Uterocalin, a lipocalin provisioning the preattachment equine conceptus: fatty acid and retinol binding properties, and structural characterization[1]

Sabine SUIRE*, Francesca STEWART*[2], Jeremy BEAUCHAMP† and Malcolm W. KENNEDY‡[3]

*Babraham Institute, Babraham, Cambridge CB2 4AT, U.K., †Protein Crystallography, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, U.K., and ‡Division of Infection and Immunity, Institute of Biomedical and Life Sciences, Joseph Black Building, University of Glasgow, Glasgow G12 8QQ, U.K.

The equine conceptus is surrounded by a fibrous capsule that persists until about day 20 of pregnancy, whereupon the capsule is lost, the conceptus attaches to the endometrium and placentation proceeds. Before attachment, the endometrium secretes in abundance a protein of the lipocalin family, uterocalin. The cessation of secretion coincides with the end of the period during which the conceptus is enclosed in its capsule, suggesting that uterocalin is essential for the support of the embryo before direct contact between maternal and foetal tissues is established. Using recombinant protein and fluorescence-based assays, we show that equine uterocalin binds the fluorescent fatty acids 11-(dansylamino)undecanoic acid, dansyl-D,L-α-amino-octanoic acid and *cis*-parinaric acid, and, by competition, oleic, palmitic, arachidonic, docosahexaenoic, γ-linolenic, *cis*-eicosapentaenoic and linoleic acids. Uterocalin also binds all-*trans*-retinol, the binding site for which is coincident or interactive with that for fatty acids. Molecular modelling and intrinsic fluorescence analysis of the wild-type protein and a Trp → Glu mutant protein indicated that uterocalin has an unusually solvent-exposed Trp side chain projecting from its large helix directly into solvent. This feature is unusual among lipocalins and might relate to binding to, and uptake by, the trophoblast. Uterocalin therefore has the localization and binding activities for the provisioning of the equine conceptus with lipids including those essential for morphogenesis and pattern formation. The possession of a fibrous capsule surrounding the conceptus might be an ancestral condition in mammals; homologues of uterocalin might be essential for early development in marsupials and in eutherians in which there is a prolonged preimplantation period.

Key words: endometrium, histotroph, lipid-binding protein, polyunsaturated fatty acids, pregnancy.

INTRODUCTION

Mammalian embryos require high levels of lipids for energy metabolism, for the construction of membranes and for signalling events involved in gene activation, pattern formation and morphogenesis [1]. For these last processes, the lipids required include polyunsaturated fatty acids (PUFAs), their derivatives (eicosanoids in particular) and retinoids [1,2]. All of these are sparingly soluble in water and are highly susceptible to degradation, so they are transported in plasma and cytosol, and across membranes, by specialized proteins [3–5]. In addition, the placenta of some mammals is known to possess transmembrane fatty acid transporters to capture certain PUFAs and concentrate them in the foetal circulation [4,6]. However, this route of acquisition is not open to those mammals in which there is no direct contact between the trophoblast and maternal cells or blood at early stages of gestation. A case in point is the horse, in which a functional allantochorial placenta providing full haemotrophic nutrition does not form until as late as 40–42 days after conception [7]. Furthermore, the equine conceptus remains unattached and surrounded by a glycoprotein capsule until approximately day 22 of gestation [8–10], so there is no direct cell–cell contact and nutrients must therefore enter and traverse an aqueous phase. Moreover, organ differentiation is in progress before the capsule breaks down and attachment/implantation occurs, so lipids essential for signalling and morphological events must be acquired intact by the embryo during the preattachment period, the only source being the uterine endometrium itself.

The growth and development of the preimplantation conceptus is supported in the horse by secretions of the endometrium, histotroph. One of the most abundant components of this secretion is a protein of approx. 19 kDa (originally termed P19, but here renamed uterocalin), which is a member of the lipocalin family of proteins [11]. Uterocalin is produced by the endometrial glandular and epithelial cells of the mare only at specific times during the oestrous cycle and after conception [11,12]. It appears in large quantities during the luteal phase of oestrus and early pregnancy, and its secretion can otherwise be induced by exogenous progesterone [11,12]. The protein was originally isolated from the embryonic capsule; it accumulates in the trophoblast cell layer and in yolk sac fluid, but mRNA encoding it can only be found in maternal endometrial cells [11]. The protein is detectable in uterine flushings for the first 14 days of oestrus and until approximately day 20 of pregnancy; the presence of the encoding mRNA is in keeping with this pattern [11]. In pregnancy, therefore, the cessation of uterocalin secretion coincides with the end of the period during which the conceptus is enclosed in its capsule. The abundance of the protein at that time suggests that it is essential for the support of the early embryo before direct contact between maternal and foetal tissues

Abbreviations used: DACA, dansyl-D,L-α-amino-octanoic acid; DAUDA, 11-(dansylamino)undecanoic acid; EUC, equine uterocalin; FABP, fatty-acid-binding protein; rEUC, recombinant bacterium-derived EUC; hEUC, horse-derived, naturally produced EUC; PDB, Protein Data Bank; PUFAs, polyunsaturated fatty acids; RBP, retinol-binding protein; retinol, all-*trans*-retinol; W150E, Trp-150 → Glu mutant of rEUC.

[1] Shortly after this paper was submitted, Cesca Stewart passed away. She had shown great dignity and courage during the preceding illness and had worked as much as she could until the end. The rest of the authors wish to dedicate this paper to her memory and lament the loss of such a good colleague and so imaginative a scientist.

[2] Deceased.

[3] To whom correspondence should be addressed (e-mail malcolm.kennedy@bio.gla.ac.uk).

is established. The histotroph of other mammals, such as pigs and dogs, contains lipocalins that bind all-*trans*-retinol (retinol) [retinol-binding proteins (RBPs)] [13,14], although that of pigs is identical with their plasma RBP [13]. However, uterocalin is quite distinct from equine plasma RBP and equine β-lactoglobulins (retinol-binding lipocalins of milk), so, whereas some species seem to have solved the problem of retinol transport by using their plasma RBP, others seem to have evolved specialized proteins for the purpose.

Lipocalins have very similar tertiary structures, but exhibit great diversity in their amino acid sequences and functions. For instance, different lipocalins have roles in olfaction, invertebrate coloration, modulation of the immune response, nutrition of infants and enzymic synthesis of prostaglandins [15]. Those about which most is known in biochemical and structural terms are transport proteins for small lipids and volatile compounds such as retinol (vitamin A), fatty acids and odorants. Examples include plasma RBP, β-lactoglobulins, tear lipocalin, olfactory (odorant-binding) protein and pheromone-binding proteins from the urine of male rodents (mice and rats) or the uterus and salivary glands of female hamsters [15,16].

Here we show that equine uterocalin (EUC) binds a range of biologically important lipids, including PUFAs, and retinol. It therefore seems to be a lipocalin that is specialized for the provisioning of the preattachment embryo of horses; homologues might also be present in other species that undergo delayed implantation or have no direct cell–cell contact between mother and foetus in the early stages. We also show that the protein has some unusual structural features that might relate to its interaction with the trophoblast surface to confer directionality on the lipid transfer.

MATERIALS AND METHODS

Natural and recombinant EUC

The native protein was isolated by flushing the uterus of mares in oestrus or early pregnancy as described previously (GenBank® accession no. X98459) [11,12]. The full-length cDNA clone for EUC was originally isolated in a lambda Uni-Zap vector (Stratagene, La Jolla, CA, U.S.A.) and then excised into a pBluescript SK plasmid [11]. Two oligonucleotide primers (5′-AAA AAA CAT ATG CTG CAC ATG GGC CCT GGG-3′ and 5′-CCT TAA GGA GAC GCC CCT GGA CCC TTG G-3′) containing an *Nde*I site and an *Eco*RI site respectively were used to amplify by PCR DNA encoding EUC without its leader sequence. The amplicon was cloned into the pTag vector (Ingenius, Abingdon, Oxon., U.K.) and the integrity of the insert was confirmed by sequencing. The insert was then subcloned into pET-15b expression vector (Novagen, Madison, WI, U.S.A.) between the *Nde*I and *Eco*RI restriction sites and used to transform *Escherichia coli* BL21(DE3)pLysS cells. Individual transformants were grown overnight at 37 °C in 5 ml of Luria–Bertani medium containing 25 μg/ml ampicillin and 34 μg/ml chloramphenicol. Large flasks containing 750 ml of Luria–Bertani medium supplemented with appropriate antibiotics were then inoculated with 5 ml of overnight cultures. When a D_{600} of 0.6 was reached, 0.4 mM of isopropyl β-D-thiogalactoside was added for 4 h to induce the synthesis of the recombinant bacterium-derived protein (rEUC) with a His_6 N-terminal tag. The cells were then collected; cell pellets were frozen and stored at −20 °C.

rEUC was purified by affinity chromatography with the HisBind Resin and buffer kit, in accordance with the manufacturer's instructions (Novagen). rEUC was further purified by chromatography on a hydroxyapatite column (type II; Bio-Rad, Hemel Hempstead, Herts., U.K.); the resulting protein preparation showed a single band of the appropriate molecular mass by SDS/PAGE and staining with Coomassie Blue. After purification, both the horse-derived, naturally produced EUC (hEUC) and rEUC preparations were passed down an Extracti-Gel D column (Pierce, Rockford, IL, U.S.A.) to remove lipid and any residual detergents.

Construction of Trp-150 → Glu (W150E) mutant of rEUC

Site-directed mutagenesis was performed in the pET15-P19 expression vector plasmid by using two synthetic oligonucleotide primers encoding the target and flanking amino acids. The primers, each complementary to opposite strands of the vector, were extended around the plasmid during 15 PCR cycles and then treated with *Dpn*I (QuickChange™ Site-Directed Mutagenesis Kit; Stratagene). The primers used to replace Trp-150 with Glu to create the W150E mutant were 5′-CCGGATGTCATGGAGATGTTTAAAAAG-3′ and 3′-GGCCTACAGTACCTCTACAAATTTTTC-5′ (the underlined nucleotide positions are those required for the mutagenesis procedure). The mutated recombinant protein was then expressed and purified in the same way as described for the wild-type rEUC.

Spectrofluorimetry and fluorescence-based ligand binding

Fluorescence binding emission spectra were recorded at 20 °C with a SPEX Fluo-Max spectrofluorimeter (Spex Industries, Edison, NJ, U.S.A.), on 2 ml samples in a silica cuvette. Correction was made for Raman scattering by solvent water where necessary. All spectra are shown uncorrected. The fluorescent fatty acids 11-(dansylamino)undecanoic acid (DAUDA) and *cis*-parinaric acid were obtained from Molecular Probes (Eugene, OR, U.S.A.). Oleic acid, arachidonic acid, docosahexaenoic acid, γ-linolenic acid, *cis*-eicosapentaenoic acid, linoleic acid, palmitic acid, retinol and dansyl-D,L-α-aminooctanoic acid (DACA) were obtained from Sigma (Poole, Dorset, U.K.). The excitation wavelengths used for DAUDA, DACA, *cis*-parinaric acid and retinol were 345, 345, 319 and 350 nm respectively. The dansylated fatty acids were stored as stock solutions of approx. 3 mg/ml in ethanol in the dark at −20 °C; they were freshly diluted in PBS [171 mM NaCl/3.35 mM KCl/10 mM Na_2HPO_4/1.8 mM KH_2PO_4 (pH 7.2)] to approx. 1 μM before use in the fluorescence experiments. Competitors of fluorescent fatty acid binding were prepared as stock solutions in ethanol at approx. 10 mM and diluted in PBS (for oleic acid) or ethanol (for other competitors). Free retinol is poorly soluble and unstable in aqueous solution, so it was dissolved and diluted in ethanol immediately before use; binding to proteins was tested by the addition of, typically, 5 μl of this directly to a cuvette containing protein in PBS. The concentrations of the proteins were estimated by A_{280} with theoretical absorption coefficients based on the amino acid compositions of the natural, recombinant and mutant proteins by using the ProtParam program through the ExPASy server (http://expasy.cbr.nrc.ca/tools/). The molar absorption coefficients at 280 nm, ϵ_{280}, were calculated from their amino acid concentrations as 13940 and 8250 $M^{-1}\cdot cm^{-1}$ for the wild-type rEUC and the W150E mutant respectively. The concentrations of both DAUDA and DACA stocks (diluted in methanol) were calculated from ϵ_{335} 4400 $M^{-1}\cdot cm^{-1}$, ϵ_{303} 76000 $M^{-1}\cdot cm^{-1}$ for *cis*-parinaric acid in ethanol and ϵ_{325} 52480 $M^{-1}\cdot cm^{-1}$ for retinol in ethanol. Fluorescence data were corrected for dilution where necessary and were fitted by standard non-linear regression techniques (with Microcal ORIGIN software) to a single non-competitive binding model to

give estimates of the dissociation constant (K_d) and maximal fluorescence intensities (F_{max}). Corrections for free retinol were made as before [17]. Similar non-linear regression methods were used to analyse results of competition experiments in which fatty acids were progressively added to a preformed complex of DAUDA or *cis*-parinaric acid and protein. A stock solution of oleic acid in ethanol was freshly diluted to 1 mM in PBS; increasing concentrations of oleic acid were added to the mixtures. For the other competitors, solutions were made in ethanol and progressively added to the complex in, typically, 5 μl portions.

Structure prediction

Protein sequence comparisons and alignments were performed with the WU-BLAST program accessed through the ExPASy server at SwissProt, or MaxHom accessed through the PredictProtein (PHD) suite of programs also on the ExPASy server at SwissProt. Secondary structure predictions and searching for structurally similar proteins were made with PHD. Searching for protein motifs such as secretory leader sequences and protein family signatures were performed with the Signalp (trained on eukaryotic sequences) and PROSITE programs. The EUC structure was modelled initially with SwissModel, refined with the CHARMm forcefield, using as templates two experimental structures identified by SwissModel as being sufficiently similar to EUC in amino acid sequence; the modelling was set to run with α_{2u}-globulin [Protein Data Bank (PDB) accession no. 2A2G] [18] and mouse urinary pheromone-carrier protein (PDB accession no. 1MUP) as composite templates, other lipocalin structures being rejected as being too dissimilar for any prediction to be adequate. Model geometries were checked with the PROCHECK suite of programs [19]. Multiple alignments were performed with MultAlin set for the Dayhoff or Blosom32 substitution matrix.

RESULTS AND DISCUSSION

Recombinant and natural uterocalin proteins

hEUC and rEUC were obtained as detailed in the Materials and methods section. Because of the short supply of the naturally produced protein, all of the ligand binding and intrinsic fluorescence experiments reported here were performed with recombinant material; however, hEUC showed essentially no difference in behaviour from recombinant material in more limited parallel experiments. The mRNA for EUC encodes a hydrophobic leader sequence which is predicted by Signalp to be cleaved between positions 18 and 19 (PHA-LH); a mass-spectroscopic analysis of hEUC provided results consistent with removal of the leader peptide (results not shown). The recombinant protein was therefore designed with this putative secretory signal removed. The amino acid sequence of EUC aligns better with those members of the lipocalin family that are transporters of small hydrophobic ligands rather than those with enzymic activities, so its ability to bind small lipids was investigated.

Fatty acid binding

The fatty acid binding activity was examined with environment-sensitive fluorescent fatty acid analogues that altered their fluorescence emission intensities and wavelength of peak emission on entry into binding proteins. Binding of DAUDA to rEUC was accompanied by a substantial increase in its fluorescence intensity and a shift in the fluorescence emission maximum from 543 to 483 nm (Figure 1A), indicative of entry into a strongly apolar environment [20]. In another fatty-acid-binding lipocalin,

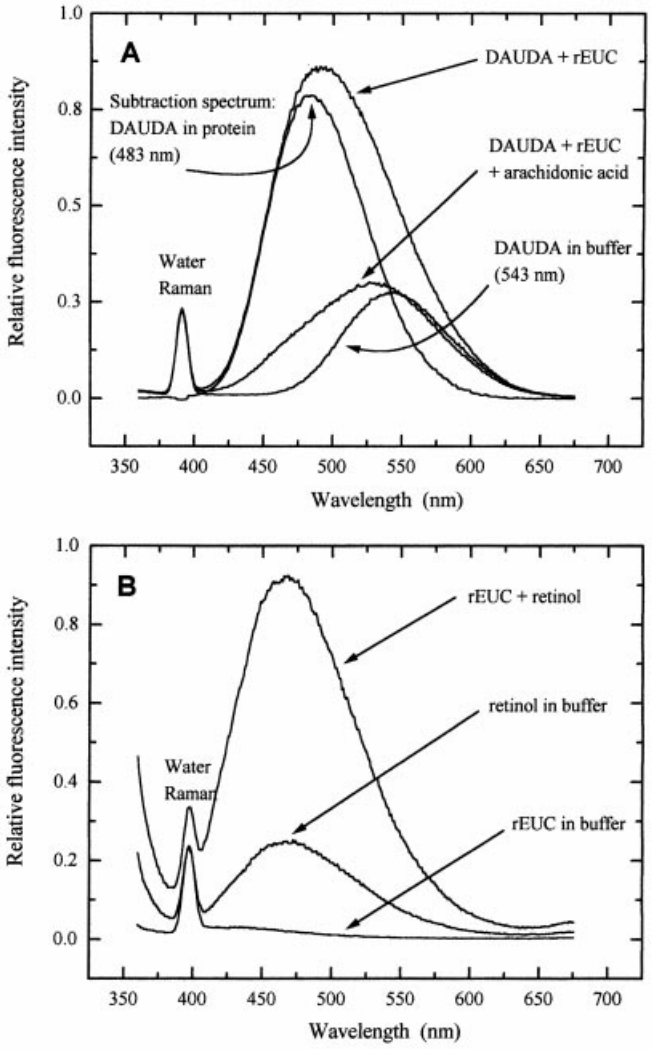

Figure 1 Fatty acid and retinol binding by rEUC

(**A**) Fluorescence emission spectra (excitation wavelength 345 nm) of 0.87 μM DAUDA alone in buffer or after the addition of 0.56 μM rEUC. The wavelength of peak emission by DAUDA when bound to rEUC is as indicated by the subtraction spectrum. Also shown is the reversal of changes in DAUDA emission by competition with arachidonic acid (0.47 μM) added to the rEUC–DAUDA complex. (**B**) Fluorescence emission spectra (excitation wavelength 350 nm) of 5.4 μM retinol in buffer or when added to 0.96 μM rEUC. Although not shown, the addition of oleic acid to the mixture resulted in a reversal of the enhancement of retinol fluorescence emission.

the blue shift in DAUDA emission is not so marked; tear lipocalin produces a shift to 490 nm [21], and lipocalins such as bovine milk β-lactoglobulin and human plasma RBP do not bind DAUDA ([22], and M. W. Kennedy, unpublished work). To place this blue shift in the context of non-lipocalin fatty acid transporters from mammals, fatty-acid-binding protein (FABP) from rat liver and from rat intestine produce shifts in DAUDA emission to 496 and 492 nm respectively [23,24]. The degree of blue shift in fluorescence emission with rEUC was therefore larger than many mammalian fatty acid transporter proteins, which presumably also means that its binding site is unusually apolar among lipocalins [20]; however, more extreme blue shifts in DAUDA emission have been observed for some novel FABPs specific to nematode worms [22,25].

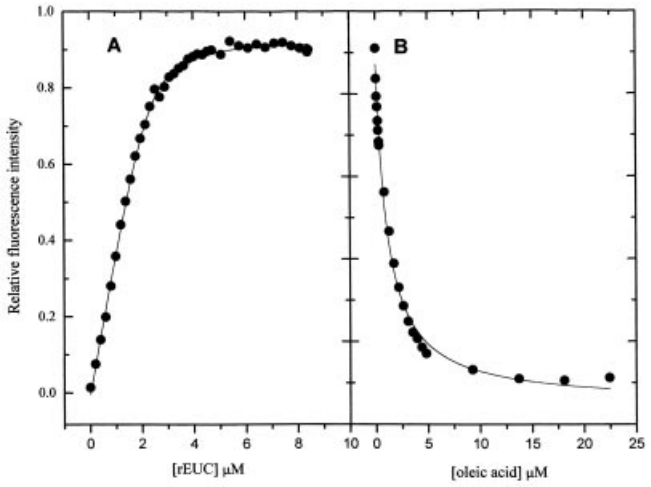

Figure 2 Titration curves for the binding of DAUDA and oleic acid to rEUC

(**A**) Change in relative fluorescence intensity at 482 nm (corrected for dilution; excitation wavelength 345 nm) of 2.34 μM DAUDA after the addition of increasing concentrations of rEUC. The solid line is the theoretical binding curve for complex formation with a K_d of 120 nM and an apparent stoichiometry consistent with one binding site per protein monomer unit. (**B**) Decrease in relative fluorescence due to the displacement of DAUDA from rEUC by oleic acid. Increasing concentrations of oleic acid were added to a mixture containing 2.34 μM DAUDA and 1.96 μM rEUC. The solid line is a theoretical curve for simple competitive binding of oleic acid in the DAUDA binding site of rEUC, with an apparent $K_i^{\text{oleic acid}}$ of 50 nM.

Fluorescence titration experiments showed that the binding of DAUDA to rEUC was saturable (Figure 2A), providing a dissociation constant, K_d, of 1.2×10^7 M^{-1}, which falls within the range recorded for other lipid transporter proteins [26,27]; the stoichiometry was consistent with 1:1 binding. In competitive titration experiments (Figure 2B), oleic acid was found to bind with an estimated $K_i^{\text{oleic acid}} \approx 50$ nM. This represents stronger binding than with DAUDA, perhaps because the dansyl reporter group interferes with entry to the binding site. Other natural fatty acids which were found to displace DAUDA from the binding site on rEUC included arachidonic, docosahexaenoic, γ-linolenic, *cis*-eicosapentaenoic, linoleic and palmitic acids, the estimated K_d values being similar to that for oleic acid (results not shown). No significant binding was observed with cholesterol, by competition with DAUDA, and rEUC also failed to alter the fluorescence emission of the intrinsically fluorescent steroid dehydroergosterol [28] or the dansyl fluorophore unattached to a fatty acid (dansylamide; results not shown).

rEUC also interacted with the fluorescent fatty acids *cis*-parinaric acid, 12-(9-anthroxyl)stearic acid and DACA, plus the general hydrophobic probe 8-anilinonaphthalene-l-sulphonic acid, all of which were displaceable with oleic or arachidonic acids (results not shown). Perhaps the most instructive of these was DACA, in which the environment-sensitive dansyl fluorophore is attached at the α-carbon rather than the methyl/ω-end of the hydrocarbon chain, as in DAUDA. The shift in emission maximum of DACA on interaction with rEUC was very similar to that occurring in DAUDA (485 and 483 nm respectively; results not shown). The fact that similar shifts in fluorescence occur regardless of which end the probe is fluorescently tagged presumably indicates that the ligand is taken into the binding site in its entirety.

Lipocalins and members of the FABP/P2/CRBP/CRABP family of β-structure-rich cytosolic fatty-acid-binding, retinol-binding and retinoic-acid-binding proteins are known from crystal and NMR structures usually to enclose their lipid ligands completely [5,29–32]. Among cytosolic lipid-binding proteins, the only exception seems to be liver FABP, which holds one fatty acid with its carboxylate exposed on the surface of the protein and another completely internalized [33]. In the lipocalin β-lactoglobulin, the carboxylate of a bound fatty acid is exposed on the surface of the protein [31], although retinol might be internalized to a greater degree [29,30]. With EUC, however, the similar fluorescence emission spectra of bound DAUDA and DACA indicates the entry of both the hydrocarbon tail and the carboxylate into a similarly apolar environment.

Fatty acid binding produces changes in the intrinsic tryptophan fluorescence emission spectrum of both β-lactoglobulin and tear lipocalin [31,34]. The tryptophan residue involved is conserved in all but one member of the lipocalin family and is positioned in the wall of the central binding cavity of these proteins [29,30]. EUC possesses two tryptophan residues, Trp-37, which is equivalent to the universally conserved Trp of lipocalins, and Trp-150, which seems to be unusually disposed on the surface of the protein (see below). Unlike with β-lactoglobulin and tear lipocalin, however, there was no detectable alteration in the tryptophan fluorescence emission spectrum of rEUC on the addition of fatty acid (results not shown). The inference is therefore that a fatty acid ligand occupies a different orientation or position within the binding cavity of EUC from that in β-lactoglobulin.

Retinol binding

Several lipocalins bind retinol, the best understood being bovine β-lactoglobulin and plasma RBP [15]. The function of the former is unclear, but it is considered to act as a possible carrier for retinol to the newborn via milk [15,30,35]; EUC might have an analagous function in delivering developmentally critical retinoids to the conceptus. Binding of retinol by rEUC was therefore investigated by exploiting the change in the intrinsic fluorescence of retinol that occurs when it binds to proteins. Because retinol is highly insoluble and unstable in water, the binding assay was modified such that retinol was added as a solution in ethanol directly to protein solution in the fluorescence cuvette [17]. This showed that rEUC bound retinol (Figure 1B), although the fluorescence enhancement was not as strong as we had observed with other RBPs [36]. Nevertheless, fluorescence titration experiments provided a dissociation constant for retinol–rEUC of 1.2×10^{-7} M^{-1}, and the stoichiometry was again consistent with 1:1 binding (results not shown). Retinol was displaceable by the addition of oleic or arachidonic acid to a retinol–rEUC complex (results not shown), indicating that the binding sites for the two types of lipid are coincident or interactive.

The wavelength of peak fluorescence emission of retinol in rEUC is 468 nm, which, under the conditions that we used, is similar to the value, 467 nm, obtained with plasma RBP (results not shown). However, with β-lactoglobulin we found the peak emission of bound retinol to be at 475 nm (results not shown); other authors have also found that β-lactoglobulin–retinol complexes emit light at longer wavelengths than with other lipocalins (for example 480 nm [34]). The unusual optical behaviour of retinoids in β-lactoglobulin is thought to be due to a lysine residue close to the ionone ring [29], which is occupied by a methionine residue in EUC and a valine residue in plasma RBP.

Structural model

The lipocalins are all rich in β structure and have eight antiparallel β-strands forming a flattened barrel, within which lies the central binding cavity for fatty acids and retinoids [15,30,31].

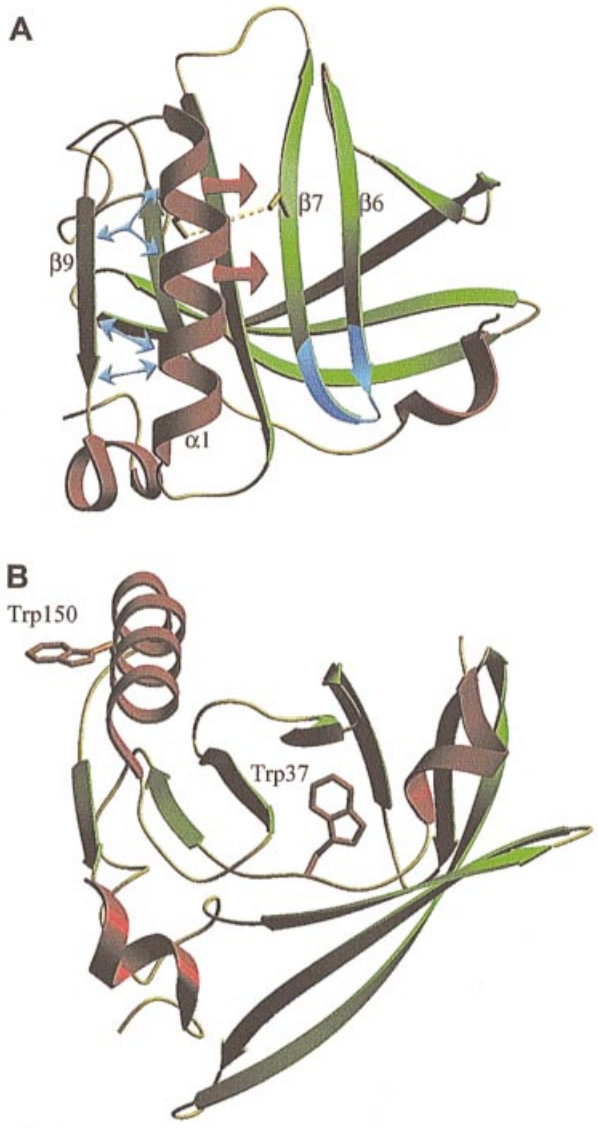

Figure 3 Structural model of EUC

(**A**) Model of uterocalin, with indications of regions of potential difference between the model and rodent major urinary protein. The loop between β-6 and β-7 is highlighted in blue, which in 90% of lipocalins contains the Thr-Asp-Tyr motif but is Val-Asp-Pro in uterocalin. Where present, the tyrosine residue forms a hydrogen bond to the backbone of a residue close to the N-terminus. In uterocalin, tyrosine is replaced by proline, which is unable to make this contact or to achieve the required φ/ψ angles. This loop is therefore likely to be substantially altered, possibly having a knock-on effect on β-6 and β-7. Also indicated (with a yellow dotted line) is a potential disulphide bond in EUC. Although the sulphur atoms are separated by 6.8 Å, a relatively small movement or rotation of the α-1 helix (indicated by red arrows) would enable the bond to be formed. Energy minimization of the structure with this disulphide bond formed indicates that the alteration is feasible. In support of this, Cys-106 on β-7 is conserved in lactoglobulin, where it makes an alternative disulphide link with a cysteine residue on β-8. Conserved salt bridges are also made in most lipocalins from α-1 to β-9 (represented by blue arrows). In uterocalin, these residues are different and the helix-positioning interactions are absent. The lack of anchoring salt-bridges and of a potential disulphide bond indicates that the α-1 helix is significantly different from the major urinary protein of rodents (PDB accession number 1MUP) and is re-positioned closer to β-6. The image was created with Setor [52]. (**B**) Alternative orientation to illustrate the relative positions of Trp-150 (occupied by Glu in the majority of lipocalins, and replaced by Glu in the W150E mutant protein), projecting directly from the main helix into the solvent, and Trp-37, which is buried within the structure of the protein and is absolutely conserved in all lipocalins.

The PHD program predicts EUC to be β-rich (14% helix, 40% extended/β structure and 46% loop/remainder) on the basis of a multiple alignment, assembled from the EUC sequence and similar lipocalins by the MaxHom routine. A CD analysis of rEUC provided estimates of 13% helix, 35% extended/β structure and 52% turn/remainder (results not shown), values approximating those predicted within the uncertainties of secondary-structure prediction from primary structure and CD. SwissModel was used to create a three-dimensional model (Figure 3) based on the coordinates of experimentally derived lipocalin structures identified by PHD as being most similar to EUC in sequence; these were α_{2u}-globulin (PDB accession no. 2A2G [37]) and mouse urinary pheromone-carrier protein (PDB accession no. 1MUP [38]). The amino acid sequence similarity between EUC and the lipocalins whose structures have been solved is high (51% and 53%), so confidence in the model is similarly high. However, inspection of the model and sequence alignments identified two regions in which EUC differs from α_{2u}-globulin and major urinary protein in potentially functionally important ways (see below and the legend to Figure 3).

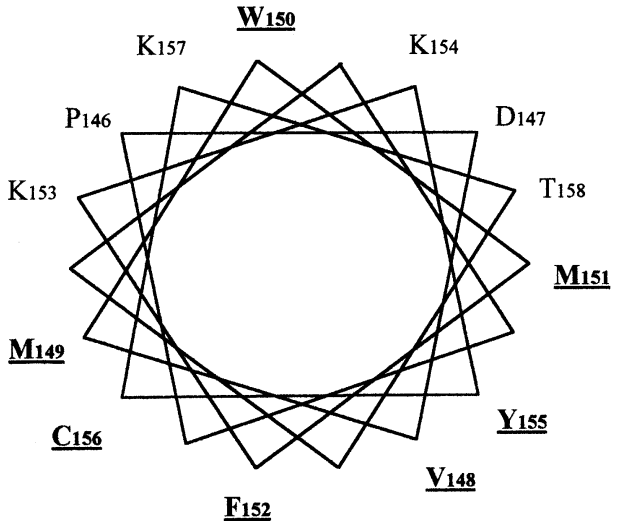

Figure 4 Amphipathicity of the EUC helix

Helical wheel projection (single-letter amino acid codes) of the main helix in EUC, illustrating both the contrasting nature of the opposing hydrophobic and charged faces of the helix and also how Trp-150 is surrounded by charged amino acids. A statistical analysis of neutral substitutions in proteins indicates that the Val, Ile, Met and Leu (aliphatic) and Phe, Tyr and Trp (aromatic) groups are interchangeable bulky hydrophobic amino acids that are commonly interchanged within a group [53,54]; analyses of crystal structures indicate that Cys, Phe, Ile, Val, Leu, Met and Trp can be grouped as amino acids that are only rarely exposed on the outside of proteins [55,56]. Those amino acids that fall into these categories are underlined and shown in **bold**.

Unusual external exposure of Trp-150

An unusual feature of the model is a tryptophan side chain (Trp-150) projecting directly out of the external helix into solvent (Figure 3B); no experimentally derived lipocalin structure has such a bulky apolar side chain in this position. Moreover, multiple alignments with all known lipocalin sequences performed with either the Dayhoff or Blosum62 substitution matrices showed that this position is usually occupied by a charged amino acid, most frequently a glutamic residue (results not shown). A helical wheel projection of the relevant section of the protein (Figure 4) illustrates the amphipathicity of this helix and shows that Trp-150 is clearly in a charged hydrophilic face. Moreover, the Ramachandran plot of the helical region of the

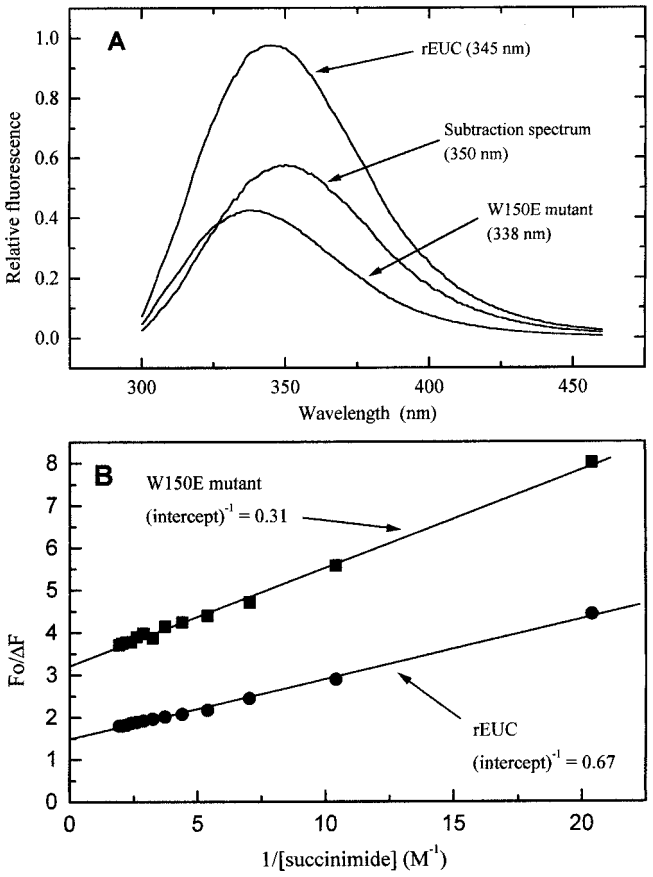

Figure 5 Unusual degree of exposure of uterocalin Trp-150 to solvent

(**A**) Intrinsic fluorescence emission spectra of equivalent quantities of rEUC protein and the W150E mutant protein, with the wavelengths of peak emission as indicated. The excitation wavelength was 290 nm. The spectrum of rEUC is a composite of the emission of Trp-37 and Trp-150, and that of the W150E protein represents the emission of Trp-37 only. The subtraction spectrum should therefore approximate the emission characteristics of Trp-150 in isolation; a peak at 350 nm is indicative of an unusual degree of exposure to solvent water. (**B**) Modified Stern–Volmer plots [41] for the progressive quenching of the intrinsic fluorescence emission of rEUC and the W150E mutant by the addition of succinimide. The reciprocal of the intercept with the $F_0/\Delta F$ axis is taken as an estimate of the proportion of tryptophan residues exposed to the soluble quenching agent, that value being 0.67 for the wild-type protein and 0.31 for the W150E mutant. This again indicates an unusually high degree of solvent exposure of Trp-150. Although not shown, the same value was obtained for hEUC as for rEUC. The excitation wavelength was 290 nm.

model (results not shown) showed that all the amino acids were in strongly favoured/ideal regions for an α-helix. Trp-150 was found to be encoded by cDNA species from endometrial tissue of two unrelated mares, and is therefore probably not the product of an uncommon allele or an error in cloning or sequencing. The exposed tryptophan residue in EUC could therefore be of functional importance; its disposition was therefore investigated directly.

A combination of intrinsic fluorescence measurements and succinimide quenching was used to discriminate between exposed and buried tryptophan side chains [39,40]. The fluorescence emission spectrum of rEUC excited at 290 nm gave a maximum at 345 nm (Figure 5A). Successive additions of succinimide to rEUC produced a blue shift in the wavelength of maximum emission to 336 nm (results not shown), indicative of differences between the solvent environments of rEUC's two tryptophan side chains. Conventional Stern–Volmer [40] analysis of the succinimide quenching data gave a non-linear plot, again indicative of heterogeneity in the environments of the two tryptophan residues. However, the modified, double-reciprocal Stern–Volmer plot [41] yielded a straight line (Figure 5B), providing an estimate for the fraction of exposed tryptophan residues of 0.67. Taken together, these results are consistent with Trp-150 indeed being distinctive in its exposure to solvent.

The disparity in the environments of the two tryptophan residues was investigated further with a mutant protein in which Trp-150 was replaced by a non-fluorescent amino acid (Glu). The replacement of Trp-37 was not performed because this position is absolutely conserved in the lipocalin family and its substitution can compromise the structure of the proteins [21]. The W150E mutant was found to be unaffected in its binding of fatty acids and retinol, so the mutant protein was taken to be structurally intact. The peak in intrinsic fluorescence of W150E occurred at 338 nm, indicating that the remaining tryptophan side chain (Trp-37) is isolated from solvent water to a high degree (Figure 5A). Subtraction of this emission spectrum from that of an equivalent concentration of the recombinant wild-type protein should approximate to the emission characteristics of Trp-150 in isolation. This gave a maximum at 350 nm, again indicating that Trp-150 is exposed to solvent to an unexpectedly high degree (Figure 5A). Quenching of the intrinsic fluorescence of the W150E mutant proceeded very slowly, although linearly, with increasing concentrations of succinimide, and with no significant alteration in the wavelength of maximum emission. These data provided a K_{sv} value of 0.22 M^{-1} (results not shown), which is consistent with a highly buried position for Trp-37; the modified Stern–Volmer plot (Figure 5B) provided an estimate of the proportion of exposed tryptophan residues as 0.31, which is lower than that for the wild-type protein (0.67). These experiments therefore confirmed the disparity in the environments of rEUC's two tryptophan residues, and indicated that Trp-150 is unusually disposed on the exterior of the protein.

What is the function of such an unusually positioned tryptophan residue? Tryptophans and other bulky hydrophobic amino acids are not usually exposed on the surface of proteins unless involved in protein–protein [42–44] and protein–membrane [45–49] interactions. For instance, a single exposed aromatic side chain accounts for 43% of the intramolecular interactions between the gp120 protein of HIV and CD4 of the lymphocyte surface [42]. Tryptophan residues are frequently positioned on the outside of membrane-interactive or integral membrane proteins and there is spectroscopic evidence that indoles can penetrate into membranes, positioning themselves interfacially in the region of the glycerol/ester region of phospholipid bilayers [48,49]. Trp-150 of EUC could therefore be involved either in the association of the protein with the trophoblast surface to deposit its cargo directly into the membrane or in the association with a specific receptor protein. Alternatively, it could be involved in a two-stage event in which the protein first associates with the trophoblast membrane. This will increase the likelihood of encounter with a receptor protein [50] that then acquires the ligand and translocates it across the membrane. EUC is known to accumulate within the trophoblast cells and in the yolk sac of the developing conceptus [12], so interaction with a surface receptor protein and subsequent internalization of protein plus ligand is an attractive possibility. Interaction with a specific receptor protein would favour the transfer of essential fatty acids and retinoids in one direction from maternal tissue to the trophoblast and avoid wasteful reabsorption by endometrial cells. However, it is also possible that the composition of the trophoblast membrane itself ensures the specificity of interaction in some way, without the need for a receptor.

EUC has the appropriate binding properties, location and timing of appearance for the provisioning of the preattachment equine conceptus with essential fatty acids and retinoids. Because it seems to be absorbed by the embryonic tissues, it could have a dual function as a transporter of essential lipids and as a protein source. EUC is distinct from equine plasma RBP, and has a broader ligand binding propensity than do plasma RBPs, so EUC might be adapted to transport a greater range of small lipids than equine plasma RBP could. It might also be that the protein needs to be of a different amino acid composition and charge from RBPs to interact with or pass through the capsule; the capsule is negatively charged, and the pI of EUC is predicted from its amino acid composition to be 9.7, whereas by the same method (ProtParam) the pI of equine plasma RBP is 5.5.

The human placenta possesses a trophoblast surface fatty acid transporter protein (p-FABP$_{pm}$) that transfers selected long-chain PUFAs into the foetal circulation [6]. This preferentially transports docosahexaenoic and arachidonic acids, which are particularly important in development, but little or no eicosapentaenoic acid, which is a growth-inhibitory eicosanoid. However, EUC seems to show no such selectivity. This could mean either that the latter fatty acid does not have the same inhibitory effect in horses or that it is excluded by the conditions under which the protein is loaded with lipids (presumably in the EUC-secreting cells).

A fibrous capsule bounding the conceptus is unusual in the eutherian placental mammals, but some marsupials display a similar feature. For instance, the conceptuses of *Didelphis virginiana* and *Monodelphis domestica* (the New World opossums) are bounded by a proteinaceous capsule that is superficially similar to that of horses [51]. It might therefore be that horses and some marsupials have retained an ancestral condition that has been lost with the evolution of advanced placentation in eutherians, and that a uterocalin-like protein is produced in marsupials but has been lost in non-equine eutherians. An alternative explanation would be that a uterocalin homologue is produced during the short preimplantation periods of other (perhaps all?) mammals, or in those in which there is delayed implantation or embryonic diapause. In either case, it is conceivable that the evolution of uterocalin was an essential step in the evolution of viviparity in mammals.

We thank Fiona McMonagle, Alan Cooper and Lindsay McDermott for their help in protein preparations and advice on the analysis of the lipid binding assays, Ross Kennedy for his help with the retinol binding assays, and Sharon Kelly and Nick Price for the CD analysis. This work was supported by The Wellcome Trust in a grant to M. W. K. and Alan Cooper (grant number 044156) and from the European Commission in a grant to S. S. (B104.CT97.5142).

REFERENCES

1 Dutta-Roy, A. K. (1997) Fatty acid transport and metabolism in the feto-placental unit and the role of fatty acid-binding proteins. Nutr. Biochem. **8**, 548–557
2 Morriss-Kay, G. M. and Sokolova, N. (1996) Embryonic development and pattern formation. FASEB J. **10**, 961–968
3 Noy, N. (2000) Retinoid-binding proteins: mediators of retinoid activity. Biochem. J. **348**, 481–495
4 Dutta-Roy, A. K. (2000) Cellular uptake of long chain fatty acids: role of membrane associated fatty acid binding/transport proteins. Cell. Mol. Life Sci. **57**, 1360–1372
5 Coe, N. R. and Bernlohr, D. A. (1998) Physiological properties and functions of intracellular fatty acid-binding proteins. Biochim. Biophys. Acta **1391**, 287–306
6 Campbell, F. M., Clohessy, A. M., Gordon, M. J., Page, K. R. and Dutta-Roy, A. K. (1997) Uptake of long chain fatty acids by human placental choriocarcinoma (BeWo) cells: role of plasma membrane fatty acid-binding protein. J. Lipid Res. **38**, 2558–2568
7 Samuel, C. A., Allen, W. R. and Steven, D. H. (1974) Studies on the equine placenta. I. Development of the microcotyledons. J. Reprod. Fertil. **41**, 441–445
8 Betteridge, K. J., Eaglesome, M. D., Mitchell, D., Flood, P. F. and Beriault, R. (1982) Development of horse embryos up to twenty two days after ovulation: observations on fresh specimens. J. Anat. **135**, 191–209
9 Betteridge, K. J. (1989) The structure and function of the equine capsule in relation to embryo manipulation and transfer. Equine Vet. J. **8** (Suppl.), 92–100
10 Oriol, J. G., Sharom, F. J. and Betteridge, K. J. (1993) Developmentally regulated changes in the glycoproteins of the equine embryonic capsule. J. Reprod. Fertil. **99**, 653–664
11 Crossett, B., Allen, W. R. and Stewart, F. (1996) A 19 kDa protein secreted by the endometrium of the mare is a novel member of the lipocalin family. Biochem. J. **320**, 137–143
12 Crossett, B., Suire, S., Herrler, A., Allen, W. R. and Stewart, F. (1998) Transfer of a uterine lipocalin from the endometrium of the mare to the developing equine conceptus. Biol. Reprod. **59**, 483–490
13 Stallingsmann, M. L., Trout, W. E. and Roberts, R. M. (1993) Porcine uterine retinol-binding proteins are identical gene-products to the serum retinol-binding protein. Biol. Reprod. **48**, 998–1005
14 Buhi, W. C., Alvarez, I. M., Shille, V. M., Thatcher, M.-J. and Harney, J. P. (1995) Purification and characterization of a uterine retinol-binding protein in the bitch. Biochem. J. **311**, 407–415
15 Flower, D. R. (1996) The lipocalin protein family: structure and function. Biochem. J. **318**, 1–14
16 Magert, H., Hadrys, T., Cieslakk, A., Groger, A., Feller, S. and Forssmann, W. (1995) cDNA sequence and expression pattern of the putative pheromone carrier aphrodisin. Proc. Natl. Acad. Sci. U.S.A. **92**, 2091–2095
17 Cogan, U., Kopelman, M., Mokady, S. and Shinitzky, M. (1976) Binding affinities of retinol and related compounds to retinol binding proteins. Eur. J. Biochem. **65**, 71–78
18 Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) The Protein Data Bank. Nucleic Acids Res. **28**, 235–242
19 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. **26**, 283–291
20 Macgregor, R. B. and Weber, G. (1986) Estimation of the polarity of the protein interior by optical spectroscopy. Nature (London) **319**, 70–73
21 Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N. and Glasgow, B. J. (1999) Binding studies of tear lipocalin: the role of the conserved tryptophan in maintaining structure, stability and ligand affinity. Biochim. Biophys. Acta **1433**, 307–320
22 Kennedy, M. W., Brass, A., McCruden, A. B., Price, N. C., Kelly, S. M. and Cooper, A. (1995) The ABA-1 allergen of the parasitic nematode *Ascaris suum*: fatty acid and retinoid binding function and structural characterization. Biochemistry **34**, 6700–6710
23 Thumser, A. E. and Wilton, D. C. (1994) Characterization of binding and structural properties of rat liver fatty-acid-binding protein using tryptophan mutants. Biochem. J. **300**, 827–833
24 Veerkamp, J. H., van Moerkerk, H. T. B., Prinsen, C. F. M. and van Kuppevelt, T. H. (1999) Structural and functional studies on different FABP types. Mol. Cell. Biochem. **192**, 137–142
25 Kennedy, M. W., Britton, C., Price, N. C., Kelly, S. M. and Cooper, A. (1995) The DvA-1 polyprotein of the parasitic nematode *Dictyocaulus viviparus*: a small helix-rich lipid-binding protein. J. Biol. Chem. **270**, 19277–19281
26 Wilkinson, T. C. I. and Wilton, D. C. (1986) Studies on fatty acid-binding proteins. The detection and quantification of the protein from rat liver using a fluorescent fatty acid analogue. Biochem. J. **238**, 419–424
27 Thumser, A. E. A., Evans, C., Worrall, A. F. and Wilton, D. C. (1994) Effect on ligand binding of arginine mutations in recombinant rat liver fatty acid-binding proteins. Biochem. J. **297**, 103–107
28 Thumser, A. E. A. and Wilton, D. C. (1996) The binding of cholesterol and bile salts to recombinant rat liver fatty acid-binding protein. Biochem. J. **320**, 729–733
29 Cho, Y., Batts, C. A. and Sawyer, L. (1994) Probing the retinol-binding site of bovine β-lactoglobulin. J. Biol. Chem. **269**, 11102–11107
30 Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C. T., Findlay, J. B. C., Sivaprasadarao, R., Jones, T. A., Newcomer, M. E. and Kraulis, P. J. (1986) The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein. Nature (London) **324**, 383–385
31 Wu, S.-Y., Perez, M. D., Puyol, P. and Sawyer, L. (1999) β-Lactoglobulin binds palmitate within its central cavity. J. Biol. Chem. **274**, 170–174
32 Sacchettini, J. C. and Gordon, J. I. (1993) Rat intestinal fatty acid binding protein: a model system for analyzing the forces that can bind fatty acids to proteins. J. Biol. Chem. **268**, 18399–18402

33 Thompson, J., Winter, N., Terwey, D., Bratt, J. and Banaszak, L. (1997) The crystal structure of the liver fatty acid-binding protein: a complex with two bound oleates. J. Biol. Chem. **272**, 7140–7150
34 Frapin, D., Dufour, E. and Haertle, T. (1993) Probing the fatty acid binding site of β-lactoglobulins. J. Prot. Chem. **12**, 443–449
35 Sawyer, L. and Kontopidis, G. (2000) The core lipocalin, bovine β-lactoglobulin. Biochim. Biophys. Acta **1482**, 136–148
36 Kennedy, M. W., Garside, L. H., Goodrick, L. E., McDermott, L., Brass, A., Price, N. C., Kelly, S. M., Cooper, A. and Bradley, J. E. (1997) The Ov20 protein of the parasitic nematode *Onchocerca volvulus*. A structurally novel class of small helix-rich retinol-binding protein. J. Biol. Chem. **272**, 29442–29448
37 Chaudhuri, B. N., Kleywegt, G. J., Bjorkman, J., Lehman-McKeeman, L. D., Oliver, J. D. and Jones, T. A. (1999) The structures of α_{2u}-globulin and its complex with a hyaline droplet inducer. Acta Crystallogr. **D55**, 753–762
38 Bocskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E. V., Cavaggioni, A., Findlay, J. B. C. and North, A. C. T. (1992) Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. Nature (London) **360**, 186–188
39 Eftink, M. R. and Ghiron, C. A. (1976) Exposure of tryptophanyl residues in proteins: quantitative determination by fluorescence quenching studies. Biochemistry **15**, 672–679
40 Eftink, M. R. and Ghiron, C. A. (1984) Indole fluorescence quenching studies on proteins and model systems: use of the efficient quencher succinimide. Biochemistry **23**, 3891–3899
41 Lehrer, S. S. (1971) Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. Biochemistry **10**, 3254–3263
42 Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. and Hendrickson, W. A. (1998) Structure of an HIV complex with the CD4 receptor and a neutralizing human antibody. Nature (London) **393**, 648–659
43 Guo, Z., Zhou, D. and Schultz, P. G. (2000) Designing small-molecule switches for protein–protein interactions. Science **288**, 2042–2045
44 Bogan, A. A. and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces. J. Mol. Biol. **280**, 1–9
45 Gelb, M. H., Cho, W. and Wilton, D. C. (1999) Interfacial binding of secreted phospholipase A_2: more than electrostatics and a major role for tryptophan. Curr. Opin. Struct. Biol. **9**, 428–432
46 Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. and MacKinnon, R. (1998) The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. Science **280**, 69–77
47 Baker, S. F., Othman, R. and Wilton, D. C. (1998) Tryptophan-containing mutant of human (Group IIa) secreted phospholipase A_2 has a dramatically increased ability to hydrolyze phosphatidylcholine vesicles and cell membranes. Biochemistry **37**, 13203–13211
48 Kachel, K., Asuncionpunzalan, E. and London, E. (1995) Anchoring of tryptophan and tyrosine analogs at the hydrocarbon polar boundary in model membrane-vesicles – parallax analysis of fluorescence quenching induced by nitroxide-labelled phospholipids. Biochemistry **34**, 15475–15479
49 Yau, W. M., Wimley, W. C., Gawrisch, K. and White, S. H. (1998) The preference of tryptophan for membrane interfaces. Biochemistry **37**, 14713–14718
50 Eigen, M. (1974) Diffusion control in biochemical reactions. In Quantum Statistical Mechanics in the Natural Sciences (Kursunoglu, B., Mintz, S. L. and Widmayer, S. M., eds.), pp. 37–61, Plenum Press, New York
51 Harder, J. D., Stonerook, M. J. and Pondy, J. (1993) Gestation and placentation in two New World opossums: *Didelphis virginiana* and *Monodelphis domestica*. J. Exp. Zool. **266**, 463–479
52 Evans, S. V. (1993) SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. J. Mol. Graph. **11**, 134–138
53 Dayhoff, M. O., Schwartz, R. M. and Orcutt, B. C. (1978) A model of evolutionary change in proteins. In Atlas of Protein and Sequence Structure, vol. 5, Suppl. 3 (Dayhoff, M. O., ed.), pp. 345–362, National Biomedical Research Foundation, Washington DC
54 Bordo, D. and Argos, P. (1991) Suggestions for 'safe' residue substitutions in site-directed mutagenesis. J. Mol. Biol. **217**, 721–729
55 Rose, G., Geselowitz, A., Lesser, G., Lee, R. and Zehfus, M. (1985) Hydrophobicity of amino acid residues in globular proteins. Science **229**, 834–838
56 Janin, J. (1979) Surface and inside volumes in globular proteins. Nature (London) **277**, 491–492

Received 13 December 2000/16 February 2001; accepted 8 March 2001