Review

Cell PRESS

Novel protein scaffolds as emerging therapeutic proteins: from discovery to clinical proof-of-concept

Thierry Wurch, Alain Pierré, and Stéphane Depil

Institut de Recherches Servier, Oncology R&D Unit, 125 Chemin de Ronde, F-78290 Croissy sur Seine, France

Recent advances in combinatorial protein engineering have made it possible to develop immunoglobulin (Ig)-based and non-Ig protein scaffolds that can potentially substitute for most whole antibody-associated properties and currently translate into biologicals with drug-like properties. During the past 10 years, the most validated scaffolds have reached the clinical development phase and, recently, one of them [Kalbitor® (Dyax)] has made it to the market, making these alternative scaffold proteins viable drug candidates in a post-antibody landscape. Interestingly, several scaffolds include an immune-active component as part of their therapeutic mode of action, which yielded spectacular clinical efficacy in some hematological malignancies. Here, we review the most recent clinical advances and analyze their benefits for patients.

Introduction

Monoclonal antibodies (mAbs) are currently the class of therapeutic molecules that yield the most significant evolution in terms of clinical success rates and financial turnover [1]. mAbs exhibit several key advantages as compared to small molecules such as increased safety, enhanced efficacy, longer plasmatic half-life and higher success rates in progression through the early clinical development phases [1]. mAbs are naturally bifunctional molecules because they are able to interact with and directly modulate their cognate target via the variable domain, while modulating several key immune responses via interaction with Fc receptors due to their conserved Fc domain [such as antibody-dependent cell-mediated cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC); see Glossary] (Figure 1) [1]. Nevertheless, certain limitations have appeared for this class of molecule including large size and steric hindrance restricting tissue penetration (solid tumors, poorly vascularized tissues) and planar binding interfaces making the binding to grooves and the catalytic sites of enzymes difficult [2]. In addition, strategic questions are raised, such as the high costs-of-goods (CoGs) associated with production/purification plants making mAb therapies expensive and complex intellectual property issues associated with technological aspects.

Corresponding author: Wurch, T. (thierry.wurch@fr.netgrs.com).
Keywords: protein engineering; therapeutic antibody; clinical trial; drug development.

To counter some of these constraints, about 50 different protein scaffolds have been discovered and documented during the past 20 years [3–5]. The term 'scaffold', as used in protein engineering, describes a single chain polypeptidic framework typically of reduced size (< 200 AA) and containing a highly structured core associated with variable portions of high conformational tolerance allowing insertions, deletions, or other substitutions. These scaffolds are based either on a conventional Ig backbone, or are derived from a completely unrelated protein [4,5].

Most of these novel scaffolds are being developed against disease targets, such as tumor necrosis factor-α (TNF-α), CD20, vascular endothelial growth factor (VEGF), CD19, and CD3, which have proven to be effective in the clinic in marketed targeted therapies in oncology and inflammatory diseases. Such validated targets were chosen in order to decrease potential safety risks in the clinic and increase the likelihood of achieving proof of concept as therapeutics in humans [1]. Nevertheless, a few original protein targets such as blood factors, not implicated in the mode of action of marketed drugs, have been selected by some companies and have successfully reached the clinic. As an example, ecallantide (DX-88, Dyax Corporation, USA) was the first engineered protein scaffold reaching the market in December 2009 as Kalbitor® [6,7]. This is a rationally designed Kunitz domain designed as a potent inhibitor of human plasma kallikrein and approved for the treatment of patients suffering from hereditary angioedema.

Currently, about 15 of these promising candidates have reached the stage of being developed as therapeutics and/or diagnostics. In this review, we describe the most clinically advanced protein scaffolds, summarize their pros and cons, and analyze their benefit in terms of therapeutic efficacy.

Non-Ig-based protein scaffolds

Anticalins, the engineered lipocalins

Lipocalins have been selected as a suitable protein backbone because they contain four exposed loops built on a rigid β-barrel structure (Figure 1a and Table 1). Libraries based on the cabbage butterfly (*Pieris*) bilin-binding protein, the human tear lipocalin, or the neutrophil-gelatinase-associated lipocalin have been developed that contain randomization of 16 accessible positions within the four exposed loops [8]. These non-natural lipocalin-based structures were named 'anticalins' and since 2001 have been commercially

0167-7799/$ – see front matter http://dx.doi.org/10.1016/j.tibtech.2012.07.006 Trends in Biotechnology, November 2012, Vol. 30, No. 11

Glossary

Adnectin: engineered protein scaffold derived from the 10th fibronectin type III domain of human fibronectin. This domain corresponds to a β-sandwich with seven β-strands and three connecting loops showing structural homologies to Ig domains without disulfide bridges (Table 1).

Ankyrin repeat: protein motif of 33 AA long, consisting of two alpha helices separated by loops, and widely distributed among living organisms including bacteria and viruses. Ankyrin repeats mediate protein–protein interactions and are among the most common structural motifs in known proteins.

Antibody-dependent cell cytotoxicity (ADCC): functional property of a conventional antibody, whereby effector cells of the immune system (mainly NK cells) actively lyse a target cell that has been bound by specific antibodies via their Fc portion.

Anticalin: engineered protein scaffold derived from lipocalins, which are storage or transport proteins. The two mostly used parental human proteins are tear lipocalin and neutrophil-gelatinase-associated lipocalin (NGAL). Their common structural features consist in a conserved β-barrel comprising eight antiparallel β-strands showing four exposed loops (Table 1).

AVIdity MultimER (Avimer): protein scaffold obtained by genetic oligomerization of small A-domains, typically 30–35 AA long and containing six cysteine residues for their proper folding, each engineered block being selected from a library of assorted A-domains for best binding to the desired target. Upon sequential panning of multiple engineered A-domains, they are associated by genetic fusion to build the final avimer candidate (Table 1).

Bispecific T cell Engager (BiTE): bispecific, monovalent diabody scaffold, based on the single chain variable domain antibody format (scFv) containing the heavy and light chain variable domains from a conventional antibody associated with flexible linkers rich in glycin and serine residues (Table 1 and Figure 2). A BiTE is composed of one single polypeptide chain. One binding module is specific for the CD3 antigen, expressed on the surface of T cells, and allowing T cell activation upon binding of the second BiTE module to a tumor-associated antigen (e.g., CD19).

Complement-dependent cytotoxicity (CDC): functional property of a conventional antibody whereby antibody binding to the C1q protein of the complement activates the classical activation cascade leading to formation of the membrane attack complex, the cytolytic end-product of the complement cascade.

Designed Ankyrin Repeat ProteIN (DARPIN): engineered protein scaffold resulting from rational design strategies (multiple sequence alignments and statistical analysis) based on human ankyrin repeat (AR) proteins, a family of protein–protein interaction regulators (Table 1).

Domain antibody (dAb): the smallest portion derived from natural human antibodies and allowing binding to a target. They consist in either a heavy (VH) or light (VL) chain variable domain, the binding interface being generated by three CDRs as opposed to six for a conventional whole antibody (Table 1). They are very similar to nanobodies (see below). Domain antibodies with specific binding to human albumin are dubbed 'AlbudAbs'.

Dual affinity re-targeting (DART): bispecific, monovalent scaffold, based on two single chain variable domain antibody fragments (scFv), the two polypeptide chains being associated by a disulfide bridge and sequences derived from antibody heavy–light chain pairing (Table 1 and Figure 2). Of interest, the VH and VL of one given variable domain are located on each side of the two different polypeptide chains (Figure 2).

Lipocalins: a family of proteins that transport small hydrophobic molecules such as steroids, bilins, retinoids, and lipids. They share limited sequence homology, but a common tertiary structure architecture based on eight antiparallel β-barrels. These proteins are widely distributed over living organisms like bacteria, vertebrate, and invertebrate cells, and in plants.

Nanobody: variable domain derived from a particular whole antibody containing only two heavy chains without the classically-associated light chains, and found in Camelids. This variable domain is called VHH and is commercially developed under the term 'nanobody'. Nanobodies carry all structural characteristics of heavy chain variable domains (VH): they contain three CDRs and a few amino acid changes to accommodate the absence of the light chain (Table 1). They are very similar to domain antibodies (see above).

Scaffold protein: as used in protein engineering, a scaffold protein refers to a polypeptide sequence typically of small size (< 200 AA) and unique, which consists of a highly structured core associated with variable domains that support modifications such as insertions, deletions, or other substitutions. These variable domains can create novel binding interfaces toward any targeted protein. The structure of protein scaffolds can be highly diverse, but usually of human origin for those developed as therapeutics.

Tetravalent tANDem AntiBody (TandAb): bispecific and bivalent scaffold, based on two scFv fragments, the two polypeptide chains being associated by non-covalent interactions in a head–tail orientation (Table 1 and Figure 2).

developed by Pieris (Munich, Germany). Anticalins are efficiently produced in *Escherichia coli* and yeast and are soluble and stable proteins [8]. Several anticalins have been selected and produced against medically relevant targets including cytotoxic T-lymphocyte antigen 4 (CTLA-4; lead compound: PRS-010 [9]), VEGF-A (lead compound: PRS-050), and the c-Met oncogene (lead compound: PRS-110) with affinities in the nanomolar to picomolar range [8]. PRS-050, Pieris' most advanced program, entered into clinical Phase I in June 2010 (Table 2) [10]. Here, the lead anticalin was linked to a 40-kDa polyethyleneglycol (PEG) moiety to enhance its plasmatic half-life [10]. The molecule binds to VEGF-A with nanomolar affinity [10] and shows powerful antagonistic activities in relevant *in vitro* and animal models, such as a rabbit model of age-related macular degeneration (AMD) and mouse xenograft models of HCT116 and U87-MG tumor cells [10,11]. Toxicology studies in rat and cynomolgus monkeys suggest that PRS-050 is well tolerated [11]. The clinical Phase I included 26 patients with progressive solid tumors, treated with doses from 0.1 to 10 mg/kg by infusion. The molecule showed a dose-dependent increase in exposure and a serum half-life of about 6 days [11]. No formal maximal tolerated dose (MTD) was reached; nevertheless, fever, chills, and hypertension were observed; the Phase II dose recommendation was 6 mg/kg [11]. The best observed clinical efficacy was stable disease [12] in nine patients with the longest duration of stabilization of 8.5 months in a melanoma patient. Biological activity associated with PRS-50 administration was observed, such as changes in circulating matrix metalloproteinase 2 (MMP2) and complex formation between PRS-050 and VEGF-A, associated with a decrease in free VEGF-A [11]. Several other anticalin projects are at the end of discovery phase [5]. One program, PRS-060, targets the interleukin-4 (IL-4) receptor α with an original route of administration by inhalation for the treatment of asthma [5].

Designed Ankyrin Repeat ProteINs (DARPins)

Ankyrin repeat (AR) proteins correspond to an abundant class of protein–protein interaction regulators in nature [13]. Designed AR proteins or DARPins are fully engineered scaffolds designed to resemble natural, human AR proteins; these are now being commercially developed by Molecular Partners (Schlieren, Switzerland). Strategies for the rational design of DARPins have been developed based on multiple sequence alignments, and statistical analysis to calculate the probability of AA usage at each position of an AR and combinatorial AR libraries have been constructed based on the 33 AA AR motif with seven randomized positions (Figure 1b and Table 1) [13]. DARPins contain typically two to four of these motifs flanked by N- and C-terminal capping motifs to shield hydrophobic regions and allow increased solubility [13]. DARPin libraries are preferentially screened using ribosome display (Box 1). Library members are well produced in *E. coli*, do not aggregate, and display high thermodynamic stability [14]. The most advanced DARPin program (MP0112; Table 2) corresponds to a highly potent VEGF-A inhibitor ($IC_{50} < 10$ pM). Two parallel open-label, non-controlled Phase I trials, together including 50 wet AMD or diabetic macular edema (DME) patients, have shown that MP0112 is safe and well tolerated when given as a single intravitreal injection of 0.04–2.0 mg [15]. The most frequent adverse event was transient, sterile inflammation [15,16]. Best corrected visual acuity (BVCA) and optical coherence tomography