J. Mol. Biol. (1993) 230, 725–732

# Functional Expression of the Uncomplexed Serum Retinol-binding Protein in *Escherichia coli*

## Ligand Binding and Reversible Unfolding Characteristics

**Holger N. Müller and Arne Skerra†**

*Max-Planck-Institut f. Biophysik, Heinrich-Hoffmann-Str. 7*
*W-6000 Frankfurt am Main 71, Germany*

*(Received 25 September 1992; accepted 24 December 1992)*

The serum retinol-binding protein solubilizes the lipophilic vitamin A alcohol and plays an important physiological role in the transport of this compound. The monomeric single-domain protein, the three-dimensional structure of which is known, constitutes a well-characterized member of the *lipocalin* family of proteins. We report here the functional expression of the apo-protein in *Escherichia coli* by secretion to the periplasm. The recombinant protein, purified in a single step by metal chelate affinity chromatography, exhibits the same ligand binding characteristics as described for the natural protein. Guanidinium chloride-induced unfolding and refolding experiments suggest that the recombinant retinol-binding protein adopts a stable conformation despite being expressed and purified in the absence of the large hydrophobic ligand. The expression system described here should also be useful for the recombinant production of other *lipocalin* proteins, thus permitting the elucidation of the structure–function relationships of ligand binding by protein engineering.

*Keywords:* RBP; *lipocalin* family; *E. coli* secretion; retinoic acid; protein folding

The serum retinol-binding protein (RBP‡) is responsible for the transport of vitamin A from the site of its storage in the liver to various peripheral target tissues (Blomhoff *et al.*, 1990). RBP serves both to solubilize the lipophilic alcohol retinol and protect this oxygen-sensitive compound against chemical degradation. RBP is synthesized in the liver hepatocyte from which it is, after uptake of the ligand, secreted to the plasma. While circulating in the blood stream, the holo-RBP is found bound to transthyretin. After delivery of retinol, this complex dissociates and the RBP becomes filtered through the kidney glomeruli and finally degraded.

Serum RBP is a monomeric protein with a molecular mass of approximately 21,000 Da and possesses a single binding site for its physiological ligand retinol. It is found in the blood plasma of all mammals (Goodman, 1984) and of various lower vertebrate species, as well (Shidoji & Muto, 1977; McKearing *et al.*, 1987; Berni *et al.*, 1992). Due to its important physiological role in the transport of the vitamin A alcohol, particularly the retinol-binding protein from mammalian sources has been subject to extensive biochemical studies (reviewed by Goodman, 1984). The cDNAs for the RBPs from man (Colantuoni *et al.*, 1983), pig (Trout *et al.*, 1991), and rat (Sundelin *et al.*, 1985) have been described.

The X-ray structure of the human serum RBP with bound vitamin A was solved and refined to high resolution (Newcomer *et al.*, 1984; Cowan *et al.*, 1990). RBP essentially consists of an eight-stranded antiparallel $\beta$-barrel which forms a central hydrophobic binding pocket for the ligand protruding deeply into the core of the protein structure. RBP represents the first member of a novel class of proteins termed the *lipocalin* family (Pervaiz & Brew, 1987). Although the proteins belonging to this family are only distantly related at the amino acid sequence level, it was established by several independent X-ray analyses that they share the highly conserved $\beta$-barrel folding motif (Papiz *et al.*,

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: RBP, retinol-binding protein; *Omp*A, outer membrane protein A; PCR, polymerase chain reaction; IPTG, isopropyl-$\beta$-D-thio-galactopyranoside; PAGE, polyacrylamide gel electrophoresis; PDI, protein disulfide isomerase; PPI, peptidyl-prolyl *cis/trans* isomerase; Gdn·HCl, guanidinium chloride; $\lambda_{Ex}$, excitation wavelength; $\lambda_{Em}$, emission wavelength; $\lambda_{Em,Max}$, maximum emission wavelength; DMF, dimethylformamide.



0022-2836/93/070725-08 $08.00/0

**OmpA coding sequence on vector:**

```
                                                                    StuI
...ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCCGGCTTCGCTACCGTAGCGCAGGcct...
   ---------+---------+---------+---------+---------+---------+----
...TACTTTTTCTGTCGATAGCGCTAACGTCACCGTGACCGgCCgAAGCGATGGCATCGCGTCCcca...

   MetLysLysThrAlaIleAlaIleAlaValAlaLeuAlaGlyPheAlaThrValAlaGlnAla
   -21                                                          -1
```

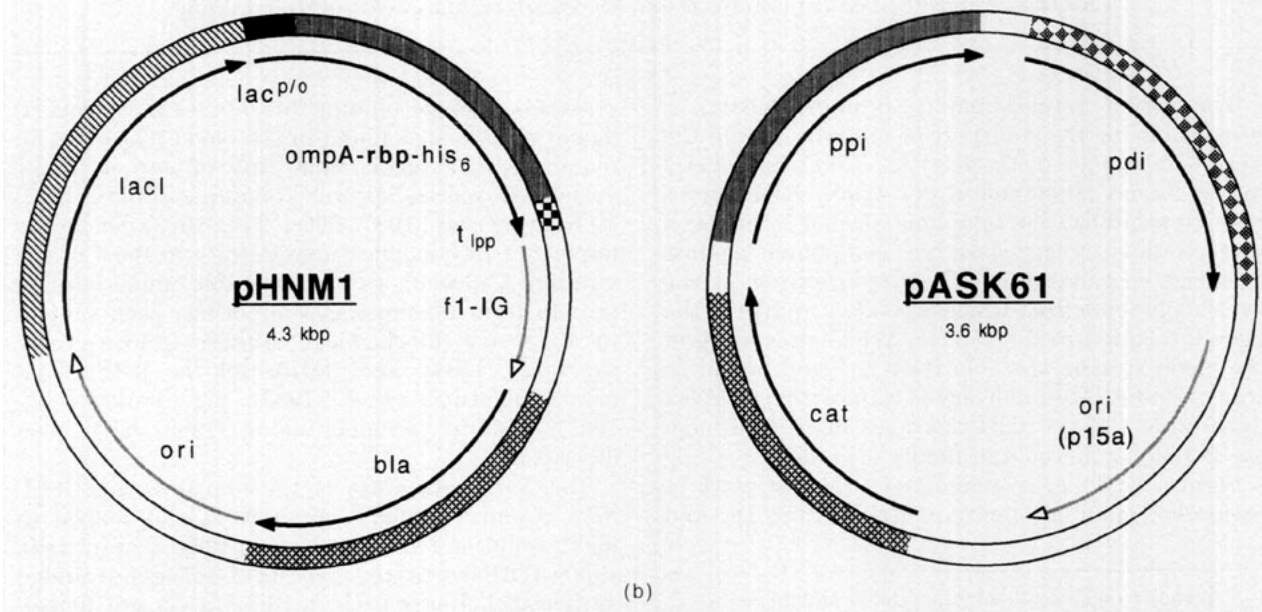
**5'-end of the coding sequence for the porcine RBP:**

```
                                          PvuII
                                    ======CAGCTG=========>
...ATGAAGTGGGTGTGGGCGCTCTTGCTGTTGGCGGCGTGGGCAGCGGCCGAGCGCGACTGCCGAGTGAGCAGCTTCCGAGTCAAGGAGAAC...
   ---------+---------+---------+---------+---------+---------+---------+---------+---------
...TACTTCACCCACACCCGCGAGAACGACAACCGCCGCACCCGTCGCCGGCTCGCGCTGACGGCTCACTCGTCGAAGGCTCAGTTCCTCTTG...

   MetLysTrpValTrpAlaLeuLeuLeuLeuAlaAlaTrpAlaAlaAlaGluArgAspCysArgValSerSerPheArgValLysGluAsn
   -16                                                  +1
```

**3'-end of the coding sequence:**

```
...ATCACTCACAATGGTTACTGTGATGGCAAATCGGAAAGAAACATTTTGTAGCAACATGG...
   --------+---------+---------+---------+---------+---------+
...TAGTGAGTGTTACCAATGACACTACCGTTTAGCCTTTCTTTGTAAAACATCGTTGTACC...
            <=======G=======C===T=========G=====GTAGTGGTGGTAGTGGTAATTCGAACTAGT
                    A       T   C                                  HindIII

   IleThrHisAsnGlyTyrCysAspGlyLysSerGluArgAsnIleLeuEnd
                            Arg          Leu   HisHisHisHisHisHisEnd (amino acid sequence
                                          183                        encoded on the primer)
```

(a)




(b)

**Figure 1.** Construction of an expression vector for the retinol-binding protein. (a) Cloning of the cDNA from pig. The previously described cDNA regions (Trout *et al.*, 1991) encoding the amino terminus and the carboxy terminus of the porcine RBP are shown together with the two PCR primers used for amplification of the structural gene (indicated above or below the corresponding cDNA sequence; with deviating nucleotides indicated an otherwise perfect match). The first primer anneals to the coding region for the amino terminus of the mature protein, thereby introducing a *Pvu*II restriction site. The blunt end obtained after restriction digest of the PCR product is compatible to the *Stu*I-generated blunt end at the 3′ terminus of the coding sequence for the OmpA signal peptide on the vector (shown above, blunt end after AGG). Thus, the original leader peptide of RBP is precisely replaced by the bacterial signal sequence. The second PCR primer anneals to the 3′ end of the RBP coding region, introducing 6 carboxy-terminal histidine residues followed by a stop codon and a *Hin*dIII restriction site for unidirectional cloning into the expression vector. Since the primer was

1986; Holden *et al.*, 1987; Huber *et al.*, 1987; Cowan *et al.*, 1990; Böcskei *et al.*, 1992) and that they have the common property of binding small water-insoluble ligand molecules (Godovac-Zimmermann, 1988).

Consequently, RBP constitutes a protein which is not only relevant from the physiological point of view but also is of interest in respect to its structural properties and binding functions. We describe here the expression and periplasmic secretion of the porcine serum RBP in *Escherichia coli* in the absence of its natural ligand retinol. By determining the ligand binding and folding characteristics of the purified recombinant protein we demonstrate that it is directly obtained in a functional form and can therefore serve as a valuable source for elucidating the structure–function relationships of RBP.

For cloning and expression of the RBP structural gene in *E. coli* the general expression and secretion vector pASK40 (Skerra *et al.*, 1991) was chosen. This plasmid encodes the bacterial signal peptide of the outer membrane protein A (OmpA) followed by a polylinker. Gene expression is tightly regulated by the *lac* promoter/operator in conjunction with the *lac* repressor gene. Serum RBP possesses three intramolecular disulfide bonds. Therefore, our strategy for its expression in the functional state relied on the secretion of the recombinant protein into the periplasm of the *E. coli* cell, mediated by the OmpA signal peptide, in order to ensure disulfide bond formation under the oxidizing conditions in this cell compartment.

The structural gene for the mature RBP was amplified by polymerase chain reaction (PCR) after first-strand cDNA synthesis from porcine liver RNA (Fig. 1(a)). The expression plasmid carrying the precise fusion of this gene to the DNA sequence encoding the OmpA leader peptide was termed pHNM1 (Fig. 1(b)). The DNA sequence of the cloned cDNA fragment was confirmed by dideoxy sequencing and proven to be identical to the sequence from pig (Trout *et al.*, 1991). There was only one exception, the codon for residue 116 of the mature protein was in our case GTG in contrast to GCG, thus, leading to the amino acid Val instead of Ala in this position. Residue 116 is also a Val in the closely related human serum RBP (Colantuoni *et al.*, 1983). The flanking regions contributed by the PCR primers were also found to be correct, and for the purpose of expression a clone was chosen which coded for an Arg residue in the partially degenerate codon number 177 (Fig. 1(a)).

The published protein sequences of the RBPs from pig and man differ by just 13 amino acid substitutions, most of which are conservative. Taking into account the substitution observed in our cDNA compared to the published sequence, the number of different amino acid side-chains is reduced to 12. Inspection of the X-ray structure of human RBP (Cowan *et al.*, 1990) revealed that all of these amino acids are located on the surface of the protein, so that the pig RBP is a suitable model for the study of ligand binding on the basis of the three-dimensional structure of the human protein. In order to reduce further the number of residues in the recombinant protein differing from the human RBP, two substitutions were introduced at positions 177 and 182 by means of the second PCR primer during cloning of the cDNA (Fig. 1(a)). The remaining ten differing amino acids of the RBP encoded on pHNM1 are (with the corresponding residues of the human RBP in parentheses): Asn50

---

designed also to anneal to the human RBP cDNA, it contained 3 twofold degenerate nucleotides. Two of them correspond to wobble positions, one leads either to a lysine codon (pig) or an arginine codon (man) for amino acid 177. Furthermore, another base exchange results in leucine (man) instead of isoleucine (pig) for amino acid 182 of RBP. For PCR cloning of the RBP structural gene, the total RNA was prepared from fresh porcine liver (Davis *et al.*, 1986) and first-strand synthesis was carried out with $d(pT)_{19-24}$ (Sigma) and Moloney murine leukemia virus RNase $H^-$ reverse transcriptase (Gibco BRL) as recommended by the manufacturer. The resulting mixture was used directly as a template for PCR with the 2 primers shown (synthesized with an Applied Biosystems automated DNA synthesizer) and *Taq* DNA polymerase (Promega) according to a standard protocol (Sambrook *et al.*, 1989). A single PCR product with the expected size was obtained. It was purified on an agarose gel, digested with *Hin*dIII and *Pvu*II and ligated to the *Stu*I/*Hin*dIII-cut vector fragment of a derivative of pASK40 (Skerra *et al.*, 1991) containing the unique *Stu*I restriction site (kindly provided by R. Glockshuber). *E. coli* K12 strain JM83 (*ara*, Δ(*lac-pro*AB), *rps*L (=*str*A), $\phi$80, *lac*ZΔM15) (Yanisch-Perron *et al.*, 1985) cells were transformed with the ligation products and positive clones were identified by restriction analysis of the isolated plasmid DNA, followed by dideoxy sequencing of the cloned gene (Tabor & Richardson, 1987). One plasmid carrying the error-free gene with an arginine codon in position 177 was designated pHNM1. DNA and RNA manipulations were performed according to standard methodology (Sambrook *et al.*, 1989), and restriction enzymes were purchased from Boehringer-Mannheim and New England Biolabs. (b) The RBP expression plasmid pHNM1 and the helper plasmid pASK61. pHNM1 carries the modified porcine RBP structural gene, fused to the *E. coli ompA* signal sequence as described above, under transcriptional control of the IPTG-inducible *lac* promoter/operator. Transcription is terminated by the strong *rho*-independent lipoprotein terminator ($t_{lpp}$). The plasmid consists of an origin of replication from the pUC family of plasmids (*ori*), an ampicillin resistance gene (*bla*), an f1 origin of replication (f1-IG) for the preparation of single-stranded plasmid DNA, and the *lac* repressor gene (*lac*I). pASK61 was derived from the pBR322-compatible vector pACYC184 (Rose, 1988) with the p15a origin of replication (*ori*) and the chloramphenicol resistance gene (*cat*) and carries the genes for the periplasmic *E. coli* enzymes peptidyl-prolyl *cis*/*trans* isomerase (*ppi*) (Liu & Walsh, 1990; Hayano *et al.*, 1991) and protein disulfide isomerase (*pdi*) (Bardwell *et al.*, 1991; Kamitani *et al.*, 1992). Both genes are under the control of their own promoters, thus leading to the simultaneous and constitutive over-expression of the two proteins by a gene-dosage effect. pASK61 was designed as a general helper plasmid promoting the catalysis of rate-limiting steps during the folding of foreign proteins in the periplasm of *E. coli* (A.S., unpublished results).

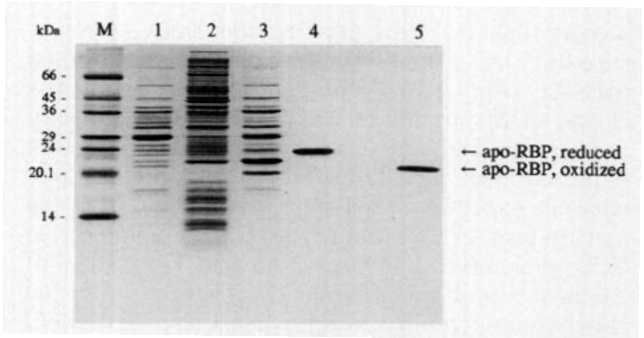



**Figure 2.** Bacterial expression and purification of the recombinant RBP. SDS/PAGE 15% of the total cell protein (lane 2) as well as the periplasmic cell fraction (lane 3) from an induced culture of *E. coli* JM83 harboring both pHNM1 and pASK61. The bands in lane 3 corresponding to PPI and PDI appear at approximate molecular masses of 20 kDa (calculated: 18,077 Da) and 22 kDa (calculated: 21,118 Da), respectively. For comparison, lane 1 shows the periplasmic cell fraction prepared under identical conditions from *E. coli* JM83 harboring the cloning vector pASK40 (Skerra *et al.*, 1991). The recombinant RBP purified by metal chelate affinity chromatography is shown in lane 4. The same protein preparation was loaded onto the protein gel without previous reduction of the disulfide bonds in lane 5. M denotes the molecular mass standard. For expression of the recombinant RBP, *E. coli* JM83 cells cotransformed with pHNM1 and pASK61 were grown in LB medium containing 100 μg ampicillin/ml and 30 μg chloramphenicol/ml at 22°C to an $A_{550}$ of approximately 0·5. Then, expression was induced by addition of 1 mM (final concentration)-IPTG (Biomol) in the presence of 5 mM (final concentration)-*N*-acetyl-L-cysteine (Merck). Induction was continued by shaking for 3 h and the cells were collected by centrifugation. For periplasmic cell fractionation the cells were resuspended in 1 M-NaCl, 1 mM-EDTA, 40 mM-sodium phosphate (pH 7·5) and kept on ice for 30 min. The spheroplasts formed during this step were separated by centrifugation and the supernatant was dialyzed overnight against 1 M-NaCl, 40 mM-sodium phosphate (pH 7·5) in order to remove the EDTA. The resulting protein solution was directly applied to a column of iminodiacetic acid Sepharose (Chelating Sepharose Fast Flow, Pharmacia) charged with $Zn^{2+}$ and the pure recombinant RBP was specifically eluted by a gradient of imidazole in the same buffer essentially as described (Skerra *et al.*, 1991). The concentration of the purified recombinant RBP was determined using a molar absorption coefficient of 40,400 $M^{-1}$ $cm^{-1}$ at 280 nm (Cogan *et al.*, 1976). SDS/PAGE was performed according to standard slab gel methodology using the buffer system of Fling & Gregerson (1986), followed by staining with Coomassie brilliant blue.

(Thr), His52 (Gln), Ile107 (Val), Gln123 (Leu), Ala138 (Ser), His142 (Asn), Phe144 (Leu), Ser145 (Pro), Val148 (Ala), Thr169 (Val).

In order to allow for the facile purification of the recombinant RBP by metal chelate affinity chromatography (reviewed by Arnold, 1991) we introduced a stretch of six histidine residues at its carboxy terminus. The carboxy terminus of the human serum RBP is far away from the retinol-binding site in the three-dimensional structure and not part of the β-barrel (Cowan *et al.*, 1990). The last few residues are not even defined in the X-ray structure so that these additional amino acid residues were not expected to interfere with the ligand-binding by RBP.

For recombinant gene expression, the *E. coli* cells harboring pHNM1 were grown at ambient temperature to a mid-log phase density and induced by addition of IPTG overnight. The recombinant RBP expressed by these cells was found to be secreted to the periplasm from which it was liberated by mild osmotic shock. However, the cells tended to lyse under these conditions and seemed to suffer from toxic effects caused by the expression of RBP. In addition, recombinant RBP isolated from these cells by metal chelate affinity chromatography was found to be contaminated by disulfide bond isomers which showed reduced mobility during SDS/PAGE under non-reducing conditions (not shown). These isomers could not be readily separated from the correctly folded protein using standard chromatographic procedures.

Both the toxicity effect and protein inhomogeneity were completely avoided when expression of the recombinant RBP was performed, under otherwise identical conditions, in the presence of a second plasmid, pASK61 (Fig. 1(b)), and with the addition of the thiol-reagent *N*-acetyl-L-cysteine during induction. pASK61 leads to the over-expression of the two periplasmic *E. coli* enzymes peptidyl-prolyl *cis/trans* isomerase (PPI: Liu & Walsh, 1990; Hayano *et al.*, 1991) and protein disulfide isomerase (PDI: Bardwell *et al.*, 1991; Kamitani *et al.*, 1992) and was constructed as a general helper plasmid to enhance the folding of foreign proteins in the periplasm of *E. coli* (A.S., unpublished results). *N*-Acetyl-L-cysteine, which is small enough to diffuse through the outer membrane porins of *E. coli* (Hancock, 1987), served as a co-substrate for PDI (M. Wunderlich & R. Glockshuber, unpublished results).

Figure 2 shows SDS/PAGE of the total cell protein as well as the periplasmic cell fraction from an induced culture of *E. coli* K12 JM83 harboring both pHNM1 and pASK61. The cell fractionation carried out with the induced cells is specific and the recombinant RBP is present as a visible band in the periplasmic fraction. In addition, the over-expression of PPI and PDI encoded on pASK61 can be seen.

The recombinant RBP was purified from the periplasmic cell fraction in a single step using metal chelate affinity chromatography as described earlier (Skerra *et al.*, 1991). The protein thus obtained is also shown on SDS/PAGE in Figure 2, both in the reduced and in the non-reduced state. The recombinant RBP appears as a single band on the gel with a size corresponding to the calculated molecular mass of the correctly processed protein. The preparation of the recombinant RBP is completely homogeneous without any additional purification steps. In particular, the protein migrates significantly faster on the SDS-containing gel if not previously reduced,

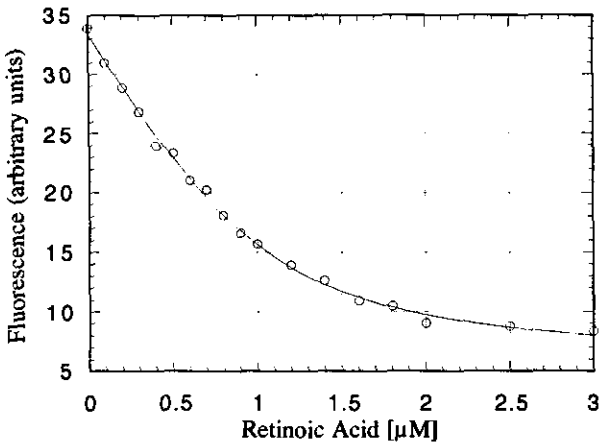



**Figure 3.** Fluorescence titration of the recombinant RBP with retinoic acid. Fluorescence measurements were carried out with a Perkin Elmer LS50 fluorimeter and a 100 μl Hellma quartz cuvette (path lengths: 10 mm and 2 mm in the directions of excitation and emission, respectively) in a thermostatted cuvette holder at 22°C. Tryptophan fluorescence was excited at 295 nm and detected (averaged over 1 s) at the emission maximum of 340 nm (excitation and emission slit widths both set to 5 nm). The solution of the purified recombinant RBP in 100 mM-NaCl, 1 mM-EDTA, 50 mM-sodium phosphate (pH 7·0) was previously cleared from particles by sterile filtration (0·22 μm, Millex-GV4, Millipore) and always pipetted with pre-lubricated tips (Roth). A stock solution of retinoic acid (Sigma) in DMF was gravimetrically prepared and then further diluted with DMF as required. For each measurement, 200 μl of the solution of RBP at a concentration of 1·0 μM was added to 2 μl of the appropriately diluted solution of retinoic acid in DMF and mixed by re-pipetting once. The solutions were incubated for 30 min at 22°C and the tryptophan fluorescence was measured. The values were corrected by a blank value measured for buffer containing 1% (v/v) DMF. A correction for the inner fluorescence filter effect caused by the chromophore of retinoic acid (Cogan *et al.*, 1976) was found not to be necessary, due to the cuvette dimensions and ligand concentrations used here. The experimental data points were fitted by non-linear regression according to the theory of bimolecular complex formation with the resulting curve shown. The calculated dissociation constant was $K'_d = 0{\cdot}21(\pm 0{\cdot}03)$ μM.

which is in agreement with a more compact structure of the SDS-denatured protein due to the presence of disulfide bonds. Gas-phase sequencing of the purified recombinant RBP revealed the amino acid sequence Glu-Arg-Asp-Cys-Arg for its amino terminus, which is identical to the sequence expected for the mature protein. Thus, the OmpA signal peptide was properly cleaved off. The yield of the purified RBP was approximately 0·15 mg/l of *E. coli* culture ($A_{550} = 1{\cdot}3$).

In order to test the functionality of the recombinant RBP, ligand binding studies with vitamin A acid were performed. It was previously shown that the serum retinol-binding proteins from various species do not only bind their natural ligand retinol but also closely related compounds, e.g. retinoic acid and retinoyl acetate, with very similar affinities (Cogan *et al.*, 1976). Apart from effects on the fluorescence properties of the ligand itself, as for retinol, the binding of these ligands always leads to a significant decrease in the intrinsic protein fluorescence of RBP due to energy transfer from excited tryptophan residues. Therefore, the ligand-binding characteristics of the serum RBP can be easily studied by fluorescence titration.

For our purpose, all-*trans* retinoic acid was chosen as the ligand molecule, which is more soluble in water than retinol (Szuts & Harosi, 1991) and chemically more stable, but possesses essentially the same apparent (i.e. neglecting the capacity for micelle-formation of the lipid-like molecule) affinity for RBP (Cogan *et al.*, 1976). Figure 3 shows the fluorescence data from an experiment in which the purified recombinant RBP was titrated with retinoic acid in a physiological buffer. By addition of increasing amounts of the ligand, the protein fluorescence dropped sharply, leading to an overall fluorescence quench of 77% ($\lambda_{Ex} = 295$ nm, $\lambda_{Em} = 340$ nm) upon saturation of the binding sites with no change of $\lambda_{Em,Max}$. For the natural human serum RBP a smaller quench of 38% was described with retinoic acid as the ligand (Cogan *et al.*, 1976), however, at $\lambda_{Ex} = 296$ nm and $\lambda_{Em} \geq 310$ nm. Since the number of tyrosine and tryptophan residues, and in particular the amino acids aligning the ligand binding pocket, are highly conserved between the human and our recombinant protein, the possibly more efficient energy transfer to the bound ligand remains to be explained in this case.

The data points obtained from the fluorescence titration experiment were fitted by non-linear regression according to the theory of bimolecular complex formation and with the assumption of one ligand binding site per protein molecule. Figure 3 reveals an excellent fit of the experimental data points to the curve over the whole range of the titration, thus, demonstrating the functional homogeneity of the protein preparation. The dissociation constant derived from the mathematical analysis was $K'_d = 2{\cdot}1(\pm 0{\cdot}3) \times 10^{-7}$ M, which is completely identical to the value for the human protein ($K'_d = 2{\cdot}1 \times 10^{-7}$ M) reported earlier (Cogan *et al.*, 1976).

Finally, the folding stabilities both of the apo-RBP directly isolated from the bacterial expression host and of the RBP–retinoic acid complex reconstituted *in vitro* were investigated in the presence of Gdn·HCl as denaturing agent (Pace, 1986). The uncomplexed recombinant RBP showed a single unfolding transition at a midpoint concentration of 2·0 M-Gdn·HCl (Fig. 4(a)), accompanied by a marked decrease of about 66% in tryptophan fluorescence ($\lambda_{Ex} = 295$ nm, $\lambda_{Em} = 320$ nm) with a change of $\lambda_{Em,Max}$ from 340 nm to 355 nm. The superposition of the data from the corresponding refolding experiment starting with the protein denatured in the presence of 5 M-Gdn·HCl (Fig. 4(a)) revealed that the unfolding transition of the apo-protein was completely reversible. Unfolding of the RBP–retinoic acid complex also resulted in a

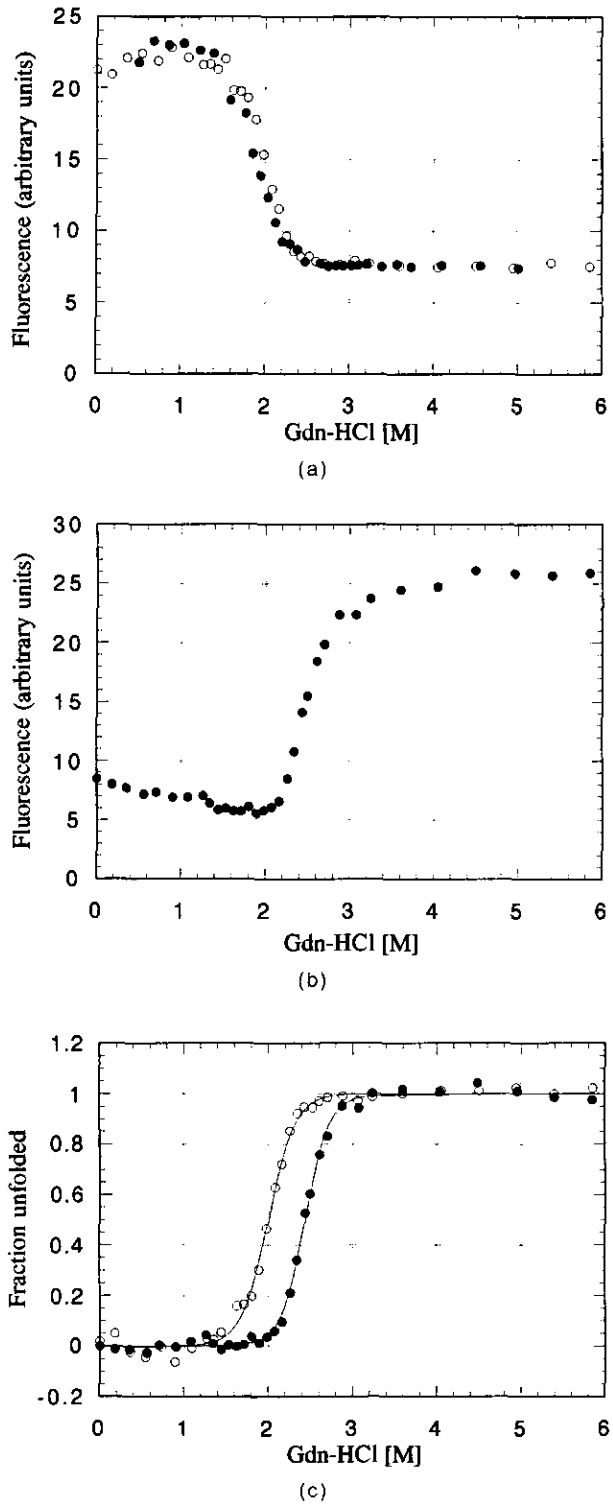



**Figure 4.** Gdn·HCl-induced unfolding and refolding of the recombinant RBP. (a) Fluorescence curve for unfolding (open circles) and refolding transitions (filled circles) of the apo-RBP. Solvent denaturation studies were conducted using the intrinsic tryptophan fluorescence emission of the recombinant RBP to monitor the ratio of native and denatured protein (Pace, 1986). Fluorescence measurements were essentially performed as described for Fig. 3, except that emission was measured at a wavelength of 320 nm (with $\lambda_{Ex} = 295$ nm), the maximum of the difference spectrum between native and denatured apo-RBP. The emission values were corrected by the fluorescence intensities measured for the buffer containing the respective concentration of Gdn·HCl. The buffer used for all unfolding and refolding experiments was 100 mM-NaCl, 1 mM-EDTA, 50 mM-sodium phosphate (pH 7·0). A stock solution of 7 M-Gdn·HCl (Sigma or USB) in this buffer was gravimetrically prepared and the concentration was confirmed by measuring the refractive index with an Abbe refractometer (Zeiss) according to Pace (1986). This stock solution was diluted with buffer to the required working concentration by use of calibrated burettes. The unfolding transition was initiated by diluting 15 µl of a 10 µM solution of apo-RBP with 135 µl of the appropriate solution of Gdn·HCl so that the final protein concentration was 1 µM. After mixing, the resulting solution was incubated at 22°C for 24 h until equilibrium was reached and fluorescence was measured. In order to monitor the refolding transition, first a concentrated solution of apo-RBP was mixed with the Gnd·HCl stock solution so that final concentrations of 5 M-Gdn·HCl and 10 µM-protein were obtained. After incubation for 24 h at 22°C the resulting solution of the unfolded RBP was mixed with diluted solutions of Gdn·HCl to obtain varying end concentrations of the denaturing agent. After incubation for additional 24 h at 22°C, tryptophan fluorescence of the solution with a final protein concentration of 1 µM was measured as before. (b) Fluorescence curve for unfolding of the RBP–retinoic acid complex. For reconstitution of the holo-protein, 750 µl of a 10 µM solution of apo-RBP in the buffer mentioned before was mixed with 7·5 µl retinoic acid (3 mM) dissolved in DMF and incubated for 30 min in the dark. The unfolding transition was essentially performed as described for (a). After incubation at 22°C for 6 days, tryptophan fluorescence of the solution with a final protein concentration of 1 µM was measured at a wavelength of 360 nm (with $\lambda_{Ex} = 295$ nm), the maximum of the difference spectrum between native and denatured holo-RBP. The emission values were corrected by the fluorescence intensities measured for the buffer containing the respective concentration of Gdn·HCl, DMF and retinoic acid. (c) Fraction unfolded *versus* concentration of Gdn·HCl both for apo- (open circles) and for holo-RBP (filled circles). The fraction unfolded for each data point from the unfolding transitions in (a) and (b) were calculated using the best-fit parameters found for the pre- and post-transitional lines using a non-linear regression method described by Santoro & Bolen (1988). The continuous line represents the corresponding least-squares fit through all the data points expressed in terms of fraction unfolded. The midpoint concentrations of Gdn·HCl for the unfolding transitions of the apo- and holo-protein were calculated to be 2·0 M and 2·4 M and the numerical results for the free energy changes of unfolding in the absence of denaturant were $\Delta G^0_{n-u} = 31{\cdot}4(\pm 2{\cdot}5)$ kJ/mol and $\Delta G^0_{n-u} = 41{\cdot}3(\pm 2{\cdot}7)$ kJ/mol, respectively. It has to be considered that, due to the low binding affinity and poor solubility of the ligand, RBP was not present at 100% in its holo-form during the unfolding experiment. Therefore, the $\Delta G$ value calculated for the reconstituted RBP–retinoic acid complex probably represents only a lower approximation for the actual unfolding stability of holo-RBP.

single transition (Fig. 4(b)), however, with an elevated midpoint concentration of 2·4 M-Gdn·HCl and a significant increase in tryptophan fluorescence ($\lambda_{Ex} = 295$ nm, $\lambda_{Em} = 360$ nm with a change of $\lambda_{Em,Max}$ from 340 nm to 355 nm). The lower fluorescence of the native complex compared to the

denatured RBP was due to the fluorescence quench effect of the bound ligand described above. Unfolding of the apo- and of the holo-protein thus follows a simple two-state model demonstrating high cooperativity for the native to unfolded transition both in the absence and presence of the ligand.

The experimental data from the GdnHCl-induced unfolding curves were analyzed by non-linear regression according to the six-parameter fit described by Santoro & Bolen (1988). The normalized unfolding curves are shown in Figure 4(c). As a result, the free energy changes of unfolding for the uncomplexed recombinant RBP and the RBP–retinoic acid complex in the absence of denaturant were calculated as $\Delta G^0_{n-u} = 31{\cdot}4(\pm 2{\cdot}5)$ kJ/mol and $\Delta G^0_{n-u} = 41{\cdot}3(\pm 2{\cdot}7)$ kJ/mol, respectively. Thus, binding of the lipophilic ligand leads to a significant increase in folding stability. This phenomenon, which was similarly observed in the case of the cellular retinoic acid-binding protein (Zhang *et al.*, 1992), clearly supports the functionality of the protein isolated from *E. coli*.

In this contribution we were able to show that it is possible to express the uncomplexed serum retinol-binding protein in a functional and stable form in a bacterial host. This was accomplished by secretion of the recombinant gene product to the periplasm of *E. coli*, which, in contrast to the cytoplasm, constitutes a cell compartment with oxidizing redox potential where disulfide bonds can form. However, the secretion strategy alone led to an inhomogeneous protein preparation and caused lysis of the bacterial cells. These problems could be circumvented by the co-expression of the periplasmic folding mediators PDI and PPI. Although this phenomenon served only for practical use in the present study, the mechanism of action, particularly of protein disulfide isomerase in conjunction with a thiol-reagent on the recombinant RBP, will be the subject of future investigations.

The recombinant RBP was purified in a single step by metal chelate affinity chromatography from the periplasmic cell fraction. The protein thus obtained was shown to possess a ligand affinity identical to that of the natural human protein in spite of differences in the amino acid sequence resulting from our cloning strategy for the cDNA. The addition of an oligo-histidine tail to the recombinant RBP seemed not to interfere with the ligand binding function and, therefore, no attempts were made to remove this affinity tag. Consequently, the metal chelate affinity purification methodology will be very useful in the future during investigations which will involve alterations of the ligand-binding pocket by site-directed mutagenesis.

The fact that it was possible to isolate the ligand-free recombinant RBP from a bacterial host organism was not entirely to be expected. Up to now serum RBP has only been isolated as a holo-protein from natural sources. It is known that vitamin A can be reversibly removed from this native complex. However, in cell culture studies it was observed that hepatocytes no longer secreted RBP during vitamin A deficiency, but accumulated it in the endoplasmic reticulum (Ronne *et al.*, 1983). Therefore, it was not quite clear whether serum RBP could be formed in the periplasm of *E. coli* as a correctly folded, soluble, and stable protein in the absence of the hydrophobic ligand.

In this regard it has to be emphasized that serum RBP possesses a rather deep and hydrophobic ligand binding pocket. Neglecting possible conformational changes, an accessible protein surface area of 251 Å$^2$ (1 Å = 0·1 nm) becomes buried upon complex formation with the ligand retinol (Cowan *et al.*, 1990). This cavity is located right at the center of the hydrophobic core of the $\beta$-barrel structure and is most probably filled with water molecules in the apo-protein. This should constitute an energetically unfavourable situation for the protein. Nevertheless, the recombinant protein folds properly, at least under optimized conditions, and yields a surprisingly stable protein with the free energy change of unfolding amounting to $31{\cdot}4(\pm 2{\cdot}5)$ kJ/mol. Deep hydrophobic binding pockets occur also in other *lipocalins* as, e.g. the bilin-binding protein (Huber *et al.*, 1987). The functional expression of serum RBP in *E. coli* described here constitutes, to the best of our knowledge, the first example for a member of the *lipocalin* family as a whole. Our findings suggest that other proteins belonging to this family may be functionally expressed with the strategy outlined here, as well. This should open the way to future understanding of the ligand-binding functions and specificities of these physiologically important polypeptides on the molecular level by protein engineering.

The authors thank F. Lottspeich for N-terminal sequencing and R. Glockshuber for helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft.

## References

Arnold, F. H. (1991). Metal-affinity separations: a new dimension in protein processing. *Bio/Technology*, **9**, 151–156.

Bardwell, J. C. A., McGovern, K. & Beckwith, J. (1991). Identification of a protein required for disulfide bond formation *in vivo*. *Cell*, **67**, 581–589.

Berni, R., Stoppini, M. & Zapponi, M. C. (1992). The piscine plasma retinol-binding protein. *Eur. J. Biochem.* **204**, 99–106.

Blomhoff, R., Green, M. H., Berg, T. & Norum, K. R. (1990). Transport and storage of vitamin A. *Science*, **250**, 399–404.

Böcskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E. V., Cavaggioni, A., Findlay, J. B. C. & North, A. C. T. (1992). Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature (London)*, **360**, 186–188.

Cogan, U., Kopelman, M., Mokady, S. & Shinitzky, M. (1976). Binding affinities of retinol and related compounds to retinol binding proteins. *Eur. J. Biochem.* **65**, 71–78.

Colantuoni, V., Romando, V., Bensi, G., Santoro, C., Costanzo, F., Raugei, G. & Cortese, R. (1983). Cloning and sequencing of a full length cDNA coding

for the human retinol-binding protein. *Nucl. Acids Res.* **11**, 7769–7776.

Cowan, S. W., Newcomer, M. E. & Jones, T. A. (1990). Crystallographic refinement of human serum retinol binding protein at 2 Å resolution. *Proteins: Struct. Funct. Genet.* **8**, 44–61.

Davis, L. G., Dibner, M. D. & Battey, J. F. (1986). *Basic Methods in Molecular Biology*. Elsevier, New York.

Fling, S. P. & Gregerson, D. S. (1986). Peptide and protein molecular weight determination by electrophoresis using a high-molarity Tris-buffer system without urea. *Anal. Biochem.* **155**, 83–88.

Godovac-Zimmermann, J. (1988). The structural motif of β-lactoglobulin and retinol-binding protein: a basic framework for binding and transport of small hydrophobic molecules? *Trends Biochem. Sci.* **13**, 64–66.

Goodman, D. S. (1984). Plasma retinol-binding proteins. In *The Retinoids* (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds), pp. 41–88, Academic Press, New York.

Hancock, R. E. W. (1987). Role of porins in outer membrane permeability. *J. Bacteriol.* **169**, 929–933.

Hayano, T., Takahashi, N., Kato, S., Maki, N. & Suzuki, M. (1991). Two distinct forms of peptidylprolyl-*cis-trans*-isomerase are expressed separately in periplasmic and cytoplasmic compartments of *Escherichia coli* cells. *Biochemistry*, **30**, 3041–3048.

Holden, H. M., Rypniewski, W. R., Law, J. H. & Rayment, I. (1987). The molecular structure of insecticyanin from the tobacco hornworm *Manducta sexta* L. at 2·6 Å resolution. *EMBO J.* **6**, 1565–1570.

Huber, R., Schneider, M., Mayr, I., Müller, R. & Deutzmann, R. (1987). Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2·0 Å resolution. *J. Mol. Biol.* **198**, 499–513.

Kamitani, S., Akiyama, Y. & Ito, K. (1992). Identification and characterization of an *Escherichia coli* gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme. *EMBO J.* **11**, 57–62.

Liu, J. & Walsh, C. T. (1990). Peptidyl-prolyl *cis-trans*-isomerase from *Escherichia coli*: a periplasmic homolog of cyclophilin that is not inhibited by cyclosporin A. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 4028–4032.

McKearing, D. M., Barton, M. C., Keller, M. J. & Shapiro, D. J. (1987). Estrogen induces transcription of the *Xenopus laevis* serum retinol binding protein gene. *J. Biol. Chem.* **262**, 4939–4942.

Newcomer, M. E., Jones, T. A., Åqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P. A. (1984). The three-dimensional structure of retinol-binding protein. *EMBO J.* **3**, 1451–1454.

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.

Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C. T., Findlay, J. B. C., Sivaprasadarao, R., Jones, T. A., Newcomer, M. E. & Kraulis, P. J. (1986). The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein. *Nature (London)*, **324**, 383–385.

Pervaiz, S. & Brew, K. (1987). Homology and structure-function correlation between $\alpha_1$-acid glycoprotein and serum retinol-binding protein and its relatives. *Fed. Amer. Soc. Expt. Biol.* **1**, 209–214.

Ronne, H., Ocklind, C., Wiman, K., Rask, L., Obrink, B. & Peterson, P. A. (1983). Ligand-dependent regulation of intracellular protein transport: effect of vitamin A on the secretion of the retinol-binding protein. *J. Cell Biol.* **96**, 907–910.

Rose, R. E. (1988). The nucleotide sequence of pACYC184. *Nucl. Acids Res.* **16**, 355.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.

Shidoji, Y. & Muto, Y. (1977). Vitamin A transport in plasma of the non-mammalian vertebrates: isolation and partial characterization of piscine retinol-binding protein. *J. Lipid Res.* **18**, 679–691.

Skerra, A., Pfitzinger, I. & Plückthun, A. (1991). The functional expression of antibody Fv fragments in *Escherichia coli*: improved vectors and a generally applicable purification technique. *Bio/Technology*, **9**, 273–278.

Sundelin, J., Laurent, B. C., Anundi, H., Träghårdh, L., Larhammar, D., Björck, L., Eriksson, U. Åkerström, B., Jones, A., Newcomer, M., Peterson, P. A. & Rask, L. (1985). Amino acid sequence homologies between rabbit, rat, and human serum retinol-binding proteins. *J. Biol. Chem.* **260**, 6472–6480.

Szuts, E. Z. & Harosi, F. I. (1991). Solubility of retinoids in water. *Arch. Biochem. Biophys.* **287**, 297–304.

Tabor, S. & Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 4767–4771.

Trout, W. E., McDonnell, J. J., Kramer, K. K., Baumbach, G. A. & Roberts, R. M. (1991). The retinol-binding protein of the expanding pig blastocyst: molecular cloning and expression in trophectoderm and embryonic disc. *Mol. Endocrinol.* **5**, 1533–1540.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.

Zhang, J., Liu, Z.-P., Jones, T. A., Gierasch, L. M. & Sambrook, J. F. (1992). Mutating the charged residues in the binding pocket of cellular retinoic acid-binding protein simultaneously reduces its binding affinity to retinoic acid and increases its thermostability. *Proteins: Struct. Funct. Genet.* **13**, 87–99.

*Edited by R. Huber*