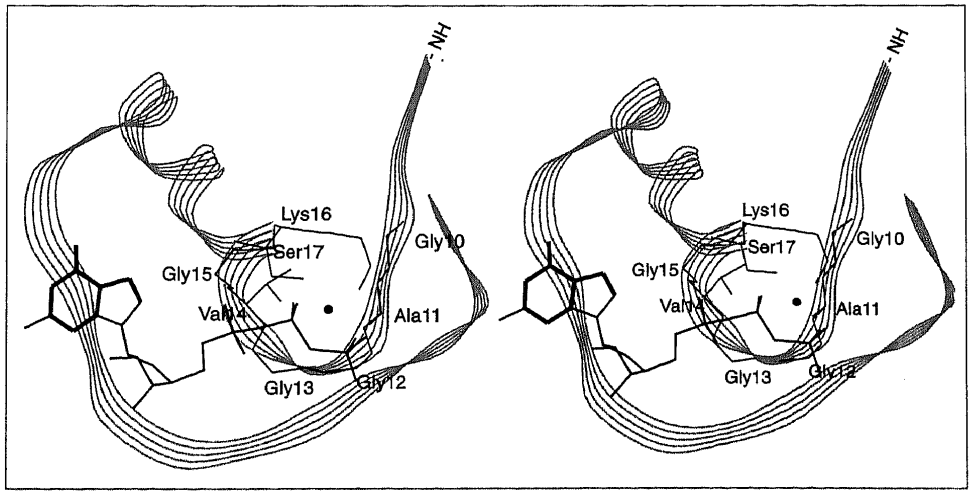

Figure 1
Stereo view of the glycine-rich sequence of the mononucleotide-binding fold of $p21^{H\text{-}ras}$. The GTP analogue GMP-PNP is shown in blue, and the Mg^{2+} ion is shown as a black circle. The coordinates of $p21^{H\text{-}ras}$ in the Brookhaven Protein Database (Accession No. PDB5P21; Ref. 16) were used to produce the figure. NH, amino terminus.

*It would exceed the scope of this article to give a thorough overview of the folding motifs found in nucleotide-binding proteins. For a detailed study, refer to the cited work and references therein.

D. Bossemeyer is at the German Cancer Research Centre, Department of Pathochemistry and Molecular Biology of Mitosis, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany.