

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

Salvador Guardiola,<sup>[a]</sup> Jesús Seco,<sup>[a]</sup> Monica Varese,<sup>[a]</sup> Mireia Díaz-Lobo,<sup>[a]</sup> Jesús García,<sup>[a]</sup> Meritxell Teixidó,<sup>[a]</sup> Laura Nevola,<sup>\*[b]</sup> and Ernest Giralt<sup>\*[a,c]</sup>

**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,<sup>[1]</sup> thereby complicating the use of conventional small-molecule drugs as inhibitors.<sup>[2]</sup> As an alternative, peptides have a higher degree of structural flexibility, which allows them to better adapt to irregular targets,<sup>[3]</sup> and they can be designed to display a variety of topologies that closely mimic the structural features found in PPIs.<sup>[4]</sup> However, peptides tend to have low serum stability and poor ability to cross physiological barriers.<sup>[5]</sup> 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.<sup>[6]</sup> 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.<sup>[7][8]</sup> To perform its function, EGF binds with high affinity ( $K_D = 1.9$  nM)<sup>[9]</sup> to its membrane receptor (EGFR),<sup>[10]</sup> triggering intracellular events that lead to uncontrolled cell growth, tumour invasion and metastasis.<sup>[8]</sup>

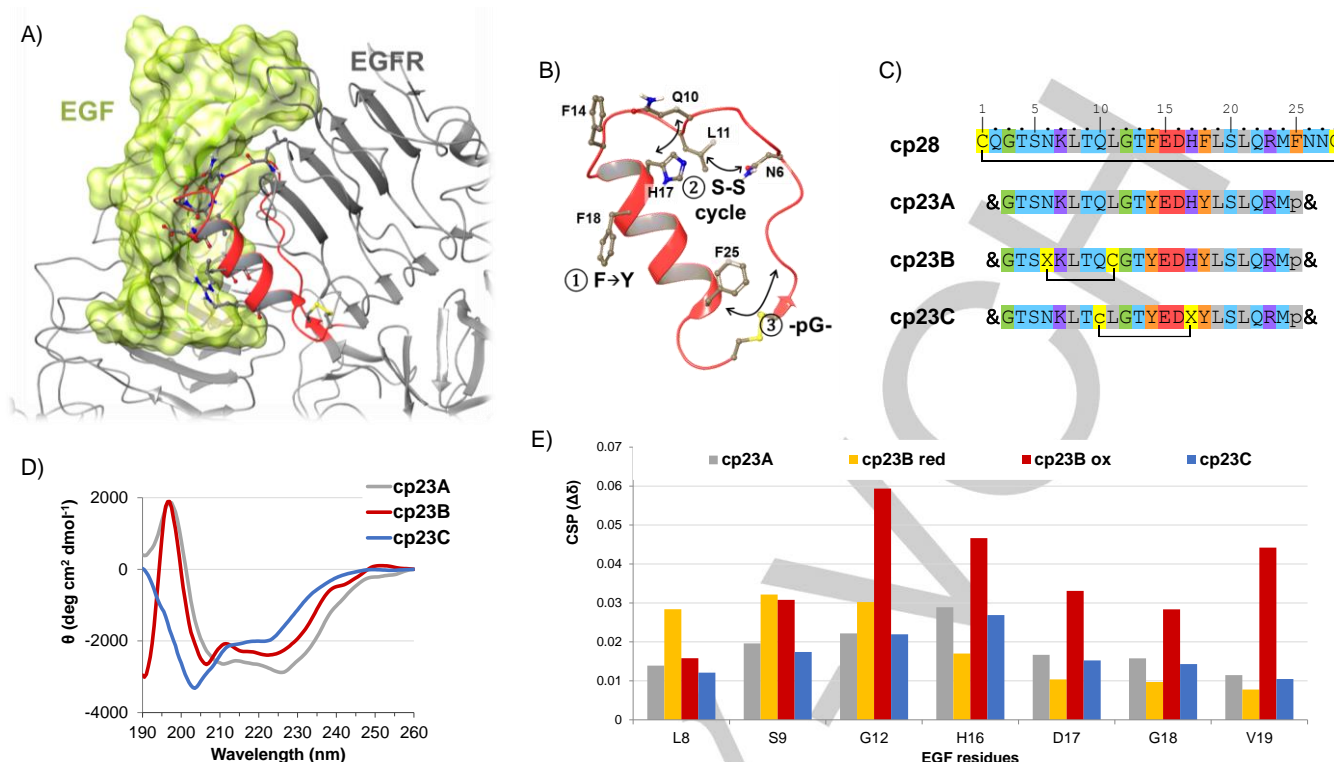
Therefore, the EGF-EGFR pathway has become a main focus for selective chemotherapeutic intervention and, as a result, two classes of EGFR inhibitors have been clinically approved, namely monoclonal antibodies<sup>[11]</sup> (cetuximab, panitumumab), which target the extracellular domain of EGFR, and small-molecule kinase inhibitors<sup>[12]</sup> (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.<sup>[13]</sup> 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.<sup>[8]</sup> 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.<sup>[14]</sup> Although similar growth factors (such as VEGF) have been successfully addressed using this approach,<sup>[15]</sup> no current drugs are able to specifically target EGF.<sup>[7]</sup> However, we recently showed that peptides represent a significant step in this direction.<sup>[16]</sup> 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.

[a] Dr. S. Guardiola, Dr. M. Díaz-Lobo, Dr. J. García, Dr. J. Seco, Dr. M. Teixidó, Prof. E. Giralt  
Institute for Research in Biomedicine (IRB Barcelona)  
The Barcelona Institute of Science and Technology  
Baldri Reixac, 10, 08028 Barcelona, Spain  
Email: ernest.giralt@irbbarcelona.org

[b] Dr. L. Nevola  
IDP Discovery Pharma SL, Barcelona Science Park  
Baldri Reixac 4, 08028, Barcelona, Spain  
Email: lnevola@idp-pharma.com

[c] Prof. E. Giralt  
Department of Organic Chemistry, University of Barcelona, Spain



**Figure 1.** A) Detail of the EGF-EGFR binding interface (EGF on green surface; EGFR in grey cartoon, the relevant residues 7–34 are highlighted in red; adapted from PDB 1ivo). B) Main modifications performed on the structure of the EGFR-derived peptide **cp28**. C) Sequence and cyclic constraints of peptide analogues **cp23A–C** (X = L-homocysteine). Residues are coloured according to their physicochemical nature: blue, polar; grey, hydrophobic; orange, aromatic; purple, basic; red, acidic; yellow, SH-residues; green, Gly. D) CD spectra of **cp23A–C** in phosphate buffer (pH 6.8, 10% TFE). E) Average CSPs induced by peptides **cp23A**, **cp23B** (reduced and oxidised) and **cp23C** on various EGF residues.

## Results and Discussion

### Design and biophysical evaluation of peptide ligands

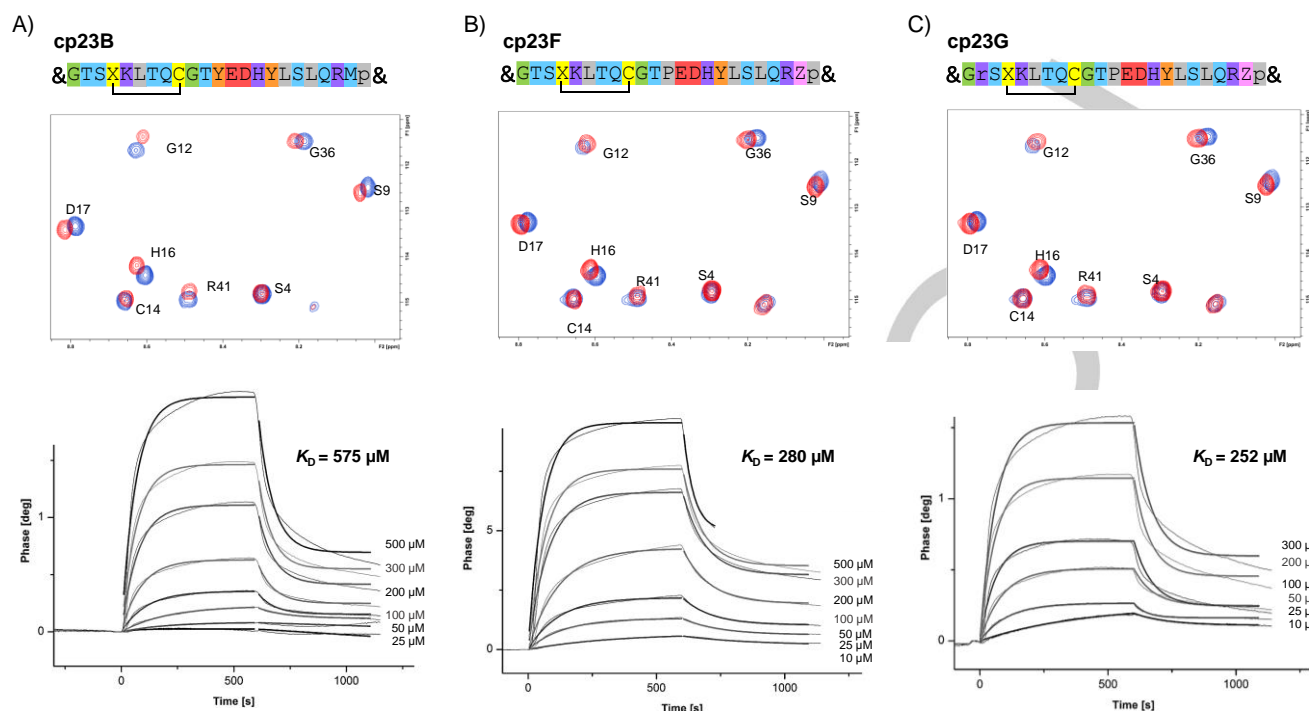
The binding of EGF to EGFR occurs through a large (1440 Å<sup>2</sup>) and complex interface, in which various domains of the receptor are involved. Nevertheless, a number of hotspots are found on residues 7 to 34 of the receptor (Figure 1A). In particular, residues 16–18 of EGFR and residues 31–33 of EGF form a short parallel  $\beta$ -sheet, stabilised by backbone hydrogen bonds. Also, the  $\alpha$ -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 ( $K_D = 286 \mu\text{M}$ ).<sup>[16]</sup> 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.<sup>[17]</sup> 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-acylurea 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



**Figure 2.** Amino acid sequences of peptides **cp23B** (A), **cp23F** (B) and **cp23G** (C) (X = L-homocysteine, Z = L-norvaline). For each, a representative region of the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of EGF is shown, in the absence (blue) and presence (red) of ligand (1 mM). Below, the SAW sensorgram at a range of concentrations and the calculated  $K_D$  values for the interaction with EGF are shown for each peptide.

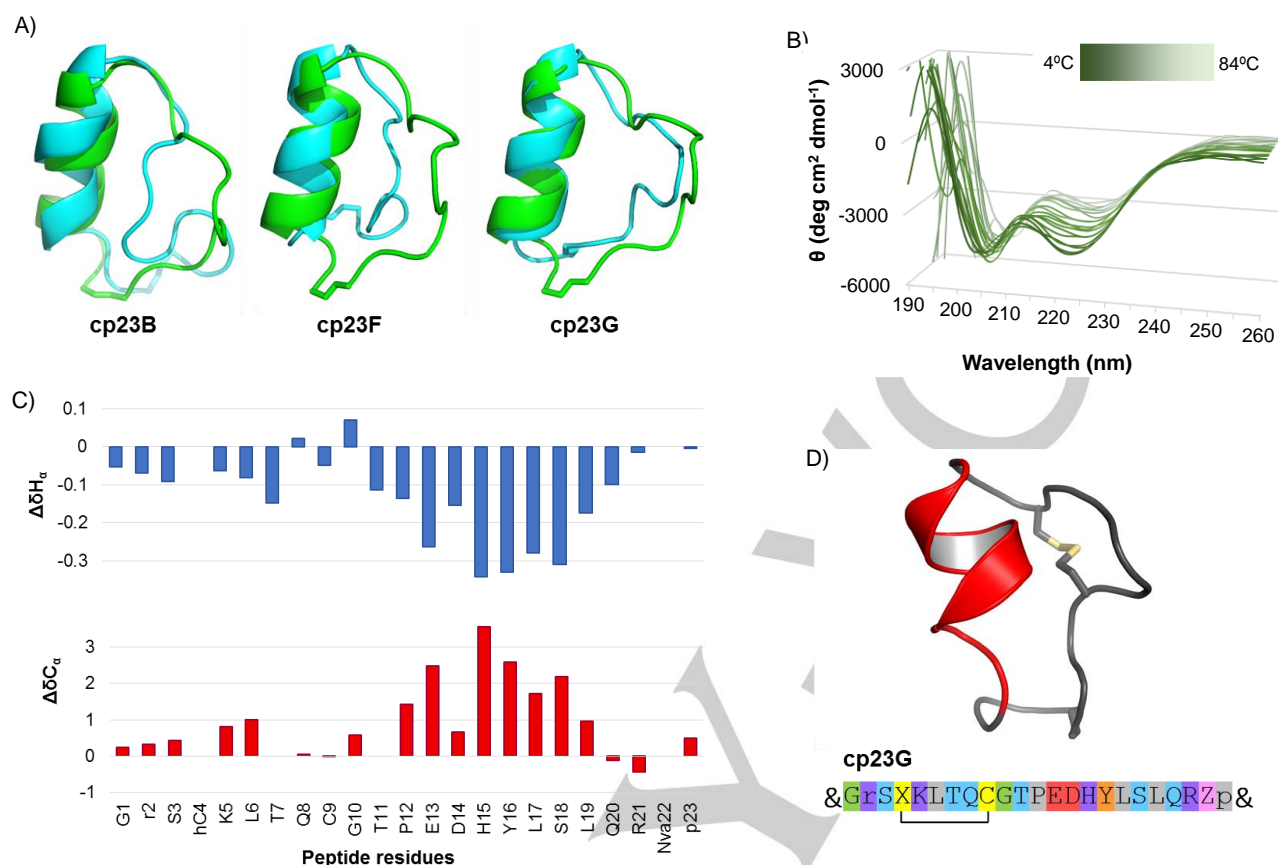
nm and a second minimum at ca. 222 nm, thereby indicating a prominent content of  $\alpha$ -helix, a motif that is crucial for the PPI. In contrast, the alternative bicyclic topology that we tested in **cp23C** had a strongly reduced helical content.

Several biophysical methods were used to assess the binding of the peptides to EGF. First, the effects of the interaction were evaluated in the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of a uniformly labelled  $^{15}\text{N}$ -EGF sample. Comparison of the average chemical shift perturbations (CSPs) induced by the three analogues revealed larger perturbations for **cp23B** than for **cp23A** and **cp23C** (Figure 1E). Notably, the oxidised **cp23B** caused greater changes in the protein signals than its reduced form. The most significant perturbations occurred on residues on the A- and B-loops of EGF (L8, S9, G12, H16-V19), which are involved in the EGF-EGFR interaction (Figure 2A).

To evaluate the strength of the peptide-EGF interactions, we used a surface acoustic wave (SAW) biosensor assay.<sup>[18]</sup> EGF was covalently immobilised on the gold surface of the chip and 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 ( $K_D = 575 \mu\text{M}$ , 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 *in silico* 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 12<sup>th</sup> position. This observation confirms the stabilizing effect of Pro as N-cap for the helix,<sup>[19]</sup> which translated



**Figure 3.** A) Replica exchange MD experiments show the most populated conformations for peptides **cp23B**, **cp23F** and **cp23G** (in blue) superimposed with the EGF bioactive conformation (in green). RMSD values calculated for the peptide backbone. B) CD spectra of **cp23G** in phosphate buffer (pH 6.8, 10% TFE) at temperatures ranging from 4 to 84°C. C) Chemical shift deviations of **cp23G**  $^1H_{\alpha}$  (top) and  $^{13}C_{\alpha}$  (bottom) from random coil values. D) NMR-derived low-energy averaged structure of **cp23G**. The  $\alpha$ -helical motif and the disulphide bond are shown in red and yellow, respectively.

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 sensorgrams 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 ( $-8.1$  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  $^{15}N$ - $^1H$  HSQC spectra of the protein confirmed their binding to the EGF-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.



Analogue	Sequence <sup>[a]</sup>	MM/GBSA (kcal/mol)	MM/PBSA (kcal/mol)	K <sub>D</sub> (μM) <sup>[b]</sup>
<b>cp23A</b>	&GTSNKLTLGTYEDH YLSLRMp&	-	-	773
<b>cp23B</b>	&GTSXKLTCQGYEDH YLSLRMp&	-31.4	6.8	575
<b>cp23C</b>	&GTSNKLTLGTYEDXY LSLRMp&	-	-	1860
<b>cp23D</b>	&GTSXKLTCQGTPADD AASAARMp&	-46.8	-12.7	1420
<b>cp23E</b>	&GTSXAETCQGYEDH YLSLRZp&	-46.6	-9.2	>2000
<b>cp23F</b>	&GTSXKLTCQGYEDH YLSLRZp&	-42.1	-2.9	280
<b>cp23G</b>	&GrSXKLTCQGYEDHY LSLRZp&	-41.4	-8.1	252
<b>cp23H</b>	&GTSXKLTCQMTPEDH YLSLRZp&	-48.1	-6.0	115

<sup>[a]</sup>X = L-homocysteine, Z = L-norvaline. Residues forming disulphide bonds are underlined. <sup>[b]</sup>Affinity values obtained by SAW.

### 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

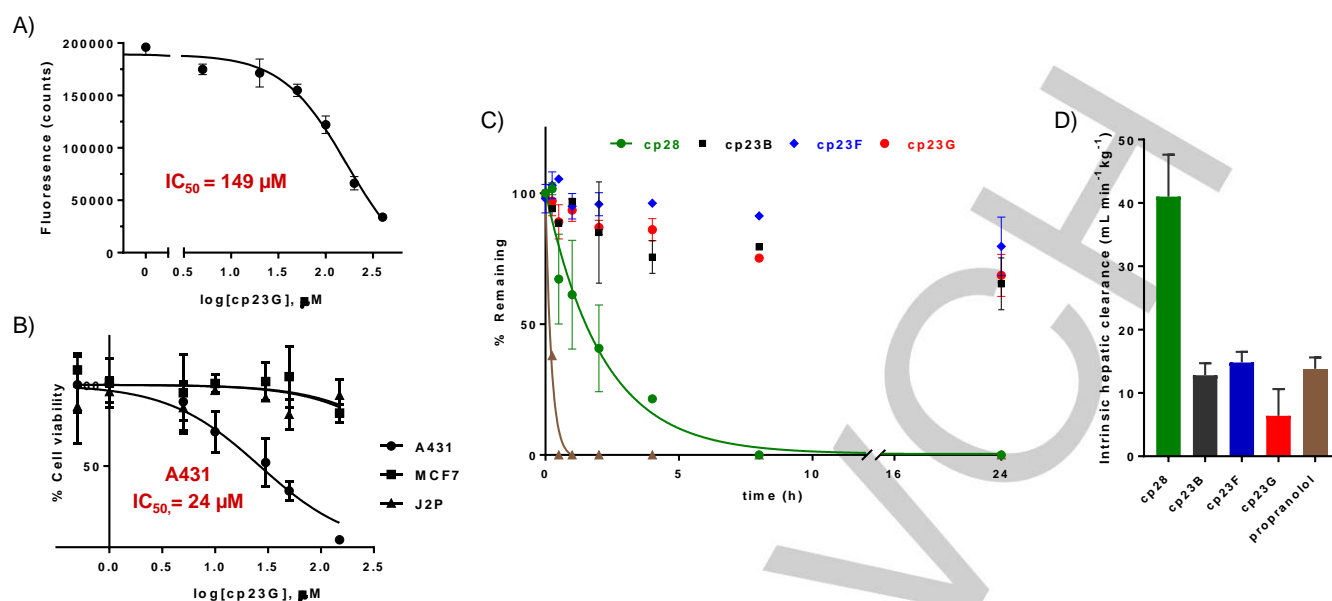
of the peptide in a buffered aqueous solution at pH 6.8. The <sup>1</sup>HN backbone chemical shifts exhibited a relatively large dispersion, which is typical of a more rigid and defined structure (Figure S7). The secondary chemical shifts of both <sup>1</sup>H<sub>α</sub> and <sup>13</sup>C<sub>α</sub> showed consecutive deviations from random coil values,<sup>[20]</sup> negative for H<sub>α</sub> and positive for C<sub>α</sub>, 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 EGF 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 EGF-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 EGF-EGFR interaction was observed (Figure 4A and S8). The **cp23G** peptide was the most potent inhibitor of the series (IC<sub>50</sub> = 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 (IC<sub>50</sub> = 2.9 μM). Under these conditions, **cp23G** had the greatest inhibitory effect (IC<sub>50</sub> = 24 μM, Figure 4B), in the same range as **cp28** (IC<sub>50</sub> = 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



**Figure 4.** A) EGF-EGFR inhibition (AlphaScreen assay) at a range of **cp23G** concentrations. B) Viability of A-431 cells (overexpressing EGFR), MCF-7 cells (normal EGFR levels) and healthy J2P fibroblasts, incubated at a range of **cp23G** concentrations. C) Stability of the peptides in human serum at 37°C (the linear peptide ACP was used as control, in brown triangles). D) Metabolic stability in human hepatic microsomes (propranolol was used as control). In all cases, experiments were performed in triplicate, and error bars represent the SD.

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).<sup>[21]</sup> 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 37°C 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.<sup>[22]</sup> 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 has been successfully targeted only using antibodies, such as cetuximab, which currently provide 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  $IC_{50}$  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.<sup>[23]</sup> All in all, the efficacy of these peptides, added to their exceptional stability in biological media, support their use as complementary compounds in multidrug cancer 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 Shanghai 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, VA). The fibroblast cell line (J2P) was kindly provided by L. Rinaldi (IRB Barcelona). The medium used for cell culture was Dulbecco's Modified Eagle's Medium (DMEM)-high glucose, obtained from Sigma-Aldrich. General material (sterile and not sterile) was acquired from Sharlab (Barcelona, Spain) and Deltalab.

**Molecular dynamics and  $\Delta G$  calculations:** Coordinates for EGF 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 EGF-peptide complex in a pre-equilibrated octahedral box of TIP3P water molecules. Chloride 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 decent method followed by 500 steps of conjugate gradient method. Thermalization of the system was performed in the NVT ensemble during 200 ps, using a time step of 1 fs and increasing the temperature from 100 to 298 K, where a force constant of  $5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  was applied to protein backbone atoms. Prior to the production run, a short MD simulation (100 ps) in the NPT ensemble was done in order to equilibrate the system density to 1 atm and 298 K. The simulations were carried out at constant pressure (1 atm) and temperature (298 K) for 20 ns, except for some peptides that were extended to 50 ns. Low harmonic constraints ( $2 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ) were used to reduce the protein mobility. Constant temperature was achieved using the Langevin thermostat with a collision frequency of  $2 \text{ ps}^{-1}$ . The SHAKE algorithm was used to keep bonds involving hydrogen atoms at their equilibrium length. The particle-mesh Ewald summation method was used to deal with long range electrostatic interactions and a cut-off of 10 Å was applied for non-bonded interactions. The simulations were conducted with the PMEMD module of the AMBER 14 program. The trajectories were analysed using frames collected every 2 ps during the production runs.

From the collected MD trajectories, the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) and Generalized Born (MM/GBSA) methods were used to compute the free energy of binding ( $\Delta G$ ) between

the peptides and EGF. A numerical solution of the Poisson-Boltzmann equation,<sup>[24]</sup> implemented in AMBER, was used to calculate the polar part of the solvation free energy, using a low ( $\epsilon = 1$ ) and high-dielectric medium ( $\epsilon = 80$ ) for the solute and solvent, respectively. The solute-solvent interface was defined by the solvent-excluded molecular surface, where a solvent probe radius of 1.4 Å was employed. The non-polar contribution to the solvation was estimated from the solvent-accessible surface area<sup>[25]</sup> using the default scaling factor in Amber ( $\gamma = 0.0072 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ). The less computationally demanding MM/GBSA approach was also computed using analogous parameters.<sup>[26]</sup> The reported values of MM/GBSA and MM/PBSA were computed for a subset of 500 equally spaced snapshots along the converged part of each simulation, when both ligand and protein reached a stable interaction state.

**Replica exchange MD:** To assess the structure of the peptides in solution, replica exchange molecular dynamics (REMD) was performed. The initial structure (identical to the one used in the MD simulation described above) was first subjected to minimization protocol consisting of 1000 steps of steep decent 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%.<sup>[27]</sup> 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.<sup>[28]</sup> 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, H<sub>2</sub>O 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 H<sub>2</sub>O/ACN (1:1) and freeze-dried. To yield the final bicyclic peptides, intramolecular disulfide bonds were formed under highly diluted conditions (20–40  $\mu\text{M}$ ), 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  $\mu$ m, 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% NaN<sub>3</sub>, pH 6.8), and spectra were recorded at concentrations of 20 and 100  $\mu$ M. 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), scanning speed (200 nm/min), response (1 s), band width (1.0 nm), and accumulation (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 ( $\lambda$  = 260 nm).

**Recombinant human EGF expression:** The nucleotide sequence encoding hEGF was cloned into a pOPIN vector<sup>[29]</sup> 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 thioredoxin fusion system.<sup>[16]</sup> Cell transformation, cell growth, protein expression and purification were performed by following the standard protocol.<sup>[16]</sup> Uniformly labelled <sup>15</sup>N-EGF was obtained by growing the *E. coli* cells in M9 minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl as the sole nitrogen source.

**NMR spectroscopy:** <sup>15</sup>N-labelled EGF was prepared in NMR buffer (20 mM sodium phosphate, 50 mM NaCl, 0.1% NaN<sub>3</sub>, pH 6.8) supplemented with 10% D<sub>2</sub>O for a final sample volume of 180  $\mu$ L, and fixed concentrations of protein (50  $\mu$ M) and peptide (1 mM). <sup>15</sup>N-<sup>1</sup>H 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-d<sub>6</sub>. Data processing was performed using TopSpin3.0. The quality of the protein sample was first evaluated by performing a <sup>15</sup>N-<sup>1</sup>H HSQC spectrum of EGF, and comparing the results with the reported data.<sup>[30]</sup> Combined NH chemical shift perturbations ( $\text{CSP} = [(\delta\text{N}/5)^2 + (\delta\text{H})^2]^{1/2}$ ) were plotted for each protein residue; only CSPs > (mean + SD) were considered significant.

For the structural analysis of **cp23G**, the peptide was dissolved in NMR buffer supplemented with 10% trifluoroethanol and 10% D<sub>2</sub>O for a final sample volume of 350  $\mu$ L. Residue specific assignments were obtained from 2D total correlated spectroscopy (TOCSY),<sup>[31]</sup> while 2D nuclear Overhauser effect spectroscopy (NOESY)<sup>[32]</sup> permitted sequence specific assignments. <sup>13</sup>C resonances were assigned from 2D <sup>13</sup>C-<sup>1</sup>H 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, Gln20-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 0.2M 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 Ca-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<sup>-1</sup> Å<sup>-2</sup>. Chirality restraints (50.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>) 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 300K 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 Ptraj module to obtain the low-energy averaged structure compatible with NMR data.

**SAW biosensor:** Affinity analyses were performed with a Sam5 Blue biosensor (SAW Instruments, Bonn, Germany). The chip surface was functionalized with EGF following the reported protocol,<sup>[16]</sup> 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 *K<sub>D</sub>* values of the interaction.

**AlphaScreen assays:** 10  $\mu$ L of EGF (His-tag functionalized, 3 nM final concentration in PBS buffer, 0.1% BSA, 0.1% Tween-20, pH 7.4) and 10  $\mu$ L of peptide (at different concentrations) were added to 96-well Optiplates (PerkinElmer) and incubated at rt for 15 min. EGFR-Fc (3 nM final concentration in the same buffer) was then added and incubated for 30 min. Next, 10  $\mu$ L of Protein A 3835 Donor beads (20  $\mu$ g/mL final concentration, 60 min incubation) and anti-His acceptor beads (10  $\mu$ L, 20  $\mu$ g/mL final concentration) were added. After 30 min incubation, the fluorescence emission was recorded at 615 nm in an EnVision Multilabel Reader. Dose-response curves were obtained with the GraphPad Prism 6.03 software, using a non-linear fit and variable slope, from which IC<sub>50</sub> values were calculated.

**Cell viability assay (sodium 2,3-bis(2-methoxy-4-nitro-5-sulphophenyl)-5-[(phenylamino)-carbonyl]-2H-tetrazolium [XTT]):** A-431 human cancer cells (2.5 x 10<sup>6</sup> EGFR/cell)<sup>[33]</sup> were seeded in a 96-well Corning microplates (2 x 10<sup>3</sup> 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  $\mu$ L, concentration range 0.5-150  $\mu$ M). Then, the peptide solution was removed and substituted with fresh medium. 50  $\mu$ L 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 IC<sub>50</sub> 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  $\mu$ M 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  $\mu$ L aliquots were extracted at different incubation times and treated with 200  $\mu$ L 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  $\mu$ L, 1  $\mu$ M final concentration) and potassium phosphate (200  $\mu$ L, 0.5 M, pH 7.4), NADPH regenerating system solution A (50  $\mu$ L) and



NADPH regenerating system solution B (10  $\mu$ L) were added. This mixture was warmed to 37°C for 5 minutes. Then, human liver microsomes (25  $\mu$ L, 0.5 mg) were added. The mixture was immediately vortexed and incubated at 37°C with orbital agitation (100 rpm). 100  $\mu$ L aliquots were taken at 0, 5, 10, 20, 30, 40, 50 and 60 minutes, and treated with 100  $\mu$ L 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

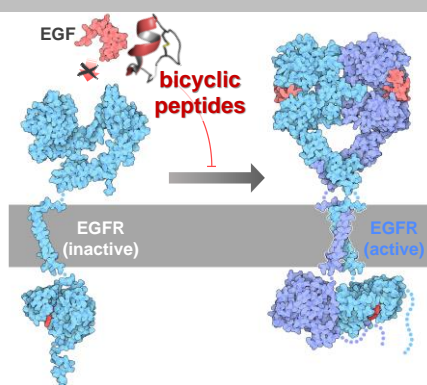
- [1] J. A. Wells, C. L. McClendon, *Nature* **2007**, *450*, 1001–9.
- [2] X. Morelli, R. Bourgeas, P. Roche, *Curr. Opin. Chem. Biol.* **2011**, *15*, 475–81.
- [3] N. London, B. Ravet, D. Movshovitz-Attias, O. Schueler-Furman, *Proteins Struct. Funct. Bioinforma.* **2010**, *78*, 3140–3149.
- [4] L. Nevola, E. Giralt, *Chem. Commun. (Camb.)* **2015**, *51*, 3302–15.
- [5] D. E. Scott, A. R. Bayly, C. Abell, J. Skidmore, *Nat. Rev. Drug Discov.* **2016**, *15*, 533–550.
- [6] M. A. Lemmon, J. Schlessinger, *Cell* **2010**, *141*, 1117–34.
- [7] C. Yewale, D. Baradia, I. Vhora, S. Patil, A. Misra, *Biomaterials* **2013**, *34*, 8690–8707.
- [8] C. R. Chong, P. A. Jänne, *Nat. Med.* **2013**, *19*, 1389–1400.
- [9] J. T. Jones, R. W. Akita, M. X. Sliwowski, *FEBS Lett.* **1999**, *447*, 227–231.
- [10] A. C. King, P. Cuatrecasas, *J. Biol. Chem.* **1982**, *257*, 3053–60.
- [11] B. You, E. X. Chen, *J. Clin. Pharmacol.* **2012**, *52*, 128–155.
- [12] A. Russo, T. Franchina, G. R. R. Ricciardi, A. Picone, G. Ferraro, M. Zanghi, G. Toscano, A. Giordano, V. Adamo, *Oncotarget* **2015**, *6*, 26814–25.
- [13] B. Zhao, L. Wang, H. Qiu, M. Zhang, L. Sun, P. Peng, Q. Yu, X. Yuan, B. Zhao, L. Wang, et al., *Oncotarget* **2017**, *8*, 3980–4000.
- [14] P. C. Rodriguez, X. Popa, O. Martínez, S. Mendoza, E. Santiesteban, T. Crespo, R. M. Amador, R. Fleytas, S. C. Acosta, Y. Otero, et al., *Clin. Cancer Res.* **2016**, *22*.
- [15] J. W. Checco, S. H. Gellman, *ChemBioChem* **2017**, *18*, 291–299.
- [16] S. Guardiola, M. Díaz-Lobo, J. Seco, J. García, L. Nevola, E. Giralt, *ChemBioChem* **2016**, *17*, 702–711.
- [17] T. A. Hill, N. E. Shepherd, F. Diness, D. P. Fairlie, *Angew. Chemie Int. Ed.* **2014**, *53*, 13020–13041.
- [18] M. Díaz-Lobo, in *Encycl. Anal. Chem.*, John Wiley & Sons, Ltd, Chichester, UK, **2016**, pp. 1–14.
- [19] J. S. Richardson, D. C. Richardson, *Science* **1988**, *240*, 1648–52.
- [20] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, *5*, 67–81.
- [21] D. Mathur, S. Prakash, P. Anand, H. Kaur, P. Agrawal, A. Mehta, R. Kumar, S. Singh, G. P. S. Raghava, *Sci. Rep.* **2016**, *6*, 36617.
- [22] K. J. Coe, T. Koudriakova, *Methods Pharmacol. Toxicol.* **2014**, 87–99.
- [23] J. Huskens, *Multivalency: Concepts, Research & Applications*, Wiley-VCH Verlag, **2018**.
- [24] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, et al., *Acc. Chem. Res.* **2000**, *33*, 889–97.
- [25] D. Sitkoff, K. A. Sharp, B. Honig, *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- [26] V. Tsui, D. A. Case, *Biopolymers* **2000**, *56*, 275–291.
- [27] A. Patriksson, D. van der Spoel, *Phys. Chem. Chem. Phys.* **2008**, *10*, 2073.
- [28] J. Weiser, P. S. Shenkin, W. C. Still, *J. Comput. Chem.* **1999**, *20*, 217–230.
- [29] N. S. Berrow, D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D. I. Stuart, R. J. Owens, *Nucleic Acids Res.* **2007**, *35*, e45.
- [30] H. W. Huang, S. K. Mohan, C. Yu, *Biochem. Biophys. Res. Commun.* **2010**, *402*, 705–710.
- [31] A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *65*, 355–360.
- [32] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546–4553.
- [33] J. Filmus, M. N. Pollack, R. Cailleau, R. N. Bick, *Biochem Biophys Res Comm* **1985**, *128*, 898–905.

## Entry for the Table of Contents (Please choose one layout)

Layout 1:

## FULL PAPER

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.



Author(s), Corresponding Author(s)\*

Page No. – Page No.

Title