Combined Use of Oligopeptides, Fragment Libraries, and Natural Compounds: A Comprehensive Approach To Sample the Druggability of Vascular Endothelial Growth Factor

Núria Bayó-Puxan, Ricard Rodríguez-Mias, Michael Goldflam, Martin Kotev, Sonia Ciudad, Christopher J. Hipolito, Monica Varese, Hiroaki Suga, Ramón Campos-Olivas, Xavier Barril, Víctor Guallar, Meritxell Teixidó, Jesús García, and Ernest Giralt

The modulation of protein–protein interactions (PPIs) is emerging as a highly promising tool to fight diseases. However, whereas an increasing number of compounds are able to disrupt peptide-mediated PPIs efficiently, the inhibition of domain-domain PPIs appears to be much more challenging. Herein, we report our results related to the interaction between vascular endothelial growth factor (VEGF) and its receptor (VEGFR). The VEGF–VEGFR interaction is a typical domain-domain PPI that is highly relevant for the treatment of cancer and some retinopathies. Our final goal was to identify ligands able to bind VEGF at the region used by the growth factor to interact with its receptor. We undertook an extensive study, combining a variety of experimental approaches, including NMR-spectroscopy-based screening of small organic fragments, peptide libraries, and medicinal plant extracts. The key feature of the successful ligands that emerged from this study was their capacity to expose hydrophobic functional groups able to interact with the hydrophobic hot spots at the interacting VEGF surface patch.

Introduction

Protein–protein interactions

Proper cellular behavior is governed by a large map of interactions between various protein partners, the so-called interactome. Research into protein–protein interactions (PPIs) has emerged as an important and challenging field in chemical biology and medicinal chemistry. New and classical screening approaches in combination with medicinal chemistry have brought more than 12 small-molecule PPI modulators into clinical development, and it is expected that the market for this class of compounds will reach 600 million EUR in the next five years. However, the design of molecules that interfere with these interactions is hindered by the size and topology of protein–protein interfaces. The surface of many proteins has evolved to interact either with other proteins or with small molecules. The contact area in a protein–protein interface is considerably larger (1500–3000 Å²) than the average contact surface in the interaction between proteins and small molecules (300–1000 Å²). In addition, protein–protein interfaces tend to be flat and often lack the grooves or cavities normally present in protein surfaces involved in the binding of small molecules. These considerations led to the concept of protein “druggability” coined by Hopkins and Groom. Despite the large size of protein–protein interfaces, early studies revealed that not all interfacial residues contribute equally to complex stability. A small subset of critical residues, called hot spots, are responsible for a large fraction of the binding free energy. It has also been pointed out that large PPIs are often dominat-
VEGF–VEGFR: a therapeutically relevant protein–protein interaction system

Vascular endothelial growth factor (VEGF or VEGF-A) is probably the most important soluble factor involved in tumor angiogenesis.[10, 16] VEGF belongs to the PDFG superfamly, which in mammals also includes VEGF-B, -C, -D, and -E, with which the former shares a varying degree of homology. VEGF is a glycoprotein expressed in various spliced isoforms ranging from 121 to 206 residues in length, and its 165-amino acid isoform is the most abundant. All these isoforms share a common homodimeric N-terminal receptor binding domain that has the ability to bind three tyrosine kinase receptors, namely, VEGFR-1 (also known as Flt-1), VEGFR-2, and VEGFR-3.[11] VEGF monomers are arranged in an antiparallel manner and are covalently linked by two symmetrical intermonomer disulfide bridges between Cys51 and Cys60. The monomer structure is characterized by a four-stranded central β sheet (i.e., strands β1, β3, β5, and β6) that is stabilized by a cystine knot motif formed by three intramolecular disulfide bonds (i.e., Cys26–Cys68, Cys57–Cys102, and Cys61–Cys104) at one end and by a small hydrophobic core at the opposite end. The structure also contains three additional short β strands (i.e., β2, β4, and β7) and two α-helical segments (i.e., α1 and α2).[12] Two symmetric receptor binding surfaces at the poles of the VEGF homodimer have been identified by mutational and structural studies.[13] The crystal structure of the VEGF–VEGFR-1 complex[14] revealed that no major conformational changes occur on VEGF upon receptor binding and that the intermolecular interaction is mediated mainly by hydrophobic contacts. Both X-ray[13] and NMR spectroscopy studies[15] suggest that the receptor binding region of VEGF has intrinsic conformational flexibility. Such flexibility may have important functional implications that allow the protein to interact with multiple receptors.

VEGF is overexpressed in most tumors and has a direct effect on sprouting angiogenesis.[16] It also acts as a survival factor by enhancing the expression of anti-apoptotic factors such as Bcl-2.[17] In general, VEGF induces receptor dimerization and activation of its kinase activity through autophosphorylation, which triggers a series of intracellular signaling pathways; it also sets off several processes common to other growth factors, namely, cell migration, survival, and proliferation.[18]

VEGF is particularly interesting as a therapeutic target because it acts directly on genetically stable endothelial cells rather than on tumor cells. Therefore, drugs that target VEGF are less prone to induce mutation selection on tumor cells and ultimately confer them with drug resistance.[19] VEGF inhibition suppresses tumor growth in animal models.[20] VEGF has also been identified as a key target to mitigate the effects of both wet age-related macular degeneration (AMD) and diabetes macular edema (DME), which are vision-limiting complications of aging and diabetes, respectively. DME affects nearly 30% of diabetes patients, whereas AMD is the main cause of permanent vision loss in the elderly.[21] High concentrations of VEGF have been correlated with chronic retinal microvascular damage. To date, 11 drugs that disrupt VEGF signaling have been approved for the treatment of several cancer types[22] and diseases related to macular degeneration.[23] Of these, seven (i.e., sunitinib,[24, 25] sorafenib,[26, 27] pazopanib,[28, 29] vandetanib,[30] axitinib,[31] regorafenib,[32] and cabozantinib[33]) are small molecules that inhibit the intracellular tyrosine kinase activity of the receptor. Of the remaining four drugs, two are monoclonal antibodies targeting either circulating VEGF (i.e., bevacizumab[34]) or the soluble extracellular fraction of the VEGF receptor (i.e., ramucirumab[35, 36]). In addition, a chimeric fusion protein (i.e., aflibercept[37]) and an aptamer (i.e., pegaptanib[38]) have been described to block VEGF.

Recently, a small molecule interacting with the VEGFR-1 D2 domain and preventing VEGF binding was reported.[39] In the present study, we explored the capacity of small molecules to directly target the VEGF protein. The results of screening peptide and fragment-based libraries and plant extracts are discussed.

Results

We tackled the discovery of VEGF ligands by using several experimental approaches, including fragment screening, design and evaluation of linear and cyclic peptide libraries, and screening of plant extracts used in traditional Chinese medicine. We focused our methodology on the use of NMR spectroscopy, including the implementation of a highly efficient chemical shift perturbation (CSP) method based on methyl-13C-Methionine selective labeling of VEGF. Computational chemistry was used to assist the screening process and was used to facilitate interpretation of the results.

Selective methyl-13C-methionine-labeled VEGF for binding studies

VEGF11–109 (from now on VEGF), a construct comprising residues 11–109 and containing the structured homodimeric N-terminal receptor binding domain, was chosen because its expression procedure and its backbone assignment have already been described.[13] By using 1H–15N HSQC NMR spectroscopy experiments, a uniformly 15N-labeled VEGF sample was used to confirm both its fold integrity and its capacity to interact with v107, a peptide that binds with high affinity to the same region of VEGF that interacts with the VEGF receptor.[40] In the presence of v107, considerable changes in more than 50% of the VEGF NH resonances were observed (results not shown). Receptor-based 1H–15N HSQC experiments provide information about the most probable localization of ligand–protein interactions, which can be correlated to their biological functional effects. However, given the dramatic effects of v107 binding in
the VEGF $^1$H-$^{15}$N HSQC spectra, the interpretation of CSPs may be misleading. Moreover, this method may be expensive and time consuming for a high-throughput screening strategy. Alternatively, in our screen of potential VEGF binders, we routinely used one-dimensional ligand-observed techniques such as saturation transfer difference (STD)\(^{(41)}\) and WaterLOGSY experiments\(^{(42)}\). T\(_2\)-relaxation edited experiments\(^{(43)}\) and \(^{19}\)F NMR spectroscopy\(^{(44)}\). However, these techniques also have some limitations. In particular, for complex mixtures, such as plant extracts, deconvolution of the contribution of each individual component may be challenging. In addition, NMR spectroscopy methods based on the existence of a fast equilibrium between the bound and free ligands may not be able to detect strong binders. Therefore, we focused our efforts on designing a cost-effective and time-efficient screening method that preserved the relevant mapping information. One way to achieve this goal is the NMR observation of specifically labeled side chain resonances of the target protein. This approach has the potential to improve both the sensitivity and spectral resolution of the NMR spectra. Moreover, as side chains are less prone to long-range CSP effects, they may be more reliable than amide signals for the extraction of structural information.

The mode of interaction of VEGF with its receptors\(^{(14, 45)}\) and several peptides\(^{(40, 46, 47)}\) was extensively studied and structurally characterized. Of the five VEGF methionine residues, two (i.e., Met18 and Met81) participate in the interaction with either v107 or VEGFR-1 and become deeply buried in the complex interface. Therefore, incorporation of selectively labeled methionine residues may enable the study of the intermolecular interactions of VEGF. The presence of three methionine residues (i.e., Met55, Met78, and Met94) outside the interface could provide extra information about the specificity of the interaction. The use of methionine as a labeling probe presents several additional advantages. Incorporation of commercially available methyl-$^{13}$C-methionine into proteins can be easily accomplished by using methionine auxotrophic E. coli strains. These strains are widely used in protein crystallography to solve the phase problem by using selenomethionine.\(^{(48)}\) The labeling procedure is cost effective as only 80 mg of methyl-$^{13}$C-methionine (300 \(\mu\)g) per liter of growth media is required. Most importantly, $^{13}$C-labeled methyl groups show high sensitivity and favorable NMR relaxation properties.\(^{(49)}\) The combined use of methionine isotopic labeling and NMR spectroscopy detection was previously described for a carboxymethylated enzyme\(^{(50)}\) and for in-cell NMR spectroscopy experiments.\(^{(51)}\)

Methyl-$^{13}$C-Met-labeled VEGF was obtained following the procedure described for uniformly $^{15}$N-labeled VEGF,\(^{(52)}\) except that the methionine auxotrophic E. coli B834 strain was used and the minimal medium for cell growth was supplemented with methyl-$^{13}$C-Methionine. The protein yield was not greatly affected by the use of auxotrophic cells. Selectively methyl-$^{13}$C-Met-labeled VEGF displayed the expected five correlation signals in the $^1$H-$^{13}$C HSQC NMR spectra (Figure 1). The addition of v107 revealed that, as expected, the most affected signals corresponded to methyl groups in Met18 and Met81. The largest change was detected on Met18, probably as a result of considerable contacts with Trp13 of v107 and the effect of its aromatic ring currents. The remaining methionine residues (i.e., Met55, Met78, and Met94) were virtually unaffected by v107.

### Evaluation of linear and cyclic oligopeptides

Oligopeptides can be considered privileged structures to modulate PPIs given that their chemical structure is identical to that involved in PPI interfaces. In addition, they are modular compounds, they cover a large chemical space, and they can be easily synthesized.

We designed two groups of all-$\alpha$ dipeptides. The first one was inspired in the retro-enantioversion\(^{(53)}\) of peptide v107, which was dissected into a series of 20 dipeptides (Table 1). The second group of 6 peptides was composed by combining a representative set of amino acids, including Ala, Ile, Trp, Glu, Arg, and Ser (Table 1).

All-$\alpha$ dipeptides were evaluated against VEGF by STD and WaterLOGSY experiments. However, after extensive exploration of a series of experimental parameters (e.g., protein and ligand concentration, number of scans, saturation time, and saturation frequency), none of the all-$\alpha$ dipeptides seemed to bind VEGF.
Table 1. All-α dipeptide library.

| Dipeptide based on the retroenantiomeric version of v107 Ac, Ile, Ser, Glu, Arg, and Trp-containing dipeptides |
|---------------------------------------------------------------|
| N term. | Dipeptideα | N term. | Dipeptideα |
| o-Leu-o-Arg | 2 mω-Pro-Trp-o-Met | o-Ala-o-Xaa |
| o-Arg-o-Glu | o-Arg-o-Ala | o-Ile-o-Xaa |
| o-Glu-o-Phe | o-Glu-o-Leu | H- |
| o-Glu-o-Trp | o-Ala-o-Ile | Ac- |
| o-Trp-o-Glu | o-Ile-o-Asp | o-Arg-o-Xaa |
| - | o-Trp-o-Xaa | - |

(a) Amide as C terminus. (b) Xaa = Ala, Ile, Ser, Glu, Arg, and Trp.

Under the hypothesis that the intrinsic flexibility of the all-α dipeptides analyzed may be detrimental for VEGF binding, we designed two symmetric cyclic hexapeptides containing chemically distinct dipeptide motifs, namely, Trp-Arg and Ile-Arg. A β-proline was included to facilitate cyclization and to confer conformational rigidity to the peptide structure. Cyclopeptides &α-Pro-Trp-Glu-β-Pro-Trp-Glu& and &α-Pro-Ile-Arg-β-Pro-Ile-Arg& were obtained by means of 9-fluorenylmethoxy-carbonyl (Fmoc)/Boc peptide chemistry on a solid support (see the Supporting Information). The capacity of the cyclic peptides to interact with VEGF was monitored by a combination of STD and protein-based CSP experiments by using a methyl-13C-Met-labeled VEGF sample. Evidence of binding was only observed for the Trp-Glu-containing cyclic hexapeptide (Figure 2). In the presence of the Trp-Glu-containing cyclic hexapeptide, selective CSPs were induced in the two VEGF methionine residues (i.e., Met78 and Met81) located in the receptor interface, whereas the remaining methionine residues were almost unaffected. Titration experiments indicated that the Trp-Glu-containing cyclic hexapeptide binds VEGF in the low-millimolar range.

These promising results encouraged us to evaluate a massive library of cyclic peptides to explore a broader range of sequences and cycle sizes. For this purpose, we used a recently developed in vitro selection technology referred to as RaPID (random nonstandard peptide integrated discovery) system.[54] The RaPID system combines an mRNA display method with flexible in vitro translation (FIT) technology used for the production of nonstandard peptides.[55, 56] With this methodology, specific sequences of nonstandard peptides can be ribosomally expressed according to their mRNA templates.[57] The FIT system has been applied to the synthesis of a wide variety of Xaa-tRNAs, for which Xaa represents proteinogenic and nonproteinogenic amino acids.[58–62] (see the Supporting Information). The selection on the mRNA display technology was performed with a mRNA library encoding for peptides from 5 to 15 amino acids with random sequences and showed enrichment of specific motifs (see the Supporting Information). The 10 macrocyclic peptides with the highest score were synthesized. Five of them were prepared as fluorescein-labeled compounds (Table 2). Macrocyclic peptides were synthesized by using Fmoc/8Bu peptide chemistry on a solid support. They were obtained in low yields (<1%), which could be attributed to the high frequency of bulky 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf)-protected arginine and β-branched amino acids and to the presence of multiple conformations of the final macrocycle. The VEGF binding ability of unlabeled macrocyclic peptides (i.e., peptides 1–5, Table 2) was evaluated by fluorescence po-

Table 2. Macrocyclic peptides designed from the mRNA display technology.

| Peptide | Sequenceα | Purity [%] |
|----------|------------|------------|
| 1 | FKPYRILVRRRKPGS | >95 |
| 2 | FRYYLVKLRRTISRCS | >95 |
| 3 | FRYYCVKVTFTVC*GS | >95 |
| 4 | FRRRKRISYYLAC*GS | >95 |
| 5 | wRR(Me)F(CiMe)Me(F)WRRPC*GS | >95 |
| 6 | wRFWRWKYKPYKPC*GS | >95 |
| 7 | wRTTTAVFLRDRGC*GS | >95 |
| 8 | wGTVFLVSKRNPCHG*GS | >95 |
| 9 | wKKTIWRRWTRRKRCAK*GS | >95 |
| 10 | wFRSYRGRGWTRWHC*GS | >95 |

[a] Small letters indicate non-natural amino acids: f = β-Phe; w = β-Trp; a = α-alanine; (Me) stands for N-methylamino acids; bold C indicates the location of the thioether bond to the N-terminal α-amino group. C* indicates the use of S-acetaminomethyl (Acim) side-chain protection.
Figure 3. Overview of the strategy used for fragment selection.
0000 0000, 0 – 0 www.chemmedchem.org 6 © 0000 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Figure 4. a) Analysis of the SiteMap computations reveals druggability mainly in the channel formed between the two chains of the VEGF homodimer (sites 1, 2, 3, and 4). Red, blue, and yellow surfaces show sites for potential H-bond acceptor, H-bond donor, and hydrophobic interactions, respectively. White dots indicate the presence of small cavities. Receptor binding sites are indicated by S. The number size correlates with the size of the binding sites. b) Representation of the XP GLIDE docking results. A selected library of approximately 500 compounds (gray carbon atoms) was docked against the VEGF homodimer (black ribbons). Sites 1 and 2 (panel a) were most populated.

Natural compounds: flavonoids from medicinal plant extracts as VEGF binders

A considerable number of drugs that are currently used in cancer therapy are natural products or their derivatives.[72,73] In this regard, traditional medicinal plants offer enormous potential in the search for new potential drugs. Cancer is among the conditions treated by traditional herbal medicine. Interestingly, some of the antitumoral plants or compounds used in traditional medicine operate through an antiangiogenic mechanism, modulating one or various targets involved in vessel formation. From the literature on traditional Chinese medicine,[74] we selected a small subset of 50 plants for their high recurrence in medical prescriptions. Dried plant material was extracted with water by means of a Soxhlet apparatus. Given that VEGF only contains five methionine residues, protein-de-protected 13C-filtered-13C-decoupled 1H NMR spectroscopy was used to detect potential VEGF binders. The spectra were obtained with a one-dimensional version of the 1H–13C HSQC pulse sequence, and they provided enough resolution to individually resolve the resonances of the VEGF methionine residues in a time-efficient manner. Identical experiments were acquired for free VEGF and for VEGF after the addition of mixtures of three different plant extracts. Two extract mixtures induced spectral changes in VEGF. Next, individual extracts from the positive mixtures were added to the protein. Small, yet dose-dependent, minor shifts were observed for Met18 and Met81 in the presence of extracts from Medulla Junci and Radix scutellariae (Figure 5). In both cases, the shifts observed were typical of weak binders in a fast exchange regime.

The chromatographic profiles of these extracts are shown in Figure 5. We pursued the isolation and characterization of the active compound(s) involved in VEGF binding. Using a combination of mass spectrometry and homonuclear (COSY and NOESY) and heteronuclear (1H–13C HSQC and HMBC) NMR spectroscopy techniques, the major component of the R. scutel lariae extract was identified as baicalin.[75] (Figure 6). We then confirmed that isolated baicalin was able to bind VEGF (Figures 5 and 6).

STD experiments were performed to identify the binding epitope of baicalin (Figure 6). Relative quantification of transferred saturation indicated that the most buried part of the molecule was the phenyl group, whereas the least saturated and consequently the most exposed part was the more hydrophilic glucuronic moiety. Finally, a STD competition assay with peptide v107 was performed. Increasing amounts of v107 were added to the baicalin–VEGF sample and subsequent STD experiments were recorded. As shown in Figure 6, these experiments revealed a stepwise reduction in the baicalin STD signals, which became undetectable in the presence of 100 μM peptide v107, which thereby indicated that baicalin and v107 compete for the same binding site.

The interaction of baicalin with VEGF was further characterized by 1H–15N HSQC experiments. Upon the addition of baicalin, the most remarkable CSPs were found in a region centered in residues 85–96. In addition, resonances corresponding to amino acids in the vicinity of positions 46–48 and 18–22 were also significantly shifted (Figure 7). Titrations were conducted and baicalin-induced CSPs in the VEGF spectra were quantified and used to determine the binding affinity; Kₐ values of 4.9 and 6.8 mM were obtained for 1H–15N HSQC and 1H–15N HSQC shifts, respectively (see the Supporting Information). Although baicalin is a weak binder, it provides a reasonable starting point to explore other structurally related, commercially available compounds.

A small set of several baicalin-related compounds were chosen (Figure 7 a). These molecules conserve the flavone scaffold and show minor variations in the number and position of several aromatic hydroxy groups. Of these molecules, the glycosylated flavone quercetin-3-β-glucoside was the only one that induced methionine perturbations on the 1H–15N decoupled 1H NMR spectra of VEGF. The interaction of baicalin with VEGF was further characterized by 1H–15N HSQC experiments. Upon the addition of baicalin, the most remarkable CSPs were found in a region centered in residues 85–96. In addition, resonances corresponding to amino acids in the vicinity of positions 46–48 and 18–22 were also significantly shifted (Figure 7). Titrations were conducted and baicalin-induced CSPs in the VEGF spectra were quantified and used to determine the binding affinity; Kₐ values of 4.9 and 6.8 mM were obtained for 1H–15N HSQC and 1H–15N HSQC shifts, respectively (see the Supporting Information). Although baicalin is a weak binder, it provides a reasonable starting point to explore other structurally related, commercially available compounds.

A small set of several baicalin-related compounds were chosen (Figure 7 a). These molecules conserve the flavone scaffold and show minor variations in the number and position of several aromatic hydroxy groups. Of these molecules, the glycosylated flavone quercetin-3-β-glucoside was the only one that induced methionine perturbations on the 1H–15N decoupled 1H NMR spectra of VEGF. The interaction of baicalin with VEGF was further characterized by 1H–15N HSQC experiments. Upon the addition of baicalin, the most remarkable CSPs were found in a region centered in residues 85–96. In addition, resonances corresponding to amino acids in the vicinity of positions 46–48 and 18–22 were also significantly shifted (Figure 7). Titrations were conducted and baicalin-induced CSPs in the VEGF spectra were quantified and used to determine the binding affinity; Kₐ values of 4.9 and 6.8 mM were obtained for 1H–15N HSQC and 1H–15N HSQC shifts, respectively (see the Supporting Information). Although baicalin is a weak binder, it provides a reasonable starting point to explore other structurally related, commercially available compounds.

A small set of several baicalin-related compounds were chosen (Figure 7 a). These molecules conserve the flavone scaffold and show minor variations in the number and position of several aromatic hydroxy groups. Of these molecules, the glycosylated flavone quercetin-3-β-glucoside was the only one that induced methionine perturbations on the 1H–15N decoupled 1H NMR spectra of VEGF. The interaction of baicalin with VEGF was further characterized by 1H–15N HSQC experiments. Upon the addition of baicalin, the most remarkable CSPs were found in a region centered in residues 85–96. In addition, resonances corresponding to amino acids in the vicinity of positions 46–48 and 18–22 were also significantly shifted (Figure 7). Titrations were conducted and baicalin-induced CSPs in the VEGF spectra were quantified and used to determine the binding affinity; Kₐ values of 4.9 and 6.8 mM were obtained for 1H–15N HSQC and 1H–15N HSQC shifts, respectively (see the Supporting Information). Although baicalin is a weak binder, it provides a reasonable starting point to explore other structurally related, commercially available compounds.

A small set of several baicalin-related compounds were chosen (Figure 7 a). These molecules conserve the flavone scaffold and show minor variations in the number and position of several aromatic hydroxy groups. Of these molecules, the glycosylated flavone quercetin-3-β-glucoside was the only one that induced methionine perturbations on the 1H–15N decoupled 1H NMR spectra of VEGF. The interaction of baicalin with VEGF was further characterized by 1H–15N HSQC experiments. Upon the addition of baicalin, the most remarkable CSPs were found in a region centered in residues 85–96. In addition, resonances corresponding to amino acids in the vicinity of positions 46–48 and 18–22 were also significantly shifted (Figure 7). Titrations were conducted and baicalin-induced CSPs in the VEGF spectra were quantified and used to determine the binding affinity; Kₐ values of 4.9 and 6.8 mM were obtained for 1H–15N HSQC and 1H–15N HSQC shifts, respectively (see the Supporting Information). Although baicalin is a weak binder, it provides a reasonable starting point to explore other structurally related, commercially available compounds.

A small set of several baicalin-related compounds were chosen (Figure 7 a). These molecules conserve the flavone scaffold and show minor variations in the number and position of several aromatic hydroxy groups. Of these molecules, the glycosylated flavone quercetin-3-β-glucoside was the only one that induced methionine perturbations on the 1H–15N decoupled 1H NMR spectra of VEGF. The interaction of baicalin with VEGF was further characterized by 1H–15N HSQC experiments. Upon the addition of baicalin, the most remarkable CSPs were found in a region centered in residues 85–96. In addition, resonances corresponding to amino acids in the vicinity of positions 46–48 and 18–22 were also significantly shifted (Figure 7). Titrations were conducted and baicalin-induced CSPs in the VEGF spectra were quantified and used to determine the binding affinity; Kₐ values of 4.9 and 6.8 mM were obtained for 1H–15N HSQC and 1H–15N HSQC shifts, respectively (see the Supporting Information). Although baicalin is a weak binder, it provides a reasonable starting point to explore other structurally related, commercially available compounds.
similarities, most of the changes produced by baicalin were concentrated around residues 86–93, located at β-strands 5 and 6 and their interconnecting loop. Conversely, quercetin-3-β-glucoside-specific perturbations were mainly located on the opposite site of VEGF, around Phe36 and the loop connecting to the second β-strand. These observations suggest that baicalin and quercetin-3-β-glucoside target overlapping but slightly different sites on VEGF owing to their distinct binding orientations. Given that the two VEGF ligands share a common flavone scaffold and differ mainly in the type and location of the glycosidic moiety and to a lesser extent in the position of various hydroxy groups, it is reasonable to place the common scaffold on the locus of Phe47, the region equally affected by these two flavonoid analogues. This model is in agreement with the experimental STD data, which suggested that the binding contribution of the flavone scaffold was stronger than that of the glycosidic moiety. The glucuronic moiety of baicalin is most likely accommodated between α-helix 1 and β-strands 5 and 6, as residues in these secondary structure elements were selectively affected by baicalin. In contrast, the sugar moiety of quercetin-3-β-glucoside is probably located on the opposite site of the protein, which induces significant shifts on Gly65 and residues on α-helix 2 (Figure 7c). Overall, the screening of plant extracts from traditional Chinese medicine allowed us to identify a flavonoid scaffold that is able to bind VEGF in the VEGFR binding interface.

Discussion
The term “druggability” is used to describe a biological target that is known to bind with high affinity to a low molecular weight drug.\(^{[5]}\) Many proteins, including several related to cancer, are considered as “undruggable” targets.\(^{[76]}\) The only VEGF ligands that have found their way into clinical practice are monoclonal antibodies\(^{[34–36]}\) and, more recently, aptamers.\(^{[38]}\) In the present study, we summarized several years of work of many members of our research group, and also of several collaborators, in an effort to explore the feasibility of a therapeutic strategy based on blocking the VEGFR interaction by using ligands—other than biologics—able to bind VEGF specifically in the region of interaction with its receptor. Herein, we explored a wide variety of approaches, including NMR spectroscopy based fragment screening, screening of peptide libraries, and activity screening of plant extracts. This extensive work provided a clear picture. First, it was difficult to identify high-affinity ligands by using these approaches. Second, the only ligands (weak) that bind to the VEGF surface patch involved in the interaction with the receptor are conformationally restrained cyclic peptides and medium-sized natural products. Finally, all the small molecular weight organic fragments able to bind VEGF interact with the protein in patches other than the therapeutically relevant one.

This behavior, in particular the difficulties encountered to identify small organic fragments targeting the VEGF interacting surface, could be related to the topological and chemical fea-
Figure 6. a) HPLC chromatogram of baicalin extracted from the Radix scutellariae plant extract. The chemical structure of baicalin is shown (right). b) STD NMR spectrum (T = 298 K) of a sample containing 1 mM baicalin and 10 μM VEGF. STD spectra of the same sample in the presence of c) 50 mM and d) 100 mM v107 peptide showing the reduction of the STD signal intensities of baicalin in the presence of v107.

Structures of the protein–protein interface. The VEGF–VEGFR-1 binding interface is flat, mainly hydrophobic, and rather large (> 800 Å² per receptor unit), which thereby suggests that targeting this interaction may be especially challenging. However, a compound interacting with the VEGF binding site in the VEGFR D2 domain and inhibiting complex formation was recently described. This observation thus indicates that this PPI may be amenable to disruption by synthetic compounds. Several VEGF segments distant in sequence generate a discontinuous binding epitope that contributes to receptor binding. The binding determinants of VEGF for VEGFR are localized on the N-terminal α-helix 1 (residues 16–27), the loop connecting β3 to β4 (residues 61–66), and strand β7 (residues 103–106) of one VEGF monomer as well as strand β2 (residues 46–48) and strands β5 and β6, and the turn connecting them (residues 79–91).

Protein–protein interfaces could be structurally classified as domain–domain or peptide–domain interactions. In domain–domain PPIs, two globular protein domains are involved in the interface, whereas in peptide-mediated PPIs a continuous polypeptide segment binds to a globular domain. The size and shape of the contact area differ considerably between the two interaction types. Peptide-mediated interfaces have relatively small contact areas with hot spot residues in close proximity. In addition, members of this class tend to contain defined pockets at the interface that may act as anchors of interacting residues. These structural features make peptide-mediated interactions more suited for inhibitors than the larger and flatter domain–domain interfaces.

An increasing collection of PPI inhibitors has been described. Representative examples of successfully targeted PPI interactions include, among others, those of the p53 activation domain with MDM2, Bcl-XL proteins with BH3 peptides, and cytokine interleukin-2 IL2 with its receptor IL2Rα. More recently, our research group reported a series of photoswitchable inhibitors of clathrin-mediated endocytosis able to disrupt the β-arrestin–β-adaptin 2 interaction. Most targeted PPIs are either peptide–domain interactions or a fraction of domain–domain PPIs that contain a continuous dominant peptide at the interaction interface that contributes most of the binding energy. In the latter case, inhibitors often target the interface region in which the dominant peptide interacts.

The reasons why targeting VEGF has proven to be extremely challenging are probably correlated with the absence of desirable druggable features in VEGF. The VEGF–VEGFR interaction is a typical domain–domain interaction. VEGF contains a relatively flat and extended interface with hot spots distributed over a large area. The distance between the Cα of Leu66 and Met18 is ~20 Å. Using several computational techniques, we predicted the lack of preformed or transient cavities or crevices with potential capacity to bind small molecules with reasonable affinity in the receptor binding surface of VEGF.

The presence of large exposed hydrophobic surface patches is one of the signatures of domain–domain PPIs. In this context, our results suggest that to disrupt domain–domain PPIs ligands would need to expose hydrophobic surfaces that are over a minimal threshold. This requirement translates, on the one hand, into the need for a minimum molecular weight. On the other hand, conformational constraints are required to prevent these relatively large hydrophobic molecules to adopt energetically favored folded conformations that would shield the hydrophobic functional groups from the water. In other words, it would be necessary to force the hydrophobic parts of the molecule to remain exposed at the surface, which would thus facilitate interaction with the hydrophobic patch at the protein surface. Our analysis could be applicable to other targets of this nature and may help in the design of inhibitors and/or screening strategies of many therapeutically relevant domain–domain PPIs.

Experimental Section

Chemistry

Protein production: VEGF11–109 (VEGF) was expressed, purified, and refolded as described in detail elsewhere. Large-scale expression was performed on a 10 L scale with a Biostat B fermentor (Sartorius). In this case, protein expression was induced once the culture reached an optical density of 1.2 at 600 nm (OD600), and cells were harvested at an OD600 of ~6.0. During the fermentation, oxygen was supplemented into gas flow and the pH was kept constant at 7.0.

Peptide synthesis: A standard solid-phase Fmoc/tBu peptide synthesis protocol was followed. Dipeptide libraries were supplied by

**These are not the final page numbers!**
Dr. F. Yraola from the Combinatorial Chemistry Unit at the Barcelona Science Park. These libraries were produced by solid-phase synthesis and were purified by HPLC to a purity 95 % or higher. For storage purposes, compounds were weighed and dissolved in [D$_6$]DMSO at a concentration of 100 mM.

For the synthesis of symmetric cyclic hexapeptide &o-Pro-Trp-Glu-o-Pro-Trp-Glu&, 2-chlorotrityl resin (CTC resin, 250 mg, 1.6 mmol g$^{-1}$) was placed in a 20 mL polypropylene syringe fitted with a polyethylene filter disc. The resin was then washed with CH$_2$Cl$_2$ (5 × 30 s) and a solution of Fmoc-l-Trp(Boc)-OH (184.3 mg, 0.7 equiv) and N,N-diisopropylethylamine (DIEA; 366 µL, 10 equiv) in CH$_2$Cl$_2$ was then added; the mixture was then stirred for 1 h at room temperature. The remaining active sites were capped with the addition of MeOH (400 µL), and the mixture was stirred for an additional 10 min at room temperature. The Fmoc-l-Trp(Boc)-O-CTC resin was subjected to the following washes: CH$_2$Cl$_2$ (5 × 30 s), DMF (5 × 30 s), and 20 % piperidine in DMF (1 × 1 min, 2 × 5 min, 1 × 10 min). Elongation of the peptide was achieved by the sequential addition of Fmoc-AA-OH (3 equiv) with 2-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyllumonomium hexafluorophosphate (HATU, 3 equiv) and DIEA (6 equiv) as coupling reagents in DMF. The coupling reaction was stirred for 30 min, and before filtration the corresponding colorimetric test (ninhydrin test for primary amines and chloranil test for secondary amines) was done to qualitatively assess if coupling reactions were completed. Removal of the Fmoc group was performed with 20 % piperidine in DMF (1 × 1 min, 2 × 10 min, 1 × 10 min). The H-Glu(OrtB)-o-Pro-Trp(Boc)-Glu(OrtBu)-o-Pro-Trp(Boc)-O-CTC peptide resin was treated with 1 % trifluoroacetic acid (TFA) in CH$_2$Cl$_2$ (5 × 30 s) to cleave the peptide from the resin. The filtrates were collected over H$_2$O (60 mL per gram of resin). CH$_2$Cl$_2$ was removed under an atmosphere of nitrogen and water was removed by lyophilization. The linear peptide was subjected to head-to-tail cyclization in solution by using (7-azabenzotriazol-1-yloxy)tris(pyrrolidino)phosphonium hexafluorophosphatate (PyAOP) and DIEA as coupling reagents in a mixture of CH$_2$Cl$_2$/DMF (98:2) at high dilution (5 mM) for 16 h at room temperature. The cyclic peptide was treated with TFA/H$_2$O/triisopropylsilane (TIS) (95:2:5:2.5, v/v/v/v) for 2 h at room temperature to remove all protecting groups. Solvent was removed under an atmosphere of ni-

Figure 7. a) Chemical structure of the flavonoid molecules tested. b) Histogram showing weighted average chemical shift changes for VEGF amide resonances between free $^{15}$N-labeled VEGF (250 µM) and $^{15}$N-labeled VEGF (250 µM) in the presence of either 7.7 mM baicalin (black) or 4.4 mM quercetin-3-β-glucoside (gray). c) Model of the VEGF–baicalin complex. VEGF is shown as a cartoon representation and baicalin as a ball and stick model.

ChemMedChem 0000, 00, 0 – 0  www.chemmedchem.org

© 0000 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

These are not the final page numbers!
trogen, and the crude product was purified by CombiFlash automated flash chromatography (Teledyne ISCO) to obtain the final peptide (39.1 mg, 47 %). H NMR (600 MHz, H2O/D2O, δ = 10.10 (s, 1 H), 8.49 (d, J = 7.6 Hz, 1 H), 7.70 (d, J = 8.1 Hz, 1 H), 7.68 (d, J = 8.8 Hz, 2 H), 7.49 (d, J = 8.5 Hz, 2 H), 7.27 (s, 2 H), 7.24 (t, J = 7.7 Hz, 2 H), 7.16 (t, J = 7.7 Hz, 2 H), 4.26–4.17 (m, 1 H), 3.71–3.62 (m, 1 H), 3.62–3.54 (m, 1 H), 3.48 (dd, J = 15.6, 4.3 Hz, 1 H), 3.22 (dd, J = 15.5, 9.5 Hz, 1 H), 2.05–1.98 (m, 2 H), 1.98–1.83 (m, 4 H), 1.82–1.73 (m, 1 H), 1.49–1.38 ppm (m, 1 H). HPLC (solvent A: H2O with 0.045 % TFA, solvent B: MeCN with 0.036 % TFA, conditions 7:3 to 0:10 in 15 min): tR = 6.2 min. Purity: > 98 %.

mRNA display-derived peptides: Chloroacetylation of the free N terminus was performed with a solution of 0.2 M CI-AC-NHS (NHS = N-hydroxysuccinimide; 2.5 mL) dissolved in N-methyl-2-pyridolone (NMP) under shaking for 40 min. For fluorescein conjugation, after chloroacetylation, the 4-methoxytrityl (Mtt) group was removed by acidic solution (TFA/TIS/CH2Cl2, 1:1:98, v/v/v) for 10 min. The peptidyl resin was then extensively washed with CH2Cl2 (3 × 1 min), NMP (3 × 1 min), 10 % DIEA in NMP (3 × 1 min), and NMP (3 × 1 min). The peptidyl resin was subjected to a reaction with a solution of 0.2 M NHS-fluorescein in 10 % DIPEA in NMP for 10 min. The crude product was redissolved in DMSO (2.5 mL) and diluted with MeCN/H2O for 3 h at room temperature. After removal of TFA, the crude prod-}

### mRNA display-derived peptides

Chloroacetylation of the free N terminus was performed with a solution of 0.2 M CI-AC-NHS (NHS = N-hydroxysuccinimide; 2.5 mL) dissolved in N-methyl-2-pyridolone (NMP) under shaking for 40 min. For fluorescein conjugation, after chloroacetylation, the 4-methoxytrityl (Mtt) group was removed by acidic solution (TFA/TIS/CH2Cl2, 1:1:98, v/v/v) for 10 min. The peptidyl resin was then extensively washed with CH2Cl2 (3 × 1 min), NMP (3 × 1 min), 10 % DIEA in NMP (3 × 1 min), and NMP (3 × 1 min). The peptidyl resin was subjected to a reaction with a solution of 0.2 M NHS-fluorescein in 10 % DIPEA in NMP for 10 min. The crude product was redissolved in DMSO (2.5 mL) and diluted with MeCN/H2O (1:1, v/v, 1000 mL). The solution was adjusted to pH 9 with Et3N to perform the head-to-side chain cyclization. The crude product was purified by HPLC. Table 3 shows the characterization data from the peptides obtained.

| Table 3. Characterization of mRNA display-derived peptides. |  |
|---|---|
| Peptide | HPLC[a] | tR [min] | Expected | Found |
| 100:0–0:100 | 4.8 | 1109.43 (z = 2) | 1109.44 (z = 2) |
| 100:0–0:100 | 4.1 | 771.77 (z = 3) | 771.77 (z = 3) |
| 100:0–30:70 | 4.9 | 796.40 (z = 3) | 796.40 (z = 3) |
| 98.5:1:0 | 4.1 | 463.60 (z = 3) | 765.40 (z = 3) |
| 98:10:50 | 4.4 | 722.05 (z = 3) | 722.05 (z = 3) |
| 80:20:50 | 7.7 | 908.44 (z = 4) | 908.44 (z = 4) |
| 80:20:50 | 10 | 899.42 (z = 2) | 899.42 (z = 3) |
| 100:0–0:100 | 4.9 | 897.47 (z = 3) | 897.47 (z = 3) |
| 1 | 100:0–0:100 | 5.5 | 647.79 (z = 4) | 647.79 (z = 4) |

[a] HPLC gradient (8 min); solvent A: H2O with 0.045 % TFA, solvent B: MeCN with 0.036 % TFA; [b] 15 min time gradient in a C8 column (1.6 mm × 15 mm).

### NMR spectroscopy

STD and WaterLOGSY experiments were conducted over a range of temperatures (37, 25, or 5 °C) with Bruker Digital Avance 600 MHz and Varian Inova 500 MHz spectrometers. Samples contained 10 μM VEGF and 0.5 or 1 μM of assayed compound(s) in 25 mM sodium phosphate (pH 7.0), 50 mM NaCl, and 90 % D2O. On-reso- nance irradiation was performed on the water signal for WaterLOGSY experiments and on the aliphatic protein region for STD experiments. The total saturation time was 2–2.5 s, and water suppres- sion was achieved with a watergate module.

In the screening experiments, fragments showing STD signals with "normal" or "strong" intensities for more than 30 % of their protons were classified as potential good binders.

Two-dimensional 1H–15N HSQC and 1H–13C HSQC experiments and 1D 13C-decoupled-1H-filtered 1H NMR spectra were usually obtained at 45 °C with a Bruker Avance III 600 MHz spectrometer equipped with a cryoprobe. Standard samples for protein-observed experiments contained 100–200 μM (1H–15N HSQC) or 30 μM (1H– 13C HSQC and 1D 13C-decoupled-1H-filtered 1H) VEGF in buffer A: 25 mM sodium phosphate (pH 7.0), 50 mM NaCl, and 90 % D2O. 2D 1H–13C HSQC experiments were typically acquired by using 256 × 96 points and 8 scans per increment with spectral widths of 960 Hz (1.6 ppm) and 453 Hz (3.0 ppm) in the 1H and 13C dimensions, re- spectively. Using this 1H spectral width, cross-signals of Met78 were folded. 13C-decoupled-1H-filtered 1H NMR spectra were recorded on 2048 points, a spectral width of 9000 Hz and 128 scans.

1H–15N HSQC spectra were acquired with 1024 complex points in the F2 (1H) dimension and 100 increments in the F1 (15N) dimension, 96 scans for each increment, and a spectral width of 7212 Hz (12.0 ppm) in F2 and 1794 Hz (29.5 ppm) in F1.

The dependence of the chemical shift changes of VEGF on the ligand concentration was used to estimate binding constants. 1H NMR and 15N NMR chemical shift changes from 1H–15N HSQC ti- trations or alternatively 1H NMR and 15N NMR shifts from 1H–13C HSQC experiments were combined as previously described.[82] For ligands interacting in the receptor-binding site of VEGF, two equal and independent binding sites were assumed.[83]

All 1H NMR spectra were acquired with a Bruker spectrometer operating at 700 MHz and equipped with a dual fluorine-proton SEF probehead and a robotic arm BACS120. The CF3 and CF cocktail samples for the 19F NMR screening were prepared at 20 and 50 μM, respectively, by diluting premixed cocktail stocks (10 μM each of the eight compounds in each stock) in [D6]DMSO in a final volume of 550 μL of buffer A.

Two 19F NMR spectra for each cocktail sample were recorded: a regular 1D and a 1D containing a CPMG T2 filter of 200/400 ms (for CF/CF3), before and after the addition of 2 μM VEGF. Spectra were recorded with a 20 Hz sample spinning using 128 scans, acquisition and recovery times of 1.2 s and 2.5 s, respectively, in 8.5–9.5 min per spectrum.

### Plant extracts from traditional Chinese medicine

Plants were obtained from Herbasin (Shenyang, China). Dried plant material (30 g) was extracted with H2O (400 mL) at reflux for 3 h by means of a Soxhlet apparatus. The volume of the aqueous extracts was reduced by evaporation under vacuum, and the extract was then freeze dried. Finally, [D5]DMSO stocks were prepared at a 200 mg mL−1 concentration.

### Acknowledgements

This work was supported by the Ministry of Economy and Competitiveness and the European Fund for Regional Development (MINECO-FEDER) (Bio2013-40716-R) and the Generalitat de Cata- lunya (XRB and 2014-SGR-521). We also thank the NMR facility at the Scientific and Technological Centre of the University of Barce- lona (CCiT UB) for their technical support.
[67] R. E. Hubbard, J. Synchrotron Radiat. 2008, 15, 227 – 230.
[68] X. Arroyo, M. Goldflam, M. Feliz, I. Belda, E. Giralt, PLOS ONE 2013, 8, e58571.
[69] M. Garav/C237s, B. L/C243pez-M/C216ndez, A. Somoza, J. Oyarzabal, C. Dalvit, A. Villa-
sante, R. Campos-Olivas, C. Gonz/C181lez, ACS Chem. Biol. 2014, 9, 1559 –
1566.
[70] Schrçdinger Release 2013-3, Schrçdinger Suite 2013 Protein Preparation
Wizard: Epik version 2.6, Impact version 6.1, Prime version 3.4, SiteMap
version 2.9, LigPrep version 2.8, Glide version 6.1; Schrçdinger LLC, New
York, NY (USA), 2013.
[71] K. W. Borrelli, A. Vitalis, R. Alcantara, V. Guallar, J. Chem. Theory Comput.
2005, 1, 1304 –1311.
[72] D. R. Yance, S. M. Sagar, Integr. Cancer Ther. 2006, 5, 9 – 29.
[73] D. J. Newman, M. C. Gordon, J. Nat. Prod. 2012, 75, 311 – 335.
[74] C. Liu, A. Tseng, S. Yang, Chinese Herbal Medicine: Modern Applications
of Traditional Formulas, CRC Press, Boca Raton, 2005.
[75] S. Ikemoto, K. Sugimura, N. Yoshida, R. Yasumoto, S. Wada, K. Yamamoto,
T. Kishimoto, Urologia 2000, 55, 951 – 955.
[76] C. M. Crews, Chem. Biol. 2010, 17, 551 – 555.
[77] C. F. Cheok, C. S. Verma, J. Baselga, D. P. Lane, Nat. Rev. Clin. Oncol. 2011,
8, 25 – 37.
[78] A. M. Petros, J. Dinges, D. J. Augeri, S. A. Baumeister, D. A. Betebenner,
M. G. Bures, S. W. Elmore, P. J. Hajduk, M. K. Joseph, S. K. Landis, D. G.
Nettesheim, S. H. Rosenberg, W. Shen, S. Thomas, X. Wang, I. Zanze, H.
Zhang, S. W. Fesik, J. Med. Chem. 2006, 49, 656 – 663.
[79] A. C. Braisted, J. D. Oslob, W. L. Delano, J. Hyde, R. S. McDowell, N. Waal,
C. Yu, M. R. Arkin, B. C. Raimundo, J. Am. Chem. Soc. 2003, 125, 3714 –
3715.
[80] L. Nevola, E. Giralt, Chem. Commun. 2015, 51, 3302 – 3315.
[81] A. Martín-Quirós, L. Nevola, K. Eckelt, S. Madurga, P. Gorostiza, E. Giralt,
Chem. Biol. 2015, 22, 31 – 37.
[82] M. Goldflam, T. Tarragó, M. Gairí, E. Giralt, Methods Mol. Biol. 2012, 831,
233 – 259.

Received: October 9, 2015
Published online on 0000

© 0000 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

These are not the final page numbers!
Master of its domain: Vascular endothelial growth factor (VEGF) is considered a relevant target for anticancer therapy with low druggability. We describe an extensive study exploring oligopeptides, fragment-based libraries, and natural products to target the receptor binding site of VEGF. The successful ligands have the capacity to expose hydrophobic groups that interact with the hydrophobic hot spots at the interacting VEGF surface patch.