Molecular Design of Peptide-Fc fusion Drugs

Lin Ning\(^a\), Bifang He\(^a\), Peng Zhou\(^a\), Ratmir Derda\(^b\), Jian Huang\(^a\)

\(^a\)Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China; \(^b\)Department of Chemistry, University of Alberta, Alberta, Canada

Abstract: Peptide-Fc fusion drugs are a category of biological therapeutics in which the Fc region of antibody is fused genetically to a peptide of interest. In this review, we summarized the key steps of peptide-Fc fusion technology and stressed the main computational resources, tools, and methods that were helpful for the rational design of peptide-Fc fusion drugs. In addition, open questions about the computer-aided molecular design of peptide-Fc fusions were raised. Answers to these questions by future research will certainly help to make the transition from peptide leads to drugs on the market more quickly and more cheaply.

Keywords: Molecular design, peptibody, mimetibody, phage display, biopanning, peptide, peptide-Fc fusion, immunoinformatics.

1. INTRODUCTION

Peptides are promising in therapeutic use. On one hand, they usually have advantages over small chemicals in terms of specificity and affinity for drug targets. On the other hand, they are smaller than antibodies, and much easier and less costly to be manufactured. At present, there are at least more than 60 peptide drugs approved all over the world and over 150 peptides in clinical trials [1]. However, to identify peptide leads for a given drug target used to be difficult; and to develop peptide leads into peptide drugs still faces great challenges. The chief obstacle affecting druggability of peptide leads comes from their small size and sensitivity to serum and tissue proteases and peptidases, which makes them notorious for a fast renal clearance and a short serum half-life [2]. Solutions to this metabolism problem of peptide leads include chemical modifications (N-terminal acetylation, C-terminal amidation, and pegylation), the use of non-natural amino acids (e.g. D-amino acids), cyclization, and fusion with scaffold proteins such as albumin and crystalized fragment of antibody [3-7].

Among the above solutions, the peptide-Fc fusion strategy is very successful. The serum half-life of peptide-Fc fusion proteins are significantly prolonged through the FcRn salvage pathway [8, 9]. In addition to pharmacokinetic benefits, the Fc fusion technology can carry functions such as ADCC and CDC, and improve other properties such as isoelectric point, stability and solubility [3-6]. Furthermore, it can also simplify the manufacturing and purification process compared with that of antibody. Hence, the peptide-Fc fusion technology becomes very popular among pharmaceutical industry and academics. The peptide-Fc fusion was called peptibody by Amgen [10], and mimetibody by Centocor (a subsidiary to Johnson & Johnson) respectively [11, 12].

The term peptibody is a fusion of peptide and antibody. It was first coined at the beginning of 2006 by a group of chemists from University of Konstanz, Germany [13]. They used the term to describe a 26-amino acid residue peptide derived from a camel heavy chain antibody (VHH). Later, a team from Netherland referred to a peptide-VHH fusion as peptibody [14]. The original meaning of peptibody was therefore similar to nanobody. Nevertheless, the scientists from Amgen borrowed this term to describe the peptide-Fc fusion proteins in the second half of 2006. Since then the term became very popular. In this review hereafter, we will use peptibody instead of peptide-Fc fusion protein. Two peptibodies have been approved by the FDA and EMA so far. One is romiplostim, for the treatment of chronic ITP [15]. Another is dulaglutide, for the treatment of type 2 diabetes [16]. At present, a considerable number of peptibodies are in preclinical studies and clinical trials [17-26].

Although the peptibody technology achieved success in terms of products approved, various indications, good market performances, and plentiful new product pipelines, its achievement is limited when compared with the huge clinical demands. Moreover, the development of peptibody drugs are still time consuming and expensive. This problem can partly be solved with the progress on computational peptidology [27]. In the following sections, we will go through the key procedures for the peptibody design and stress the relevant bioinformatics resources, tools, and methods that can help to design peptibody drug more rationally and efficiently.

2. OBTAINMENT OF PEPTIDE LEADS

The first step to design a peptibody is to get peptide leads that can bind to the drug target of interest. These peptide leads can be obtained from native ligands, biopanning, and computational design or prediction.

2.1. Native ligands

If a native ligand for a drug target is small, it can be used as a peptide lead directly. If it is huge in size, we can split the whole binding domain of the native ligand into a panel of
overlapping peptides and then find the high affinity binders as peptide leads [28]. Let’s take dulaglutide as an example. As shown in Figure 1, this peptibody consists of three parts: bioactive peptide, linker, and the Fc fragment of IgG1. The active peptide was actually an analog optimized from human glucagon-like peptide 1, which is the native ligand for GLP-1 receptor. Compared with the peptide lead, i.e. the native human GLP-1, the analog part of dulaglutide has three site-directed mutations (A8G, G22E, and R36G), which can improve solubility and increase protection against proteases [4]. Not only small peptide ligands, but also large domains of native protein receptors can be selected to make another type of Fc fusion drugs such as aflibercept, conbercept, etanercept, and so on [12, 29]. Unfortunately, the natural partners for a given drug target, either as proteins or as small peptides, are limited in number. The huge clinical demands require more alternative peptide leads.

Figure 1 Schematic diagram of dulaglutide

2.2. Biopanning

With the technical advances on combinatorial chemistry and molecular biology, random peptide libraries with high capacity have become available [30, 31]. The screening of these chemical or biological libraries is termed as biopanning. Among all kinds of biopannings, phage display technology is most widely used. The complexity of a phage-displayed random peptide library is usually about 10^14, providing plentiful unique peptides for selection. It is convenient to obtain a set of peptides binding to a drug target [32-35]. These peptides or the common sequence derived from them can be used as peptide leads for the design of corresponding peptibodies [33, 36].

Romiplostim is another successful example [32]. As shown in Figure 2, it consists of the Fc fragment of IgG1, a 5-glycine linker, a 14-residue bioactive peptide, another linker with 8 glycines, and again another copy of the same bioactive peptide with the sequence IEGPTLRQWLARA. In fact, this bioactive peptide was the common sequence derived from several sets of biopanning results, where a pool of phage-displayed peptide libraries was screened against thrombopoietin receptor (TPOR). Its sequence was not found in the primary sequence of thrombopoietin, the natural ligand of TPOR. However, the peptide and TPO bound to TPOR in a competitive manner. The peptide showed a very high affinity to TPOR and stimulated the proliferation and maturation of megakaryocytes efficiently [32]. Hence it was used as the bioactive peptide part to make the peptibody drug candidates.

Figure 2 Schematic diagram of romiplostim

Phage display is a powerful experimental technology that can obtain peptide leads for any drug target efficiently and conveniently. However, it suffers a lot from all kinds of false positive peptides, which do not bind to or only bind to the target non-specifically. These peptides are so-called target-unrelated peptides (TUP) [37-40]. With recent advances on bioinformatics studies in this field, there are currently a series of computational resources that can help to deal with this problem [41]. Firstly, the BDB database which is short for biopanning data bank provides not only an information portal to panning data, but also an evidence-based platform for verifying the specificity of experimental results [42-44]. Secondly, a suite of online tools called SAROTUP is built, which can scan experiment results and report possible TUPS [45-48]. In this web services suite, the TUPScan tool is based on knowledge of known TUP motifs and sequences [45]. PhD7Faster, SABinder, and PSBinder are based on machine learning and data mining of the BDB database. They can predict peptides with growth advantage in the Ph.D.-7 library [46], peptides bound to streptavidin [47], and peptides bound to polystyrene respectively [48]. All the resources discussed above can certainly help to identify peptide leads more reliably through biopanning. With the rapid accumulation of biopanning data, we coined the term in silico phage display [49], where a virtual peptide library is panned against a virtual target and the yielded top peptides were then synthesized and validated by wet lab. Compared to the traditional strategy with experiment only, in silico phage display will be a paradigm shift in this filed. It is expected that we can obtain reliable peptide leads with less time and less cost in future.

2.3. Rational Design

With the booming of bioinformatics, there are now many methods and tools that can help to design or predict peptide leads for a given drug target. These approaches can be grouped into two major categories: structure-based peptide design and structure-free peptide design [50].
Design of Peptide-Fc fusion

Computational methods such as molecular modeling, 3D-QSAR, molecular docking, and molecular dynamics analysis have been widely used in structural bioinformatics for general purposes [51-54]. These methods are also proper for structure-based peptide design [55-57]. However, these methods are usually time-consuming and computation-extensive. With the accumulation of the complex structures of protein-protein interaction, London et al found that more than half of such interactions were dominated by one hot linear binding segment at the interface and built a protocol to derive inhibitory peptides based on the complex structures [58]. This group later automated this protocol and provided a simple web server called Peptiderive [59], which made the structure-based peptide design easier and faster. Recently, a case report compared two peptibodies experimentally [60]. One peptibody was 814-Fc. Another was TA-Fc. The BLyS binding peptide 814 was obtained from phage display library and the peptide TA was derived from computer-aided design on BLyS-TACI interaction. The results showed that both peptibodies had bioactivities. However, 814-Fc showed two-fold inhibition effect on BLyS binding to TACI compared with that of TA-Fc [60]. This case report indicates that the rational peptide design based on 3D structures needs to be improved further. Nevertheless, bioinformatics studies on structures are still valuable. For example, the structural bioinformatics study on peptide obtained from phage display and the following peptibody design can certainly provide new insights [33, 36]. However, we still lack for an automatic and Peptiderive-like protocol, pipeline and server for this issue.

Though more and more 3D structures of proteins have been solved, there are still many proteins that lack structural information. For instance, most membrane proteins lack 3D structures due to their difficulties in crystallization, though they are usually very important drug targets. In this situation, structure-free peptide design is a good alternative solution. With the rapid increasing of available bioactive peptide database and the advance of artificial intelligence, predicting peptide leads based on sequence information has become popular [61-74]. However, there is a big problem for the current structure-free peptide design methods. As the bioactive peptide data obtained from a given drug target is not sufficient enough at present, most available predictors were trained by pooled data with similar function. Therefore, most current methods can predict peptide leads with special bioactivity. Their drug targets however are not known. We expect that the bioactive peptide data will become bigger and more elaborate soon and new computational tools for specific and important drug targets will be available in near future.

3. SELECTION OF FC

The second step to design a peptibody is to select the proper Fc region from different classes or subclasses of immunoglobulin or engineered Fc mutants. This selection is mainly up to the development purpose.

Theoretically, Fc from the IgA molecule should be selected if the peptibody is designed to function at mucosa and Fc from the IgE molecule will be more appropriate if the peptibody is designed to inhibit allergic responses [75]. As the main goal of peptibody is usually to prolong the serum half-life of the bioactive peptide, Fc from the IgG molecule is the best choice, which can target the fusion protein to the FcRn salvage pathway and recycle the peptibody [9]. There are four subclasses of human IgG, which named IgG1, IgG2, IgG3, and IgG4 respectively. They are numbered according to their serum abundances. In addition, the four subclasses are also different in terms of affinity to FcRn, half-life, hinge length, and various biological roles in host responses [76]. In practice, IgG3 has never been chosen for the development of monoclonal antibody or peptibody drugs due to several considerations [76]. It is widely accepted that Fc region of IgG1 is suitable for the design of peptibody against cancer and those of IgG2 or IgG4 are fit for the treatment of a metabolic dysfunction [4]. This is because the Fc region of IgG1 displays high ADCC and CDC which can improve the cytotoxic activity of the peptibody compared with those of IgG2 and IgG4. The convention seems true as the Fc region of the approved peptibody drug dulafludotide belongs to IgG4 subclass and it is used to treat type 2 diabetes. However, the Fc region of another approved peptibody drug romiplostim is IgG1 while its function is to stimulate the production of platelets rather than killing the cell. In a very recent report, the Fc region of IgG2a was selected to link a peptide that bound many if not all forms of amyloid [26]. Including peptibodies in preclinical and clinic studies, the Fc region of IgG1 is the most popular choice.

There are now many known mutants of Fc region as Fc engineering has been studied for many years [77]. These mutants show modulated features in both serum half-life and relevant effector functions such as ADCC and CDC. For example, Fc mutant (N434A) showed a twofold extension of serum half-life, a double mutant (T250R, M428L) with a 2.8-fold prolonged serum half-life, and a triple mutant (M252Y, S254T, T256E) with a 4-fold improvement in serum half-life [77]. These mutants might be better choices for the design of peptibody. In the Fc engineering field, there are a lot of experimental data dispersed in the published papers. As far as we know, there is no special database in this field, which makes it difficult to carry out further bioinformatics studies on Fc engineering. We believe the computational methods can help to find better Fc which will then benefit the design of both peptibody and antibody.

4. WAYS OF FUSION

The third step to design a peptibody is to fuse the peptide leads and Fc together genetically. Where to put the peptide leads? How many peptide leads to use? What kind of linker to choose? As shown in previous figures, linkers usually consist of a stretch of glycines, which can eliminate steric hindrance and improve the flexibility and accessibility of bioactive peptide. There are at least three places to put the peptide leads [5]. One is at the N-terminal of Fc, like dulafludotide shown in Figure 1. The second place is at the C-terminal of Fc, like romiplostim shown in Figure 2. The third place is the loop region of CH2 or CH3 domain of Fc region. Such kind of fusion is call FcLoop peptibody [5]. Compared with those terminal peptibodies, FcLoop peptibody showed
significant improvement of pharmacokinetics [5]. However, the bioactive peptides in terminal peptibodies are more accessible and flexible. For a certain peptide lead, different ways of fusion might significantly affect its bioactivity. In an extreme case, a NGF blocking peptide fused to C-terminal of Fc displayed a NGF neutralizing activity 1500 times higher than that of N-terminal version and the reasons were still unknown [5].

The number of peptide leads used in the peptibody will also affect the bioactivity of the whole fusion protein. As shown in previous figures, each chain of dulaglutide has only one copy of bioactive peptide, where in romiplostim, two copies spaced by a linker. AMG386 is an anti-angiogenic peptibody that neutralizes angiopoietins 1 and 2. It has two bioactive peptides per Fc strand. The experimental data showed that the bioactivity of AMG386 is ten times greater than that of the equivalent fusion with a single peptide per Fc strand [5]. Moreover, the peptide leads number per Fc strand also has impact on its serum half-life [9]. Molecular weight is one of the main factors that affect the renal clearance of biological drugs. If a protein is larger than 60 KD, the renal filtration will contribute little in its clearance. The ideal size of peptibody is similar to that of albumin (69 KD). Using more copies of bioactive peptides will help to improve its pharmacokinetics. However, too much repeats may raise risk of immunogenicity issues.

5. EVALUATION OF IMMUNOGENICITY

The fourth step to design a peptibody is to evaluate its immunogenicity based on the sequence of the peptibody. An essential concern in developing biological drugs is their immunogenicity [78]. The peptibody design is no exception. The adverse immune reaction including neutralizing antidual antibodies can not only affect the efficacy of the drug but also its safety. Let’s take peginesatide as an example. It is a pegylated bioactive peptide obtained from panning against EPOR. The drug was approved by FDA for the treatment of anemia in patients with chronic kidney disease. However, it was soon withdrawn due to serious adverse drug reactions [79]. The good news for peptibody is that the Fc region of human IgG can modulate the immunogenicity of the peptide leads [75]. Even so, the design of peptibody should follow the FDA guidance on immunogenicity. If not, the designed peptibody may have a higher risk to fail. AMG 819 is a NGF-neutralizing peptibody that can release pain. Its peptide lead was 20-amino acids long and obtained from biopanning against human NGF. It showed promising results in animal experiments. However, its phase I clinic trial was terminated due to the immunogenicity issue. After a single injection of this peptibody, 37% of all subjects developed detectable antibodies [5]. Later, immunoinformatics analysis suggested that a potential promiscuous T-helper epitope for a broad range of HLA alleles located in the peptide leads [80]. Hence, the immunogenic evaluation of primary sequence of fusion proteins at the design stage is especially important, which is also stated in the FDA guidance.

There are a lot of databases and computational tools available for the prediction of B cell epitopes, cytotoxic T cell epitopes, and helper T cell epitopes [81-89]. The designer of the peptibody should choose relevant tools to analyze the whole sequence of peptibody. Focus on the peptide lead if it is obtained from biopanning; Be careful about the joining areas among bioactive peptide, linker, and Fc region where neoantigen might form due to fusion; Pay attention to Fc region if the Fc mutant is used. Because the importance of T cell epitopes and higher reliability in T cell epitope prediction, we suggest that the evaluation of immunogenicity should stress on T cell epitopes, especially those covered major HLA alleles. For biopanning results, we suggest building an algorithm and method to evaluate the degree of humanness of obtained peptide sequences.

CONCLUSION

Peptibody is an important type of biodrugs fused by Fc and bioactive peptides. As computational methods and tools are rapidly being developed and applied in each key steps of the molecular design of peptibody, we can expect more and better peptibody drugs will be approved in future.

LIST OF ABBREVIATIONS

- ADCC: antibody-dependent cell-mediated cytotoxicity
- BLYS: B lymphocyte stimulator
- CDC: complement-dependent cytotoxicity
- EMA: European medicines agency
- EPOR: erythropoietin receptor
- Fc: fragment crystallizable of antibody
- FcRn: neonatal Fc receptor
- FDA: food and drug administration
- GLP-1: glucagon-like peptide 1
- HLA: human leukocyte antigen
- ITP: immune thrombocytopenic purpura
- NGF: nerve growth factor
- TACI: transmembrane activator and CAML interactor
- TPO: thrombopoietin
- TPOR: thrombopoietin receptor
- TUP: target-unrelated peptides
- VH: variable domain of heavy chain antibody

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.
ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China [61571095 and 31671361] and the Fundamental Research Funds for the Central Universities of China [ZYGX2005Z006 and ZYGX20161186].

REFERENCES

Design of Peptide-Fc fusion


Received: November 18, 2017 Revised: January 18, 2018 Accepted: January 22, 2018