Identification of possible adenine nucleotide-binding sites in nitrogenase Fe- and MoFe-proteins by amino acid sequence comparison

R.L. Robson

AFRC Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, England

Received 21 May 1984

Published amino acid sequences for nitrogenase component polypeptides were compared with those of other proteins which also bind adenine nucleotides. Three sequences which might contribute to an adenine nucleotide-binding domain were found for the Fe-protein component of nitrogenase. The β-subunit of the MoFe-protein (*nifK* gene product) contains a sequence which is similar to other proteins which exhibit ATPase activity. No similarities were observed for the α-subunit of this component. The findings are discussed in relation to the experimental data on adenine nucleotide binding and the proposed role of ATP in the enzyme mechanism.

Amino acid sequence *Adenine nucleotide binding* *Nitrogenase* *Fe-protein* *MoFe-protein*

1. INTRODUCTION

The enzyme complex responsible for the reduction of N_2 to ammonia comprises two component proteins. The larger contains two different polypeptides of about 50–60 kDa, with an $\alpha_2\beta_2$ subunit structure, 2 atoms of Mo, approx. 32 atoms of Fe and 24–30 atoms of labile S. This component is thought to be the site of N_2 binding and reduction and will be referred to here as the MoFe-protein. The smaller component comprises two identical subunits, each approx. 30 kDa, and is generally considered to contain 4 atoms each of Fe and labile S. It acts as a specific donor of low-potential electrons to the MoFe-protein and will be termed the Fe-protein (see [1]).

For activity, nitrogenase requires a low-potential electron donor, MgATP and an anoxic environment. The role of ATP in the mechanism of nitrogenase has been the subject of considerable study. Both protein components are required for hydrolysis of ATP which occurs at significant rates in the absence of reductant. In the conversion of N_2 to 2 molecules of NH_4^+, reductant is required and 16 ATPs are hydrolysed concomitantly with electron transfer from the Fe-protein to the MoFe-protein complex (see [1]). The site of ATP hydrolysis is not known and a dual role for ATP is indicated [18], e.g., MgATP may bridge the two components in the complex [7]. In general, Fe-proteins bind 2 molecules of MgATP and either one or two molecules of MgADP [2–4] but these stoichiometries may be influenced by the redox state of the protein [4]. ATP induces a conformational change and lowers the mid-point redox potential of the Fe-protein and also produces changes in various spectra and increased sensitivity to several chemical agents including O_2 (see [1]).

Reports on the binding of adenine nucleotides to the MoFe-protein are generally negative. However, in highly purified protein from *Klebsiella pneumoniae*, 4 low-affinity ATP-binding sites were detected as was labilization of the terminal phosphate groups of both ATP and ADP [5]. Also, 4 binding sites for Mn^{2+} and MnATP were inferred from NMR studies of proton relaxation enhancement [6]. However, ATP binding by MoFe-protein from *Azotobacter vinelandii* was not detected [4].

Published by Elsevier Science Publishers B.V.

Amino acid sequences for nitrogenase components have been obtained for a number of organisms either by direct determination or by deduction from DNA sequences of the genes which code for them: *nifH* for the Fe-protein polypeptide and *nifD* and *nifK* for the α- and β-subunits, respectively, of the MoFe-protein (see [8]). Recently, there has been considerable interest in comparing amino acid sequences for proteins which have similar capabilities. Similarities in sequences of proteins which bind adenine nucleotides may indicate regions which could be involved in binding of those molecules [9]. Here, amino acid sequences for various ATP, ADP or AMP binding proteins were compared with those for nitrogenase proteins. The findings are discussed in the light of experimental evidence regarding the possible binding ATP sites in nitrogenase and the role of ATP in the mechanism of nitrogenase.

2. EVIDENCE

2.1. *The Fe-protein*

Amino acid sequence data are available for the Fe-proteins from 7 organisms. Sequences are remarkably conserved between these proteins and those areas most highly conserved are probably functionally important regions (see [10]). Fig.1 shows the first 120 N-terminal amino acid residues of all nitrogenase Fe-proteins published. Similar sequences from a number of other ATP-binding proteins are aligned for comparison. Three comparable regions can be observed. The first (A) was found within the first 30 N-terminal residues of the Fe-proteins and is a highly conserved sequence. In adenylate kinase the comparable sequence is also close to the N-terminal end and X-ray crystallographic studies have shown that it forms a flexible loop which enfolds AMP [31,32].

Region B of the Fe-proteins is also highly conserved in all but one species and shows a particularly striking similarity to residues 437–470 of *Escherichia coli* ATPase α-subunit and in particular a 22-residue sequence between residues 442 and 464. Nine amino acids are identical and others are conservative replacements. The DNA sequence data for this region of the ATPase subunit show that it is 50% homologous to all Fe-protein sequences reported. In the Fe-protein from *Clostridium pasteurianum* there is a marked difference: to fit this sequence to the other Fe-proteins in this region, a 2 amino acid deletion must be assumed [10], whereupon -Leu-Arg- replaces -Ala-Ala found 5 residues towards the N-terminal end. If this region were important for MgATP or MgADP binding then some differences might be predicted between this protein and the other Fe-proteins.

Region C of the Fe-protein is less similar to sequences of other adenine nucleotide-binding proteins than is region B. Nevertheless, this sequence compares well with regions found to be similar in a number of different adenine nucleotide-binding proteins [9] but which also bear similarity to region A for the Fe-protein. The most striking similarity is to *E. coli* ATPase β-subunit where 7 residues in 21 are identical. Similarities, some at different residues, can be identified in the other proteins when aligned with this region.

2.2. *The MoFe-protein*

Data indicating a binding site for adenine nucleotide on the MoFe-protein are inconsistent. The search for such a domain in subunits of this protein could help to resolve this problem. Published sequence data for this protein are limited to the α-subunit of the MoFe-protein of *C. pasteurianum* [23], part of the α-subunit from *K. pneumoniae* [13], and cysteinyl peptides from MoFe-protein of *A. vinelandii* [24]. The complete DNA sequences for this component from *Anabaena* have been reported [25,26].

A search of the available sequence data with sequences implicated in nucleotide binding in other proteins revealed one region of similarity in the β-subunit (*nifK* product) from *Anabaena* (fig.2). Region 74–96 is similar to the *N,N'*-dicyclohexylcarbodiimide (DCCD)-binding region of the β-subunit of ATPase [27]. In a recent review of ATPase it was pointed out that this region is similar to residues 259–281 of *E. coli recA* gene product which also exhibits ATPase activity [28]. Comparison of the *recA* sequence with the *Anabaena* sequence shows an even more striking similarity, particularly when conservative residues are included. When the cysteinyl peptides for the *A. vinelandii* protein were aligned along the *Anabaena* sequence, the 'α-3' and 'α-1' cysteinyl peptides form a contiguous sequence in the region of

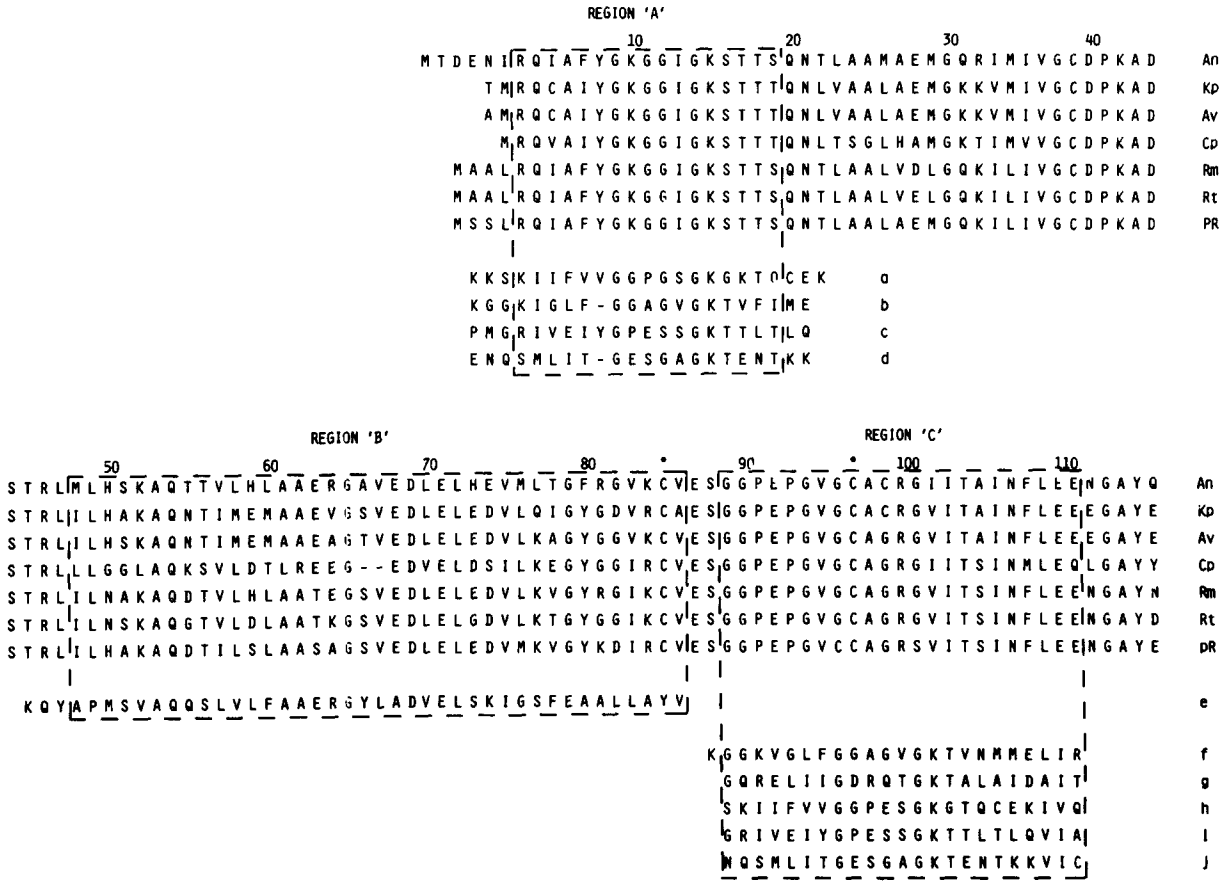

Fig.1. Similarities between amino acid sequences of nitrogenase Fe-proteins and other adenine nucleotide-binding proteins. Fe-proteins: An, *Anabaena* [14]; Kp, *K. pneumoniae* [12,13]; Av, *A. vinelandii* [10]; Cp, *C. pasteurianum* [11]; Rm, *Rhizobium meliloti* [15]; Rt, *R. trifolii* [16]; PR, *Parasponium rhizobium* [17]. Comparative sequences from other proteins: (a) porcine adenylate kinase residues 6–27 [22]; b, bovine ATPase β-subunit residues 149–169 (see [9]); c, *E. coli* Rec A protein residues 57–78 [23]; d, nematode myosin residues 161–181 (see [9]); e, *E. coli* ATPase α-subunit residues 432–473 [19,20]; f, *E. coli* ATPase β-subunit residues 150–173 [21]; g, *E. coli* ATPase α-subunit residues 162–184 [19,20]; h, porcine adenylate kinase residues 8–30 [22]; i, *E. coli* Rec A protein residues 60–82 [23]; j, nematode myosin residues 171–193 (see [9]). Boxes indicate identical or conservative residues. * marks Cys 85 and 97 of Fe-protein. Amino acid residues are denoted by the single-letter amino acid code.

interest. Comparison with the part sequence of *K. pneumoniae nifD* suggests that these peptides actually derive from the β-subunit (*nifK* product) of the *A. vinelandii* MoFe-protein. Fig.2 shows that if these peptides are indeed contiguous and similar to the *Anabaena* region, then those residues similar to the *recA* gene product and ATPase β-subunit are conserved. In addition, in the *A. vinelandii* peptide tyrosine is found at an equivalent position to residue 272 of Rec A protein whereas this is a conserved residue in the protein from *Anabaena*.

Analysis of data available for MoFe-protein α-subunits revealed no obvious similarities to any of the sequences previously mentioned. Sequence comparison of MoFe protein α- and β-subunits revealed no marked similarity between the polypeptides as indicated before [26].

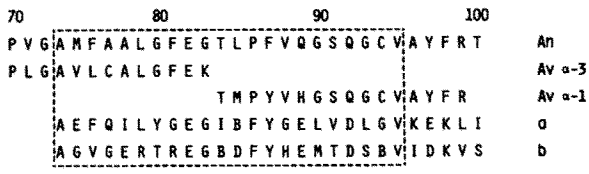

Fig.2. Similarities between amino acid sequences of nitrogenase MoFe-protein (β-subunit) and other adenine nucleotide-binding proteins. MoFe-proteins: An, *Anabaena* residues 71–100 [26]; Av α-3, *A. vinelandii* α-3 cysteinyl peptide [24]; Av α-1, *A. vinelandii* α-1 cysteinyl peptide [24]; a, *E. coli* Rec A protein residues 259–286 [23]; b, *E. coli* ATPase β-subunit residues 177–202, DCCD-sensitive region [27]. Amino acids are denoted by the single-letter amino acid code.

3. DISCUSSION

Gross similarities in amino acid composition showed that nitrogenase was similar to ATPase [33]. Sequence comparison now indicates that similar sequences are present in nitrogenase and the α- and β-subunits of ATPase. Similarities were also observed with other adenine nucleotide binding proteins. In these other proteins, some of these sequences have been identified as important in binding adenine nucleotides by crystallographic studies or by chemical modification of residues.

All Fe-proteins of nitrogenase contain at their N-terminal regions a sequence, similar to the N-terminal end of adenylate kinase, that forms a loop which is part of a domain important for binding AMP, but which under certain conditions can accommodate MnATP. The flexible properties of this loop in adenylate kinase are important in the conformational changes which occur in this protein [31,32]. Therefore, it is likely that this region forms part of an adenylate-binding domain in the Fe-protein.

Region B of nitrogenase Fe-proteins contains a sequence 22 residues long which is remarkably similar to residues 437–470 of the ATPase α-subunit, suggesting that this sequence might also be involved in forming the adenylate-binding domain of the Fe-protein. Region B is identical in all Fe-proteins except for that from *C. pasteurianum*. If this region is involved in nucleotide binding then the domain may be slightly different in this protein. Indeed, nitrogenase from this organism is less sensitive to inhibition by nucleotides other than ADP than the Fe-proteins from *K. pneumoniae* and *A. vinelandii* [29]. Adjacent to region B in the Fe-protein is a region (C) some 21 residues long which is most similar to residues 51–71 of the *E. coli* ATPase β-subunit and which in turn is similar to other ATP-binding proteins as reported in [9]. Region C is centred around Cys 97, which in the protein from *A. vinelandii* may provide a ligand for an Fe-S cluster [30]. Also, Cys 85 is involved with, or close to, the binding site of MgATP because it is protected by this nucleotide from carboxymethylation by iodoacetic acid. MgADP protects this residue but to a lesser degree [30]. Cys 85 lies between regions B and C. Thus there is experimental evidence to support the contention that regions B and C identified by sequence comparison may participate in nucleotide binding.

The 3 sequences identified here may contribute to the formation of a single domain. However, it is possible that MgATP and MgADP do not bind at equivalent sites, since these nucleotides exhibit complex competitive binding kinetics [2,4] which has led to the suggestion of a 'regulatory' site [4]. This could explain why so much sequence might be involved in nucleotide binding.

ATP hydrolysis occurs only after complex formation between the Fe- and the MoFe-proteins. The site of ATP hydrolysis is not known and may be associated with the Fe-protein but the absence of any ATPase activity in the purified component would argue that such a site would only be induced on complex formation. However, the site of hydrolysis might be on the MoFe-protein as low rates of labilization of the terminal phosphates of both ATP and ADP were observed for this component from *K. pneumoniae* [5]. The comparative sequence data support this view in that the *nifK* product (β-subunit of the MoFe-protein) from *Anabaena* and possibly *A. vinelandii* contains a sequence very similar to the DCCD-sensitive site of the β-subunit of ATPase and to a region of the Rec A protein of *E. coli* [28]. It has not been demonstrated that the DCCD-sensitive site is the site of ATP hydrolysis, however the glutamic acid residues modified by this reagent are thought to be important for binding MgATP [27]. Sequence data for other MoFe-proteins may confirm that this sequence has been highly conserved. A site on the MoFe-protein which would bind MgATP with low affinity, or contribute residues important for ATP

hydrolysis after complex formation, would be consistent with the suggested dual role for ATP in the mechanism of nitrogenase [18]. ATPase activity of nitrogenase can be blocked by *o*-phthalic aldehyde, which is thought to react with an NH_2 group present on the MoFe-protein but not the Fe-protein and this inhibition is partially eliminated in the presence of ATP [34]. It is also interesting to note the similarities between ATPase and nitrogenase. In both, ATP hydrolysis requires complex formation between the component proteins, one of which binds nucleotides strongly (Fe-protein, and the α-subunit of ATPase) but has no ATPase activity, and a second (of 3 components) which may contain residues important for hydrolysis but which does not provide a high-affinity site for nucleotides (MoFe-protein and β-subunit of ATPase).

There is no evidence from sequence comparison to suggest a possible site of ATP or ADP binding or hydrolysis on the α-subunit of the MoFe-protein, and sequence comparisons of subunits of these proteins indicate no general similarities between them.

ACKNOWLEDGEMENTS

I would like to thank Dr R. Eady and Professor J.R. Postgate for reading the manuscript and making many helpful suggestions. I would also like to thank Miss Beryl Scutt for typing, and Media Services, University of Sussex for photographic work.

REFERENCES

[1] Mortenson, L.E. and Thorneley, R.N.F. (1979) Annu. Rev. Biochem. 48, 387–418.
[2] Tso, M.Y.W. and Burris, R.H. (1973) Biochim. Biophys. Acta 309, 263–270.
[3] Hageman, R.V., Orme-Johnson, W.H. and Burris, R.H. (1980) Biochemistry 19, 2333–2342.
[4] Cordewener, J., Haaker, H. and Veeger, C. (1983) Eur. J. Biochem. 132, 47–54.
[5] Miller, R.W., Robson, R.L., Yates, M.G. and Eady, R.R. (1980) Can. J. Biochem. 58, 542–548.
[6] Kimber, S.J., Bishop, E.O. and Smith, B.E. (1982) Biochim. Biophys. Acta 705, 385–395.
[7] Eady, R.R., Imam, S., Lowe, D.J., Miller, R.W., Smith, B.E. and Thorneley, R.N.F. (1980) in: Nitrogen Fixation (Stewart, W.D.P. and Gallon, J.R. eds) pp.19–35, Academic Press, London.
[8] Robson, R.L., Kennedy, C.K. and Postgate, J.R. (1983) Can. J. Microbiol. 29, 954–966.
[9] Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) EMBO J. 1, 945–951.
[10] Hausinger, R.P. and Howard, J.B. (1982) J. Biol. Chem. 257, 2483–2490.
[11] Tanaka, M., Haniu, M., Yasonobu, K.T. and Mortenson, L.E. (1977) J. Biol. Chem. 252, 7093–7100.
[12] Sundaresan, V. and Ausubel, P.M. (1981) J. Biol. Chem. 256, 2808–2812.
[13] Scott, K.F., Rolfe, B.G. and Shine, J. (1981) J. Mol. Appl. Genet. 1, 71–81.
[14] Mevarech, M., Rice, D. and Haselkorn, R. (1980) Proc. Natl. Acad. Sci. USA 77, 6476–6480.
[15] Torok, I. and Kondorosi, A. (1981) Nucleic Acids Res. 9, 5711–5723.
[16] Scott, K.F., Rolfe, B.G. and Shine, J. (1983) DNA 2, 149–155.
[17] Scott, K.F., Rolfe, B.G. and Shine, J. (1983) DNA 2, 141–148.
[18] Smith, B.E., Thorneley, R.N.F., Eady, R.R. and Mortenson, L.E. (1976) Biochem. J. 157, 439–447.
[19] Kanazawa, H., Kayano, T., Mabuchi, K. and Futai, M. (1981) Biochem. Biophys. Res. Commun. 103, 604–612.
[20] Gay, N.J. and Walker, J.E. (1981) Nucleic Acids Res. 9, 2187–2194.
[21] Saraste, M., Gay, N.J., Eberle, A., Runswick, M.J. and Walker, J.E. (1981) Nucleic Acids Res. 9, 5287–5296.
[22] Heil, A., Müller, G., Noda, L., Pinder, T., Schirmer, H., Schirmer, I. and Von Zabern, I. (1974) Eur. J. Biochem. 43, 131–144.
[23] Horii, T., Ogawa, T. and Ogawa, H. (1980) Proc. Natl. Acad. Sci. USA 77, 313–317.
[24] Lundell, D.J. and Howard, J.B. (1981) J. Biol. Chem. 256, 6385–6391.
[25] Mazur, B.J. and Chui, C.-F. (1982) Proc. Natl. Acad. Sci. USA 79, 6782–6786.
[26] Lammers, P.J. and Haselkorn, R. (1983) Proc. Natl. Acad. Sci. USA 80, 4723–4727.
[27] Yoshida, M., Allison, W.S., Esch, F.S. and Futai, M. (1982) J. Biol. Chem. 257, 10033–10037.
[28] Futai, M. and Kanazawa, H. (1983) Microbiol. Rev. 47, 285–312.
[29] Weston, M.F., Kotake, S. and Davis, L.C. (1983) Arch. Biochem. Biophys. 225, 809–817.
[30] Hausinger, R.P. and Howard, J.B. (1983) J. Biol. Chem. 258, 13486–13492.
[31] Sachsenheimer, W. and Schulz, G.E. (1977) J. Mol. Biol. 144, 23–36.
[32] Pai, E.F., Sachsenheimer, W., Schirmer, R.H. and Schulz, G.E. (1977) J. Mol. Biol. 144, 37–45.
[33] Kleiner, D., Littke, W., Bender, H. and Wallenfels, K. (1976) J. Mol. Evol. 7, 159–165.
[34] Gvozdev, R.L., Sadkov, A.P. and Mitsova, I.Z. (1979) Biochem. SSR 44, 1696–1698.