Isolation of a cDNA for the rat heavy neurofilament polypeptide (NF-H)

Philip A. Robinson, Didier Wion* and Brian H. Anderton+

Department of Immunology, St. George's Hospital Medical School, Cranmer Terrace, London SW17 ORE, England

Received 16 October 1986

We have isolated from a rat brain λgtll expression library two overlapping cDNA clones of sizes 2.5 and 3.0 kb corresponding to the heavy neurofilament polypeptide (NF-H). The 2.5 kb insert apparently represents virtually the whole of the C-terminal tail, the 3.0 kb insert also encodes the conserved epitope for the monoclonal antibody, anti-IFA. The identity of the cDNAs was established by comparison of the predicted amino acid sequence with the known partial amino acid sequence of porcine NF-H. A repeat peptide sequence that may be a multiphosphorylation site was identified in the C-terminal non-helical tail.

Neurofilament (Rat brain) cDNA library Cytoskeleton

1. INTRODUCTION

In mammals, neurofilaments are composed of three major polypeptides with apparent M_r values of approx. 70000, 160000 and 200000 (NF-L, NF-M and NF-H, respectively) [1]. Neurofilament function and metabolic regulation are poorly understood but it has been found using specific antibodies that NF-L and NF-M appear in the immature neurone before NF-H [2–5]. Cloning of intermediate filament cDNA and genes will enable a greater understanding of their genetic regulation, both during development and in the fully differentiated organism. Recently, two laboratories have isolated cDNA probes for NF-L [6,7]. We now describe the isolation of cDNA probes for NF-H.

2. MATERIALS AND METHODS

2.1. *Preparation of a rat brain cDNA library*

Total rat brain poly(A^+) mRNA was isolated by a modified [8] LiCl/urea extraction procedure [9] and chromatography on poly(U)-Sepharose [10]. cDNA was prepared [11] and cDNA greater than 500 bp in length was inserted into the *Eco*RI site of the expression vector, λgt11 [12].

2.2. *Preparation of antisera and screening of the library*

Rabbit antisera to dephosphorylated NF-L, NF-M and NF-H [13] were prepared and stripped of anti-*E. coli* antibodies by incubating a 1:20 dilution of antiserum in foetal calf serum (FCS) with 0.05 vol. of an overnight culture of plating cells and 0.1 vol. of a lysed culture of BNN97 cells [12] for 16 h at 4°C. After centrifugation the supernatants were diluted 1:5 with FCS and used for screening the library [12], substituting protein A-peroxidase for iodinated protein A [14].

2.3. *Sequencing*

The inserts were subcloned into the *Eco*RI site of M13 mp8 and sequenced by the dideoxy chain termination method.

2.4. *Western blotting of fusion proteins*

Lysogens in *E. coli* strain Y1089 were grown to

+ To whom correspondence should be addressed
* Present address: INSERM U298, Faculté de Médecine, Rue Haute de Reculée, 49045 Angers Cédex, France

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

mid-log phase and fusion-protein production induced with IPTG [12]. The cell pellets after centrifugation were boiled in SDS sample buffer and analysed by Western blotting using NF-H monoclonal antibodies (mabs) RS18, 147 and 1215 and the pan intermediate filament mab, anti-IFA [15,16].

3. RESULTS AND DISCUSSION

A library of 5×10^6 independent plaques was generated and approx. 5×10^5 plaques screened. Two strongly positive plaques were cloned. The sizes of the inserts were determined by agarose gel electrophoresis to be 2.5 and 3.0 kb. The 2.5 kb insert was subcloned into the plasmid, pAT153, DNA prepared, nick-translated and used to probe a Southern blot of the two inserts [17]. Both gave strongly positive signals (not shown), indicating that they were similar. A Northern blot of rat brain mRNA with this same probe gave a strong band of 4.5 kb (not shown) [17].

Western blots of the fusion proteins generated by the two inserts showed that both were labelled by NF-H mabs, RS18, 147 and 1215 but that only the 3.0 kb fusion protein was labelled by the mab, anti-IFA (not shown). The epitope for anti-IFA is probably present in all intermediate filament structural polypeptides and is located at the highly conserved C-terminal end of α-helix 2 of the general intermediate filament structure [18]. The epitopes for the three neurofilament mabs are located in the non-helical C-terminal tail of NF-H [15]. Thus, the anti-IFA epitope appears to be encoded in the first 500 bp of the 5′-end of the 3.0 kb insert and the other three epitopes are encoded in the overlapping 3 ends of both inserts since sequence analysis showed that the 3′-ends both have poly(A^+) tails.

The labelling pattern with the monoclonal antibodies enabled us to predict that the 5′-end of the 3.0 kb insert would correspond approximately to helix 2 of NF-H which is a segment of known amino acid sequence for NF-L, NF-M and NF-H [6,18–20]. A comparison of a sequence of 61 amino acid residues revealed only two differences between that predicted from our rat cDNA sequence and the published amino acid sequence of porcine NF-H but there were 30 differences between the predicted rat sequence and porcine NF-M and 35 differences with rat NF-L (fig.1a).

```
NF-Lr  -    HEEEIAELQAQIQ-----YAQISVEMDVSSKPDLSAALKDIRAQYEKLA
NF-Mp 238   HEEEVADLLAQIQ-----ASHITVERKDYLKTDISSALKEIRSQLECHS
NF-Hp  -    HQEE_____________AQAEARDALKCDVTSALREIRAQLEGHA
NF-Hr  -    HQEEVGELLGQIQGCGAAQAQAQAEARDALKCDVTSALREIRAQLEGHT

NF-Lr  -    AKNMQNAEEWFKSRFTVLTESA A K NTDAVRAAKDEVSESR
NF-Mp 282   DQNMAQAEEWFKCRYAKLTEAA(Q,E)NKEAIRSAKEEIAEYR
NF-Hp  -    VQSTLQQEEWFRVRLDRLSEAA K V NTDAM_____________
NF-Hr  -    VQSTLQSEEWFRVRLDRLSEAA K V NTDAMRSAQEEITEYR
```

Fig. 1a.

```
NF-L     CACGAGGAAGAGATCGCCGAGCTGCAGGCTCAGATCCAG
NF-H     CACCAGGAGGAGGTGGGCGAGCTGCTCGGTCAGATTCAGGGCTGC

NF-L              TATGCTCAGATCTCCGTGGAGATGGACGTGTCCTCC
NF-H     GGTGCCGCGCAGGCGCAGGCTCAGGCCGAGGCTCGGGACGCCCTC

NF-L     AAGCCCGACCTCTCCGCCGCTCTCAAGGACATCCGCGCTCAGTAC
NF-H     AAGTGCGACGTGACGTCGGCGCTGCGGGAGATCCGCGCGCAGCTC

NF-L     GAGAAGCTGGCCGCCAAGAATATGCAGAATGCCGAAGAGTGGTTC
NF-H     GAAGGACACACGGTGCAGAGTACGCTGCAGTCAGAGGAGTGGTTC

NF-L     AAGAGCCGCTTCACGGTGCTAACCGAGAGCGCCGCCAAGAACACC
NF-H     CGAGTGAGATTGGACCGACTCTCAGAGGCAGCCAAAGTGAACACG

NF-L     GACGCAGTGCGCGCTGCCAAGGACGAGGTGTCGGAAAGCCGC
NF-H     GATGCTATGCGCTCTGCCCAAGAGGAGATAACTGAGTACCGG
```

Fig. 1b.

```
  K  E  A  Q  G  E  E  E  E  E  A  E  E  G  G  E  E  A
AAAGAGGCCCAAGGGGAGGAAGAAGAAGAGGCAGAAGAGGGAGGAGAAGAAGCA

  A  T  T  S  P  P  A  E  E  A  A  S  P  E  K  E  T  K
GCAACTACGTCTCCCCCTGCAGAAGAGGCTGCATCTCCAGAAAAGGAAACCAAG

  S  P  V  K  E  E  A  K  S  P  A  E  A  K  S  P  A  E
TCTCCTGTGAAAGAAGAGGCCAAGTCCCCAGCTGAGGCCAAGTCCCCAGCTGAG

  A  K  S  P  A  E  A  K  S  P  A  E  V  K  S  P  A
GCCAAGTCACCAGCTGAGGCCAAGTCCCCAGCTGAGGTCAAATCTCCAGCT
```

Fig. 1c.

Fig.1. (a) 5′-end nucleotide and deduced amino acid sequences of rat 3.0 kb NF-H cDNA compared with the known partial amino acid sequences of porcine NF-H and NF-M [18–20] and the deduced sequence of rat NF-L [6]. r, rat; p, porcine; – – –, deletions allowing for better alignment of homologous sequences; ——, regions of unknown sequence; non-homologous residues are underlined in NF-Hr; amino acid residue position is given in column 2. (b) Rat NF-L [6] and NF-H DNA homology; the gap enables maximum alignment for homology and corresponds to the extra 5 amino acids in NF-H. (c) 5′-end DNA and deduced amino acid sequences of the 2.5 kb NF-H cDNA.

Hence, when comparing the same set of amino acid residues, the 3.0 kb insert is seen to be an NF-H cDNA. Limited DNA homology between rat NF-H and rat NF-L was found in this region (fig.1b). There also appears to be an additional five amino acids in the rat NF-H inserted immediately after the end of helix 1b corresponding to residue 251 in porcine NF-M (fig.1a); the protein sequence for this segment in porcine NF-H has not been reported.

We postulate that the 5′-end of the 2.5 kb cDNA (fig.1c) probably corresponds to the stretch

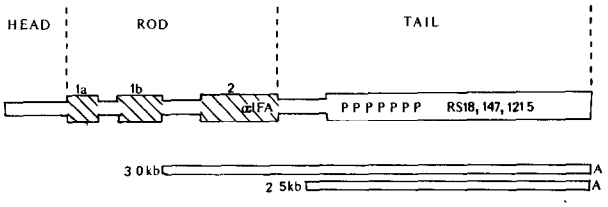

Fig.2. Relative positions of overlapping fragments of NF-H cDNAs in relation to the proposed protein secondary structure. Cross-hatched areas 1a, 1b and 2 are the α-helical domains; P, phosphorylation site; α-IFA, RS18, 147 and 1215 are the locations of the epitopes for these mabs; A, poly(A) tail.

of the NF-H C-terminal tail that is known to be heavily phosphorylated [21]. The sequence of this region reveals a peptide repeat unit -Glu.(Ala,Val).Lys.Ser.Pro.Ala- that is repeated at least five times (fig.1c). Four of these units contain alanine as the second amino acid and the other contains valine; this represents a conservative difference. Centrally placed in this repeat unit is a serine residue in a charged environment. This could well represent a phosphorylation site and thus the repeat of this sequence may form a multiphosphorylation site in NF-H.

Fig.2 shows the probable relationship of the two cDNAs to the known protein sequences, secondary structure motif and epitope and phosphorylation site locations of NF-H [18–20].

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council and The Wellcome Trust. We thank Therese Loviny for invaluable advice on DNA sequencing.

REFERENCES

[1] Traub, P. (1985) Intermediate Filaments, Springer, Berlin.
[2] Calvert, R. and Anderton, B.H. (1982) FEBS Lett. 145, 171–175.
[3] Pachter, J.S. and Liem, R.K.H. (1984) Dev. Biol. 103, 200–210.
[4] Shaw, G. and Weber, K. (1982) Nature 298, 277–279.
[5] Willard, M. and Simon, C. (1983) Cell 35, 551–559.
[6] Julien, J.-P., Ramachandran, K. and Grosveld, F. (1985) Biochim. Biophys. Acta 825, 398–404.
[7] Lewis, S.A. and Cowan, N.J. (1985) J. Cell Biol. 100, 843–850.
[8] Darmon, M.C. and Paulin, D.J. (1985) J. Neurochem. 44, 1672–1678.
[9] Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303–314.
[10] Shapiro, D.J. and Schimke, R.T. (1975) J. Biol. Chem. 250, 1759–1764.
[11] Watson, C.J. and Jackson, J.F. (1985) in: DNA Cloning. A Practical Approach, vol.1 (Glover, D.M. ed.) pp.79–88, IRL Press, Oxford.
[12] Huynh, T.V., Young, R.A. and Davis, R.W. (1985) in: DNA Cloning. A Practical Approach, vol.1 (Glover, D.M. ed.) pp.49–78, IRL Press, Oxford.
[13] Carden, M.J., Schlaepfer, W.W. and Lee, V.M.-Y. (1985) J. Biol. Chem. 260, 9805–9817.
[14] Hawkes, R., Niday, E. and Gordon, J. (1982) Anal. Biochem. 119, 142–147.
[15] Miller, C.C.J., Brion, J.-P., Calvert, R., Chin, T.K., Eagles, P.A.M., Downes, M.J., Flament-Durand, J., Haugh, M., Kahn, J., Probst, A., Ulrich, J. and Anderton, B.H. (1986) EMBO J. 5, 269–276.
[16] Pruss, R.M., Mirsky, R., Raff, M.C., Thorpe, R., Dowding, A.J. and Anderton, B.H. (1981) Cell 27, 419–428.
[17] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, NY.
[18] Geisler, N., Kaufmann, E., Fischer, S., Plessmann, U. and Weber, K. (1983) EMBO J. 2, 1295–1302.
[19] Geisler, N., Fischer, S., Vandekerckhove, J., Plessmann, U. and Weber, K. (1984) EMBO J. 3, 2701–2706.
[20] Geisler, N., Fischer, S., Vandekerckhove, J., Van Damme, J., Plessmann, U. and Weber, K. (1985) EMBO J. 4, 57–63.
[21] Julien, J.-P. and Mushynski, W.E. (1983) J. Biol. Chem. 258, 4019–4025.