

## 1 Introduction

Peptides and proteins play a crucial role in all biological systems and are involved in a wide variety of biochemical processes such as catalysis, signal transduction, cell-cell communication, gene regulation, and the control of physiological functions like metabolism, the hormone system, and immune response.<sup>1</sup> Furthermore, these biopolymers represent important structural components in cells and organisms, e.g., in filamentous and fibrous molecules. Consequently, many biological disorders and human diseases are either caused by defective proteins or mediated by protein-protein interactions during viral or bacterial infection. A prominent example is represented by the infection process of human cells with HIV-1, the virus that causes the Acquired Immunodeficiency Syndrome (AIDS). After binding of the viral envelope glycoprotein complex (gp120/gp41) to receptors on the surface of the target cell, the fusion of the viral and cellular membranes is mediated by a series of conformational changes of gp120/gp41, including the formation of a six-helix bundle.<sup>2</sup> The inhibition of such protein-protein interactions by peptide- and protein-based inhibitors appears as a promising pharmaceutical concept for medicinal therapy. The entry of HIV-1 particles into human cells can thus be prevented by the application of peptides that mimic components of the six-helix bundle and, therefore, inhibit the refolding process of the envelope protein.<sup>3,4,5</sup> The manifold functionalities of peptides and proteins are based on the physical and chemical properties of specific amino acid side chains and on their three-dimensional structure, the folding.<sup>6</sup> The complete information about this higher ordered conformation is programmed in the primary sequence of the polymer, which means the alignment of the amino acid residues.<sup>7,8</sup> Therefore, the key to controlling the biological functionality and therapeutical efficacy of peptide-based drugs is provided by the systematic amino acid composition and the directed incorporation of building blocks that introduce the desired physical, chemical, and structural properties. Thereby, any knowledge about the structure of the target protein supports the rational design of potential binders to specific regions of the molecule.<sup>9</sup> However, the development of peptide-based drugs based on an alignment of the twenty canonical amino acids is limited by profound factors that almost zero out the pharmacological potency of an amino acid polymer.<sup>10</sup> One major reason is the very low bioavailability that is mainly caused by a fast metabolic degradation *in vivo* by proteases, a slow passive transport of these polar molecules in blood, and a negligible membrane penetration level.<sup>11,12</sup> A further drawback of the applicability of native peptides as drugs is the high flexibility of their backbones. Since the stability of specific secondary folds is determined by tertiary interactions between residues that are distant in sequence, and bioactive peptides are usually too small to adopt any tertiary structure, these molecules are mostly unstructured

before binding to their target protein.<sup>13</sup> This drastically worsens the peptides pharmacokinetical properties. Considering these issues and the fact that intra- and intermolecular protein-protein interactions that can be found in nature are optimized by evolutionary selection processes,<sup>14</sup> their inhibition with peptides that are exclusively made of the same repertoire of building blocks can hardly be efficient. Consequently, the application of non-natural building blocks for peptide design and protein engineering has been of growing interest and focused on in the last two decades.<sup>15</sup> These artificial amino acid surrogates can be designed to increase metabolic stability as well as to induce and/or stabilize folding elements. Furthermore, they can introduce new functional groups, such as heteroatoms that are not found in the natural amino acid pool, and, therefore, enable new specific peptide-protein interactions.<sup>16</sup> Newly developed techniques such as advanced coupling strategies in solid phase peptide synthesis,<sup>17</sup> convergent peptide synthesis,<sup>18</sup> enzymatic methods,<sup>19</sup> the methods of segment condensation,<sup>20</sup> native chemical peptide ligation,<sup>21,22,23,24</sup> expressed protein ligation,<sup>25,26</sup> and biosynthetic insertion via an aminoacylated suppressor-tRNA<sup>27,28</sup> extend the application of non-proteinogenic building blocks to peptides and proteins of higher molecular mass. Besides the incorporation of C<sup>α,α</sup>-dialkylated amino acids, prevention from proteolytic degradation can be very efficiently achieved by the replacement of native amide bonds with surrogates<sup>29,30</sup> and/or the reversal of the peptide bond (NH-CO instead of CO-NH) giving rise to the so-called retro- and partially modified retro (PMR) peptides.<sup>31</sup> In retro-inverso analogues, the stereochemistry of one or more amino acids of the reversed segment is inverted.<sup>32</sup> However, this peptidomimetic concept demands advanced synthetical methods and cannot be realized by applying standard peptide synthesis procedures. A synthetically more convenient approach to increase metabolic stability of peptides is represented by the incorporation of N-alkylated amino acid analogues and the design of peptoids.<sup>33</sup> This sort of peptidomimetics is composed of derivatives of amino acid residues that have the native side chain shifted from the chiral α-carbon atom to the achiral backbone nitrogen (N-alkylated glycines).<sup>34</sup> The resultant loss in chirality within the peptoids will consequently have a major impact on the folding properties and, therefore, on the biological activity of the altered peptide. Different sorts of peptide and protein building blocks can initiate and stabilize the folding of specific secondary structure elements. Aminoisobutyric acid (Aib), the simplest C<sup>α,α</sup>-dialkylated amino acid, is a strong inducer of helical structures<sup>35,36</sup> even in peptide denaturing solvents.<sup>37</sup> An analogous behavior was described for several Aib derivatives with larger side chains<sup>38,39</sup> and C<sup>α<sub>i</sub></sup> ↔ C<sup>α<sub>i</sub></sup> cyclizations.<sup>40</sup> As a consequence, C<sup>α</sup>-methylated analogues of natural amino acids can be used in bioactive peptides to increase structural and metabolic stability without losing the functionality of native side chains. An analogous approach to the induction and stabilization of turn structures is represented by the application of amino acid proline chimera

that possess both the conformational rigidity of proline and the desired side chain.<sup>41,42</sup> While lots of strategies have been invented to date in order to strengthen structural properties and stability of peptides, the introduction of new non-natural amino acid side chain interactions to the design of peptide-based inhibitors requires the incorporation of functional groups that are not found in native peptide-protein and protein-protein interactions. Such functionalities can be further very successfully applied in probing the structure and function of proteins. In this context, amino acids that contain fluorine have become very prominent.

Fluorine in fluorocarbon compounds, the so-called “organic fluorine” is beyond doubt an outstanding element in organic chemistry. Besides the synthetic potential of fluorine and of reagents derived from it, numerous fluorinated compounds found in materials science to medicinal chemistry, have great impact on our daily life. In drug development, the incorporation of fluorine has become a very prominent strategy to overcome serious problems of newly designed active compounds such as an insufficient bioavailability caused by fast *in vivo* metabolism and poor transportation or diffusion through the body.<sup>43</sup> These features let fluorine emerge as the “magic element” in medicinal chemistry in respect to small organic molecules and make it extremely valuable for the improvement of drugs that are based on peptide structure. Furthermore, this heteroatom introduces a new functionality into the design of artificial peptide-protein interactions. It has been shown that even one fluorocarbon amino acid, which was incorporated instead of its hydrocarbon analogue, can have a strong impact on protein structure.<sup>44</sup> Although recent studies of different research groups demonstrated that fluoroalkyl substituted amino acids can even tremendously increase peptide and protein stability by forming very strong non-natural side chain interaction cores, the applicability of this sort of non-proteinogenic building blocks for rational peptide design and protein engineering remains strongly limited to date. The main reason for this limitation is that the nature of the highly interesting impact that fluorine has on amino acid side chain interactions in proteins as well as the basic physical and chemical properties of fluoroalkyl groups are still far from understood. Several aspects like the role of fluorine in hydrogen bonding, the space filling of fluoroalkyl groups, and their interactions with hydrophobic domains are controversially discussed issues that remain to be studied systematically. Therefore, the general consequences that the site directed incorporation of a fluorinated amino acid has on the structure and function of a specific protein or peptide is unpredictable. Undoubtedly, a systematic evaluation of the interaction properties of fluoroalkyl substituted amino acids in native protein environments will provide new potent tools for the rational design of peptide and protein interaction sites.