Lipocalins are promising drug candidates, either based on their natural ligand-binding functions or as engineered 'anticalins' with novel specificities.

Keynote review:

Lipocalins in drug discovery: from natural ligand-binding proteins to 'anticalins'

Steffen Schlehuber and Arne Skerra

Lipocalins are a widespread family of small, robust proteins that typically transport or store biological compounds which are either of low solubility or are chemically sensitive, including vitamins, steroid hormones, odorants and various secondary metabolites. There are approximately ten different lipocalins in the human body, with the plasma retinol-binding protein being the most well known. Some lipocalins have a pathophysiological role, which opens possibilities for their use in medical applications. Furthermore, lipocalins from blood-sucking insects have evolved as scavengers for mediators of inflammation. As well as using the natural ligand-binding function, lipocalins have also been recruited as scaffolds for the design of artificial binding proteins termed 'anticalins'®. These novel proteins have potential applications as antidotes, antagonistic protein therapeutics or as target-recognition modules in a new generation of immunotoxins.

The lipocalins represent a family of functionally diverse, small proteins that comprise 160–180 amino acid residues and have weak sequence homology but high similarity at the tertiary structural level [1,2]. Members of this family have important biological functions in a variety of organisms, from bacteria to humans. The majority of lipocalins are responsible for the storage and transport of compounds that have low solubility or are chemically sensitive, such as vitamins, steroids and metabolic products [3]. The human plasma retinol-binding protein (RBP) was the first lipocalin for which a 3D structure was elucidated [4,5]; RBP transports the poorly soluble and oxidation-prone vitamin A from the liver, where it is stored as a fatty acid ester, to several target tissues. Some lipocalins appear to have a more specialized role in vertebrates [6], participating, for example, in olfaction and

Steffen Schlehuber
PIERIS Proteolab AG,
Freising-Weihenstephan,
Germany
Arne Skerra*
Lehrstuhl für Biologische Chemie,
Technische Universität München,
Freising-Weihenstephan,
Germany
*e-mail: skerra@wzw.tum.de

STEFFEN SCHLEHUBER
Steffen Schlehuber was born in Fulda, Germany, and studied chemistry at the Technical University of Darmstadt, where he specialized in biochemistry. He went on to complete a doctoral thesis in the laboratory of Arne Skerra at the Technical University of Munich, Germany, obtaining his PhD in 2001. During his doctoral study, Schlehuber was involved in the development of anticalins, which are engineered ligand-binding proteins derived from natural lipocalin proteins. Steffen Schlehuber is cofounder and CSO of PIERIS Proteolab AG, a biotechnology company situated in Freising-Weihenstephan, Germany. Founded in 2001, PIERIS focuses on the development and commercialization of anticalins for therapeutic and diagnostic uses, predominantly in the area of oncology and cardiovascular diseases.

ARNE SKERRA
Arne Skerra was born in Wiesbaden, Germany, and studied chemistry at the Technical University of Darmstadt. In 1989, he received his PhD at the Ludwig-Maximilians University, Munich, where he had performed, under the supervision of Andres Plückthun and Ernst-Ludwig Winnacker, important research on the bacterial expression of functional antibody fragments. After spending one year as a postdoctoral research fellow with Greg Winter and Cesar Milstein at the MRC Laboratory of Molecular Biology in Cambridge, UK, he joined the department of Hartmut Michel at the Max-Planck-Institute of Biophysics in Frankfurt am Main. In 1994, Skerra became Professor of Protein Chemistry at the Technical University of Darmstadt. Four years later he moved to the Technical University of Munich, where he was appointed a Full Professor to the Chair of Biological Chemistry at the Life Science Campus, Weihenstephan. Skerra is Chairman of the study group on protein engineering and design at the Society for Biochemistry and Molecular Biology and a Board Member of the biochemistry section of the Society of German Chemists. In 2001, he cofounded the biotechnology start-up company PIERIS Proteolab AG.

1359-6446/04/$ – see front matter PII: S1359-6446(04)03294-5

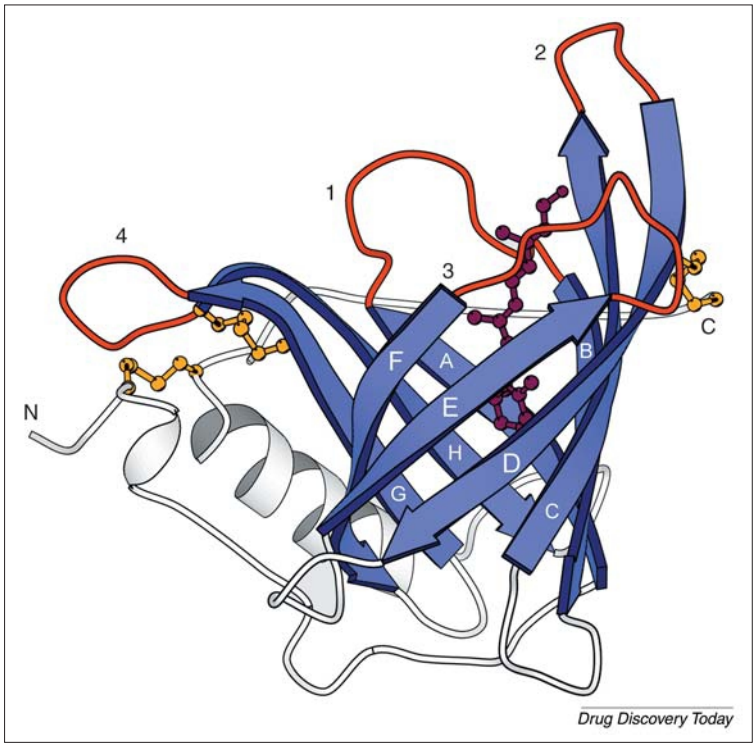

FIGURE 1

General structure of human retinol-binding protein, a prototypic lipocalin. Ribbon diagram (MolScript software) of the crystal structure of RBP (Protein Data Bank entry 1RBP) with the bound ligand retinol (ball and stick representation in magenta). The eight antiparallel strands of the conserved β-barrel structure are shown in blue with labels A to H, and the four loops, which are highly variable among the lipocalin family, are coloured red and numbered. The typical α-helix that is attached to the central β-barrel in all lipocalins, the loops at the closed end and the N- and C-terminal peptide segments are shown in grey. The three disulfide bonds of RBP are depicted in yellow.

regulation of the immune response, and even acting as an enzyme in the formation of prostaglandin D_2.

Despite their extremely poor sequence homology, the lipocalins share a structurally conserved β-barrel as their central folding motif, which consists of eight antiparallel β-strands that are arranged in a cylindrical manner (Figure 1). At the N-terminal of the protein, the β-sheet region is preceded by a coiled polypeptide segment, whereas at the C-terminal the β-sheet is appended by a characteristic α-helix and an amino acid stretch that is in an extended conformation. Within the β-barrel, the antiparallel strands (assigned A to H) are arranged in a $(+1)_7$ topology. These strands coil in a right-handed and conical manner around a central axis such that part of the backbone of strand A can form hydrogen bonds with strand H.

One end of the β-barrel is closed by the N-terminal peptide segment that traverses the base of the barrel between two short loops, which connect strands B to C and F to G, before entering into β-strand A. Dense packing of side chains in this region and within the adjacent part of the barrel leads to the formation of a hydrophobic core. At the opposite end, the conical β-barrel is typically open to solvent and provides access to a cavity. Four loops connect the strands in a pair-wise fashion and form the entrance to a cup-shaped pocket where the cognate hydrophobic ligand is bound. Hence, the term 'lipocalin' (derived from the Greek and Latin word 'calyx' meaning drinking vessel) was coined for this protein family [7]. In contrast to the highly conserved β-barrel topology, there is considerable variation in this loop region among individual members, with differences evident in the amino acid composition, conformation and length of the contributing polypeptide segments [2], which consequently gives rise to the variety of ligand specificities observed.

Indeed, the binding site of the lipocalins can adopt extremely different shapes. The binding site can form a wide, funnel-like opening to the solvent, as in the case of neutrophil gelatinase-associated lipocalin (NGAL) [8]. Alternatively, the loops can close over the cavity within the β-barrel and fully encapsulate the ligand, as in the mouse major urinary protein (MUP) [9]. The ligand pocket can extend deep within the hydrophobic core of the β-barrel, as is observed for RBP where the β-ionone ring of retinol becomes buried at the bottom of a narrow channel that harbours the oligo-isoprene side chain [4], or the ligand pocket can form an extended cave with several lobes close to the hydrophobic core, as noted in the structure of human tear lipocalin (Tlc) [10,11].

Because lipocalins comprise a single, small polypeptide chain that adopts a simple fold (Figure 1), this family of proteins affords several benefits for protein engineering and mass production, as well as applications in biotechnology and medicine. Many lipocalins exhibit significant thermostability and thus represent robust proteins [12]. Most natural lipocalins are of a monomeric globular nature and, particularly in the human body, are abundant components of plasma, tissue and secretory fluids. Few members of the family are glycosylated, thus lipocalins can be easily produced as recombinant proteins using bacterial expression systems [13].

Lipocalins are typically secretory proteins and as such they have a varying number of disulfide bonds. Although only one disulfide bridge is conserved in human lipocalins, connecting the C-terminal polypeptide segment to the β-barrel (Figure 1), this crosslink appears to be generally dispensable for formation of the lipocalin fold because this intramolecular bridge is absent in the *Escherichia coli* lipocalin Blc [14]. Furthermore, several human lipocalins exhibit a single unpaired Cys residue. Although the function of this residue has not always been elucidated, this amino acid occasionally forms an intermolecular covalent link to another protein. For example, apolipoprotein D (ApoD) associates with apolipoprotein A-II [15] and NGAL links to neutrophil gelatinase (also called matrix metalloproteinase IX) to varying extents [8]. However, this additional Cys residue can usually be removed by site-directed mutagenesis without a resultant loss of ligand-binding activity [16].

The binding specificity for low molecular weight compounds is well-characterized for many lipocalins. Some lipocalins have high ligand specificity, for example, RBP