Toward a novel drug to target the EGF-EGFR interaction: design of metabolically stable bicyclic peptides


Abstract: In cancer, proliferation of malignant cells is driven by overactivation of growth signaling mechanisms, such as the EGF-EGFR pathway. Despite its therapeutic relevance, this extracellular protein-protein interaction (PPI) has remained elusive to inhibition by synthetic molecules, mostly due to its large size and lack of binding pockets and cavities. Designed peptides, featuring cyclic motifs and other structural constraints, have the potential to modulate such challenging PPIs. Here we present the structure-based design of a series of bicyclic constrained peptides that mimic an interface domain of EGFR and inhibit the EGF-EGFR interaction by targeting the smaller partner (EGF). This design process was guided by the integrated use of in silico methods and biophysical techniques, such as NMR and SAW. The best analogues were able to selectively reduce the viability of EGFR+ human cancer cells. In addition to their efficacy, these bicyclic peptides are endowed with exceptional stability and metabolic resistance—two features that make them suitable candidates for in vivo applications.

Introduction

Protein-protein interactions (PPIs) play a relevant role in all cellular processes. The selective modulation of PPIs has emerged as a new strategy for therapeutic intervention. PPIs are typically characterised by large, flat and featureless interfaces,[1] thereby complicating the use of conventional small-molecule drugs as inhibitors.[2] As an alternative, peptides have a higher degree of structural flexibility, which allows them to better adapt to irregular targets,[3] and they can be designed to display a variety of topologies that closely mimic the structural features found in PPIs.[4] However, peptides tend to have low serum stability and poor ability to cross physiological barriers.[5] In this regard, these limitations need to be addressed if in vivo applications of these molecules are to be pursued.

At the onset of cancer, cells acquire the capacity to self-proliferate through the uncontrolled production of molecules that promote cell growth and differentiation.[6] Among these, epidermal growth factor (EGF) has been shown to participate in many types of solid tumour, including head and neck, breast, colon, ovarian and non-small-cell lung cancer.[7] To perform its function, EGF binds with high affinity (Kd = 1.9 nM)[5] to its membrane receptor (EGFR),[10] triggering intracellular events that lead to uncontrolled cell growth, tumour invasion and metastasis.[8]

Therefore, the EGF-EGFR pathway has become a main focus for selective chemotherapeutic intervention and, as a result, two classes of EGFR inhibitors have been clinically approved, namely monoclonal antibodies[11] (cetuximab, panitumumab), which target the extracellular domain of EGFR, and small-molecule kinase inhibitors[12] (gefitinib, erlotinib), which block the intracellular phosphorylation of the receptor. Despite impressive initial response rates to these drugs, most patients end up developing drug resistance—a major limitation that reduces the long-term efficacy of the therapy.[13] The overexpression of EGF and/or EGFR, as well as activating mutations on the receptor, have been shown to underlie the main drug resistance mechanisms.[14] Thus, novel approaches to block the EGF-EGFR pathway and targeted combination therapies are in high demand.

In this regard, the direct inhibition of EGF is considered a promising strategy.[14] Although similar growth factors (such as VEGF) have been successfully addressed using this approach,[15] no current drugs are able to specifically target EGF.[7] However, we recently showed that peptides represent a significant step in this direction.[16] Herein, we present the structure-based design of bicyclic-constrained peptides that adopt a well-defined structure and mimic one of the interacting domains of EGFR. In this process, computational and biophysical methods were used as complementary approaches to guide the design of more stable and active analogues. In addition to studying their interaction with EGF at the molecular level, we confirmed their potential to block the EGF-EGFR interaction in a specific receptor-ligand assay and also in human carcinoma cells, which constitutively overexpress EGFR. Finally, the stability of the best designs was validated in biologically aggressive media, such as serum and liver hepatocytes, thereby confirming the feasibility of our lead compounds for further in vivo application.
Results and Discussion

Design and biophysical evaluation of peptide ligands

The binding of EGF to EGFR occurs through a large (1440 Å³) and complex interface, in which various domains of the receptor are involved. Nevertheless, a number of hotspots are found on residues 7 to 34 of the receptor (Figure 1A). In particular, residues 16–18 of EGF and residues 31–33 of EGFR form a short parallel β-sheet, stabilised by backbone hydrogen bonds. Also, the α-helical motif in this EGFR domain features several hotspots (Asp22, Leu25, Arg29) that establish key contacts with EGF. As we recently reported, a cyclic peptide (cp28) comprising this EGFR region was able to recapitulate the binding of the receptor to EGF (Kd = 286 μM). However, the large size and hydrophobic character of this peptide, in addition to its all-L composition, translated into poor solubility and biological stability, thus limiting its in vivo application and potential development as a drug candidate.

Using a rational design approach, our goal was to mimic the bioactive conformation of the receptor with constrained peptides, which are endowed with superior stability and drug-like properties. Initially, we designed three peptides (named cp23A, B and C, Figure 1B and C), each with a turn-inducing D-Pro-Gly motif that replaced part of the long cp28 loop. This loop plays a structural role in EGFR but is not directly involved in interactions with EGF. In addition, we conservatively mutated the Phe residues at the helix for Tyr, in order to improve the overall water solubility of the construct. Finally, peptides cp23B and C bore a second cyclic constraint, in the form of a disulphide bridge, which aimed to restrain the flexibility of the peptide loop. In these last designs, we inspected the peptide sequence for side chains that would allow the introduction of a disulphide crosslink, without altering the overall conformation of the peptide. In cp23B, Asn6 and Leu11 were replaced by L-homocysteine and L-Cys, respectively. Alternatively, in cp23C, Gln10 and His17 were modified by a D- and L-configured building blocks (Figure 1C).

The synthesis of the bicyclic scaffolds was achieved using the native chemical ligation strategy (Figure S1). This approach involves Fmoc-based solid-phase peptide synthesis (SPPS) of linear peptides featuring a C-terminal N-acrylurea moiety, which acts as thioester precursor during the ligation. Treatment of the unprotected peptide in solution with a thiol additive (4-mercaptophenol) then catalyses in situ the intramolecular ligation, forming the head-to-tail cyclic peptides in under an hour at pH 7 (see the Experimental Section for details). Finally, the thiol groups are readily oxidised to yield the desired bicyclic peptides.

The circular dichroism (CD) spectra of monocyclic cp23A and bicyclic cp23B (Figure 1D) showed a negative band at ca. 208
Interaction (MD) simulations are able to extensively sample the theoretical mean free energy of binding events occurring on this functionalised surface, which directly affect the acoustic wave parameters of the biosensor, were measured. In this assay, the bicyclic peptide cp23B displayed the best affinity of the series ($K_D = 575 \, \mu M$, Figure 2A and Table 1), consistent with the changes observed in the NMR spectra of EGF. As shown by these results, the helical motif is a required element for stabilising the entire peptide backbone. The bicyclic architecture of cp23B—but not of cp23C—further constrains the peptide and enables the folding of the loop in the bioactive conformation.

Having proved that we could retain the main requisites for binding in a smaller, more rigid, and synthetically accessible scaffold, we used computational methods to enhance the folding (in terms of resemblance to the EGFR bioactive conformation) and binding affinity of the peptides. Compared to docking methods, which treat the protein as a rigid or semi-rigid body and are not convenient for flexible proteins such as EGF, molecular dynamics (MD) simulations are able to extensively sample the interacting complex in a solution environment. We subjected the cp23B-EGF complex to free MD simulation and, from regular snapshots taken during the first part of the simulation, we calculated the theoretical mean free energy of binding ($\Delta G = -31.4 \, \text{kcal/mol}$, according to the MM/GBSA method). In the last part of the simulation, larger fluctuations occurred on the peptide backbone, causing its dissociation from EGF (Figure S2).

We then performed virtual mutations on the weaker points—from a design point of view—of the cp23B sequence. For instance, several replacement options were tested for Met22, as oxidation of its side chain often occurred during experimental manipulation. Also, we explored the introduction of a second Pro residue at the N-terminus of the $\alpha$-helix, with the aim to stabilise this structural motif. In total, we generated and evaluated 28 new peptide structures (Table S1). As revealed by the calculated $\Delta G$ values, in most cases the Y→P mutation produced an 8-10 kcal/mol gain of interaction energy, especially when Pro was at the 12th position. This observation confirms the stabilizing effect of Pro as N-cap for the helix, which translated to conformational smoothing in the vicinity of the Pro residue.
into less backbone fluctuations and improved peptide-EGF contacts during the MD (Figure S2). Also, a conservative M22Nva mutation (as in cp23E and F) was suitable for replacing the unstable side chain of this Met residue.

To test the new computational designs, we selected three of the best analogues, namely cp23D, cp23E and cp23F, for chemical synthesis and evaluation (Table 1). NMR revealed that all three peptides produce clear CSPs on the same part of the protein (L8, S9, G12, H16-V19) as cp23B, thus confirming the same mode of binding (Figure 2B). In the quantitative SAW experiment, the interaction of cp23F produced well-defined association and dissociation curves, which translated into a higher affinity ($K_D = 279 \mu M$) than cp23B. In contrast, the sensorsgrams for both cp23D and cp23E were low-intensity and imprecise, thus indicating loss of affinity (Figure S4). These observations suggest that the introduction of radical modifications on the peptide sequence, as in these last designs, impair binding to EGF.

Given the abovementioned results, we implemented a final cycle of computational modelling for the further optimisation of cp23F. To this end, we generated a set of heterochiral mutants by introducing L- and D-amino acids in suitable positions of the sequence. Not surprisingly, most new mutations had a detrimental effect on the $\Delta G$ (Table S2). Nevertheless, the substitution of Gly10 for large apolar residues, such as Met (as in cp23H) or Tyr, formed new hydrophobic contacts with the B-loop of EGF (Val29, Val30), thus improving the $\Delta G$ score. The introduction of D-amino acids in position 2 of the sequence also led to theoretically stronger interactions, with the T2r mutant (cp23G) having the lowest MM/PBSA score of the entire set ($-8.1 \text{ kcal/mol}$). As seen in the MD simulation, the guanidine group of the D-Arg formed a highly stable salt bridge with Asp42 in the C-loop of EGF (Figure S2), and enhanced the hydrophilic character of the peptide.

Experimentally, the affinities of cp23G ($K_D = 252 \mu M$) and cp23H ($K_D = 115 \mu M$), as measured by SAW, were superior to that of cp23F, although not proportionally to the predicted $\Delta G$ values (Figure 2C and Table 1). The induced shifts on the $^{15}$N-$^1$H HSQC spectra of the protein confirmed their binding to the EGFR-binding epitope of EGF. However, the poor solubility of cp23H prompted us to select cp23G for further studies.

**Table 1.** Sequence of all cp23 analogues. For cp23B analogues, the $\Delta G$ values are shown, using either PB or GB solvation methods.
Structural characterisation of the bicyclic peptides

To gain a molecular-level understanding of the differences between these analogues, we simulated their minimal energy structures by replica exchange MD (Figure 3A). Compared to cp23B, the helical domain in cp23F was better structured and more stable, thanks to the constraining effect of the Pro residue in position 12. Nevertheless, backbone RMSDs of about 5 Å (compared to the bioactive conformation, PDB 1ivo) were obtained for both analogues, due to the poor folding of the flexible loop of the peptide. The new D-Arg mutation in cp23G partly solved this problem. Indeed, a close inspection of its lowest energy structure showed how the hydrocarbon part of the Arg side chain fills the empty space at the central cavity of the peptide, thus preventing the collapse of the flexible loop. Overall, this renders a conformation that closely mimics that of EGFR (RMSD = 2.7 Å).

To confirm the computational predictions, we investigated the structural characteristics of cp23G in solution. First, we performed CD spectroscopy of the peptide at a range of temperatures (Figure 3B). The CD spectra at 25°C showed two negative bands at ca. 208 and 222 nm, thereby confirming the existence of the α-helical motif. Notably, when the sample was heated to 84°C, the CD spectrum was nearly identical to the former, which reveals a high conformational rigidity for cp23G, given its small size. Second, we performed NMR measurements of the peptide in a buffered aqueous solution at pH 6.8. The 1H backbone chemical shifts exhibited a relatively large dispersion, which is typical of a more rigid and defined structure (Figure S7). The secondary chemical shifts of both 1H2 and 13Cα showed consecutive deviations from random coil values, negative for H2 and positive for Cα, along the Pro12-Gln20 region, indicative of α-helix conformation (Figure 3C). This hypothesis was further confirmed by the presence of short- and medium-range NOEs in this part of the sequence (Figure S7). Using these experimental constraints, we modelled a 3D-structure of cp23G in solution (Figure 3D), which resulted very similar to that of the previous replica exchange simulation (RMSD = 1.8 Å). This structure exhibits the features that have been key in this design process, namely precise folding of the peptide backbone, which is stabilised by the α-helix and the intramolecular disulphide bridge.

PPI inhibition assays

To produce a biological outcome, peptides binding to EGFR must disrupt the interaction with its receptor. Therefore, they must compete with the very strong affinities that naturally govern this process. On the other hand, such molecules benefit from acting on the extracellular space, where higher effective concentrations can be reached. We first assessed the capacity of our peptides to inhibit the EGFR-EGFR interaction in an AlphaScreen bead-based assay, which mimics the protein setting found in cells. For all the designs, a concentration-dependent inhibition of the EGFR-EGFR interaction was observed (Figure 4A and S8). The cp23G peptide was the most potent inhibitor of the series (IC50 = 149 µM), an effect that correlates with the binding affinities obtained by SAW.

We next evaluated the effect of our peptides on the viability of A-431 cells, which represent a particular subtype of skin carcinoma, characterised by abnormally high expression levels of EGFR. Cell viability was measured by XTT after a 72-h incubation with EGF-containing media and a range of inhibitor concentrations. To validate the assay conditions, erlotinib was used as positive control (IC50 = 2.9 µM). Under these conditions, cp23G had the greatest inhibitory effect (IC50 = 24 µM, Figure 4B), in the same range as cp28 (IC50 = 30 µM). Consistent with the previous results, cp23B and cp23F were not able to fully block cell proliferation (Figure S9). To confirm that the inhibitory effect was specific for the EGFR pathway, a parallel assay was performed on two EGFR-independent cells, namely MCF7 (a breast cancer line that overexpresses oestrogen receptor, but normal levels of EGFR) and healthy fibroblasts. In both lines, lack of toxicity was observed for all the peptides tested (Figure 4B and S9), a result that is in agreement with the principle of targeted chemotherapy.

Stability studies in biological media

Analogue | Sequence<sup>1</sup> | MM/GBSA (kcal/mol) | MM/PBSA (kcal/mol) | Ke (µM)<sup>2</sup> |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cp23A</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-</td>
<td>-</td>
<td>773</td>
</tr>
<tr>
<td>cp23B</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-31.4</td>
<td>6.8</td>
<td>575</td>
</tr>
<tr>
<td>cp23C</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-</td>
<td>-</td>
<td>1860</td>
</tr>
<tr>
<td>cp23D</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-46.8</td>
<td>-12.7</td>
<td>1420</td>
</tr>
<tr>
<td>cp23E</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-46.6</td>
<td>-9.2</td>
<td>&gt;2000</td>
</tr>
<tr>
<td>cp23F</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-42.1</td>
<td>-2.9</td>
<td>280</td>
</tr>
<tr>
<td>cp23G</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-41.4</td>
<td>-8.1</td>
<td>252</td>
</tr>
<tr>
<td>cp23H</td>
<td>&amp;GTSNLK Tolgyedh YLSLRQMPa &amp;</td>
<td>-48.1</td>
<td>-6.0</td>
<td>115</td>
</tr>
</tbody>
</table>

<sup>1</sup>X = L-homocysteine, Z = L-norvaline. Residues forming disulphide bonds are underlined. <sup>2</sup>Affinity values obtained by SAW.
One of the main challenges faced by therapeutic peptides is that they typically present short half-life in serum, rapid clearance from circulation, and poor tissue distribution. Being an all L-peptide, our first EGF-EGFR inhibitor cp28 was predictably labile in vivo. In fact, in human serum, cp28 showed a half-life of 89 min, which is reasonable for a natural peptide (the half-life of a linear peptide is typically less than 30 min). To identify the main cleavage sites on the sequence of cp28, we performed MS/MS analysis of the serum sample after 2 h of incubation. Interestingly, the three main metabolites resulted from hydrolysis and/or oxidation of Arg23 and Met24 (Figure S10), a region in which the peptide chain adopts a rather extended conformation and is thus more accessible to the active site of proteolytic enzymes.

On the contrary, the new bicyclic analogues, which present a more constrained structure and no free N- and C-terminal ends, were expected to present greater resistance to proteolytic degradation. To validate this hypothesis, we analysed the concentration of peptides in human serum at 37ºC for 24 h. Remarkably, no degradation was appreciated after 8 h, and cp23F and cp23G remained more than 60% unaltered after 24 h (Figure 4C). In these peptides, norvaline replaces the more vulnerable Met24 and this region is further protected from hydrolysis by an artificial D-Pro-Gly motif adjacent to it.

The greater stability of the bicyclic peptides was further confirmed in human hepatic microsomes, which represent a more aggressive environment than the serum. Indeed, liver microsomes contain a diversity of metabolising enzymes, such as proteases, oxidases, dehydrogenases and acyl-/alkyl-transferases, all of which contribute to peptide metabolism, especially in hydrophobic sequences. Experimentally, the bicyclic cp23 analogues, cp23G in particular, showed significantly lower clearance values than cp28 (Figure 4D), a metric that correlates with the in vivo metabolic clearance.\(^{[22]}\)

The difference in stability can be explained not only by the presence of unnatural residues, which hamper the recognition of enzymes, but also by the lower hydrophobicity of cp23G. As a result, the structural modifications introduced in this new bicyclic design not only guarantee the target affinity and cellular efficacy—already shown in the first cyclic inhibitor—but add relevant features such as increased solubility and superior biostability.

**Conclusions**

The EGF-EGFR interaction involves the formation of high-affinity dimeric complexes—mediated by several discontinuous epitopes—which extend along a large and irregular binding interface. Due to these challenging features, this PPI is the main chemotherapeutic approach in several types of cancer. However, poor bioavailability and the development of acquired drug resistance reduce the long-term success of antibody-based therapies. Designed peptides have emerged as a new class of therapeutics that combine the surface recognition properties of antibodies and the pharmacological behaviour of small molecules.

Using computational algorithms and biophysical techniques, we have implemented a rational and iterative process that has allowed the generation of a series of bicyclic constrained peptides that mimic the bioactive conformation of EGFR upon binding to EGF. The binding of these molecules to EGF has been quantified by SAW and characterised by NMR, and their inhibitory effect on the EGF-EGFR interaction has been...
confirmed in a specific PPI-disruption assay. Moreover, these peptides selectively decrease the viability of human cancer cells overexpressing EGFR, with mid-micromolar IC₅₀ values—only one order of magnitude higher than the FDA-approved inhibitor erlotinib. Although weak affinities are sometimes considered a handicap, they are not an exception for PPI inhibition and provide the opportunity to selectively target biomarkers (like EGFR) that are overexpressed in cancer cells but also present in most healthy tissues, thus reducing toxicity. Indeed, as a result of the harmful effects of chemotherapy, there is renewed interest in the use of weak binders, especially as part of a multivalent strategy.[23] All in all, the efficacy of these peptides, added to their exceptional stability in biological media, support their use as complementary compounds in multivale drug therapy.

**Experimental Section**

**Materials:** All amino acids, resins, solvents and reagents were purchased from Bachem AG (Bubendorf, Switzerland), Iris Biotech (Marktredwitz, Germany), Scharlab (Barcelona, Spain), Sigma-Aldrich (Milwaukee, WI), GL Biochem Shangai Ltd. (Shanghai, China), and Fluka Chemika (Buchs, Switzerland). The human carcinoma (A-431) and breast cancer (MCF7) cell lines were acquired from American Type Culture Collection (Manassas, USA). The fibroblast cell line (J2P) was kindly provided by L. Rinaldi (IRB Barcelona). The medium used for cell culture was Dubecco’s Modified Eagle’s Medium (DMEM)-high-glucose, obtained from Sigma-Aldrich. General material (sterile and not sterile) was acquired from Scharlab (Barcelona, Spain) and Deltalab.

**Molecular dynamics and ΔG calculations:** Coordinates for EGFR and cp28 (residues 7 to 34) were extracted from the Protein Data Bank (PDB ID: 1ivo). For the bicyclic analogues, the peptide structure was manually modified and the local clashes were relieved by performing a short minimization. The Amber Parm99SB force field was used; the Leap module of the AMBER package was used to immerse the EGFR-peptide complex in a pre-equilibrated octahedral box of TIP3P water molecules. Chlorine or sodium ions were added to obtain an electrostatically neutral system. The initial complex structure was first subjected to a minimization protocol consisting of 1000 steps of steepest descent method followed by 500 steps of conjugate gradient method. The optimized structure was gradually heated to 300 K in 200 ps. The final structure was chosen as the initial structure for all the 16 replicas in REMD simulations. Temperatures were set in a range from 300 to 600 K with an exchange probability of 30%.[27] Generalized Born model with an effective salt concentration of 0.2 M was deployed to mimic the solvation effect. Nonpolar solvation term was approximately represented by surface area term.[28] Integral time step was set to 1 fs. Temperature was regulated using Berendsen thermostat with a coupling time constant of 1 ps. SHAKE algorithm was used to constrain all the covalent bonds involving hydrogen atoms. Swaps were attempted every 2 ps and MD simulations were extended to 200 ns for each replica.

**Solid-phase peptide synthesis:** All bicyclic peptides were synthesized on Dawson Dbz AM resin with a substitution of 0.4-0.5 mmol/g. The first amino acid was manually coupled. The N-Fmoc-protected amino acid (3 equiv) was preactivated with HATU (3 equiv) and DIPEA (3 equiv) in the minimal volume of DMF and was added to the resin. The mixture was allowed to react in an orbital shaker with stirring for 30 min. Then, peptide chains were elongated on a CEM Liberty Blue microwave peptide synthesizer. Fmoc deprotection was carried out using 10% (w/v) piperazine and 0.1 M OxymaPure in a 9:1 mixture of NMP and EtOH. The N-Fmoc-protected amino acids (5 equiv, 0.2 M in DMF) were added with OxymaPure (5 equiv, 1 M in DMF) and DIC (5 equiv, 0.5 M in DMF) to the resin. The mixtures were stirred for 3 min at 90°C, except for cysteines, histidines and arginines, which were coupled at 50°C for 10 min. The N-terminal residue was introduced using a Boc-amino acid. After chain elongation, the resin was washed extensively with DMF and of 1,2-dichloroethane. 4-nitrophenyl chloroformate was dissolved in the least amount of 1,2-dichloroethane, added to the resin and left to gently agitate for 1 h at room temperature. The resin was washed with DCM and DMF, and reacted for 30 min with a 0.5 M solution of DiPEA in anhydrous DMF. Finally, peptides were cleaved with concomitant removal of the side-chain protecting groups, using TFA, H₂O and TIS (92.5:5:2.5).

**Native chemical ligation:** The ligation buffer (6M guanidinium hydrochloride, 200 mM sodium phosphate, 20 mM TCEP, 100 mM 4-mercaptophenol, pH 7) was freshly prepared and bubbled with nitrogen. The peptide was dissolved at a 2-3 mM concentration and the solution was stirred at rt for 4 h. Then, the reaction was acidified, extracted with TBME (2 x 50 mL) and loaded on a PoraPak™ C18 Cartridge for desalting. The guanidinium salts were washed with buffer, while the peptide was eluted at the end in H₂O/ACN (1:1) and freeze-dried. To yield the final bicyclic peptides, intramolecular disulfide bonds were formed under highly diluted conditions (20-40 µM), by stirring an aqueous solution of the peptides (pH 8) under air oxygen for 24 h. Peptides were purified by semi-preparative HPLC on a Waters 2700 sample manager.
equipped with a Waters 2487 dual-wavelength absorbance detector, a Waters 600 controller, a Waters fraction collector and Masslynx software by using a Sunfire C18 column (150 x 10 mm x 3.5 µm, 100 Å, Waters), flow rate 6.6 mL/min, solvent A=0.1% trifluoroacetic acid in water; solvent B=0.1% trifluoroacetic acid in acetonitrile.

Circular dichroism: Circular dichroism spectra were recorded using a Jasco 810 UV-Vis spectropolarimeter, equipped with a CDF 426S/426L peltier. Peptide samples were dissolved in NMR buffer (20 mM sodium phosphate, 50 mM NaCl, 0.1% NaN₃, pH 6.8), and spectra were recorded at concentrations of 20 and 100 µM. A third reading was done after adding 10% trifluoroethanol (TFE) to the sample. The following parameters were used: sensitivity (standard, 100 mdeg), start (260 nm), end (190 nm), data pitch (0.5 nm), scanning mode (continuous), accumulating (3). A blank spectrum of the buffer was subtracted from all recordings, and the resulting spectra were smoothed using the Savitzky-Golay method (convolution width = 21) and taken to zero at the far-UV region (A = 260 nm).

Recombinant human EGF expression: The nucleotide sequence encoding hEGF was cloned into a pOPIN vector[29] as a SUMO fusion protein, which also included a N-terminal His-tag and a SUMO protease cleavage site. This construct produced higher yields (2 mg/L) than the previously reported thioferredoxin fusion system.[30] Cell transformation, cell growth, protein expression and purification were performed by following the standard protocol.[31] Uniformly labelled 15N-EGF was obtained by growing the E. coli cells in M9 minimal medium containing 15NH₄Cl as the sole nitrogen source.

NMR spectroscopy: 15N-labelled EGF was prepared in NMR buffer (20 mM sodium phosphate, 50 mM NaCl, 0.1% NaN₃, pH 6.8) supplemented with 10% D2O for a final sample volume of 180 µL, and fixed concentrations of protein (50 µM) and peptide (1 mM). 15N-1H HSQC spectra were recorded at 25°C using a Bruker 600 MHz NMR spectrometer equipped with a cryogenic probe. Spectra were acquired with 128x2048 complex points with a total of 8 transients per increment. The x-carrier frequency was referenced to internal DSS-D6. Data processing was performed using TopSpin3.0. The quality of the protein sample was first evaluated by performing a 13N-1H HSQC spectrum of EGF, and comparing the results with the reported data.[32] Combined NH chemical shift perturbations (CSP = [(δN/5) + (δH)2]1/2) were plotted for each protein residue; only CSs > (mean + 2SD) were considered significant.

For the structural analysis of ep22G, the peptide was dissolved in NMR buffer supplemented with 10% trifluoroethanol and 10% D2O for a final sample volume of 350 µL. Residue specific assignments were obtained from 2D total correlated spectroscopy (TOCSY),[31] while 2D nuclear Overhauser effect spectroscopy (NOESY)[32] permitted sequence specific assignments. 13C resonances were assigned from 2D 13C-1H HSQC spectra. The TOCSY and NOESY mixing times were 70 and 200 ms, respectively. The D1 relaxation delay was 1.5 s.

Simulated annealing protocol: main NOEs among peptide residues (Tyr16-Glu13, Arg2-Leu19, Arg21-Leu17, Glu20-Tyr16, His15-Gln8) were deployed through a set of 500 independent simulated annealing MD simulations where NMR restraints were applied by using the Sander module in conjunction with the AMBER99 force field. The most populated conformation in the REMD was used as the starting structure in each annealing simulation. The generalized Born implicit solvent model with an effective salt concentration of 0.2 M was applied. A total of 100 ps of MD were run with a 1 fs time step. A long-range cutoff of 10.0 Å was used for non-bonded interactions. The upper and lower distance CO-Ca atom restraint limits were set to 10.0 and 4.5 Å, respectively, to sample a broader conformational according to NOE data, with a force constant of 10.0 kcal mol⁻¹ Å⁻². Chirality restraints (50.0 kcal mol⁻¹ Å⁻²) were used to avoid chirality flipping at high temperatures in the simulated annealing protocol. Temperature was increased to 700 K in the former 20 ps of simulation to be then gradually cooled down to 300 K in the remaining simulation time. Different seed numbers were used in each simulation. For each independent simulation, the lowest AMBER energy structure conformation was selected as representative of the conformational space explored. The 500 simulated annealing conformations were finally clustered using P traj module to obtain the low-energy averaged structure compatible with NMR data.

SAW biosensor: Affinity analyses were performed with a SamS Blue biosensor (SAW Instruments, Bonn, Germany). The chip surface was functionalized with EGF following the reported protocol,[34] and peptides were injected at a range of micromolar concentrations selected to cover the binding equilibrium constant. Sensorgrams were analysed using the Origin Pro 7.5 and FitMaster software. Briefly, the signal at the association equilibrium were represented for each ligand concentration, and non-linear fitting of the data was performed to obtain the Kd values of the interaction.

AlphaScreen assays: 10 µL of EGF (His-tag functionalized, 3 mM final concentration in PBS buffer, 0.1% BSA, 0.1% Tween-20, pH 7.4) and 10 µL of peptide (at different concentrations) were added to 96-well Opisplates (PerkinElmer) and incubated at rt for 15 min. EGFR-Fc (3 mM final concentration in the same buffer) was then added and incubated for 30 min. Next, 10 µL of Protein A 3835 Donor beads (20 µg/mL final concentration, 60 min incubation) and Anti-His acceptor beads (10 µL 20 µg/mL final concentration) were added. After 30 min incubation, the fluorescence emission was read at 615 nm in an EnVision Multitabel Reader. Dose-response curves were obtained with the GraphPad Prism 6.03 software, using a non-linear fit and variable slope, from which IC50 values were calculated.

Cell viability assay (sodium 2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-5-[[(phenylamino)carbonyl]-2H-tetrazolium [XTT]): A-431 human cancer cells (2.5 x 10⁴ EGFR/cell)[35] were seeded in a 96-well Corning microplates (2 x 10³ cells/well). After 24 h of incubation at 37°C, cells were treated with the peptides dissolved in Dulbecco’s Modified Eagle’s Medium (high glucose, 10% FBS containing EGF and growth factors) for 72 h (100 µL, concentration range 0.5-150 µM). Then, the peptide solution was removed and substituted with fresh medium. 50 µL of activated-XTT solution (0.1 mL of the activation reagent in 5.0 mL of the XTT reagent, mixed 15 min before use) was added to each well and incubate for 4 h at 37°C. The absorbance at 450 nm was measured in a PowerWave X reader (Bio-Tek, INC), where the positive control was cells treated with DMSO and the negative control was untreated cells. Each experiment was performed in triplicate, and the IC50 was calculated by linear regression fitting (concentration vs. response curve) using GraphPad Prism. The same conditions were used for studying MCF7 and J2P fibroblast viability.

Stability in human serum: Peptides at a final concentration of 150 µM were dissolved in Hank’s balanced salt solution and incubated at 37°C in the presence of 90% human serum during 24 h. Then, 50 µL aliquots were extracted at different incubation times and treated with 200 µL of MeOH to precipitate serum proteins. After 30 min centrifugation at 4°C, the supernatant was filtered and analysed by HPLC to calculate the percentage of intact peptide in the sample. The linear peptide ACP was used as positive control.

Metabolic stability in human microsomes: Peptides were dissolved in water (713 µL, 1 µM final concentration) and potassium phosphate (200 µL, 0.5 M, pH 7.4). NADPH regenerating system solution A (50 µL) and...
NADPH regenerating system solution B (10 µL) were added. This mixture was warmed to 37°C for 5 minutes. Then, human liver microsomes (25 µL, 0.5 mg) were added. The mixture was immediately vortexed and incubated at 37°C with orbital agitation (100 rpm). 100 µL aliquots were taken at 0, 5, 10, 20, 30, 40, 50 and 60 minutes, and treated with 100 µL of ACN to precipitate serum proteins. The extracted samples were centrifuged and analysed by UPLC-MS.

Acknowledgements

The authors thank the Barcelona Supercomputing Center - Red Española de Supercomputación (BSC-RES) for access to computational resources and technical support. We also thank the Mass Spectrometry Core Facility (IRB Barcelona) and the NMR facility at the Scientific and Technological Centre of the University of Barcelona (CCiT UB). This study was funded by the Ministry of Economy and Competitiveness (MINECO) and the European Fund for Regional Development FEDER (BIO 2016-75327-R, RTC-2015-4336-1, PCIN-2015-052); the Generalitat de Catalunya (XRB and 2014SGR-521). IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from MINECO (Government of Spain).

Keywords: EGF • bicyclic peptides • protein-protein interactions • rational design • metabolic stability

Starting from a key binding domain of EGFR, we have used computational and biophysical methods to design a series of bicyclic constrained peptides—highly resistant to metabolic degradation—that disrupt the EGF-EGFR interaction and block EGFR-dependent cancer cell proliferation.