Biochem. J. (1991) 278, 365–368 (Printed in Great Britain)



# Synthesis of a gene for rat liver fatty-acid-binding protein and its expression in *Escherichia coli*

Andrew F. WORRALL, Carol EVANS and David C. WILTON*
Department of Biochemistry and SERC Centre for Molecular Recognition, University of Southampton, Southampton SO9 3TU, U.K.

A gene coding for rat liver fatty-acid-binding protein (FABP) has been constructed by the single-step ligation of ten synthetic oligonucleotides. The gene has been cloned into the bacterial expression vector pKK223-3. Induction of protein synthesis from this gene results in over expression of an FABP that is indistinguishable in its structure and binding properties with that isolated from rat liver.

## INTRODUCTION

Rat liver fatty-acid-binding protein (FABP) belongs to a family of genetically related low-molecular-mass cytosolic proteins that bind fatty acid and/or other non-polar ligands (Kaikaus *et al.*, 1990). All these proteins are presumed to enhance the uptake of their respective ligands from the extracellular medium and to facilitate ligand transport and metabolism within the cell. However, their precise physiological significance remains obscure.

Ten members of this family have been identified to date, which include liver FABP (L-FABP), heart FABP ('H-FABP'), intestinal FABP (I-FABP), adipocyte lipid-binding protein ('ALBP'), cytosolic retinol binding proteins ('CRBP') I and II, cytosolic retinoic acid binding protein ('CRABP'), gastrophin, mammary-derived growth inhibitor and P2 myelin protein (Kaikaus *et al.*, 1990). The most studied protein in terms of ligand-binding properties is L-FABP; however, only the high resolution crystal structures for the P2 protein from myelin (Jones *et al.*, 1988) and I-FABP (Sacchettini *et al.*, 1988, 1989*a*) are available. Several of the proteins have been cloned and expressed in *Escherichia coli*, and this has provided a source of protein for both X-ray crystal-structure determinations and n.m.r. studies.

The crystal structures have revealed a characteristic $\beta$-clam structure, and this basic feature is shared by other small, but unrelated, non-polar ligand-binding proteins, including serum retinol-binding protein (Newcomer *et al.*, 1984) and a bacterial photoreceptor protein (McRee *et al.*, 1989). In the case of I-FABP the detailed structure of both the apo (Sacchettini *et al.*, 1989*b*) and the holo form of the protein with bound palmitate (Sacchettini *et al.*, 1989*a*) is available and has revealed that the rigid $\beta$-clam structure is maintained in the absence of ligand and that the ligand is replaced in the central hydrophobic cavity by six ordered water molecules in the apoprotein.

We have been studying the L-FABP for several years and have successfully made use of a fluorescence approach using the fluorescent fatty acid analogue 11-(dansylamino)undecanoic acid (DAUDA). This fluorescent probe has the unique advantage of binding to L-FABP with a 60-fold fluorescence enhancement, which makes it very effective for detecting, quantifying and studying the ligand-binding properties of the protein (Wilkinson & Wilton, 1986, 1987*a*,*b*; Sheridan *et al.*, 1987). In addition, we have used chemical modification to probe the nature of the ligand-binding site of L-FABP (Wilton, 1989; Evans & Wilton, 1990). However, potentially the most important benefit from the use of the fluorescence approach to FABP structure and function has been the development of a continuous fluorescence displacement assay for enzymes that directly or indirectly release long-chain fatty acids (Wilton, 1990, 1991). The assay may also be used to quantify subnanomolar amounts of substrates for fatty-acid-releasing enzymes (Wilton, 1991).

The potential interest in rat liver FABP, both as a well-defined protein for studying the molecular recognition between ligand and protein and also as the basis for fatty acid-linked assays, made the protein an obvious target for gene cloning and overexpression in *E. coli*. Although the rat liver FABP had previously been cloned (Lowe *et al.*, 1984) we decided that, because of its small size (14 kDa), the protein was particularly amenable to an approach involving total-gene synthesis. Several genes have been constructed by total-gene synthesis (Groger *et al.*, 1988), and this technique is now well established. In particular, we have found it to be a rapid and efficient method for the cloning of proteins for which the amino acid sequence has been determined (Worrall & Connolly, 1990).

The present paper describes the synthesis of a gene which codes for rat liver FABP in which the DNA sequence is derived from its amino acid sequence. We demonstrate that this protein may be overexpressed in *E. coli* by the insertion of the gene in the expression vector pKK223-3 (Brosius & Holy, 1984), which carries the isopropyl $\beta$-D-thiogalactoside (IPTG)-inducible *tac* promoter. A comparison of various properties of the protein from rat liver and *E. coli* is also described and established that both proteins are structurally and functionally identical.

## EXPERIMENTAL

All restriction endonucleases and DNA-modifying enzymes were obtained from Boehringer-Mannheim (Lewes, East Sussex, U.K.) or Northumbria Biologicals Ltd. (Cramlington, Northumberland U.K.) and were used according to the manufacturers' protocols. *E. coli* strains used were: TG1 (Carter *et al.*, 1985); Δ(*lac-proAB*), *supE*, *thi*, *hsd*Δ5, *F'* [*traD*36, *proAB*, *lacI*$^q$, *lacZ*ΔM15]), JM105 (Yanisch-Perron *et al.*, 1985, Δ(*lac-proAB*), *thi*, *strA*, *endA*, *sbcB*15, *hsdR*4, *F'* [*traD*36, *proAB*, *lacI*$^q$*Z*Δ*M*15]. pKK223-3 was obtained from Pharmacia (Uppsala, Sweden) and M13mp19 from Boehringer-Mannheim. Bacteria containing plasmids were grown in Luria–Bertani (LB) broth (Maniatis *et al.*, 1982) containing ampicillin (50 $\mu$g/ml) at 37 °C. All

Abbreviations used: FABP, fatty-acid-binding protein (I-, intestinal; L-, liver); DAUDA, 11-(5-dansylamino)undecanoic acid; DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid); IPTG, isopropyl $\beta$-D-thiogalactoside; NAG, naphthoylaminodecyl-agarose; LB, Luria–Bertani.

* To whom correspondence should be sent.

standard DNA manipulations and electrophoresis procedures were carried out as previously described (Maniatis *et al.*, 1982).

The gene was designed (Fig. 1*a*) with the help of commercially available software (Unger, 1986). Oligonucleotides were synthesized on a 0.2 micromolar scale on an Applied Biosystems 381A DNA synthesizer using standard phosphoramidite chemistry. They were purified as previously described (Worrall & Connolly, 1990). Oligonucleotides were 5′-phosphorylated on the DNA synthesizer by using a solid-phase chemical phosphorylating reagent (Cruachem Ltd., Glasgow, Scotland, U.K.). Oligonucleotides 1 and 10 (Fig. 1*a*) were not phosphorylated, since they carry the self-complementary 5′-overlaps of the complete gene. This prevents concatemer formation of the gene during ligation. Annealing and ligation of the ten oligonucleotides was achieved in a single step, and the resultant crude gene fragment was used directly for cloning into M13mp19, previously cut with *Eco*R1 and *Hin*dIII.

This recombinant M13 DNA was used to transform *E. coli* TG1, which were then grown on H-plates (containing, per litre: Bactotryptone, 10 g; NaCl, 8 g; agar, 12 g) in the presence of IPTG and 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside as an indicator for recombinant phage. Double-stranded DNA was prepared from white plaques (Messing, 1983), and this DNA was cleaved with *Eco*RI and *Hin*dIII in an attempt to reisolate the 414-bp fragment. Where successful, the plaque was used to prepare single-stranded DNA (Messing, 1983) for use as a template in a dideoxy sequencing (Sanger *et al.*, 1977) using Sequenase enzyme (U.S. Biochemical Corp.).

Mutagenesis of the Shine–Dalgarno (1974) sequence of the gene was carried out by cleaving M13 double-stranded DNA with *Eco*R1 and *Sac*1 and re-ligation with a synthetic duplex to reform the circular DNA. After this experiment, the resulting clone was resequenced through the whole length of the new insert, including the re-formed restriction sites. Subcloning of the gene into the expression vector pKK223-3 was achieved by ligating the 414-bp fragment from the recombinant M13 into the vector previously cut with the same restriction enzymes and treated with calf intestinal phosphatase. The resultant plasmid was used to transform *E. coli* JM105, which were then grown on LB agar containing 50 μg of ampicillin/ml.

Large-scale cultures (1–5 litres) in LB broth were grown to an $A_{550}$ of 0.3–0.5 at 37 °C, then induced with IPTG (2 mM) for 4–5 h. After harvesting, the cells were resuspended in 0.1 M-potassium phosphate buffer, pH 7.4, containing 5 mM-EDTA and sonicated until clear. Purification of the cell-free supernatant was by modifications of the method of Wilton (1989). An initial $(NH_4)_2SO_4$ fractionation was performed, first to 50% saturation, then, after centrifugation, to 70% saturation. The supernatant, after dialysis against 0.1 M-potassium phosphate buffer, pH 7.4, was applied to a naphthoylaminodecyl-agarose (NAG) column (Wilton, 1989). The purified protein from the NAG affinity column was, after concentration, subjected to a final fractionation on a column (26 mm × 1000 mm) of Sephadex G-75, which also removed the ethanol required for elution from the NAG affinity column. Typically 10 mg of pure protein was obtained from 1 litre of culture.

DAUDA-binding studies were performed as previously described (Wilkinson & Wilton, 1986, 1987*b*), as were the 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) studies (Wilton, 1989).

The sequence of the *N*-terminus of the FABP was determined by using an Applied Biosystems 477A protein sequencer.

## RESULTS AND DISCUSSION

### Gene synthesis and expression of *E. coli*

We have designed a synthetic gene which codes for the rat liver FABP (Fig. 1). We have used *E. coli*-preferred codons in this design (Grosjean & Fiers, 1982) wherever possible, and have inserted unique restriction sites in the gene by utilizing the redundancy in the genetic code, to enable simple gene manipulation in the future. We designed the gene with its own Shine–Dalgarno sequence (Fig. 1*a*) taken from the highly expressed A protein of phage R17 (Shine & Dalgarno, 1974). Cloning of the 414-bp synthetic DNA fragment into M13mp19 yielded many recombinant phages from which the fragment could be re-isolated. Only one was sequenced, and this proved to have a single point mutation in which the codon for L-50 had changed from the designed CTG to TTG. Since TTG is still a leucine codon, we did not alter the sequence isolated. This M13 clone was designated 'M13mp19FABP1'. Subcloning this gene into pKK223-3 yielded a new plasmid designated 'pFABP1'.

Protein-induction experiments using pFABP1 produced amounts of FABP equivalent to about 10 mg/litre of original culture. As a result of these protein-induction experiments it was decided to try to increase the yield of FABP from the system by altering the Shine–Dalgarno sequence and its position relative to the ATG start codon of the gene. The sequence and spacing chosen (Fig. 1*b*) was that which occurs in the translation start

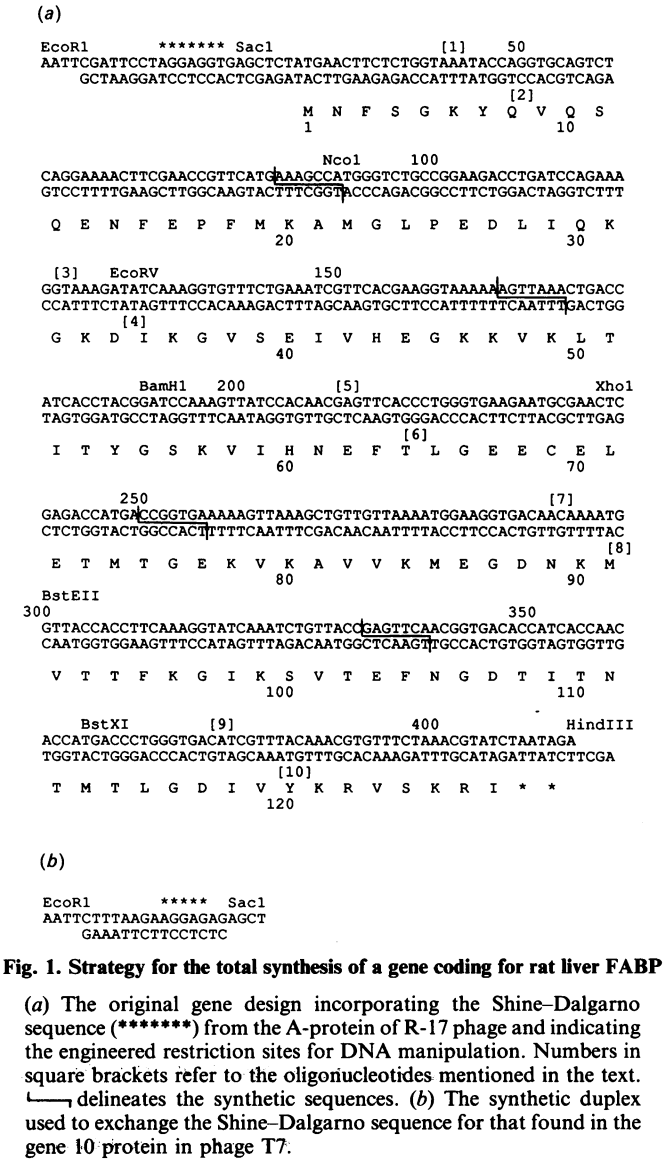


**Fig. 1. Strategy for the total synthesis of a gene coding for rat liver FABP**

(*a*) The original gene design incorporating the Shine–Dalgarno sequence (*******) from the A-protein of R-17 phage and indicating the engineered restriction sites for DNA manipulation. Numbers in square brackets refer to the oligonucleotides mentioned in the text. ⌊⌋ delineates the synthetic sequences. (*b*) The synthetic duplex used to exchange the Shine–Dalgarno sequence for that found in the gene 10 protein in phage T7.

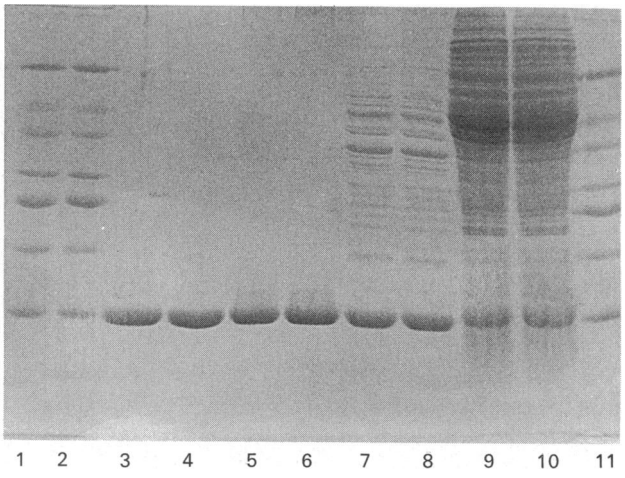


**Fig. 2. SDS/PAGE of FABP during purification from *E. coli***

Lanes 1, 2 and 11, molecular-mass standards (66, 50, 45, 36, 29, 20 and 14 kDa); lanes 3 and 4, FABP eluted from Sephadex G-75; lanes 5 and 6, protein eluted from a NAG column with 25% ethanol; lanes 7 and 8, supernatant after precipitation with 70%-satd. $(NH_4)_2SO_4$; lanes 9 and 10, soluble fraction from *E. coli*.

**Table 1. Comparison of the DAUDA-binding properties of hepatic FABP isolated from rat liver and *E. coli***

The DAUDA-binding properties of FABP from rat liver and *E. coli* were measured in 0.1 M-potassium phosphate buffer, pH 7.4. Wavelength maxima and fluorescence-intensity measurements involved 2 $\mu$M-FABP and 1 $\mu$M-DAUDA. $K_d$ and stoichiometry measurements involved 1 $\mu$M-FABP and 0.25–2.5 $\mu$M-DAUDA. Fluorescence measurements were made at room temperature and are uncorrected.

| Parameter | Rat liver FABP | *E. coli* FABP |
|---|---|---|
| Wavelength of fluorescence emission maximum (nm) | 498 | 498 |
| Fluorescence intensity/nmol of FABP (arbitrary units) | 330 | 329 |
| Dissociation constant ($K_d$) of DAUDA binding ($\mu$M) | $0.06 \pm 0.02$ | $0.10 \pm 0.02$ |
| Stoichiometry of DAUDA binding to FABP (mol/mol) | 0.95 | 1.14 |

sequence for the gene-10 protein of phage T7 (Olins *et al.*, 1988). This sequence is reported as being useful for enhancing expression of foreign genes in *E. coli*. Thus M13mp19FABP1 was cleaved with *Eco*R1 and *Sac*1, and a replacement synthetic duplex reformed the phage DNA, now designated 'M13mp19FABP2'. Subcloning as described above yielded the construct pFABP2.

Induced cultures of *E. coli* containing pFABP2 were harvested and sonicated cell-free extracts assayed for rat liver FABP using the DAUDA-binding assay (Wilkinson & Wilton, 1986). After the optimal induction period (4–5 h), functional FABP was detected at a concentration equivalent to 20 mg/litre of original culture, whereas essentially no fluorescence enhancement could be detected in the supernatant from cultures containing the parent plasmid pKK223-3. Similarly, uninduced cultures of JM105[pFABP2] yielded no fluorescence enhancement in this assay. After cell disruption by sonication, analysis of the cell-free supernatant by gel-permeation f.p.l.c. revealed a single DAUDA-binding fluorescent peak at a position corresponding to that of authentic rat liver FABP.

The *E. coli*-derived FABP was purified from the cell-free supernatant and a final yield of pure protein equivalent to about 10 mg/litre of original culture was obtained. SDS/PAGE analysis of the protein after various stages of purification are shown in Fig. 2. The purification procedure involves the use of NAG in a hydrophobic affinity column, and this step should also remove any fatty acid. This yield of protein compares very favourably with that reported using the plasmid pJBL2, where approx. 1 mg of purified liver FABP was obtained from 1 litre of culture (Lowe *et al.*, 1987). We also obtained relatively low expression of protein using pJBL2 (C. Evans & D. C. Wilton, unpublished work).

### Properties of the purified FABP from *E. coli*

In order to confirm that the *E. coli*-derived FABP was identical with that isolated from rat liver, a comparison of the binding properties and other parameters related to structure was performed. We have already used the fluorescent fatty acid probe DAUDA to investigate the binding properties of the protein from rat liver (Wilkinson & Wilton, 1986, 1987*b*), and the results summarized in Table 1 confirm that the *E. coli*-derived protein shows binding properties identical with those of this protein. Thus the fluorescence intensity and wavelength of maximum emission were identical for both proteins, as was the $K_d$ for DAUDA binding [as determined by the method of Cogan *et al.* (1976)]. Similarly DAUDA was displaced by long-chain fatty acids with both proteins (results not shown), confirming the use of the *E. coli*-derived FABP in the fluorescence displacement assay for lipases. In addition, both proteins showed similar DTNB-reactivity with the expected one thiol, cysteine-69, per molecule.

The *N*-terminal six residues of the protein were sequenced and, like the cloned protein that has been expressed in *E. coli* (Lowe *et al.*, 1984), the terminal methionine residue was still present and unblocked, whereas the protein from rat liver is blocked by acetylation at the *N*-terminal (Takahashi *et al.*, 1982).

In conclusion, total gene synthesis provided a very rapid and convenient method for overexpressing rat liver FABP in *E. coli*, and the purified protein could be obtained in high yield. The protein that is expressed in *E. coli* from the synthetic gene appears to be structurally and functionally identical with the protein from rat liver, apart from having a free *N*-terminal methionine residue. Therefore this *E. coli* system will provide a convenient source of FABP, both for structure–functions studies, using site-directed mutagenesis, and for use in the fluorescent displacement assays involving the release of long-chain fatty acids.

Protein sequencing was performed by the protein sequencing unit of the SERC Centre for Molecular Recognition at Southampton. FABP gene synthesis and expression was supported by a grant from the British Technology Group. Financial support from the Wellcome Trust is gratefully acknowledged.

## REFERENCES

Brosius, J. & Holy, A. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 6929–6933
Carter, P., Bedouelle, H. & Winter, G. P. (1985) Nucleic Acids Res. **13**, 4431–4443
Cogan, U., Kopelman, M., Mokady, S. & Shinitzky, M. (1976) Eur. J. Biochem. **65**, 71–78
Evans, C. & Wilton, D. C. (1990) Mol. Cell. Biol. **98**, 135–140
Groger, G., Ramalho-Ortigao, F., Steil, H. & Seliger, H. (1988) Nucleic Acids Res. **16**, 7763–7771
Grosjean, H. & Fiers, W. (1982) Gene **18**, 199–209
Jones, T. A., Bergfors, T., Sedzik, J. & Unge, T. (1988) EMBO J. **7**, 1594–1604
Kaikaus, R. M., Bass, N. M. & Ockner, R. K. (1990) Experientia **46**, 617–630

Lowe, J. B., Strauss, A. W. & Gordon, J. I. (1984) J. Biol. Chem. **259**, 12696–12704
Lowe, J. B., Sacchettini, J. C., Laposota, M., McQuillan, J. J. & Gordon, J. I. (1987) J. Biol. Chem. **262**, 5931–5937
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
McRee, D. E., Tainer, J. A., Meyer, T. E., Van Beeumen, J., Cusanovich, M. A. & Getzoff, E. D. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 6533–6537
Messing, J. (1983) Methods Enzymol. **101**, 20–78
Newcomer, M. E., Jones, T. A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P. A. (1984) EMBO J. **3**, 1451–1454
Olins, P. O., Devine, C. S., Rangwala, S. H. & Kavka, K. S. (1988) Gene **73**, 227–235
Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. (1988) J. Biol. Chem. **263**, 5815–5819
Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. (1989*a*) J. Mol. Biol. **208**, 327–339
Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. (1989*b*) Proc. Natl. Acad. Sci. U.S.A. **86**, 7736–7740
Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Sheridan, M., Wilkinson, T. C. I. & Wilton, D. C. (1987) Biochem. J. **242**, 919–922
Shine, J. & Dalgarno, L. (1974) Proc. Natl. Acad. Sci. U.S.A. **71**, 1342–1346
Takahashi, K., Odani, S. & Ono, T. (1982) FEBS Lett. **140**, 63–66
Unger, S. H. (1986) Recombinant Toolkit, Version 2.0, Elsevier Scientific Publishers, Amsterdam
Wilkinson, T. C. I. & Wilton, D. C. (1986) Biochem. J. **238**, 419–424
Wilkinson, T. C. I. & Wilton, D. C. (1987*a*) Biochem. J. **242**, 913–917
Wilkinson, T. C. I. & Wilton, D. C. (1987*b*) Biochem. J. **247**, 485–488
Wilton, D. C. (1989) Biochem. J. **261**, 273–276
Wilton, D. C. (1990) Biochem. J. **266**, 435–439
Wilton, D. C. (1991) Biochem. J. **276**, 129–133
Worrall, A. F. & Connolly, B. A. (1990) J. Biol. Chem. **265**, 21889–21895
Yanisch-Perron, C., Viera, J. & Messing, J. (1985) Gene **33**, 103–119

Received 29 January 1991/19 March 1991; accepted 4 April 1991