Overlapping synthetic peptides as a tool to map protein-protein interactions − FSH as a model system of nonadditive interactions.

Tomás A. Santa-Coloma

Laboratory of Cellular and Molecular Biology, Institute for Biomedical Research (BIOMED), School of Medical Sciences, Pontifical Catholic University of Argentina (UCA), and the National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina.

Running title: Overlapping synthetic peptides and nonadditive interactions

Corresponding author: Tomás A. Santa Coloma, Institute for Biomedical Research (BIOMED), School of Medical Sciences, Pontifical Catholic University of Buenos Aires, Alicia Moreau de Justo 1600, Buenos Aires C1107AFF, Argentina. Tel/Fax: 5411-4338-0886. E-mail: tsantacoloma@gmail.com, tomas_santacoloma@uca.edu.ar

Keywords: synthetic peptides, receptors antagonists, protein-protein interactions, agonist peptide, antagonist peptide, free-energy paradoxes.

Abstract

In earlier work, we used partially overlapped synthetic peptides as a tool to find regions of interaction between the human FSH hormone and its receptor, aiming to find possible antagonists or agonists. Years later, the FSH and FSH receptor 3D structures were reported by other laboratories. The 3D results were in close agreement with the interacting regions predicted by using synthetic peptides. These earlier studies are reviewed here, and the predicted regions of interaction compared to the FSH and FSH receptor 3D structures to illustrate the usefulness of the synthetic peptide strategy to find binding regions. Different contact regions contribute multiplicatively to the high affinity of the entire ligand; thus, peptides covering a fraction of the anchor sites and with low free energy density cannot reach the affinity of the entire molecule. The earlier use of multiple linear regression to find the relevant predictors for effective binding, and a new way to estimate $\Delta G°$ and nonadditive interactions for the synthetic peptides in solution, by using the buried surface area (BSA), will be discussed.

1. Introduction

Synthetic peptides have been used extensively to map interacting regions between proteins, including different types of ligands, receptors, and antibodies [1-4], or to identify interacting partners [5]. Also, to produce synthetic agonists or antagonists with therapeutic purposes [6]. In this review, the usefulness of using overlapping synthetic peptides as a tool to map interacting regions will be discussed, using the glycoprotein hormone FSH as a model system [7].

The glycoprotein hormones FSH, LH, hCG, TSH are essential for normal growth and reproductive function [8, 9]. FSH contains two subunits, FSH-α and FSH-β [7, 10, 11]. The

α-subunit is common to all the gonadotropins and the β-subunit differs, conferring the binding specificity towards its receptor. Its primary, secondary, tertiary, and quaternary (heterodimer) structures are illustrated in Figure 1. The sequences of the mature subunits are shown. The sequence of the FSH-β, 111 amino acids, is longer than the FSH-α, which has 92 amino acids (Fig. 1A). The secondary structure (Fig. 1B) was derived from the X-ray data 1FL7 (Protein Data Bank, PDB, entry 1FL7, rcsb.org). The sequences shown lack a few amino acids at the N and C-terminal regions that could not be assigned in the 3D structure. Fox et al. [12], by using X-ray diffraction, were the first to obtain the three-dimensional structure of FSH using a βThr26Ala mutant of human follicle-stimulating hormone (FSH); it was with a 3.0 Angstrom resolution. The mutation Thr26Ala avoided the glycosylation at Asn24, making a protein suitable for crystallization, without compromising its receptor binding and activity. Figure 1C shows the 3D FSH molecule obtained by Fox et al. [12] corresponding to 1FL7. For a detailed discussion and features of the FSH structure, see the work of Fox et al. [12], Fan et al. [13-15], and Jiang et al. [16-18]. Also, important structural features of the FSHR have been recently reviewed by Ulloa-Aguirre et al. [19].

We focused our earlier work on the β subunit of FSH since it is unique for each gonadotropin and determines their specificity. We first study the possible interacting regions of FSH-β with the FSH receptor (FSHR), and then, the possible contact regions between the FSH-α and β subunits. These earlier studies regarding the mapping of the interacting regions of FSH are reviewed here, now that the 3D of the FSH and the FSH::FSH receptor complexes are known, and we can compare the predicted binding sites assessed by using synthetic peptides with the actual binding regions in the 3D structures.

Near three decades ago, using synthetic peptides, we found two receptor-binding regions in the FSH-β subunit, FSH-β-(33-53) and FSH-β-(81-95) [20-22]. However, when the

3D structure of FSH was determined [12, 19], these two regions were mistakenly disregarded as possible binding regions, given the wrong impression that the overlapping peptide approach was useless. Thus, all future work neglected these earlier findings. However, later the 3D structure of the FSH-FSHR complex was determined [15, 18], and, as it will be discussed here, the regions FSH-β-(33-53) and FSH-β-(81-95) actually include the main receptor-binding regions of the FSH-β subunit, as the synthetic peptide approach predicted more than 20 years earlier [21, 22].

The strategy of joining two separated binding regions to increase the affinity by mimicking the protein surface, used in earlier work [23], will be also discussed together with the limits that ΔG imposes on the ideal affinity of synthetic peptides. ΔG paradoxes, in the sense that they show a situation that is true but seems difficult to understand because it has opposite characteristics according to the value and sign of the ΔG terms enthalpy and entropy, will be also discussed. In addition, earlier attempts to define the relevant parameters for binding and to predict the relative inhibitory ability by using multiple linear regressions will be discussed. Noteworthy, in these regressions, hydropathicity was not a good predictor as the structural Garnier's parameters and average flexibility of the peptide. Finally, additive and nonadditive effects on the free energy of binding for these small synthetic fragments of FSH-β will be discussed together with a strategy to calculate their ΔG° values by using the buried surface area (BSA).

These observations reinforce the idea of using overlapping synthetic peptides to map interacting regions of proteins, develop possible antagonists or agonists, and study additive and nonadditive interactions.

2. Finding the receptor-binding regions of FSH-β using synthetic overlapping peptides

The studies aiming to map the receptor-binding regions of FSH were initiated more than three decades ago, in the laboratory of Leo E. Reichert Jr. at The Albany Medical College (Albany, NY). The goal was to find the interacting regions of FSH with its receptor and to evaluate the possibility of developing synthetic agonists or antagonists of FSH by using synthetic peptides. We focused on the FSH-β subunit since this subunit determined the binding specificity. This did not mean that the alpha-subunit was not important for productive binding. It was already known that both subunits of LH, TSH, and hCG were able to interact with the receptor and have intrinsic biological activity [11]. Using a similar approach, Leng et al. [24] later found a region of interaction between the alpha subunit of gonadotropins and the FSH receptor, located between amino acids GPH-α-(32-46). Thus, it was clear at that time that both subunits contributed to keeping the correct structure and binding to the receptor.

Initially, it was observed that epidermal growth factor (EGF) was able to compete with the binding of FSH to its receptor, although with much less affinity [25]. When the EGF and the FSH-β sequences were compared by using dot plots (similarity matrixes) [26], two tetrapeptides in common were found, TRDL (aa 34-37) and KTCT (aa 49-52). The KTCT sequence was within the pro-peptide of EGF, thus it was not involved in the binding competition. However, noteworthy, KTCT was also located near TRDL in the FSH-β, and synthetic peptides corresponding to these two sequences were both able to inhibit FSH binding to the receptor [25]. Meanwhile, a computer program developed by Stanley Krystek, working at that time in the recently established laboratory of Thomas Andersen, showed that regions covered by these tetrapeptides were hydrophilic and probably exposed to the surface [27, 28]. Intriguingly, preliminary results [20, 29] also showed a cross-talk between the

EGF signaling and the FSH signaling, reflected in reduced general protein phosphorylation in Sertoli cells pretreated with EGF before FSH stimulation, suggesting the activation of phosphatases by EGF, and a cross-talk among the FSH and EGF signaling. This crosstalk was later studied by other laboratories [30], including the possible involvement of the dual phosphatase DUSP6, which expression is upregulated by EGFR through ERK1/2 signaling [31].

Since these two peptides were close to each other in the FSH-β sequence (Fig. 1A), in preliminary work, Andersen et al. [32] synthesized a peptide including both regions TRDL and KTCT, and the sequence between them as a bridge, the peptide FSH-β-(33-53) (Fig. 2A), and tested its inhibitory capacity [32]. We later reported that FSH-β-(33-53) had ~100-fold improved affinity (Kd = 1 x 10^{-4} mol L^{-1}, Ka=1 x 10^{4} $mol^{-1}L$; Kd=1/Ka) compared to TRDL (FSH-β-(34-37), Kd= 1.25 x 10^{-2} mol L^{-1}, Ka= 0.8 x10^{2} $mol^{-1}L$) or KTCT alone (FSH-β-(49-52), Kd= 0.71 x 10^{-2} mol L^{-1}, Ka 1.41 x 10^{2} $mol^{-1}L$) [21] (Fig, 2A). These encouraging results prompted us to map the entire FSH-β subunit by using the 15 mer overlapping peptides shown in Figure 2B (5 overlapping amino acids at each extreme), a strategy previously used by Charlesworth et al. [33] to map the receptor-binding regions of the alpha subunit of gonadotropins, using thyroid-stimulating hormone (TSH) as a model system [11, 34, 35]. Thus, using these peptides to compete for ^{125}I-FSH binding to the receptor, we found that peptides FSH-β-(31-45) and FSH-β-(81-95) had the strongest inhibitory ability [22]. FSH-β-(31-45) included the region FSH-β-(34-37), TRDL, that, as mentioned above, was previously identified as a receptor-binding region [25, 36]. The peptide FSH-β-(81-95) produced the strongest inhibition, suggesting that it formed another receptor binding region. FSH-β-(81-95) includes the "determinant loop" (amino acids Cys87-Cys94 in FSH, 93-100 in hLH, hCG, and 88-95 in hTSH), a region previously thought to be determinant

for the interaction of glycoprotein hormones to receptors [37]. In conclusion, FSH-β-(33-53) and FSH-β-(81-95) were proposed to contain the receptor-binding regions of FSH-β [22].

3. Regions of interactions between the alpha and beta subunits of FSH

The same 15 mer overlapping peptides were then used to identify possible regions of interaction between the FSH-α and the FSH-β subunits [38]. The peptides were adsorbed to nitrocellulose and their ability to bind ^{125}I-FSH-α was tested. The results suggested the presence of several regions of interaction. The ^{125}I-FSH-α subunit was able to bind primarily to the synthetic peptides FSH-β-(11-25), (41-55), (51-65), (101-111). The regions of interaction with the beta subunit in the alpha subunit were not studied. A difficulty with this solid-phase approach is that some peptides might have stronger binding to the nitrocellulose than others. Also, the binding to the nitrocellulose can alter the peptide structure, and the bovine serum albumin used to block empty spaces could also have shielded some peptides, at least partially. The peptides were attached to the nitrocellulose surface by using their hydrophobic regions, thus hindering these regions from the interaction. In retrospect, these results should have been confirmed by adsorbing the FSH-α subunit to nitrocellulose and evaluating the ^{125}I-FSH-β binding in the presence of the different peptides, and then analyzing if the results had an agreement with those obtained by attaching the peptides to the membrane. It should be also considered that another difficulty for assays with synthetic peptides in solution is that some peptides may be partially insoluble or completely insoluble. Other laboratories later improved this initial solid-phase approach, and at present, the peptides can be labeled with fluorochromes, biotin, etc. for quantification, and attached to the membrane by using linkers or BSA, or even custom made on chips, avoiding the above-mentioned difficulties by using a variety of different strategies [39-47].

Now that we have available the FSH 3D structure, we can analyze whether the strategy of using overlapping synthetic peptides to map the interface regions gave a reasonable approximation, as it will be discussed below.

4. The FSH 3D structure and the location of the predicted regions of interaction

We performed an initial work regarding the FSH 3D structure in collaboration with John W. Crabb from the W. Alton Jones Cell Science Center (Lake Placid, NY). It corresponded to a fragment of FSH, the peptide FSH-β-(33-53), obtained by using NMR[48]. Residues FSH-β-(38-40) and FSH-β-(46-48) were seen to form an antiparallel β-sheet, linked by a short loop formed by amino acids (41-45); thus, the peptide structure of FSH-β-(33-53) showed the basic features of a β-hairpin at residues V38-Q48. Interestingly, β-hairpins have been recognized as an important structural motif for the design of protein epitope mimetics [49, 50].

As shown in Figure 3, the regions FSH-β-(33-53) and FSH-β-(81-95) form a continuous surface in the middle of the FSH-β subunit, as was previously postulated when the overlapping synthetic peptide approach was used [23]. These regions are exposed to the surface and form possible anchor sites that could easily interact with the FSH receptor, as was later demonstrated when the structure of the FSH-FSHR complex was determined (discussed in the next section).

On the other hand, the β-hairpin structure that we found by NMR at V38-Q48, using the synthetic peptide FSH-β-(33-53) in solution [48], was confirmed in the 3D FSH structure obtained by Fox et al. [12]; these authors named the FSH-β-(38-48) hairpin as the βL2 loop. Despite the strong evidence that the region FSH-β-(33-53) includes a receptor-binding region [7, 21-23, 48, 51], and similar earlier work by Keutmann et al. with hCG and LH [11, 52, 53],

Fox et al. [12], and Roth and Dias [54], the last authors, using alanine scanning, thought that this β-hairpin loop was not essential for FSH binding and that it rather stabilized the heterodimer association. Roth found that mutation of 48QKTCT52 to 48AAACA52 produced an FSHβ-subunit that could not associate with the α-subunit to form the heterodimer. However, they did find that mutation of 37LVY39 to 37AAA39 caused a 20-fold reduction in receptor binding. The mutations 34TRDL37 to 34AAAA37 or 44RPKI47 to 44APAA47 caused lesser but measurable effects. The strongest inhibitory effect was in region FSH-β-(37-39). This was in agreement with our earlier results obtained with the synthetic peptide FSH-β-(31-45), which showed a high inhibition of FSH binding to the receptor [22], and with the previous work with the peptide FSH-β-(33-53) [7, 21, 23, 48, 51]. Thinking that the binding reduction was actually due to a dissociation between the subunits, Fox et al. and Dias et al. erroneously thought that the entire region FSH-β-(33-53) was not important for binding to the receptor; they thought that it was only relevant for interaction with the α-subunit [12, 54, 55]. However, later Fan & Hendrickson [13-15] concluded that the C-terminal segments of both FSH-α and β-subunits, as well as the αL2 and βL2 loops, actually formed receptor binding sites. The reason for the controversy about the region (33-53) was that these amino acids interact with both the α-subunit and the receptor, as will be discussed in the next section. Thus, the βL2 loop (V38-Q48), included in the peptide FSH-β-(33-53), was indeed involved in the interaction with the FSH receptor, as the synthetic peptide approach showed originally [21, 25, 32].

Regarding the second binding region FSH-β-(81-95), its synthetic peptide showed the strongest inhibition of ^{125}I-FSH binding to FSHR [22]. It includes the "cystine noose" or "determinant loop", which constitutes a short loop between two disulfide-linked cysteines β87 and β94, known to be important for receptor binding and activity in hCG [56] and FSH [7, 22, 23, 51, 57, 58]. The regions of interaction buried in the interface of FSH-β-(81-95) with the

α-subunit can be found by using the Protein Interfaces, Surfaces, and Assemblies (PISA) server (ebi.ac.uk/pdbe/pisa). The PISA service analyzes the free energy of formation, solvation energy gain, interface area, hydrogen bonds, salt bridges across the interface, and hydrophobic specificity to analyze the 3D structure and make a prediction of the possible stable complex. The amino acids of FSH-β-(81-95) shown by the PISA results to be involved in the interaction with FSH-α are C87, S89, D90, S91, T92, D93, C94, T95 (Table 1). Therefore, the region (81-95) has interactions with both the α-subunit and the receptor. This can be observed comparing Tables 1 and 2 (the complete data obtained from the PISA server are in Tables S1 and S2, suppl. file). This will be further discussed in the next section when considering the FSH::FSHR complex.

Concerning the FSH-α subunit, a region implicated in binding to the receptor, GPH-α-(32-46), was first found by Charlesworth [33], and further studied by Grasso et al. [24]. It also forms a continuous surface region next to FSH-β-(81-95), as shown in Figure 3. Details of the interacting sites between the FSH-α and the FSH-β subunits can be found in the work by Fox, Dias, Fan, and collaborators, already mentioned [12-15, 19, 54, 57-61].

The predicted interacting regions between the FSH-α and FSH-β subunits, obtained by using synthetic peptides, FSH-β-(11-25), (41-55), (51-65), and (101-111) [38], are illustrated in Figure 4. As it can be observed, these peptides embrace the FSH-α subunit (GPH-α). However, some regions of interaction were missed by our solid-phase approach. The PISA server data of Table 1 (complete data in Table S1), obtained from the PDB (protein data bank) entry 1FL7, show buried surface areas (BSA) at amino acid residues FSH-β-(8), (10-11), (13), (15), (27-41), (44-53), (56), (58), (74), (76-77), (87), (89-101), (103), (105), and (107-109). Thus, the residues at amino acids (8), (10), (27-41), (74), (76-77), and (89-101) were not predicted by the synthetic peptide approach. The most important regions not predicted to be part of the interface between FSH-α and FSH-β were the regions

(27-41) and (89-101). The first region, (27-41), partially overlaps with the FSH-β-(33-53) region (38-48 is the βL2 loop involved in receptor binding, with residues 42-43 not buried). The second region, (89-101), partially overlaps with the synthetic peptide FSH-β-(81-95). Probably the missing residues were hindered by the attachment to the nitrocellulose. Nevertheless, all the peptides predicted to have interaction indeed corresponded to interacting sites, which include the amino acids FSH-β-(11), (13), (15), (44-53), (56), (58), (101), and (107-109).

Thus, the overall prediction of the interacting regions by using the synthetic overlapping peptides was good, although not perfect, with few interactions missed when the solid-phase assay was used by attaching the overlapping peptides to the nitrocellulose membrane. A very good agreement was found when the peptides were used as competitors in solution instead of using the attachment to nitrocellulose. As will be discussed below, it was later determined the 3D structure of the FSH::FSHR complex, showing with a robust basis that the regions predicted by the synthetic peptide approach were indeed involved in receptor binding.

5. Comparison between the predicted binding regions of FSH-β with the FSH-β 3D structure bound to the ectodomain of its receptor.

Earlier attempts to identify the FSH binding regions in the FSH receptor have been reviewed in the doctoral thesis dissertation of Johann Sohn and previous work [62-64]. These authors, using alanine-scanning and photoaffinity labeling of peptides, concluded that both the receptor endodomain (membrane inserted) and ectodomain, participate in the interaction with FSH, and signal transduction. They also concluded that the FSH-β subunit was close to the N-terminal region of the FSHR, and the α subunit should be projected

toward the exoloop 3 of the endodomain [62-64]. However, the spatial FSH orientation was somehow controversial until Fan and Hendrickson [13-15] determined the 3D structure of FSH bound to the extracellular domain of its receptor. The structure obtained from the PDB database (1XWD), plotted by using PyMol, is shown in Figure 5A. The sequences of interaction with the FSH receptor, previously assessed by using synthetic peptides, are highlighted. As shown, FSH-β-(33-53) and FSH-β-(81-95) are both parts of the surface of interaction with the FSH receptor. FSH-β-(33-53) includes the βL2 loop, residues V38-Q48. The regions of FSH interaction with the FSH receptor endodomain remain to be found.

Regarding the βL2 loop, Fan and Hendrickson only considered the tip of the βL2 loop (40-45) as a region of interaction [15]. Interestingly, Jiang et al. [16, 18], further adding a region between the endo and ectodomain of the receptor (PDB entries 4AY9 [18] and 4MQW [16]), later showed that the βL2 loop constituted part of a hydrophobic pocket crucial for binding and activation of the receptor, through binding of a sulfated tyrosine Y335 from the receptor, as it was previously suggested by Costagliola et al. [65]. This interaction with Y335, illustrated in Figure 5B (using the data from the PDB entry 4MQW), induces a conformational change essential for receptor activation [18]. Thus, the βL2 loop, included in the synthetic peptide FSH-β-(33-53) is crucial for receptor binding and activation.

The second binding region that we found by using synthetic peptides, FSH-β-(81-95) was also correct. Fan & Hendrickson found that amino acids FSH-β (89), (90), (93-99), (103), and (105) were buried in the receptor-ligand interface. They did find that the actual binding site was extended beyond amino acid (95) and included (96-99), (103), and (105). Thus, it is immediate that a peptide including the amino acids (96-99) could be an even better inhibitor than FSH-β-(81-95), found in our initial screening. Interestingly, recently Prabhudesai et al. [66] designed the peptide FSH-β-(89-97) (the “seat-belt” region), which

behaves as an FSH antagonist, as FSH-β-(81-95) did. The two peptides were not measured in the same work, but the Kd values reported in each case were 0.4 x 10^{-3} mol L^{-1} for FSH-β-(81-95) [23]and 5.8 x 10^{-3} mol L^{-1} for FSH-β-(89-97) [66]. The lower affinity (higher Kd) of FSH-β-(89-97) could be due to the smaller surface of interaction and capacity to emulate the native structure of a shorter peptide (15 mer vs 9 mer). Interestingly, Prabhudesai et al. also showed that the residues 91-STDC-94 could be substituted with alanine (AAAA), suggesting that the central residues of FSH-β-(89–97) are not critical for FSH-binding [67]; this alanine analog had less inhibitory activity. These results agree with the PISA server data of Table 2, showing that amino acids 91 and 92 do not contribute to the buried surface area (BSA) in the interaction with FSHR, and therefore, are not directly involved in binding. However, amino acids FSH-β-(93) and FSH-β-(94) add 42.46 $Å^2$ and 8.35 $Å^2$ to the BSA respectively, contributing in some degree to ΔG° of binding. Now, it would be interesting to evaluate the inhibition ability of FSH-β-(88-106), which includes the entire buried ligand-receptor interface of this region (Table 2). Later in this review, the affinity of FSH-β-(34-47), FSH-β-(88-106), and FSH-β-(34-47)-(88-106), which include all the buried surface areas of FSH-β, will be calculated by using an empirical approach based on the buried surface area (Table 3).

Again erroneously, most of the region FSH-β-(81-95) was also not considered involved in binding to the FSH receptor by Lindau-Shepard et al. [58], since residues Asp 88 and Asp 90 were not found to be involved in the interaction, by using the alanine-scanning method. On the other hand, as shown in Table 2, the rest of the binding region defined by Lindau-Shepard, FSH-β-(93-99), was remarkably correct. As shown in Tables 1 and 2 (and Tables S1 and S2, suppl. file), the region Glu81–Asp88 does not form part of the interface FSH-β::FSHR (except Cys 87 which is partially buried in the interface), confirming that Asp88 was not involved. However, residues FSH-β-(89-95), including Asp90 that forms a

salt bridge and excluding S91 and T92, are buried in the interface between the FSH-β and the FSH receptor, and constitute a binding region, as it was previously predicted by the inhibition obtained with the synthetic peptide FSH-β-(81-95). On the other hand, biochemical data also assign a role for the C-terminal region of FSH-β in binding to the receptor [12]. The corresponding interface and buried residues are illustrated in Table 2 (and Table S2, suppl. file), obtained from the PISA server. Grasso et al. [68] reported *in vivo* effects of hFSH-β-(81-95) and its subdomain hFSH-β-(90-95) on the mouse estrous cycle. In addition, the synthetic peptide hFSH-β-(90-95) inhibited the binding of ^{125}I-hFSH to bovine calf testis membranes, antagonized FSH-stimulated estradiol biosynthesis by primary cultures of rat Sertoli cells, and prolonged vaginal estrus in normally cycling mice. On the other hand, a synthetic peptide corresponding to hFSH-beta-(81-86), was inactive in vitro and had no effect on the mouse estrous cycle, reinforcing the role of region hFSH-β-(90-95) in FSH binding. This is also in agreement with the PISA data in Table 2 (complete data in Table S2) showing that hFSH-beta-(81-86) has no contribution to the buried surface area. Thus, FSH-β-(81-95) included a receptor-binding region, as predicted by the synthetic peptide approach [22].

Interestingly, the residues FSH-β-(42-43) not buried in the α::β interface of 1FL7 (Table 1; obtained from the PISA server by using the PDB 1FL7 entry) are the most buried in the FSHR::FSH-β interface of the 4MQW crystal (Table 2, PISA server, 4MQW entry), demonstrating the importance of the βL2 loop in the interaction FSH-β::FSHR. It is also notorious how far from the FSHR are the loops βL1 and βL3, formed by amino acids FSH-β-(11-25) and FSH-β-(62-76) respectively (Fig. 5A) [12]. These regions might be involved in a trimer interaction [16]. In this regard, the FSH binding to its receptor was partially inhibited (~50%) by peptides FSH-β-(11-25) and FSH-β-(21-35) and FSH-β-(61-75) (~25% inhibition), suggesting certain interaction of loops βL1 and βL3 with the FSH

receptor. These suggested that additional receptor binding regions might be present in these loops, in particular in the N-terminal region FSH-β-(1-15), as it was postulated by Steward for luteinizing hormone (LH) [69]. Interestingly, the peptide FSH-β-(1-15) was able to bind Ca^{2+}, a cation also necessary for high-affinity binding to the receptor [70]. The possible significance of this finding is unknown since the 3D structure does not appear to form an E-F hand in that region.

The 15 mer peptides of FSH-β predicted to interact with FSHR were (> 60% inhibition) FSH-β-(31-45), (71-85), and (81-95). Analyzing the actual interactions in the 3D structure by using the PISA server, the regions of interaction are FSH-β-(35-46), (89-90), (93-99), (103), (105), as shown in Table 2 (complete data in Table S2, suppl. file). Only the residues (96-99), (103), and (105) were not predicted to bind the receptor by using the synthetic peptide approach. Also, the peptide FSH-β-(71-85) was able to inhibit ^{125}I-FSH binding to the receptor by more than 60% but does not appear in the interface between the FSH-β and FSHR, according to the PISA server results. However, in the trimeric structure reported by Jiang et al. [16, 18], βLeu73 is in the interface with another monomer of the proposed FSH-FSHR trimer, interacting with a second neighboring receptor monomer. Interestingly, mutation βLeu73Glu at the potential exosite also promotes signaling [18]. Therefore, it seems that interference with this binding region between monomers increases the FSH-FSHR affinity. This also agrees with the results obtained by Jiang with a second amino acid in the interface between monomers, βThr60, since its mutation increases binding to the receptor. Probably the monomers are more active than the dimers or trimers and this mutation produces a repulsion due to the added negative charges that dissociate the trimer into monomers. The residues βThr60 and βLeu73 are buried between the monomer interfaces (interacting FSH-β of one monomer with the FSHR of another monomer) in the trimeric structure of FSH::FSHR (PDB entry 4MQW). For these

interactions, in the receptor, the buried residues at the interface are Arg59, Val60, Glu84, Val85, Glu87, Tyr110, and His134; in FSH-β, Tyr58, Glu59, Thr60, Arg62, Ala70, Asp71, Ser72, Leu73, and Thr75 (results of PISA server with structure 4MQW, interface #15, structures Z and B). Later Jian et al. provided further evidence for a functional trimer [16]. These findings agree with our earlier results showing that the peptide FSH-β-(71-85), including βLeu73, inhibited the FSH binding to the receptor by more than 70% [22].

Figure S1 (suppl. file) shows the FSH-β::FSHR interactions with more detail, obtained by using PyMol with the interfaceResidues.py plugin (Fig. S1A) and LigPlot+ v.2.2.5 (Fig. S1B). A detailed analysis of all the interactions between the FSH-α, FSH-β, and the FSHR is out of the scope of this work, which aimed to illustrate the usefulness of synthetic peptides to find putative binding regions. An initial analysis can be found in the review by Leo E. Reichert Jr. [71], and later in the work of Fox et al. [12], Fan and Hendrickson [13-15], Jian et al. [16-18], and the reviews by Ulloa Aguirre et al. [19, 72]. We can conclude that the use of overlapping synthetic peptides is a good strategy as an initial screening to find putative receptor-binding regions and possible agonists/antagonists when the structures of the ligand-receptor complexes are unknown.

6. **Serine analogs of FSH-β**-(33-53) and FSH-β-(81-95) **- Induced fit or conformational selection?**

To avoid possible cross-linking difficulties with the free Cys [21], serine (Ser, S) analogs of FSH-β-(33-53) and FSH-β-(81-95) were synthesized. These peptides also inhibited the FSH-β binding to its receptor, with a slightly improved affinity [51]. However, we have a surprise and contrary to the Cys-containing peptides, the Ser analogs were not able to induce estradiol (E2) biosynthesis.

Let us first briefly review the different models of ligand-receptor interactions for a better understanding of the results obtained with the serine analogs (Figure 6). It has occurred a long evolution in the ideas since the rigid key-lock model of the ligand-receptor interaction of Fisher was postulated (Fig. 6A)[73, 74]. It was then postulated the idea of a ligand-induced fit into the receptor structure leading to activation (Fig. 6B) and later a mutual fit in protein-protein interactions (Fig. 6C) [75]. Alternatively, it was envisioned that the receptor might exist in multiple conformations and that the active receptor conformation is eventually selected upon binding, in a process called "conformational selection" (Fig. 6D) [75, 76], which may include also mutual fitting.

Sticking to an "induced fit" model [77, 78], one might think that the entire ligand molecule should be present to effectively change the receptor conformation, allowing the signal to be transduced with full activity. Since the peptides that we used (15 mer) cover a small area and are flexible, an "induced fit" in the entire receptor molecule was not likely to be present. However, to our surprise, the peptides FSH-β-(33-53) and (81-95) partially induced estradiol (E2) biosynthesis [21, 22]. Therefore, we originally thought that a "conformational selection" model could explain these results. According to the conformational selection model (Fig. 6D), the receptor should dynamically change from a series of conformational states from the fully active state to the completely inactive state. Then, the presence of a peptide corresponding to a receptor-binding region should increase the probability of activation by changing the equilibrium, so the fraction of receptor molecules in the active conformation increases. This should occur during a period enough to give productive signaling. However, since the serine analogs were not able to induce E2 biosynthesis, the idea of a conformational selection was not supported by the results obtained, in which the Ser analogs behave only as antagonists, unable to activate the receptor. Perhaps the Cys-containing peptides were able to activate the receptor disrupting

the S-S bridges in the hinge region, a region that maintains the receptor in an inactive conformation in the absence of the hormone [19, 79]. Therefore, at least for FSH, the induced fit theory is more adequate. We thought that the peptides were highly flexible and having only a subset of the anchoring sites might not be enough to obtain the adequate conformational change needed to activate the receptor. Thus, the lack of an "induced fit" and conformational change in the receptor due to the small region covered by these peptides or the lack of Cys residues to reduce disulfide bridges might explain why the Ser analogs did not induce E2 biosynthesis. On the other hand, the Cys-containing peptides might have a nonspecific redox activity due to the free R-SH groups, able to activate the E2 biosynthesis by indirectly modifying the aromatase activity [80] or through some other nonspecific mechanism.

Now that the 3D structure of FSH and the FSH::FSHR are known, the reciprocally induced fit model is supported by evidence. Thus, we can have a better idea of the reasons why the Ser analogs do not produce receptor activation. The receptor might in theory change dynamically from diverse active states; however, the ectodomain, and in particular the hinge region (296-331), precludes a basal activity, maintaining the receptor in an inactive conformation [81]. Thus, a "conformational selection" is very unlikely to occur with the FSHR.

It has been already discussed that FSH binds to the ectodomain and then reshapes its conformation to form a binding pocket for a sulfated receptor Tyr (sTyr335, PDB structure 4AY9) [18]. Several other changes in the FSH were previously detailed by Fan and Hendrickson [14, 15]. If the interfaces between α and β subunits are compared between the free (PDB 1FL7) and bound FSH (PDB 4MQW) by using the PISA server data, clear changes in the buried amino acids can be observed, particularly in the already buried areas of the free FSH (results not shown). These interactions eventually lead to receptor

activation, which also showed extensive conformational changes [14-16, 18]. Therefore, an "induced fit" in both the FSH and the FSHR is evident in the FSH::FSHR interaction. This mutually induced fit suggests that the binding of a semi-rigid FSH induces a conformational change in the receptor that leads to receptor activation. On the contrary, a flexible peptide, like the synthetic Ser analogs, does not have enough rigidity to produce the induced fit needed to reshape the receptor and induce its activation.

Interestingly, the FSH has a thioredoxin-like activity [82, 83], attributed to a CXXC motif in the FSH-β-(81-95) region, and glutathione and NEM affect FSH binding to receptor [84]. However, the disulfide bridges in the FSH molecule, at first sight, are far away from the disulfide groups in the hinge region of the receptor (or any other Cys residue in the receptor), and any disulfide interchange seems unlikely unless the actual 3D structure of the FSH::FSHR complex in solution differs from the PDB data obtained from FHS::FSHR crystals. The glutathione and NEM effects on FSH binding might rather be due to the disruption of the disulfide bridges in the hinge region. NEM effects suggest that some S-S groups can dynamically change their redox status. Is it possible that the FSH structure in solution be enough flexible so that the CXXC region can reach the disulfides in the hinge region of the receptor? NMR or electron crystallography data with the entire receptor in its environment might be needed to study this possibility. Alternatively, since the synthetic peptides FSH-β-(33-53) and FSH-β-(81-95) do have free sulfhydryl groups, these R-SH groups could disrupt the S-S bridges in the rigid hinge structure that maintains the FSH receptor inhibited, producing a partial activation by relaxing the structure, allowing the induction of the basal activity of E2 biosynthesis. This also may explain why Ser analogs were unable to produce receptor activation. Further studies are needed to understand the reasons why Ser analogs do not induce E2 biosynthesis. In vitro, binding of these Cys-containing peptides can be optimized by using appropriate redox buffers that

include glutathione, DTT, β-mercaptoethanol, or similar reducing reagents; however, this strategy cannot be used in vivo, where the Ser analogs might be more appropriate.

7. Mimicking two surface-adjacent binding regions

The hydrophobic pattern of FSH-β [22] suggested that the regions FSH-β-(33-53) and FSH-β-(81-95) might form a continuous surface on the native protein. Thus, we decided to join them in one synthetic peptide and see if the affinity could be increased, as shown in Figure 7. This strategy was different from the earlier one using a peptide comprising the regions FSH-β-(34-37) (TRDL) and FSH-β-(49-52) (KTCT), the peptide FSH-β-(33-53) since these two peptides were close to each other in the FSH sequence and formed a defined FSH-β-(38-48) hairpin, the βL2 loop (Fig. 7A). Instead, the peptide FSH-β-(33-53)-(81-95) comprised two binding regions separated by a long region (residues 54-80) on the primary FSH sequence, that was deleted (Fig. 7B). As expected, the association constant of FSH-β-(33-53)-(81-95) ($K_a = 2 \times 10^4$ mol^{-1} L) increased compared to the Ka value of its individual components, $K_{aA} = 1.0 \times 10^4$ mol^{-1} L for peptide FSH-β-(33-53) and $K_{aB} = 2.5 \times 10^3$ mol^{-1} L for peptide FSH-β-(81-95) [23]. These results suggested that the FSH-β-(33-53)-(81-95) peptide mimicked the receptor-binding region on the FSH surface. As proved now by the X-ray structure of FSH (Figure 3), this was correctly predicted. However, its affinity towards the FSH receptor, $K_a = 2 \times 10^4$ mol^{-1} L, was only slightly improved compared to the affinity of each peptide alone; it was still far away from the affinity of the entire β-subunit, $K_a = 1.1 \times 10^7$ mol^{-1} L [23], or from the affinity of the native FSH hormone (α+β subunits), $K_a = 1.02\ 10^{10}$ mol^{-1} L [85].

To increase the affinity, one may try to synthesize the entire region between the two peptides, an optimal bridge, but this would be close to synthesizing the entire FSH-β. A better approach would be to synthesize these two regions with a bridge designed with the

aid of the 3D structure. Most importantly, examining the buried surface area it was noted now that at least four amino acids were missing at the end of the binding region B (amino acids 96-99). These residues might contribute significantly to the free energy of binding, and it will be later considered.

The affinity might be improved, at least partially, by adding bridges/linkers [86], or by restricting selected conformations making the peptide more rigid [87, 88]. However, there are also limitations with this approach; in some cases, one can get the opposed result restricting the conformations, which constitutes a known ΔG paradox, the "paradox of conformational constraints" [89]. In some cases, certain flexibility is needed for best interaction by mutual induced fitting [90-92], and fixing a suboptimal conformation may be detrimental. Another paradox is the presence of enthalpic/entropic compensations [93]. When the entropy is changed to optimize binding, often the enthalpy changes in a way that ΔG° is maintained almost constant, and the same may occur when the enthalpy is modified. Therefore, some authors postulate that it is better to optimize ΔG° instead of trying to optimize entropy or enthalpy separately [93-95].

The idea of joining two discontinuous binding regions to "mimic" the ligand-receptor binding interface [23] was then applied by other laboratories to different proteins [6, 96, 97], a strategy improved with the increased availability of 3D data. Joined peptides, or bound to a scaffold structure, may be used if they do not interfere with each other. Otherwise, it would be preferable to use them at the same time as separate molecules. However, the effects on the affinities are not multiplicative if the fragments are not joined. To have unjoin ligands would be equivalent to having two inhibitors bound independently to different sites in the receptor. Joining discontinuous epitopes can be useful as agonists/antagonists, but also as better antigens, with increased affinities for their

antibodies; the synthesis of tri- and tetracyclic FSH/hCG mimics is a good example of this strategy [98].

8. **Multiple linear regressions**.

After measuring the FSH binding inhibition capacity of each overlapping FSH-β peptide, a multiple linear regression (MLR) was performed to analyze the results [22]. The contribution to the regression of the different predictors was used to evaluate their relative relevance in binding inhibition. The MLR and the different predictors were estimated by using the GCG package of the Wisconsin University [99], located in a remote computer at SUNNY and accessed from the Albany Medical College (NY, USA) by modem (those old times without internet!). The values of the different predictors were taken as the average for each 15 mer peptide. The antigenic index was highly correlated with the other parameters since it was derived from the same predictors, and therefore it was not included in the final MLR. From the different synthetic overlapping peptides used, three peptides that were quite insoluble were left outside of the regression set and then used to see if the model provided a reasonable estimation of their values.

Surprisingly, *hydrophobicity* was not a significant predictor, and it was left out of the regression after an initial calculation. The multiple linear regression was statistically significant ($P < 0.0001$), and a high coefficient of correlation was obtained ($r = 0.93$). The obtained MLR regression was:

$$BI = 526.8 + 2.178\,(S) - 381.6\,(F) + 40.8\,(H) - 48.6\,(E) - 40.9\,(T)$$

where BI was the "binding index" of ^{125}I-FSH in the presence of the different peptides (% ^{125}I-FSH binding), S the surface probability, F the flexibility, and H, E, and T were the % of the helix, extended and turn conformations according to the Garnier's method [100]. Finding the BI was equivalent to predicting the binding affinity or the binding free energy

of each peptide. The contributions to the R^2 were T = 0.35, E = 0.23, H = 0.15, F = 0.05, and S = 0.02. The coil conformation, C, and the hydrophilicity did not contribute significantly to the regression.

It was a counterintuitive result that the *hydrophobicity* showed no relevance. Instead, the more relevant predictors were those related to Garnier's structural parameters and the peptide flexibility. Note the sign and magnitude of the corresponding coefficients in the regression. H and S >0, thus, an increased helix content or surface probability was detrimental to binding inhibition (BI). On the other hand, F, E, and T were <0; then, high flexibility or increased turn or extended conformations (beta sheets), favored the ability of these synthetic peptides to inhibit FSH binding to the receptor. The negative influence of surface probability was also unexpected. This suggested that the highly charged residues that are normally on the surface do not favor a strong affinity or that they are located away from the binding interface, but still have influence in the binding, as was later observed by Kastritis et al. in another model system [101]. All this was reasonable except for the lack of contribution of the hydrophilicity (or the inverse, the hydrophobicity).

Although we did not evaluate this method with other ligands, the results suggested that MLR strategy could be applied to characterize other model systems. Years later Chen et al. [102], by using a dataset of 113 protein-protein interactions, applied multiple linear regression to predict binding affinities. Noteworthy, they also found that hydrophilicity was not a relevant predictor. This lack of correlation with hydrophilicity was attributed to the fact that the binding regions can be hydrophobic, hydrophilic, or both [102-104].

As mentioned, another significant predictor found by our earlier regression model was peptide flexibility. The importance of peptide flexibility has been later observed in several reports [90-92]. It could be related to the ability of flexible peptides to better adapt to an "induced fit" and to increase the area of interaction, the buried surface area. However,

this will work only if we are looking for inhibitors. Analogs that activate the receptor might need a more rigid structure to be able to force a conformational change in the receptor, by induced fit or conformational selection. Figure S2 (suppl. file) shows the flexibility along the FSH-β sequence, obtained by using the online server CABSFlex 2.0 (biocomp.chem.uw.edu.pl/CABSflex2)[105]. Interestingly, the binding regions FSH-β-(33-53) and FSH-β-(81-95) are the most flexible regions and a notorious reduction in flexibility is observed upon binding, particularly in the binding region FSH-β-(33-53). On the other hand, the flexibility changes in the region FSH-β-(50-80), which is not a receptor-binding region, are small. Finally, the relevance of the helix stability for the binding affinity has also been later reported [106].

Another interesting finding came up when the MLR approach was applied to the thioredoxin-like activity of these synthetic peptides. Since the FSH had a thioredoxin-like activity [82, 83], I thought that some of the 15 mer peptides could also have similar activity and that these overlapping peptides could help to confirm the active redox region. In addition, we could also find in this way the smallest enzyme-like activity ever described. And this was the case: several peptides, those with Cys residues, had thioredoxin-like activity. Noteworthy, as a predictor for high redox activity, the structural parameter H (% helix conformation) was even more important than the number of cysteines in the peptide. The relative contributions to the R^2 were H> No. of Cys > T > S > F, all with $p < 0.03$. According to the regression model, a low H, T, and S content (low structured regions), and a high number of cysteines and surface probability (S), favor the redox activity. These results were quite different from those required for binding to the FSH receptor with high affinity, previously mentioned. All of them are reasonable since the redox CxxC motif must be exposed to the surface and be flexible to reach the target redox pair. These results showed that an enzymatic-like activity could be present in a noteworthy small peptide such

as FSH-β-(81-95) and that the redox activity of the R-SH residues can be modulated by the surrounding residues and structural motifs.

Although these regression models were based on one sequence, they show that this kind of analysis could be a powerful tool to predict/analyze binding affinities and enzymatic activities. Recently, other regression models based on a large dataset of over 100 proteins and different predictors have been developed, able to calculate $\Delta G°$ or K_a. One model often used is the PPA-prep2 server of Yugandhar and Gromiha [107]. More recently, using ISLAND (island.pythonanywhere.com) [108], which included a multiple linear regression over several structural and thermodynamical parameters, taken from a large dataset of ligand-receptor complexes, a more precise prediction for K_d could be obtained. ISLAND is based on the earlier approach of Abbasi et al. [109]. Both methods used only the amino acid sequences of the ligand and the receptor, without the need for 3D structural data. The sequences should have at least 20 amino acids.

By using PPA-pred2, the predicted binding parameters for FSH-β were $\Delta G° = -11.32$ kcal/mol and $K_d = 0.5 \times 10^{-8}$ mol L^{-1}. By using ISLAND, the values predicted for FSH-β were $\Delta G° = -10.87$ and $K_d = 1.06 \times 10^{-8}$ mol L^{-1}, both close to the experimental value of 9.1×10^{-8} mol L^{-1}. The K_d value is computed from predicted $\Delta G°$ at room temperature (25 °C) by using the equation $\Delta G° = (1.9858775 \times 10^{-3}$ kcal mol^{-1} $K^{-1}) \times ((273.15 + 25)$ K) x ln (K_d). With ISLAND, not only the sequence is needed but also the protein class. However, with either PPA-pred2 or ISLAND, if only a fragment of the FSH-β sequence is used, for example, the sequence of FSH-β-(33-53), the results show near the same value as those obtained with the entire FSH-β, which is clearly wrong.

On the other hand, if we know the 3D structure, PRODIGY can calculate the binding affinity of protein-protein complex, and the affinity of proteins with small ligands [110]. For binding of FSH-β to FSHR, the results were $\Delta G° = -10.4$ kcal/mol and $K_d = 2.6 \times 10^{-8}$ mol

L^{-1}, close to the measured value of K_d = 9.1 x 10^{-8} mol L^{-1}. However, it cannot calculate the ΔG° for the synthetic peptides representing the different binding regions of FSH.

Thus, these programs, as they are right now, only allow [6, 96, 97] estimations of entire proteins, not their fragments. This is a limitation that should be resolved, ideally for 15 mer peptides (or smaller), so we can predict the affinities of small overlapping peptides. However, to be able to do this, non-additive terms must be considered in ΔG° estimations, as will be discussed in the following sections. Abbassi et al. [109] recognized that even the state-of-the-art sequence-only predictors of binding affinities are far from satisfactory, as yet. A promising approach to build the 3D structure from the amino acid sequence is DeepMind's AlphaFold 2 software [111]. It could be also useful to predict the buried surface areas (BSA) of interactions for complexes with unknown 3D structures, and the BSA and ΔG° of synthetic peptides corresponding to their binding regions, as it will be discussed in the next sections.

9. The principle of ΔG additivity – Limitations without nonadditivity terms

Remembering a series of lectures given by Gregorio Weber during one of his visits to Buenos Aires in 1986, and his previous work [112-115], I decided to test whether the individual binding regions of FSH-β behaved additively. Before, William P. Jencks had described the principle of ΔG additivity and cooperativity applied to proteins [116], based on Weber's ΔG of coupling [112], or the earlier and equivalent Wyman's ΔG of interactions [117].

This principle of additivity was applied often over the years, not always having present that it is only an approximation when ΔG° of coupling/interactions is not included [118]. As it was noted by Wyman [117], Weber [112], Jencks [116], and Rashin [119], entropy has a key contribution that cannot be neglected. Even more, it usually corresponds to entropic changes in the entire system that cannot be distributed as a sum of individual TΔS°

components (non-additivity in ΔS°), as it was proven by Mark and van Gunsteren [120]. Let us see now how these ideas may apply to the binding of FSH-β and FSH-β synthetic peptides to the FSH receptor.

Through binding competition, we estimated the association constants $K_a = 1/K_d$ of the synthetic peptides corresponding to the different FSH binding regions (Fig. 7 and Table 3)[23]. As previously reported in that work, the relation between the ΔG° values and the dissociation constant K_a is given by Gibbs´s equation (see Appendix 1, suppl. File, for details),

$$\Delta G^\circ = -RT \ln K_a = RT \ln K_d$$

For binding of *i* ligands to receptor, it can be shown that (Appendix 2),

$$\Delta G^\circ = \left(\sum_i \Delta G^\circ_i\right) + \Delta G^\circ_{int} \quad [1]$$

Equation [1] summarizes the principle of free energy additivity: the free energy of binding ΔG° is the sum of the free energy of each binding region, anchor sites, hot spots, or residues at the interface, ΔG°_i, plus the free energy ΔG°_{int}, due to interactions among binding sites in the ligand and the receptor. This last term also includes all the interactions with solvent, conformational entropy changes [119], and any other non-additive contribution to the free energy, including entropic changes of the entire system.

Although this equation was originally thought for two or more independent ligands bound to different receptor binding regions or subunits (i.e., O_2 binding to hemoglobin), it also stands when two synthetic peptides corresponding to different binding regions are bound separately or are joined together and constitute a new molecule with two or more anchor sites [23].

If we consider a simplified "ideal" condition in which $\Delta G°_{int} = 0$, then equation [1] becomes

$$\Delta G° = \sum_i \Delta G°_i \quad [2] \quad \text{Model 1}$$

where $\Delta G°_i$ corresponds to the free energy of association for each binding region; applying the Gibbs equation to [2] (details in Appendix 2),

$$K_a\ (total) = \prod_i K_{a_i} \quad [3] \quad \text{Model 1}$$

Applying equation [3] to the FSH-β binding regions FSH-β-(33-53), with K_{aA} = 1.0 x 10^4 mol^{-1} L, and FSH-β-(81-95), with K_{aB}= 0.25 x 10^4 mol^{-1} L, the association constant predicted for the joint peptide FSH-β-(33-53)-(81-95) is $K_{aA}.K_{aB}$ = 2.5 x 10^7 mol^{-1} L (Fig. 7). This predicted value is close to the affinity of the entire FSH-β-subunit, 1.1 x 10^7 mol^{-1} L. However, as already mentioned, the joint peptide FSH-β-(33-53)-(81-95), actually had a K_a = 2 x 10^4 mol^{-1} L, three orders of magnitude less than the predicted value from the product $K_{aA}.K_{aB}$ (Model 1). At first, we thought that this notorious difference might be due to incorrect folding of the joint peptide. However, if we consider a more realistic approach in which nonadditive interactions are included ($\Delta G°_{int} \neq 0$), then the value seen experimentally is not that far from the predicted value. We will show the effects of including nonadditive interactions in the following sections (Models 2 and 3).

10. A more realistic approach: including nonadditive interactions

First, instead of adding $\Delta G°_{int}$ to the total free energy in equation [1], it is possible to include a coefficient "*c*" of correction. Since $(\sum x) + b \cong c \sum x$; then, $\Delta G° = (\sum_i \Delta G°_i) + \Delta G°_{int} \cong c \sum_i \Delta G°_i$. Thus

$$\Delta G° = c \sum_i \Delta G°_i \quad [4]$$

The coefficient "c" can be obtained empirically from the ratio,

$$c = \frac{\Delta G°}{\sum_i \Delta G°_i}$$

where $\Delta G°$ is the free energy of binding and $\sum_i \Delta G°_i$ is the sum of the observed free energies of each binding region. From [4], $c = 1 + \frac{\Delta G°_{int}}{\sum_i \Delta G°_i}$. If $|\Delta G°_{int}| \ll |\sum_i \Delta G°_i|$ then $c \rightarrow 1$. Here, $\sum_i \Delta G°_i < 0$ since it corresponds to the free energy of association, $\Delta G°_a = -RT \ln K_a$. Thus, for example, in the presence of negative cooperativity among binding regions (negative interactions), $\Delta G°_{int} > 0$ [113] and $c < 1$. In the presence of positive cooperativity, $\Delta G°_{int} < 0$ and $c > 1$.

Wyman [117] defined a similar coefficient, "ω", that instead of the ratio of $\Delta G°$s, corresponds to the ratio of the association constants K_a[121],

$$\omega = \frac{K_{a_{AB}}}{K_{a_A} K_{a_B}}$$

Since $\Delta G^\circ = (\sum_i \Delta G^\circ_i) + \Delta G^\circ_{int}$ then $\Delta G^\circ_{int} = \Delta G^\circ - (\sum_i \Delta G^\circ_i)$, and replacing in terms of the equilibrium constants, $\Delta G^\circ = -RT \ln K$, then $\Delta G^\circ_{int} = -\mathrm{RT} \ln \frac{K_{a_{AB}}}{K_{a_A} K_{a_B}}$ and,

$$\boxed{\Delta G^\circ_{\mathrm{int}} = -\mathrm{RT} \ln \omega}$$

If there are no coupling/interactions and both regions behave independently then *c* and ω = 1. If positive cooperativity exists *c* and ω > 1 and in the presence of negative cooperativity *c* and ω < 1.

In this regard, as was already mentioned, the peptide FSH-β-(33-53)-(81-95) has a K_{AB} = 2 x 10^4 mol^{-1} L, really far from the predicted value of $K_A.K_B$ = 2.5 x 10^7 mol^{-1} L for additive interactions (without nonadditive interactions, Model 1). In this case *c* = 0.58 and ω = 8 x 10^{-4} (Appendix 4). These values of *c* and ω suggest a strong contribution of ΔG°_{int} to the free energy (alternatively, suggest that an interacting region K_C is missing). The coefficient *c* not only includes cooperativity effects but also the free energy corresponding to all the interactions represented in the term ΔG°_{int}.

Now, in the term ΔG°_{int}. is also possible to separate additive interactions, ΔG°_{ad}, from nonadditive interactions, ΔG°_{na}.

$$\boxed{\Delta G^\circ_{\mathrm{int}} = \Delta G^\circ_{ad} + \Delta G^\circ_{na}}$$

From equation [1], $\Delta G^\circ = (\sum_i \Delta G^\circ_i) + \Delta G^\circ_{ad} + \Delta G^\circ_{na}$. Again, since $(\sum_i x_i) + b \cong c \sum_i x_i$, we can define a new coefficient of correction c´ that will include the additive term ΔG°_{ad}. Thus,

$$\Delta G^\circ = c' \sum_i \Delta G^\circ_i + \Delta G^\circ_{na} \quad [5]$$

On the other hand, if we have a predictor *x* that contributes linearly to the free energy observed for each component, then $\Delta G^\circ_i = f(x)$ and $c' \sum_i \Delta G^\circ_i = c' \sum_i f(x)$ Replacing this term in [5]

$$\Delta G^\circ(x) = c' \sum_i f(x) + \Delta G^\circ_{na} \quad [6]$$

Since the sum of a linear function is a linear function, $c' \sum_i f(x) = c' F(x)$

and replacing this term in [6],

$$\Delta G^\circ(x) = c' F(x) + \Delta G^\circ_{na} \quad [7]$$

If some interactions are non-additive or are independent of *x*, then $\Delta G^\circ_{na} \neq 0$. Equation [7] is not a linear function; it is an affine function, a linear function plus a translation represented by ΔG°_{na}. A graph *ΔG°(x)* vs *F(x)* will be still a line but with a different slope *c´*, and ordinate at the origin ΔG°_{na}. On the other hand, if all the interactions are additive and depend on *x*, then $\Delta G^\circ_{na} = 0$, and equation [7] becomes equation [8]:

$$\Delta G^\circ(x) = c' F(x) \quad [8]$$

Since [8] is now a linear function, a graph ΔG° vs *x* will be a line passing through the origin (0,0) with slope c´. Equations [7] and [8] will be applied to FSH in the next section, in which the buried surface area (BSA) at the interface ligand-receptor is the predictor *F(x)* and *c'* is the coefficient γ, the free energy density. The linear function represented by [8] can be observed in Figure 9, as Model 2, with ordinate to the origin equal to zero. The affine function of equation 7 is shown as Model 3 in the same figure.

11. ΔG°, the free energy density γ, and the buried surface area (BSA).

The concept of accessible surface area was introduced by Lee and Richards [122]. Then Chothia and Janin emphasized the relevance of the buried surface area (BSA) at the ligand-receptor interface to predict the free energy of binding, $\Delta G°_{binding}$ [102, 123-128],

$$\Delta G°_{binding} = \gamma\, BSA \quad [9]$$

Equation [9] is a linear function, equivalent to equation [8]; here F(x) is the buried surface area BSA and c´= γ is the slope of the regression $\Delta G°$ vs BSA; γ is the *free energy density* (kcal mol^{-1} $Å^{-2}$). Some authors obtained the γ values considering only the hydrophobic residues of the BSA at the interface, other authors use the total BSA of the ligand, and others the total BSA of the ligand (L) plus the receptor (R), divided or not by two. For this reason, the γ values may differ from author to author. For example, Vallone et al. [129] obtained an empirical value of -15 ± 1.2 cal mol^{-1} $Å^{-2}$ using the burying hydrophobic residues. In more recent work, using average data from proteins and small ligands, Houk et al. [127] obtained a value of γ = -7 cal mol^{-1} $Å^{-2}$, considering only the area of the ligand, and Chen et al. -16 to -4 cal mol^{-1} $Å^{-2}$ depending on the BSA range, taken as BSA the buried areas of the ligand plus the receptor at the interface [102]. A value of -25 cal mol^{-1} $Å^{-2}$ was obtained when only the hydrophobic buried surface area of the total complex was considered [128] or when the total BSA is considered to estimate ΔG of folding in studies of protein stability [119]. We will use here averages BSA values from the 3 complexes of FSH::FSH receptor co-crystal 4MQW (Appendix 3, suppl. file), and BSA for FSH and FSH-β will be calculated as the buried surface area of the ligand (L) only. Since we know the equilibrium constant K_a for FSH, 1.02 ± 0.3 x 10^{10} mol^{-1} L (Triton X-100 solubilized

receptor)[130], which corresponds to a ΔG°= -13.65 kcal mol^{-1}, we can estimate a γ value for FSH, in our experimental conditions. From equation [9],

$$\gamma' = \Delta G°_{FSH}/BSA \quad [10]$$

The BSA of both subunits α+β is 1882 ± 91 (n=3) Å^2 (Appendix 3, suppl. file). In our experimental conditions with ΔG° = -RT ln K_a = -0.5925 x ln (1.02 x 10^{10} mol^{-1} L) = -13.65 kcal mol^{-1}. Then, from [10], γ′= -13.65/1882 = -0.00726 kcal mol^{-1} Å^{-2},

$$\gamma' = -0.00726\ kcal\ mol^{-1}Å^{-2}$$

Noteworthy, this γ′ value is identical to the γ value obtained by Houk et al. [127] by using the average BSA and K_a values for several proteins and ligands of different sizes, γ = -0.00726 kcal mol^{-1} Å^{-2} (recalculated from Table 7 of Houk´s paper to add decimals). This value was calculated only from the BSA area of the ligand at the interface, as we did above.

The surface free energy density γ for FSH is then 7.26 cal mol^{-1}Å^{-2}. Thus, the BSA that produces a change of 1 kcal/mol in FSH-β binding to receptor is,

$$\Delta G° = -1\frac{kcal}{mol} = -0.00726\frac{\text{kcal}}{\text{mol Å}^2}\ \text{x BSA}$$

and,

$$\text{FSH BSA}_{1\ \text{kcal/mol}} = 138\ Å^2$$

This area is larger than the average area reported by Houk et al. of 67 Å^2 [127]. A large area on the interface compensates for the lack of hot spots (HS), which are residues or small regions of high binding energy[131]. The presence of an HS can be predicted by using the SPOTON online software of Bonvin's laboratory at the server alcazar.science.uu.nl [132, 133]. In agreement with the high value of 138 Å^2 per 1 kcal/mol of free energy (low free energy

density), the SPOTON results show that FSH-β does not have HS residues. However, the FSHR does have six HSs located at Leu55, Gln79, Arg101, Lys104, Tyr124, and Asn129. Several null spots (NS) were also identified by SPOTON, which are indicated in Appendix 6 (suppl. file). On the other hand, according to the results obtained by using LigPlot+, the HS receptor residues Leu55, Gln79, Arg101, and Lys104 are involved in the interaction with FSH- β (Fig. S1B), and Tyr124 and Asn129 with FSH-α (PDB 4mqw, chain D). Due to the lack of HS residues in FSH, the energy of binding is spread along the FSH interacting surface, including regions of both subunits. In these conditions, it would be difficult to design a small peptide mimicking FSH effects or even to obtain a strong antagonist if the entire interface (BSA of FSH-α and FSH-β) is not somehow included in the synthetic peptide.

12. Free energy prediction from BSA assuming additivity without interactions, as a first approach.

From equation [8], [9] and [10],

$$\Delta G°(x) = \gamma' BSA(x) \quad [11] \quad \text{Model 2}$$

This is a linear function, a line in the graph ΔG° vs BSA with ordinate zero at the origin and slope $\gamma' = -0.00726$ kcal mol^{-1} Å^{-2} for FSH, as shown in Figure 9. The closed triangles represent the predicted ΔG° values from the BSA values of FSH, FSH-β, and the different FSH-β peptides shown in Table 3, using Model 2 (equation 11). The actual experimental values are also shown in Figure 9, as closed circles.

Thus, for each binding region of FSH-β, such as A (33-53) and B (81-95), we can use the corresponding γ′.BSA(x) value to estimate $\Delta G°_i$, and compare them with the experimental values. This is equivalent to using equation [2] $\Delta G° = \sum_i \Delta G°_i$ or [3],

$K_a\ (total) = \prod_i K_{a_i}$. As shown in Figure 9 and Table 3, the $\Delta G°_i$ values predicted for the synthetic peptides by using this approximation (Model 2) are far from the observed values. The differences increase at lower BSA values. This is in agreement with the results already shown for the joint peptide FSH-β-(33-53)-(81-95), in which the observed value was far from the predicted value of $K_{aA}.K_{aB}$ (Model 1).

The problem here is that equation [11] only works when all the interactions are additive in ΔG° terms and the nonadditive interactions can be neglected. This is the case of ligands and receptors without strong conformational changes upon binding or without additional entropic effects. In other words, when $\Delta G°_{na} = 0$. However, this is true only if the relationship is linear, additive, with ordinate at the origin zero (or near to zero), as occurs with the proteins and small ligands of Figure 29 in the work of Houk et al. [127]. If nonadditive interactions are present, the system is no longer linear, as occurs when synthetic peptide fragments are used, or when the ligand or the receptor suffer strong conformational changes upon binding. In those cases, the system is rather represented by an affine function with a smaller slope and ordinate at the origin different from zero (equation [7]), as shown in Figure 9 (Model 3). This is due to the strong hydration and diversity of conformations that the synthetic peptides have in solution, which imply additional hydrophobic effects and significant conformational entropy changes upon binding. Other effects can also contribute to $\Delta G°_{na}$ such as disulfide bridges and long-distance interactions. It is interesting to note here that the model of Chothia and Janin [125] does not apply for proteins in which strong conformational changes are present upon binding [128], for the same reasons: those systems are not linear and the contributions of the nonadditive interactions, $\Delta G°_{na}$, cannot be neglected.

This lack of additivity expressed in the ordinate at the origin different from zero and in the reduced slope, explains why the predicted $K_{AB} = K_A.K_B$ value (Model 1 and 2) was

far from the observed value in the peptide FSH-β-(33-53)-(81-95). In the next section, we will consider nonadditive effects, which gives a much better approximation to correctly predict the observed $\Delta G°_i$ values.

13. Free energies of the individual binding regions ($\Delta G°_i$) taken into account the nonadditive contributions ($\Delta G°_{na}$)

As already mentioned, when the observed values (ob, closed circles) were added to Figure 9, a different slope γ'' was noted. It reflects the additional conformational entropy of the free peptides in solution, which are only partially in their native states and partially unfolded and hydrated. According to equation [7], this corresponds to an affine function with slope γ'' and ordinate at the origin $\Delta G°_{na}$,

$$\Delta G°_{ob}(x) = \gamma'' BSA(x) + \Delta G°_{na} + \varepsilon(x)$$

Where $\varepsilon(x)$ is the error or the vertical distance from each point to the fitted linear line, which corresponds to the interactions and properties of each residue not considered in the model. The fitted linear equation corresponding to the observed values (Fig. 9A) is:

$$\Delta G°(x) = \gamma'' BSA(x) + \Delta G°_{na} \quad [12] \quad \text{Model 3}$$

The ordinate at the origin accounts for the interactions that do not have a dependency on BSA, plus all the nonadditive interactions (conformational entropy, hydration, etc.), including entropy changes in the entire system. Thus, the predicted $\Delta G°$ for a protein/peptide region from residue *i* to residue *j* is,

$$\Delta G° = \sum_i^j \gamma'' BSA_i + \Delta G°_{na} = \sum_i^j \Delta G°_i + \Delta G°_{na} \quad [13]$$

where $\gamma''$is the predicted slope from equation [12], and γ'' BSA_i are the predicted $\Delta G°_i$ values for each residue in the region, from *i* to *j*; $\Delta G°_{na}$ corresponds to the ordinate at the origin of the fitted line [12]. Thus, Model 3 is similar to Model 2 but with a different slope, γ'', and the addition of the term $\Delta G°_{na}$. In other words, we can calculate the free energy of a fragment as the sum of each calculated additive free energy value $\Delta G°_i$, adding the term $\Delta G°_{na}$ at the end (equation [13]). Figure 9B and Table S3 show the predicted $\Delta G°_i$values for each FSH-β residue.

According to the PISA server data (Appendix 3, suppl. file), in the complex FSH::FSHR, the buried surface area at the interface is BSA = 3608 ± 167 (n=3) $Å^2$; in FSH alone is 1882 ± 91 (n=3) $Å^2$, and in FSH-β 852 ± 37 (n=3) $Å^2$; these values correspond to those shown in the PISA server data for each chain (Table S2), averaged. The BSA for each peptide used in the regression was calculated as the mean values of the PISA server BSA values for each amino acid residue of the 3 chains in the 3D structure 4MQW (interfaces #4 to #9 in the PISA server, values shown in Table S2, suppl. file).The association constant for FSH is K_a = 1.02 x 10^{10} mol^{-1} L [130], and for FSH-β is K_a = 1.1 x 10^7 mol^{-1} L [23].

Fitting the observed values of Figure 9 (and Table 3) to a line, we obtained a new slope γ´´ and ordinate to the origin $\Delta G°_{na}$, in agreement with equation [12], Model 3:

$$\Delta G°(x) = (-0.00587\ BSA(x) - 3.125)\ kcal\ mol^{-1} \quad [14] \quad \text{Model 3}$$

where BSA is the buried surface area (ligand only) at the interface. This affine function [14] gives a more accurate prediction of ΔG° for each binding region than equation [11] (ΔG°(x) = -0.00726 * BSA(x)), which is a linear function that assumes additivity in all the interactions, without including nonadditive interactions. The results show that the observed free energy for the joint peptide FSH-β-(33-53)-(81-.95), $\Delta G°_{ob}$ = -5.9 kcal mol^{-1}, was not

that far from the predicted value $\Delta G° = -6.4$ kcal mol^{-1} (Table 3 and Fig. 9, Model 3). In terms of K_a, 2 x 10^4 mol^{-1} L (observed value) and 5.1 x 10^4 mol^{-1} L (predicted value with Model 3) respectively. Before, without considering the nonadditive interactions, the predicted value was K_a = 9.6 x 10^2 mol^{-1} L (predicted values with Model 2), and 2.5 x 10^7 mol^{-1} L (model 1). Thus, linear regression [12] (or [14] in numbers, model 3), provides a better prediction for the observed $\Delta G°$ values.

Since the difference between the K_a 5.1 x 10^4 mol^{-1} L and the K_a of FSH-β 1.1 x 10^7 mol^{-1} L was yet of three orders of magnitude, the rest of the sequence was analyzed to see if some other regions could contribute to binding. As already mentioned, examining the BSA values of Table 2 it was observed that in the joint peptide FSH-β-(33-53)-(81-95) an important region of the BSA was missing, corresponding to amino acids (96-99). These residues add an average of 238 $Å^2$ to the BSA. Thus, by adding the peptides (96-99) and (100-106) and reducing (33-53) to (34-47), the peptide FSH-β-(34-47)-(88-106), according to the regression [14](Model 3), we might have a theoretical affinity of 0.9 x 10^6 mol^{-1} L, which is now closer to the value of FSH-β, 1.1 x 10^7 mol^{-1} L (Table 3). This predicted value will be even closer to the observed value if we force the fitted line to pass through the value of FSH-β by using a weighted regression (not shown).

This illustrates the usefulness of the $\Delta G°_i$ predictions. Interestingly, Prabhudesai et al. [66] tested the peptide FSH-(89-97), which inhibits FSH binding but still misses the residues 98 and 99 that should provide additional 98 $Å^2$ to the interface and an increased affinity. By using this empirical method, we can explore the entire molecule searching for the smallest regions with the higher affinity, which will be in principle that with the highest buried surface area (Fig. 9B and Table S3). In conclusion, a peptide FSH-β-(34-47)-(88-106), as shown in Table 3, might behave similarly to FSH-β, provided that the correct

conformation can be achieved by using an optimized bridge between residues FSH-β-(34-47) and FSH-β-(88-106).

Thus, the ΔG additivity allows us to predict the affinities of the individual amino acids residues and binding regions, as long as the BSA of the interface can be calculated and we have some experimental data with synthetic peptides to estimate γ´´ and the ordinate at the origin from the regression ΔG° vs BSA (Fig. 9A). It would be especially useful to have a program similar to ISLAND, but able to calculate the affinity of the different synthetic peptides or binding regions. However, such a program should consider nonadditive effects (i.e. conformational entropy), which is not the case so far.

In recent years, molecular dynamic methods have also allowed the calculation of binding affinities; these methods require access to supercomputing capabilities [106, 134, 135]. It would be interesting to test their capabilities to estimate the binding affinity of synthetic peptides corresponding to different binding regions of a protein in solution. In this way, we could obtain a simulated regression similar to $\Delta G° = -0.00587 * BSA - 3.125$ kcal mol^{-1} without the need for experimentally obtained results. However, we should keep in mind the limitations of these approaches since, as was discussed, the entropy of the entire system cannot be decomposed into individual contributions, and the nonadditive effect must be considered. The nonadditivity is now an emerging issue [136-138], and it is a problem without a solution, as yet. Until the nonadditivity issues are solved, the use of synthetic overlapping peptides could be of major help to better understand these systems a to develop ligand agonists or antagonists. More accurate molecular simulations are needed, able to calculate the individual free energies of each anchoring region in the native protein and the free energy of interaction of corresponding synthetic peptides in solution, including enthalpic and entropic terms and nonadditive interactions, even in the absence of X-ray data.

14. Concluding remarks

Remarks

- The earlier mapping of regions FSH-β-(33-53) and FSH-β-(81-95) as FSH receptor-binding regions, by using overlapping synthetic peptides, was correct.
- Solid-phase assays should be used with caution to map binding regions due to possible steric hindrance in the interactions, heterogeneity in the attachment and other issues. The soluble assays should also be used with caution due to restrictions to adopt the correct conformation or due insolubility difficulties with certain synthetic peptides. Ideally, both methods should be used to confirm the results.
- The presence of free sulfhydryl groups in synthetic peptides may produce nonspecific biological responses due to their redox activity. Also, the concentration might be reduced due to polymerization by oxidation of the cysteine residues. Ser analogs may solve these problems in cases where Cys residues are not mandatory for an effective interaction or functional role.
- The redox activity of synthetic peptides not only depends on the number of Cys residues. The structure, flexibility and surface probability also influence the redox activity. This could be used to modulate the redox activity of synthetic peptides or to design specific redox buffers.
- Improved affinities using native sequences may be obtained by simultaneously using peptides corresponding to different anchor sites, and even more if the binding regions can be joined by using an optimized bridge or scaffold.
- Due to the lack of a unique hot spot region in FSH, the ΔG° energy of binding is spread on the FSH molecule, and it would not be possible to design a minimal peptide with an affinity equivalent to the entire FSH, at least with the native sequence.
- As a first approximation, the free energy of binding can be decomposed as the sum of the energies of each binding region, which is reflected in the product of their affinity constants. However, a high free energy of non-additive interactions, $\Delta G°_{na}$, can significantly affect these values.
- If the 3D structure is known or can be predicted, a linear regression $\Delta G°(x) = \gamma' BSA(x)$ can be used to predict the affinity of new designed ligands covering different areas of the binding interface. A more precise linear regression can be obtained if empirical ΔG° data are measured by using synthetic peptides covering fragments from different binding regions and BSA values, allowing estimation of the free energy due to nonadditive or BSA-independent interactions, $\Delta G°_{na}$. For FSH fragments, $\Delta G°_{na}$ is near -3 kcal mol^{-1},

$$\Delta G°(x) = \gamma'' BSA(x) + \Delta G°_{na}$$

$$\Delta G°(x) = (-0.00587 \text{ x } BSA(x) - 3.125) \text{ kcal mol}^{-1}$$

The results described here show that the use of synthetic overlapping peptides is a useful approach to map protein-protein interacting regions and to find possible receptor antagonists/agonists. The originally predicted FSH-β binding regions were later found on the FSH 3D structure, creating an almost continuous binding surface. Later, the 3D structure of the FSH::FSHR co-crystal showed that certainly FSH-β-(33-53) and FSH-β-(81-95) include the most relevant receptor-binding regions of FSH-β.

ΔG° values suggest that the entire ligand molecule is needed for fully productive interaction, selectivity, and signaling. Nevertheless, small peptides can function as antagonists provided that enough concentration is used, and several anchor sites are included in the synthetic peptide. The ΔG° values predict that it would not be possible to obtain a synthetic peptide with an affinity similar to the native ligand if only a few regions of the ligand are included; the entire interface is needed to obtain a full affinity, comparable to the native ligand.

Improved results with the fragments of the native sequences may be obtained by simultaneously using peptides corresponding to different anchor sites or including the different anchor sites in the same peptide. However, if the binding region of the ligand is highly structured (i.e. mostly alpha-helix), and covers most of the binding interface, as occurs with PTH analogs [139], then the receptor could be fully activated by a fragment. This is not the case with the complex heterodimeric FSH molecule, which does not have a unique hot spot region and the free energy density is low. On the other hand, the FSHR does have six HS residues, interacting with FSH-α and FSH-β, which should be considered (Appendix 6, suppl. file). There is still much to learn about synthetic peptides able to mimic the FSH-β subunit or other ligands. For example, the region FSH-β-(81-95) might be extended to cover the entire interface of interaction; the resulting FSH-β-(33-53)-(81-99)

or (34-47)-(88-106) joint peptides might provide increased affinities, and could be even improved by using an optimized bridge/linker between these two binding regions [86].

The affinity of the fragments can be predicted by using the BSA, provided that the 3D structure has been determined or predicted by homology modeling or similar approaches. The linear function $\Delta G° = \gamma$. BSA is only a gross approximation when considering fragments of a ligand, or proteins/receptors that undergo extensive conformational changes upon binding. The $\Delta G°$ prediction can be improved if experimental K_a values are obtained for the entire ligand and different synthetic peptides, that can be used to estimate γ'' and $\Delta G°_{na}$ from the function $\Delta G°(x) = \gamma'' BSA(x) + \Delta G°_{na}$ (equation [12]).

It would be interesting to see if the value of $\Delta G°_{na} = -3.12$ kcal mol^{-1}, corresponding to BSA = 0 in equation [12], is similar when it is measured using synthetic peptides from different ligand-receptor complexes. Determination of the nonadditive contributions that are expressed in $\Delta G°_{na}$ needs empirical values, by now. Machine learning, molecular dynamics (MD), and other approaches that rely on free energy additivity are remarkably good in calculating the free energy of the entire system. However, these methods cannot predict the binding affinity of the individual regions in solution with accuracy, unless the nonadditivity effects are considered [136-138]. In the future, improved algorithms might allow predicting $\Delta G°$ values for small fragments and $\Delta G°_{na}$ values without the need for experimental data.

The empirical method shown here was focused on the design of peptides with better affinity, independently of their possible agonist or antagonist activity. To predict how the synthetic peptides will behave regarding their activity as agonists or antagonists is far more complex. In this regard, considering the design of antagonists, it was mentioned that high

flexibility may be important since it will minimize a possible induced fit in the receptor, followed by receptor activation. As was mentioned, serine analogs may be used in the design of FSH antagonists to avoid possible nonspecific redox activities that can reduce the effective concentration, due to a random cross-linking [21], and produce some agonist activity. However, it should also be considered that in some cases Cys residues may be key residues for binding interactions and structural stability; free cysteines can also add functional negative charges besides possible cross-linking reactions, which may be important in ligand-receptor interactions, and in particular for the FSH-FSHR interaction [21, 83, 84]. Regarding the design of agonists, the problem with FSH is that the agonist activity of its synthetic peptides might be the result of non-specific redox activity of the Cys residues since the Ser analogs showed no agonist activity. Perhaps the use of more Cys residues and a more rigid molecule might help in obtaining a better agonist response, but before considering how to improve this agonist response, we need to know the precise role of the redox activity of each Cys residue and the mechanism by which this redox activity activates the FSH-like response in cells.

The analysis or design of synthetic peptides has been described in detail in the earlier work by William F. Degrado [140], and there are many interesting reviews now on the subject [141-144]. Once the binding regions are identified, the affinity of the synthetic peptides could be increased by adding the different interacting regions in one peptide, optimizing the bridges, and introducing chemical modifications to optimize $\Delta G°$, $\Delta H°$, or $\Delta S°$.

Conflict of interest:

The author declares no conflicts of interest.

Data and Software Availability

Data: All the data used is available. The data have been previously published or obtained from the online servers indicated in the text and added to suppl. file. If further help is needed it will be provided by the author.

Software: All software used is from third parties.

-Microsoft Office 365 with Word and Excel (microsoft.com).

-ESPript 3.0 (espript.ibcp.fr). Extracting and rendering sequence and 3D information from atomic structures of proteins. Online free software[145].

-Open-source PyMol version 2.5, running under Python 3.8 for Windows (pymolwiki.org; pymol.org). Molecular visualization system. Free software [146].

-GIMP 2.10 (gimp.org). To compose and final edit graphics as tiff files. Free software [147].

-RStudio 2021.09.1 Build 372 (rstudio.com). Integrated Development Environment for R. Free software [148].

-LigPlot+ v.2.2.5 (ebi.ac.uk/thornton-srv/software/LigPlus). Automatic generation of 2D ligand-protein interaction diagrams. Free academic version[149].

-CABSFlex 2.0 (biocomp.chem.uw.edu.pl/CABSflex2). Automatically predict and visualize protein structure dynamics Free online server [105].

-SpotOn (alcazar.science.uu.nl). High Accuracy Identification of Protein-Protein Interface Hot-Spots [132, 133].

Acknowledgments

The administrative help of Diego Battiato and Romina D'Agostino is acknowledged. This work was supported by The National Scientific and Technical

Research Council of Argentina (CONICET) [grants numbers: PIP 2015-2017 11220150100227CO and PUE 2016 22920160100129CO], The National Agency for the Promotion of Science and Technology (ANPCYT) [grants number: PICT 2018-04429], and the Pontifical Catholic University of Argentina (UCA).

References

1. Kucinskaite-Kodze, I.; Simanavicius, M.; Dapkunas, J.; Pleckaityte, M.; Zvirbliene, A., Mapping of Recognition Sites of Monoclonal Antibodies Responsible for the Inhibition of Pneumolysin Functional Activity. *Biomolecules* **2020**, 10, 1009.
2. Brunetti, J.; Falciani, C.; Bernini, A.; Scali, S.; Bracci, L.; Lozzi, L., Molecular definition of the interaction between a tumor-specific tetrabranched peptide and LRP6 receptor. *Amino Acids* **2020**, 52, 915-924.
3. Zhao, R.; Xiao, Q.; Li, M.; Ren, W.; Xia, C.; Liu, X.; Li, Y.; Tan, T.; Wu, D.; Sun, L., Rational design of peptides for identification of linear epitopes and generation of neutralizing monoclonal antibodies against DKK2 for cancer therapy. *Antibody Therapeutics* **2020**, 3, 63-70.
4. Tribbick, G., Multipin peptide libraries for antibody and receptor epitope screening and characterization. *J. Immunol. Methods* **2002**, 267, 27-35.
5. Meyer, K.; Selbach, M., Peptide-based Interaction Proteomics. *Mol. Cell. Proteomics* **2020**, 19, 1070-1075.
6. Groß, A.; Hashimoto, C.; Sticht, H.; Eichler, J., Synthetic Peptides as Protein Mimics. *Frontiers in bioengineering and biotechnology* **2016**, 3, 211-211.
7. Reichert, L. E., Jr.; Dattatreyamurty, B.; Grasso, P.; Santa-Coloma, T. A., Structure--function relationships of the glycoprotein hormones and their receptors. *Trends Pharmacol. Sci.* **1991**, 12, 199-203.

8. Costagliola, S.; Urizar, E.; Mendive, F.; Vassart, G., Specificity and promiscuity of gonadotropin receptors. *Reproduction* **2005**, 130, 275-81.

9. Banerjee, A. A.; Joseph, S.; Mahale, S. D., From cell surface to signalling and back: the life of the mammalian FSH receptor. *FEBS J.* **2020**.

10. Pierce, J. G.; Faith, M. R.; Giudice, L. C.; Reeve, J. R., Structure and structure-function relationships in glycoprotein hormones. *Ciba Found. Symp.* **1976**, 41, 225-50.

11. Ryan, R. J.; Charlesworth, M. C.; McCormick, D. J.; Milius, R. P.; Keutmann, H. T., The glycoprotein hormones: recent studies of structure-function relationships. *FASEB J.* **1988**, 2, 2661-9.

12. Fox, K. M.; Dias, J. A.; Van Roey, P., Three-dimensional structure of human follicle-stimulating hormone. *Mol. Endocrinol.* **2001**, 15, 378-89.

13. Fan, Q. R.; Hendrickson, W. A., Assembly and structural characterization of an authentic complex between human follicle stimulating hormone and a hormone-binding ectodomain of its receptor. *Mol. Cell. Endocrinol.* **2007**, 260-262, 73-82.

14. Fan, Q. R.; Hendrickson, W. A., Comparative structural analysis of the binding domain of follicle stimulating hormone receptor. *Proteins* **2008**, 72, 393-401.

15. Fan, Q. R.; Hendrickson, W. A., Structure of human follicle-stimulating hormone in complex with its receptor. *Nature* **2005**, 433, 269-77.

16. Jiang, X.; Fischer, D.; Chen, X.; McKenna, S. D.; Liu, H.; Sriraman, V.; Yu, H. N.; Goutopoulos, A.; Arkinstall, S.; He, X., Evidence for Follicle-stimulating Hormone Receptor as a Functional Trimer. *J. Biol. Chem.* **2014**, 289, 14273-82.

17. Jiang, X.; Dreano, M.; Buckler, D. R.; Cheng, S.; Ythier, A.; Wu, H.; Hendrickson, W. A.; el Tayar, N., Structural predictions for the ligand-binding region of

glycoprotein hormone receptors and the nature of hormone-receptor interactions. *Structure* **1995**, 3, 1341-53.

18. Jiang, X.; Liu, H.; Chen, X.; Chen, P. H.; Fischer, D.; Sriraman, V.; Yu, H. N.; Arkinstall, S.; He, X., Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109, 12491-6.
19. Ulloa-Aguirre, A.; Zarinan, T.; Jardon-Valadez, E.; Gutierrez-Sagal, R.; Dias, J. A., Structure-Function Relationships of the Follicle-Stimulating Hormone Receptor. *Front. Endocrinol. (Lausanne)* **2018**, 9, 707.
20. Reichert, L. E.; Grasso, P.; Boniface, J. J.; Santa Coloma, T. A.; Zhang, S.-B., Studies on the mechanism of follitropin-receptor intreraction and signal transduction in testis In: Structure-function relationship of gonadotropins. *Serono symposia publications from Raven Press* **1989**, 65, 237-246.
21. Santa-Coloma, T. A.; Dattatreyamurty, B.; Reichert, L. E., Jr., A synthetic peptide corresponding to human FSH beta-subunit 33-53 binds to FSH receptor, stimulates basal estradiol biosynthesis, and is a partial antagonist of FSH. *Biochemistry* **1990**, 29, 1194-200.
22. Santa-Coloma, T. A.; Reichert, L. E., Jr., Identification of a follicle-stimulating hormone receptor-binding region in hFSH-beta-(81-95) using synthetic peptides. *J. Biol. Chem.* **1990**, 265, 5037-42.
23. Santa-Coloma, T. A.; Crabb, J. W.; Reichert, L. E., Jr., A synthetic peptide encompassing two discontinuous regions of hFSH-beta subunit mimics the receptor binding surface of the hormone. *Mol. Cell. Endocrinol.* **1991**, 78, 197-204.
24. Leng, N.; Grasso, P.; Deziel, M. R.; Reichert, L. E., Jr., A synthetic peptide corresponding to glycoprotein hormone alpha subunit residues 32-46 inhibits gonadotropin binding to receptor. *Pept. Res.* **1995**, 8, 272-7.

25. Sluss, P. M.; Krystek, S. R., Jr.; Andersen, T. T.; Melson, B. E.; Huston, J. S.; Ridge, R.; Reichert, L. E., Jr., Inhibition of iodine-125-labeled human follitropin binding to testicular receptor by epidermal growth factor and synthetic peptides. *Biochemistry* **1986**, 25, 2644-9.

26. Gibbs, A. J.; McIntyre, G. A., The diagram, a method for comparing sequences. Its use with amino acid and nucleotide sequences. *Eur. J. Biochem.* **1970**, 16, 1-11.

27. Krystek, S. R., Jr.; Reichert, L. E., Jr.; Andersen, T. T., Analysis of computer-generated hydropathy profiles for human glycoprotein and lactogenic hormones. *Endocrinology* **1985**, 117, 1110-24.

28. Krystek, S. R., Jr.; Dias, J. A.; Reichert, L. E., Jr.; Andersen, T. T., Prediction of antigenic sites in follicle-stimulating hormones: difference profiles enhance antigenicity prediction methods. *Endocrinology* **1985**, 117, 1125-31.

29. Santa Coloma, T. A.; Reichert, L. E., In; 1989, p 166.

30. Casarini, L.; Crépieux, P., Molecular Mechanisms of Action of FSH. *Front. Endocrinol. (Lausanne)* **2019**, 10.

31. Queipo, M. J.; Gil-Redondo, J. C.; Morente, V.; Ortega, F.; Miras-Portugal, M. T.; Delicado, E. G.; Pérez-Sen, R., P2X7 Nucleotide and EGF Receptors Exert Dual Modulation of the Dual-Specificity Phosphatase 6 (MKP-3) in Granule Neurons and Astrocytes, Contributing to Negative Feedback on ERK Signaling. *Front. Mol. Neurosci.* **2017**, 10, 448.

32. Andersen, T.; Sluss, P.; Quirk, S.; Krystek Jr, S.; Reichert Jr, L., Abstract 235. Inhibition of 125l-hFSH binding to receptor by synthetic peptide fragments of human FSH-beta subunit. Annual meeting of the Society for Study of Reproduction. 20–23 July, University of Illinois at Urbana–Champaign, Urbana, Illinois, USA. *Biol. Reprod.* **1987**, 36, 125.

33. Charlesworth, M. C.; McCormick, D. J.; Madden, B.; Ryan, R. J., Inhibition of human choriotropin binding to receptor by human choriotropin alpha peptides. A comprehensive synthetic approach. *J. Biol. Chem.* **1987**, 262, 13409-16.

34. Morris, J. C., 3rd; Jiang, N. S.; Charlesworth, M. C.; McCormick, D. J.; Ryan, R. J., The effects of synthetic alpha-subunit peptides on thyrotropin interaction with its receptor. *Endocrinology* **1988**, 123, 456-62.

35. Erickson, L. D.; Rizza, S. A.; Bergert, E. R.; Charlesworth, M. C.; McCormick, D. J.; Ryan, R. J., Synthetic alpha-subunit peptides stimulate testosterone production in vitro by rat Leydig cells. *Endocrinology* **1990**, 126, 2555-60.

36. Strynadka, N. C.; Redmond, M. J.; Parker, J. M.; Scraba, D. G.; Hodges, R. S., Use of synthetic peptides to map the antigenic determinants of glycoprotein D of herpes simplex virus. *J. Virol.* **1988**, 62, 3474-83.

37. Ward, D. N.; Moore, W. T., In; Harper and Row, Baltimore, MD: 1979.

38. Santa-Coloma, T. A.; Reichert, L. E., Jr., Determination of alpha-subunit contact regions of human follicle-stimulating hormone beta-subunit using synthetic peptides. *J. Biol. Chem.* **1991**, 266, 2759-62.

39. Lee, Y.; Kang, D. K.; Chang, S. I.; Han, M. H.; Kang, I. C., High-throughput screening of novel peptide inhibitors of an integrin receptor from the hexapeptide library by using a protein microarray chip. *J. Biomol. Screen.* **2004**, 9, 687-94.

40. Andresen, H.; Grötzinger, C.; Zarse, K.; Birringer, M.; Hessenius, C.; Kreuzer, O. J.; Ehrentreich-Förster, E.; Bier, F. F., Peptide microarrays with site-specifically immobilized synthetic peptides for antibody diagnostics. *Sens. Actuators B Chem.* **2006**, 113, 655-663.

41. Frank, R., High-density synthetic peptide microarrays: emerging tools for functional genomics and proteomics. *Comb. Chem. High Throughput Screen.* **2002**, 5, 429-40.

42. Foong, Y. M.; Fu, J.; Yao, S. Q.; Uttamchandani, M., Current advances in peptide and small molecule microarray technologies. *Curr. Opin. Chem. Biol.* **2012**, 16, 234-42.

43. Thiele, A.; Stangl, G. I.; Schutkowski, M., Deciphering enzyme function using peptide arrays. *Mol. Biotechnol.* **2011**, 49, 283-305.

44. Köhn, M., Immobilization strategies for small molecule, peptide and protein microarrays. *J Pept Sci* **2009**, 15, 393-7.

45. Uttamchandani, M.; Yao, S. Q., Peptide microarrays: next generation biochips for detection, diagnostics and high-throughput screening. *Curr. Pharm. Des.* **2008**, 14, 2428-38.

46. Winkler, D. F.; McGeer, P. L., Protein labeling and biotinylation of peptides during spot synthesis using biotin p-nitrophenyl ester (biotin-ONp). *Proteomics* **2008**, 8, 961-7.

47. Chersi, A.; Sezzi, M. L.; Romano, T. F.; Evangelista, M.; Nista, A., Preparation and utilization of fluorescent synthetic peptides. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1990**, 1034, 333-336.

48. Agris, P. F.; Guenther, R. H.; Sierzputowska-Gracz, H.; Easter, L.; Smith, W.; Hardin, C. C.; Santa-Coloma, T. A.; Crabb, J. W.; Reichert, L. E., Jr., Solution structure of a synthetic peptide corresponding to a receptor binding region of FSH (hFSH-beta 33-53). *J. Protein Chem.* **1992**, 11, 495-507.

49. Zerbe, K.; Moehle, K.; Robinson, J. A., Protein Epitope Mimetics: From New Antibiotics to Supramolecular Synthetic Vaccines. *Acc. Chem. Res.* **2017**, 50, 1323-1331.

50. Robinson, J. A., Max Bergmann lecture protein epitope mimetics in the age of structural vaccinology. *J Pept Sci* **2013**, 19, 127-40.

51. Santa-Coloma, T. A.; Crabb, J. W.; Reichert, L. E., Jr., Serine analogues of hFSH-beta-(33-53) and hFSH-beta-(81-95) inhibit hFSH binding to receptor. *Biochem. Biophys. Res. Commun.* **1992**, 184, 1273-9.

52. Keutmann, H. T.; Mason, K. A.; Kitzmann, K.; Ryan, R. J., Role of the beta 93-100 determinant loop sequence in receptor binding and biological activity of human luteinizing hormone and chorionic gonadotropin. *Mol. Endocrinol.* **1989**, 3, 526-31.

53. Keutmann, H. T.; Charlesworth, M. C.; Mason, K. A.; Ostrea, T.; Johnson, L.; Ryan, R. J., A receptor-binding region in human choriogonadotropin/lutropin beta subunit. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, 2038-42.

54. Roth, K. E.; Dias, J. A., Scanning-alanine mutagenesis of long loop residues 33-53 in follicle stimulating hormone beta subunit. *Mol. Cell. Endocrinol.* **1995**, 109, 143-9.

55. Roth, K. E.; Dias, J. A., Follitropin conformational stability mediated by loop 2 beta effects follitropin-receptor interaction. *Biochemistry* **1996**, 35, 7928-35.

56. Campbell, R. K.; Dean-Emig, D. M.; Moyle, W. R., Conversion of human choriogonadotropin into a follitropin by protein engineering. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, 88, 760-4.

57. Dias, J. A.; Zhang, Y.; Liu, X., Receptor binding and functional properties of chimeric human follitropin prepared by an exchange between a small hydrophilic intercysteine loop of human follitropin and human lutropin. *J. Biol. Chem.* **1994**, 269, 25289-94.

58. Lindau-Shepard, B.; Roth, K. E.; Dias, J. A., Identification of amino acids in the C-terminal region of human follicle-stimulating hormone (FSH) beta-subunit involved in binding to human FSH receptor. *Endocrinology* **1994**, 135, 1235-40.

59. Dias, J. A., Human follitropin heterodimerization and receptor binding structural motifs: identification and analysis by a combination of synthetic peptide and mutagenesis approaches. *Mol. Cell. Endocrinol.* **1996**, 125, 45-54.

60. Schmidt, A.; MacColl, R.; Lindau-Shepard, B.; Buckler, D. R.; Dias, J. A., Hormone-induced conformational change of the purified soluble hormone binding domain of follitropin receptor complexed with single chain follitropin. *J. Biol. Chem.* **2001**, 276, 23373-81.

61. Krystek, S. R., Jr.; Dias, J. A., Glycoprotein hormones tied but not tethered like other cysteine-knot cytokines. *Trends Pharmacol. Sci.* **2005**, 26, 439-42.

62. Sohn, J. Identification and Characterization of Contact Sites Between Human Follicle Stimulating Hormone and the Follicle Stimulating Hormone Receptor. University of Kentucky Doctoral Dissertations 270 (2005). 2005.

63. Sohn, J.; Youn, H.; Jeoung, M.; Koo, Y.; Yi, C.; Ji, I.; Ji, T. H., Orientation of follicle-stimulating hormone (FSH) subunits complexed with the FSH receptor. Beta subunit toward the N terminus of exodomain and alpha subunit to exoloop 3. *J. Biol. Chem.* **2003**, 278, 47868-76.

64. Sohn, J.; Ryu, K.; Sievert, G.; Jeoung, M.; Ji, I.; Ji, T. H., Follicle-stimulating hormone interacts with exoloop 3 of the receptor. *J. Biol. Chem.* **2002**, 277, 50165-75.

65. Costagliola, S.; Panneels, V.; Bonomi, M.; Koch, J.; Many, M. C.; Smits, G.; Vassart, G., Tyrosine sulfation is required for agonist recognition by glycoprotein hormone receptors. *The EMBO journal* **2002**, 21, 504-513.

66. Prabhudesai, K. S.; Raje, S.; Dhamanaskar, A.; Modi, D.; Dighe, V.; Contini, A.; Idicula-Thomas, S., Identification and in vivo validation of a 9-mer peptide derived from FSHβ with FSHR antagonist activity. *Peptides* **2020**, 170367.

67. Prabhudesai, K. S.; Raje, S.; Desai, K.; Modi, D.; Dighe, V.; Contini, A.; Idicula-Thomas, S., Central residues of FSHβ (89-97) peptide are not critical for FSHR binding: Implications for peptidomimetic design. *Bioorg. Med. Chem. Lett.* **2021**, 44, 128132.

68. Grasso, P.; Rozhavskaya, M.; Reichert, L. E., Jr., In vivo effects of human follicle-stimulating hormone-related synthetic peptide hFSH-beta-(81-95) and its subdomain hFSH-beta-(90-95) on the mouse estrous cycle. *Biol. Reprod.* **1998**, 58, 821-5.

69. Stewart, M.; Stewart, F., Constant and variable regions in glycoprotein hormone beta subunit sequences: implications for receptor binding specificity. *J. Mol. Biol.* **1977**, 116, 175-179.

70. Santa-Coloma, T. A.; Grasso, P.; Reichert, L. E., Jr., Synthetic human follicle-stimulating hormone-beta-(1-15) peptide-amide binds Ca2+ and possesses sequence similarity to calcium binding sites of calmodulin. *Endocrinology* **1992**, 130, 1103-7.

71. Reichert, L. E., Jr., The functional relationship between FSH and its receptor as studied by synthetic peptide strategies. *Mol. Cell. Endocrinol.* **1994**, 100, 21-7.

72. Ulloa-Aguirre, A.; Zariñán, T., The Follitropin Receptor: Matching Structure and Function. *Mol. Pharmacol.* **2016**, 90, 596-608.

73. Fischer, E., Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **1894**, 27, 2985-2993.

74. Tripathi, A.; Bankaitis, V. A., Molecular Docking: From Lock and Key to Combination Lock. *Journal of molecular medicine and clinical applications* **2017**, 2, 10.16966/2575-0305.106.

75. Feixas, F.; Lindert, S.; Sinko, W.; McCammon, J. A., Exploring the role of receptor flexibility in structure-based drug discovery. *Biophys. Chem.* **2014**, 186, 31-45.

76. Paul, F.; Weikl, T. R., How to Distinguish Conformational Selection and Induced Fit Based on Chemical Relaxation Rates. *PLoS Comp. Biol.* **2016**, 12, e1005067-e1005067.

77. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R., Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **2006**, 49, 534-553.

78. Carlson, H. A., Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, 6, 447-52.

79. Mueller, S.; Jaeschke, H.; Günther, R.; Paschke, R., The hinge region: an important receptor component for GPHR function. *Trends Endocrinol. Metab.* **2010**, 21, 111-22.

80. Winnett, G.; van Hagen, D.; Schrey, M., Prostaglandin J2 metabolites inhibit aromatase activity by redox-sensitive mechanisms: potential implications for breast cancer therapy. *Int. J. Cancer* **2003**, 103, 600-5.

81. Agrawal, G.; Dighe, R. R., Critical involvement of the hinge region of the follicle-stimulating hormone receptor in the activation of the receptor. *J. Biol. Chem.* **2009**, 284, 2636-47.

82. Boniface, J. J.; Reichert, L. E., Jr., Evidence for a novel thioredoxin-like catalytic property of gonadotropic hormones. *Science* **1990**, 247, 61-4.

83. Grasso, P.; Santa-Coloma, T. A.; Boniface, J. J.; Reichert, L. E., Jr., A synthetic peptide corresponding to hFSH-beta-(81-95) has thioredoxin-like activity. *Mol. Cell. Endocrinol.* **1991**, 78, 163-70.

84. Santa-Coloma, T. A.; Grasso, P.; Reichert, L. E., Jr., Sulfhydryl groups are involved in the interaction of FSH with its receptor. *Biochem. Biophys. Res. Commun.* **1991**, 176, 1256-61.

85. Cheng, K. W., Properties of follicle-stimulating-hormone receptor in cell membranes of bovine testis. *The Biochemical journal* **1975**, 149, 123-132.

86. Soler, M. A.; Fortuna, S., Influence of Linker Flexibility on the Binding Affinity of Bidentate Binders. *The Journal of Physical Chemistry B* **2017**, 121, 3918-3924.

87. Timmerman, P.; Puijk, W. C.; Meloen, R. H., Functional reconstruction and synthetic mimicry of a conformational epitope using CLIPS™ technology. *Journal of Molecular Recognition: An Interdisciplinary Journal* **2007**, 20, 283-299.

88. Schmidt, M.; Toplak, A.; Quaedflieg, P. J. L. M.; Ippel, H.; Richelle, G. J. J.; Hackeng, T. M.; van Maarseveen, J. H.; Nuijens, T., Omniligase-1: a powerful tool for peptide head-to-tail cyclization. *Advanced Synthesis & Catalysis* **2017**, 359, 2050-2055.

89. Kumar, E. A.; Chen, Q.; Kizhake, S.; Kolar, C.; Kang, M.; Chang, C.-e. A.; Borgstahl, G. E. O.; Natarajan, A., The paradox of conformational constraint in the design of Cbl(TKB)-binding peptides. *Sci. Rep.* **2013**, 3, 1639.

90. Amaral, M.; Kokh, D. B.; Bomke, J.; Wegener, A.; Buchstaller, H. P.; Eggenweiler, H. M.; Matias, P.; Sirrenberg, C.; Wade, R. C.; Frech, M., Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat Commun* **2017**, 8, 2276.

91. Tuffery, P.; Derreumaux, P., Flexibility and binding affinity in protein-ligand, protein-protein and multi-component protein interactions: limitations of current computational approaches. *Journal of the Royal Society, Interface* **2012**, 9, 20-33.

92. Forrey, C.; Douglas, J. F.; Gilson, M. K., The Fundamental Role of Flexibility on the Strength of Molecular Binding. *Soft matter* **2012**, 8, 6385-6392.

93. Khrapunov, S., The Enthalpy-entropy Compensation Phenomenon. Limitations for the Use of Some Basic Thermodynamic Equations. *Current protein & peptide science* **2018**, 19, 1088-1091.

94. Chodera, J. D.; Mobley, D. L., Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu Rev Biophys* **2013**, 42, 121-42.

95. Sharp, K., Entropy-enthalpy compensation: fact or artifact? *Protein Sci.* **2001**, 10, 661-7.

96. Eichler, J., Peptides as protein binding site mimetics. *Curr. Opin. Chem. Biol.* **2008**, 12, 707-13.

97. Demolombe, V.; de Brevern, A. G.; Molina, F.; Lavigne, G.; Granier, C.; Moreau, V., Benchmarking the PEPOP methods for mimicking discontinuous epitopes. *BMC Bioinformatics* **2019**, 20, 1-17.

98. Smeenk, L. E. J.; Timmers-Parohi, D.; Benschop, J. J.; Puijk, W. C.; Hiemstra, H.; van Maarseveen, J. H.; Timmerman, P., Reconstructing the Discontinuous and Conformational β1/β3-Loop Binding Site on hFSH/hCG by Using Highly Constrained Multicyclic Peptides. *ChemBioChem* **2015**, 16, 91-99.

99. Womble, D. D., GCG: The Wisconsin Package of sequence analysis programs. *Methods Mol. Biol.* **2000**, 132, 3-22.

100. Garnier, J.; Osguthorpe, D. J.; Robson, B., Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **1978**, 120, 97-120.

101. Kastritis, P. L.; Rodrigues, J. P. G. L. M.; Folkers, G. E.; Boelens, R.; Bonvin, A. M. J. J., Proteins Feel More Than They See: Fine-Tuning of Binding Affinity by Properties of the Non-Interacting Surface. *J. Mol. Biol.* **2014**, 426, 2632-2652.

102. Chen, J.; Sawyer, N.; Regan, L., Protein-protein interactions: general trends in the relationship between binding affinity and interfacial buried surface area. *Protein Sci.* **2013**, 22, 510-5.

103. Bogan, A. A.; Thorn, K. S., Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, 280, 1-9.

104. Ofran, Y.; Rost, B., Protein-protein interaction hotspots carved into sequences. *PLoS Comput. Biol.* **2007**, 3, e119.

105. Kuriata, A.; Gierut, A. M.; Oleniecki, T.; Ciemny, M. P.; Kolinski, A.; Kurcinski, M.; Kmiecik, S., CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res.* **2018**, 46, W338-W343.

106. Modi, V.; Lama, D.; Sankararamakrishnan, R., Relationship between helix stability and binding affinities: molecular dynamics simulations of Bfl-1/A1-binding pro-apoptotic BH3 peptide helices in explicit solvent. *J. Biomol. Struct. Dyn.* **2013**, 31, 65-77.

107. Yugandhar, K.; Gromiha, M. M., Protein-protein binding affinity prediction from amino acid sequence. *Bioinformatics* **2014**, 30, 3583-9.

108. Abbasi, W. A.; Yaseen, A.; Hassan, F. U.; Andleeb, S.; Minhas, F. U. A. A., ISLAND: in-silico proteins binding affinity prediction using sequence information. *BioData Min.* **2020**, 13, 20.

109. Abbasi, W. A.; Asif, A.; Ben-Hur, A.; Minhas, F., Learning protein binding affinity using privileged information. *BMC Bioinformatics* **2018**, 19, 425.

110. Vangone, A.; Schaarschmidt, J.; Koukos, P.; Geng, C.; Citro, N.; Trellet, M. E.; Xue, L. C.; Bonvin, A., Large-scale prediction of binding affinity in protein-small ligand complexes: the PRODIGY-LIG web server. *Bioinformatics* **2019**, 35, 1585-1587.

111. Skolnick, J.; Gao, M.; Zhou, H.; Singh, S., AlphaFold 2: Why It Works and Its Implications for Understanding the Relationships of Protein Sequence, Structure, and Function. *J. Chem. Inf. Model.* **2021**, 61, 4827-4831.

112. Weber, G., Energetics of ligand binding to proteins. *Adv. Protein Chem.* **1975**, 29, 1-83.

113. Weber, G., Ligand binding and internal equilibria in proteins. *Biochemistry* **1972**, 11, 864-78.

114. Jameson, D. M. The seminal contributions of Gregorio Weber to modern fluorescence spectroscopy. In *New Trends in Fluorescence Spectroscopy*; Springer: 2001, pp 35-58.

115. Weber, G., *Protein interactions*. Springer, Netherlands.: 1992.

116. Jencks, W. P., On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences* **1981**, 78, 4046-4050.

117. Wyman Jr, J., Linked functions and reciprocal effects in hemoglobin: a second look. *Adv. Protein Chem.* **1964**, 19, 223-286.

118. Dill, K. A., Additivity principles in biochemistry. *J. Biol. Chem.* **1997**, 272, 701-704.

119. Rashin, A. A., Buried surface area, conformational entropy, and protein stability. *Biopolymers* **1984**, 23, 1605-20.

120. Mark, A. E.; van Gunsteren, W. F., Decomposition of the free energy of a system in terms of specific interactions. Implications for theoretical and experimental studies. *J. Mol. Biol.* **1994**, 240, 167-76.

121. Cattoni, D. I.; Chara, O.; Kaufman, S. B.; González Flecha, F. L., Cooperativity in Binding Processes: New Insights from Phenomenological Modeling. *PLoS One* **2016**, 10, e0146043.

122. Lee, B.; Richards, F. M., The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **1971**, 55, 379-400.

123. Miller, S.; Lesk, A. M.; Janin, J.; Chothia, C., The accessible surface area and stability of oligomeric proteins. *Nature* **1987**, 328, 834-6.

124. Chothia, C., Hydrophobic bonding and accessible surface area in proteins. *Nature* **1974**, 248, 338-9.

125. Chothia, C.; Janin, J., Principles of protein-protein recognition. *Nature* **1975**, 256, 705-8.

126. Chothia, C., Structural invariants in protein folding. *Nature* **1975**, 254, 304-8.

127. Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X., Binding affinities of host–guest, protein–ligand, and protein–transition-state complexes. *Angewandte Chemie International Edition* **2003**, 42, 4872-4897.

128. Kastritis, P. L.; Bonvin, A. M., On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *J R Soc Interface* **2013**, 10, 20120835.

129. Vallone, B.; Miele, A. E.; Vecchini, P.; Chiancone, E.; Brunori, M., Free energy of burying hydrophobic residues in the interface between protein subunits. *Proceedings of the National Academy of Sciences* **1998**, 95, 6103-6107.

130. Dattatreyamurty, B.; Schneyer, A.; Reichert, L. E., Jr., Solubilization of functional and stable follitropin receptors from light membranes of bovine calf testis. *J. Biol. Chem.* **1986**, 261, 13104-13.

131. Bogan, A. A.; Thorn, K. S., Anatomy of hot spots in protein interfaces11Edited by J. Wells. *J. Mol. Biol.* **1998**, 280, 1-9.

132. Moreira, I. S.; Koukos, P. I.; Melo, R.; Almeida, J. G.; Preto, A. J.; Schaarschmidt, J.; Trellet, M.; Gümüş, Z. H.; Costa, J.; Bonvin, A., SpotOn: High Accuracy Identification of Protein-Protein Interface Hot-Spots. *Sci. Rep.* **2017**, 7, 8007.

133. Melo, R.; Fieldhouse, R.; Melo, A.; Correia, J. D. G.; Cordeiro, M. N. D. S.; Gümüş, Z. H.; Costa, J.; Bonvin, A. M. J. J.; Moreira, I. S., A Machine Learning Approach for Hot-Spot Detection at Protein-Protein Interfaces. *Int. J. Mol. Sci.* **2016**, 17, 1215.

134. Lazim, R.; Suh, D.; Choi, S., Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems. *Int. J. Mol. Sci.* **2020**, 21.

135. Wan, S.; Bhati, A. P.; Zasada, S. J.; Coveney, P. V., Rapid, accurate, precise and reproducible ligand-protein binding free energy prediction. *Interface Focus* **2020**, 10, 20200007.

136. Boulton, S.; Van, K.; VanSchouwen, B.; Augustine, J.; Akimoto, M.; Melacini, G., Allosteric Mechanisms of Nonadditive Substituent Contributions to Protein-Ligand Binding. *Biophys. J.* **2020**, 119, 1135-1146.

137. Calabrò, G.; Woods, C. J.; Powlesland, F.; Mey, A. S.; Mulholland, A. J.; Michel, J., Elucidation of Nonadditive Effects in Protein-Ligand Binding Energies: Thrombin as a Case Study. *J. Phys. Chem. B* **2016**, 120, 5340-50.

138. Kramer, C.; Fuchs, J. E.; Liedl, K. R., Strong nonadditivity as a key structure-activity relationship feature: distinguishing structural changes from assay artifacts. *J. Chem. Inf. Model.* **2015**, 55, 483-94.

139. Zhao, L.-H.; Ma, S.; Sutkeviciute, I.; Shen, D.-D.; Zhou, X. E.; de Waal, P. W.; Li, C.-Y.; Kang, Y.; Clark, L. J.; Jean-Alphonse, F. G.; White, A. D.; Yang, D.; Dai, A.; Cai, X.; Chen, J.; Li, C.; Jiang, Y.; Watanabe, T.; Gardella, T. J.; Melcher, K.; Wang,

M.-W.; Vilardaga, J.-P.; Xu, H. E.; Zhang, Y., Structure and dynamics of the active human parathyroid hormone receptor-1. *Science* **2019**, 364, 148-153.

140. Degrado, W. F., Design of peptides and proteins. *Adv. Protein Chem.* **1988**, 39, 51-124.
141. Ernst, P.; Plückthun, A., Advances in the design and engineering of peptide-binding repeat proteins. *Biol. Chem.* **2017**, 398, 23-29.
142. Hamley, I. W., Small Bioactive Peptides for Biomaterials Design and Therapeutics. *Chem. Rev.* **2017**, 117, 14015-14041.
143. Gutte, B.; Klauser, S., Design of catalytic polypeptides and proteins. *Protein Eng. Des. Sel.* **2018**, 31, 457-470.
144. Wang, G.; Narayana, J. L.; Mishra, B.; Zhang, Y.; Wang, F.; Wang, C.; Zarena, D.; Lushnikova, T.; Wang, X., Design of Antimicrobial Peptides: Progress Made with Human Cathelicidin LL-37. *Adv. Exp. Med. Biol.* **2019**, 1117, 215-240.
145. Gouet, P.; Robert, X.; Courcelle, E., ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **2003**, 31, 3320-3.
146. Schrodinger, LLC *The PyMOL Molecular Graphics System, Version 2.5 (Open-Source PyMol)*, 2020.
147. Team, T. G. D., GIMP, Available at: https://www.gimp.org. **2019**.
148. Team, R., RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA (rstudio.com). **2021**.
149. Laskowski, R. A.; Swindells, M. B., LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **2011**, 51, 2778-86.

Tables

Table 1: FSH-β buried residues in the interface FSH-α/FSH-β. All the residues in the interface FSH-α/FSH-β #1 are shown (yellow). Only a few residues with solvation energy effect = 0 (solvent-accessible residues) are shown since all have the same BSA value (zero)(blue). The full table can be found as supplementary data, Table S1, corresponding to the values of the PISA interfaces #1, 2 and 3 in the PDB 1FL7 crystal. HSDC: Residues making Hydrogen/Disulphide bond, Salt bridge or Covalent link; ASA: Accessible Surface Area (Å^2); BSA: Buried Surface Area (Å^2); Δ^iG: Solvation energy effect (kcal/mol); |||: Buried area percentage, one bar per 10%.

FSH-β	#	HSDC	ASA	BSA	BSA%	Δ^iG
B:ASN	7		113.67	0.00		0.00
B:ILE			48.10	12.55	|||	0.20
B:THR	9		54.17	0.00		0.00
B:ILE	10		25.45	20.60	||||||||	0.33
B:ALA	11		25.35	5.03	||	-0.06
B:ILE	12		3.92	0.00		0.00
B:GLU	13		58.21	33.41	||||||	-0.11
B:LYS	14		17.63	0.00		0.00
B:GLU	15	HS	103.95	38.92	||||	-0.01
B:GLU	16		88.03	0.00		0.00
B:ALA	26		23.88	0.00		0.00
B:TRP	27		52.70	52.70	||||||||||	0.77
B:CYS	28	H	24.69	23.33	|||||||||	-0.17
B:ALA	29		47.75	31.06	|||||||	0.50
B:GLY	30	H	47.27	34.22	|||||||	-0.24
B:TYR	31		194.21	84.14	|||||	1.12
B:CYS	32	H	43.32	37.55	|||||||||	-0.11
B:TYR	33		177.11	75.21	|||||	0.48
B:THR	34	H	69.70	65.94	|||||||||	-0.38
B:ARG	35		163.42	28.34	||	0.45
B:ASP	36	HS	120.19	98.82	|||||||||	-0.22
B:LEU	37		72.95	19.48	|||	0.28
B:VAL	38	H	154.01	150.22	|||||||||	1.72
B:TYR	39		177.47	80.89	|||||	1.11
B:LYS	40	HS	137.35	73.25	||||||	0.29
B:ASP	41		78.52	9.98	||	-0.11
B:PRO	42		126.53	0.00		0.00
B:ALA	43		72.57	0.00		0.00
B:ARG	44		204.84	50.00	|||	-0.19
B:PRO	45		89.48	19.89	|||	0.12
B:LYS	46		95.07	10.62	||	-0.08
B:ILE	47		106.25	27.94	|||	0.15
B:GLN	48		57.58	31.44	||||||	0.38
B:LYS	49		107.77	3.67			-0.04
B:THR	50	H	55.42	47.20	|||||||||	-0.05
B:CYS	51	H	32.55	32.55	|||||||||	-0.02
B:THR	52		18.18	15.35	|||||||||	0.09
B:PHE	53		35.35	7.31	|||	0.12

FSH-β	#	HSDC	ASA	BSA	BSA%	Δ^iG
B:LYS	54		104.29	0.00		0.00
B:GLU	55		96.97	0.00		0.00
B:LEU	56		56.03	7.60	||	0.12
B:VAL	57		80.61	0.00		0.00
B:TYR	58		138.57	27.49	||	-0.16
B:GLU	59		71.26	0.00		0.00
B:LEU	73		134.25	0.00		0.00
B:TYR	74		100.92	21.21	|||	-0.24
B:THR	75		77.39	0.00		0.00
B:TYR	76		16.89	2.68	||	0.04
B:PRO	77		49.69	40.74	|||||||||	0.61
B:VAL	78		14.01	0.00		0.00
B:LYS	86		127.23	0.00		0.00
B:CYS	87		33.33	6.53	||	0.10
B:ASP	88		72.43	0.00		0.00
B:SER	89	H	88.61	21.52	|||	-0.25
B:ASP	90		136.24	3.11			0.05
B:SER	91	H	80.36	44.98	||||||	-0.16
B:THR	92	H	44.27	43.04	||||||||||	0.53
B:ASP	93	H	120.07	77.87	|||||||	0.27
B:CYS	94		65.19	16.84	|||	0.27
B:THR	95	H	73.08	40.28	||||||	-0.37
B:VAL	96		157.53	83.13	||||||	0.81
B:ARG	97		212.15	17.07			0.08
B:GLY	98	H	45.83	25.76	||||||	0.14
B:LEU	99	H	140.39	64.68	|||||	0.95
B:GLY	100	H	39.61	29.56	||||||||	0.15
B:PRO	101		83.27	62.13	||||||||	0.83
B:SER	102		68.06	0.00		0.00
B:TYR	103		105.97	45.37	|||||	-0.06
B:CYS	104		23.53	0.00		0.00
B:SER	105		24.56	17.02	|||||||	0.13
B:PHE	106		118.65	0.00		0.00
B:GLY	107		29.87	0.98			-0.01
B:GLU	108	S	100.08	54.14	||||||	-0.29
B:MET	109	H	218.69	34.29	||	0.19

Table 2: FSH-β buried residues in the interface FSHR/FSH-β. All the residues in the interface FSHR/FSH-β #7 are shown (yellow). Only a few residues with solvation energy effect = 0 (solvent-accessible residues) are shown since all have the same BSA value (zero)(blue). The full table can be found as supplementary data, Table S2, corresponding to the values of the PISA interfaces #7, 8, and 9 in the PDB 4MQW crystal. HSDC: Residues making Hydrogen/Disulphide bond, Salt bridge or Covalent link; ASA: Accessible Surface Area (Å^2); BSA: Buried Surface Area (Å^2); Δ^iG: Solvation energy effect (kcal/mol); |||: Buried area percentage, one bar per 10%.

FSH-β	#	HSDC	ASA	BSA	BSA%	ΔiG
H:THR	34		70.08	0.00		0.00
H:ARG	35		169.87	26.43	||	-0.34
H:ASP	36		119.22	1.59	|	-0.02
H:LEU	37		50.31	26.93	|||||	0.43
H:VAL	38	H	150.20	13.41	|	0.12
H:TYR	39	H	166.26	26.73	||	0.33
H:LYS	40	H	120.38	27.77	|||	0.09
H:ASP	41		46.93	10.68	|||	-0.10
H:PRO	42		126.34	90.14	|||||||	0.95
H:ALA	43	H	94.54	82.89	||||||||	0.82
H:ARG	44		64.22	29.27	|||||	-0.29
H:PRO	45		95.65	43.25	|||||	0.69
H:LYS	46	HS	124.33	53.96	|||||	-0.35
H:ILE	47		117.62	0.00		0.00
H:ASP	88		73.57	0.00		0.00
H:SER	89	H	88.38	40.02	|||||	0.03
H:ASP	90	S	140.52	36.62	|||	-0.18
H:SER	91		77.03	0.00		0.00
H:THR	92		43.28	0.00		0.00
H:ASP	93	HS	121.94	42.46	||||	-0.40
H:CYS	94		64.33	8.35	||	-0.10
H:THR	95	H	84.75	36.87	|||||	-0.03
H:VAL	96		123.24	31.60	|||	0.50
H:ARG	97	HS	200.20	110.92	||||||	-0.98
H:GLY	98		59.92	11.99	|||	0.11
H:LEU	99		140.75	86.25	|||||||	1.37
H:GLY	100		43.73	0.00		0.00
H:PRO	101		75.93	0.00		0.00
H:SER	102		51.67	0.00		0.00
H:TYR	103	H	99.55	51.23	||||||	-0.01
H:CYS	104		14.50	0.00		0.00
H:SER	105		66.45	0.34	|	0.01
H:PHE	106		77.61	0.00		0.00

Table 3: Predicted and Observed K_a values for FSH-β synthetic peptides. Predicted values using Model 2 were obtained from $\Delta G°(x) = \gamma' BSA(x)$ with γ′ = -0.00726 kcal mol^{-1} Å^{-2}. Predicted values using Model 3 were obtained from $\Delta G°(x) = \gamma'' BSA(x) - 3.124895$, with γ′′= -0.0058697 kcal mol^{-1} Å^{-2}. ΔG° = -RT ln Ka; Ka= e- $^{(\Delta G°/RT)}$;RT=0.5925 kcal mol^{-1} ; Ka=1/Kd.

Peptide	**BSA** Å^2	**Observed K_a** mol^{-1} L	**Observed $\Delta G°_i$** kcal/mol	**Predicted Model 2 $\Delta G°_i$** kcal/mol (from γ′. BSA(x))	**Predicted Model 2 K_a** mol^{-1} L (from Predicted Model 2 $\Delta G°_i$)	**Predicted Model 3 $\Delta G°_i$** kcal/mol (from regression of observed values)	**Predicted Model 3 K_a** mol^{-1} L (from predicted Model 3 $\Delta G°_i$)
FSH-β-(49-52)	0	1.4 x 10^2	-2.9	0	1	-3.1	2.0 x 10^2
FSH-β-(34-37)	38	0.8 x 10^2	-2.6	-0.3	1.6	-3.3	2.8 x 10^2
FSH-β-(81-95)	157	2.5 x 10^3	-4.6	- 1.1	6.8	-4.0	9.2 x 10^2
FSH-β-(33-53)	404	1.0 x 10^4	-5.5	- 2.9	1.4 x 10^2	-5.5	1.1 x 10^4
FSH-β-(33-53)-(81-95)	561	2.0 x 10^4	-5.9	- 4.1	9.6 x 10^2	-6.4	5.1 x 10^4
FSH-β	852	1.1 x 10^7	- 9.6	-6.2	3.4 x 10^4	-8.1	0.9 x 10^6
FSH	1882	1.0 x 10^{10}	-13.6	-13.6	1.0 x 10^{10}	-14.2	2.4 x 10^{10}
FSH-β-(89-97)	297	1.7 x 10^2	-3.05	-2.2	3.8 x 10^1	-4.9	3.7 x 10^3
FSH-β-(96-99)	238	ND	ND	- 1.7	1.8 x 10^1	-4.5	2.1 x 10^3
FSH-β-(34-47)	404	ND	ND	- 2.9	1.4 x 10^2	-5.5	1.1 x 10^4
FSH-β-(88-106)	447	ND	ND	-3.2	2.4 x 10^2	-5.7	1.6 x 10^4
FSH-β-(34-47)-(88-106)	852	ND	ND	-6.2	3.4 x 10^4	-8.1	0.9 x 10^6

Observed Ka values: FSH-β-(49-52), FSH-β-(34-37), FSH-β-(81-95), FSH-β-(33-53), FSH-β-(33-53)-(81-95) and FSH-β from Santa-Coloma et al.[23]; FSH from Dattatreyamurty, et al. [130]; FSH-β-(89-97) from Prabhudesai et al. [66].

Legend to Figures

Figure 1. FSH structure. A: Primary structure of FSH-α and FSH-β subunits, mature forms; the processed subunits without the signal peptide are shown. B: Secondary structure of FSH-α and FSH-β subunits according to the 3D structure reported in the PDB entry 1FL7; a few amino acids are missing at the N and C-terminal ends in this structure; it was drawn using ESPript 3.0 (espript.ibcp.fr)[145]. C: FSH 3D structure corresponding to the PDB entry 1FL7[12]. 3D images in all figures were built by using the open-source PyMol version 2.5, running under Python 3.8 for Windows (pymolwiki.org; pymol.org) [146].

Figure 2: Synthetic peptides used to map the receptor-binding regions of FSH-β. A: The first FSH-β binding regions found, by homology with EGF, were TRDL and KTCT. Then, the FSH-β-(33-53) peptide, holding both tetrapeptides plus the interconnecting region, was evaluated. The equilibrium constants of dissociation K_d and association K_a are shown ($K_a = 1/K_d$). B: Overlapping synthetic peptides that were used to further map the different regions of interaction in the FSH-β subunit. The predicted FSH-receptor binding regions FSH-β-(31-45), included in the FSH-β-(33-53) peptide, and FSH-β-(81-95), are labeled as "R". The predicted regions of interaction with the FSH-α subunit (GPH-α) corresponded to the peptides FSH-β-(11-25), (41-55), (51-65), and (101-111), are labeled as "α".

Figure 3. FSH 3D structure showing the predicted receptor binding regions in the FSH-β. The predicted binding regions between FSH-β and its receptor are shown, corresponding to FSH-β-(33-53) and FSH-β-(81-95), mapped by using the 15 mer overlapping synthetic peptides shown in Fig. 2B. The regions FSH-β-(33-53) and FSH-β-(81-95) form a continuous surface on the β subunit, as previously suggested without knowing the 3D structure [23]. The figure corresponds to PDB entry 1FL7. The binding region FSH-α-(32-46) is also shown.

Figure 4. Regions of FSH-β predicted to interact with FSH-α. The 15 mer overlapping peptides were also used to find possible binding regions of FSH-β interacting with the α subunit. For this purpose, a solid phase binding assay over nitrocellulose was used, over which the different 15 mer peptides were adsorbed, and the binding of ^{125}I-FSH-β was measured. The predicted binding regions FSH-β-(11-25), (41-65), and (101-111), form a continuous region that partially embraces the α subunit. The actual interacting regions at the interface are shown in Table 1 and with details in the supplementary Table S1, corresponding to the 1FL7 crystal, derived from the PISA server (ebi.ac.uk/pdbe/pisa).

Figure 5. The 3D structure of the FSH-FSHR complex and the predicted binding sites using synthetic peptides. The three-dimensional FSH structure is shown together with the ectodomain of the FSH-receptor. The X-ray coordinates were obtained from the PDB database and plotted by using PyMol. A: The FSH-β-(33-53) and FSH-β-(81-95) regions are shown, corresponding to the synthetic peptides able to inhibit FSH binding to the receptor. The 3D structure corresponds to the PDB entry 1XWD. B: Interactions of the FSH-β-(38-48) loop with the FSH receptor (PDB entry 4MQW, chains A, B, X). The sulfated tyrosine 335 (sY335) is shown (blue) in the hydrophobic pocket of the βL2 loop (V38-Q48). The rest of the interactions are listed in Table 2 and Table S2, in supplementary materials.

Figure 6: The ligand-receptor models of interaction. Several models have been postulated to explain how a ligand interacts with its receptor. A) Lock and Key model. This was the first and more simple view of the possible interaction. Both the ligand and the receptor function as rigid molecules, fully complementary, as occurs with a lock and a key. B) Induced fit. A rigid ligand induces a conformational shift in the receptor, adapting its interacting surface to the ligand. C) The induced fit may be also bidirectional, with both the ligand and the receptor adapting their structures to each other for best binding. D)

Conformational selection: the receptor exists in multiple conformations in equilibrium. The ligand binds to the complementary one and displaces the equilibrium towards the optimal conformation; a mutually induced fit can also occur.

Figure 7: Joining two binding regions. A: The amino acids TRDL, KTCT, and the bridge sequence among these receptor binding regions, were joined together in the synthetic peptide FSH-β-(33-53). B: the regions FSH-β-(33-53) and FSH-β-(81-95) were joined without the bridge sequence in the synthetic peptide FSH-β-(33-53)-(81-95). Following the principle of additivity of ΔG, the expected K_d value ($K_d = 1/K_a$) for the joint peptide was similar to the value measured for FSH-β. The joint peptide had an increased affinity (lower dissociation rate) compared with both peptides separately; however, it was far from the expected value $K_{AB} = K_{aA}.K_{aB}$ (Model 1).

Figure 8: The additive nature of free energy. The free energy of interaction results from the addition of the individual free energies of each binding component. These regions can be mimicked by synthetic peptides, with limitations imposed by synergistic or antagonistic interactions, reflected in the $\Delta G°_{int}$ term. It can be replaced by a correction factor "c". If 3D data are available or can be predicted by homology modeling or a similar approach, the $\Delta G°_i$ for each binding region can be estimated by using the buried surface areas (BSA) since $\Delta G°(x) = \gamma'. BSA(x)$ (Model 2). Considering the more realistic presence of nonadditive interactions, the free energy can be predicted from the affine function $\Delta G°(x) = \gamma''. BSA(x) + \Delta G°_{na}$ (Model 3). The slope γ'' can be calculated empirically from the observed values by linear regression and $\Delta G°_{na}$ from the ordinate at the origin, as shown in Figure 9.

Figure 9: Predicted and observed ΔG° values as a function of the buried surface area (BSA). A: Predicted and observed ΔG° values for the different FSH-β peptides,

corresponding to the values shown in Table 3. Closed triangles (▲) show the predicted $\Delta G°$ values by using $\Delta G°(x) = \gamma'$. BSA(x) (Model 2), with $\gamma' = -0.00726$ kcal mol^{-1} Å^{-2}; closed circles (●) show the experimentally observed $\Delta G°$ values of the peptides, in solution. The fitted line of the observed values shown corresponds to the function $\Delta G° (x) = \gamma''$. BSA(x) + $\Delta G°_{na}$ (Model 3), which is $\Delta G° (x) = (-0.0058697$ kcal mol^{-1} Å^{-2}) BSA(x) - 3.124895 kcal mol^{-1}; it was used to calculate the predicted Model 3 values of Table 3. The gray area shows the 95% confidence interval. The graph was plotted by using Rstudio and the ggplot2 library as shown in the suppl. file. B: Additive values of free energy $\Delta G°_{i} = \gamma''$. BSA_i for each FSH-β residue *i*, where γ'' corresponds to the slope of the regression $\Delta G° (x) = \gamma''$. BSA(x) + $\Delta G°_{na}$ (panel A, model 3, $\gamma'' = -0.0058697$ kcal mol^{-1} Å^{-2}; ordinate at the origin $\Delta G°_{na} = -3.124895$ kcal mol^{-1}). The total free energy of a protein fragment, from amino acid *i* to *j*, can be calculated as $\Delta G°_{i \to j} = \left(\sum_i^j \Delta G°_i\right) + \Delta G°_{na.}$ The first term corresponds to the additive part of the $\Delta G°$, explained by the BSA values, and the last term corresponds to the nonadditive part, including values of BSA-independent properties and entropy changes; the corresponding BSA and $\Delta G°_i$ values are indicated in Table S3 (suppl. file). Two "hot regions" appear between amino acid residues FSH-β-(34-47) and FSH-β-(88-106) (Table S3), with a total $\Delta G°_i$ of -2.37 ± -0.18 kcal mol^{-1} and -2.63 ± -0.05 kcal mol^{-1}, respectively, plus $\Delta G°_{na} = -3.124895$ kcal mol^{-1}.

Figures

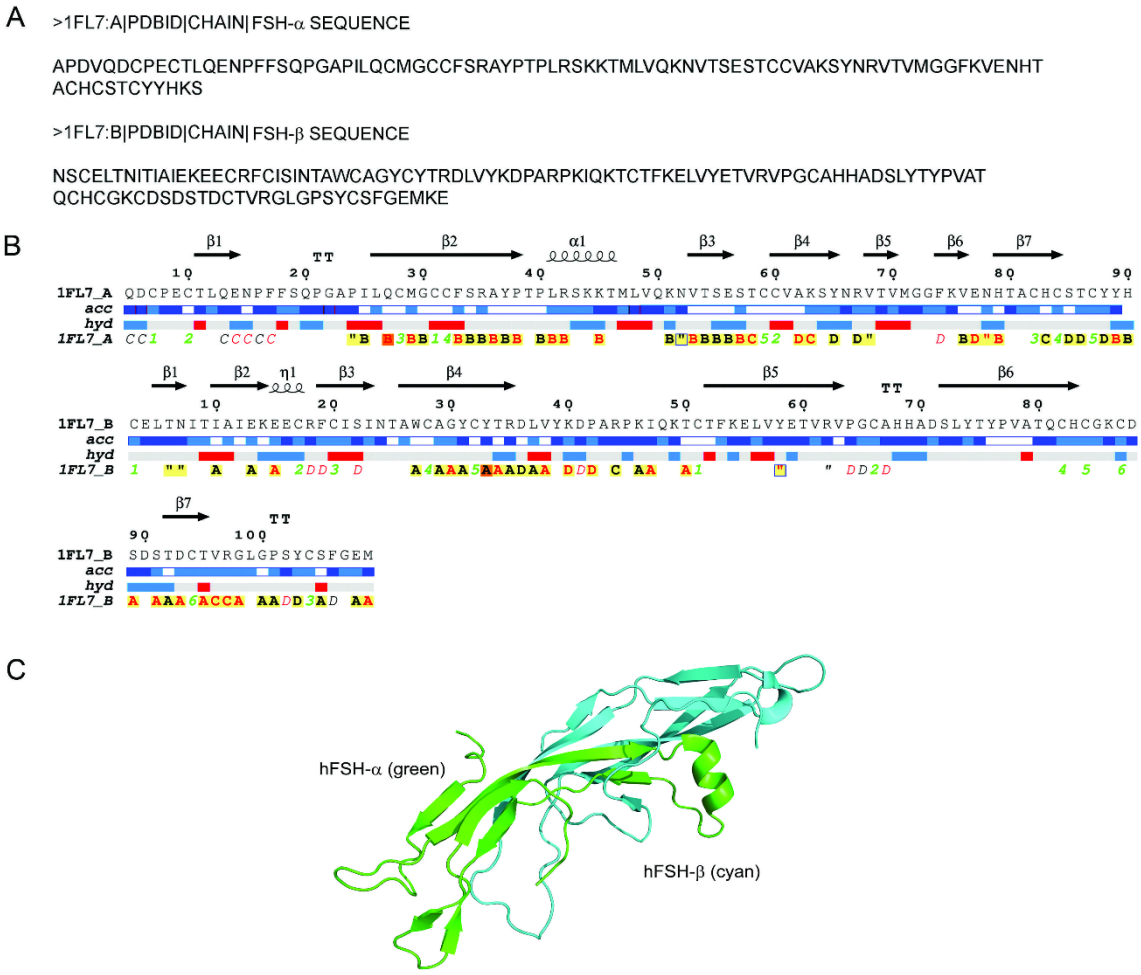

Figure 1

A

34-TRDL-37	Kd = 1.25 x 10^{-2} mol L^{-1}	Ka = 0.80 x 10^{2} mol^{-1} L
49-KTCT-52	Kd = 0.71 x 10^{-2} mol L^{-1}	Ka = 1.41 x 10^{2} mol^{-1} L
33-YTRDLVYKDPARPKIQKTCTF-53	Kd = 1.00 x 10^{-4} mol L^{-1}	Ka = 1.00 x 10^{4} mol^{-1} L

B FSH-β 15 mer overlapping synthetic peptides

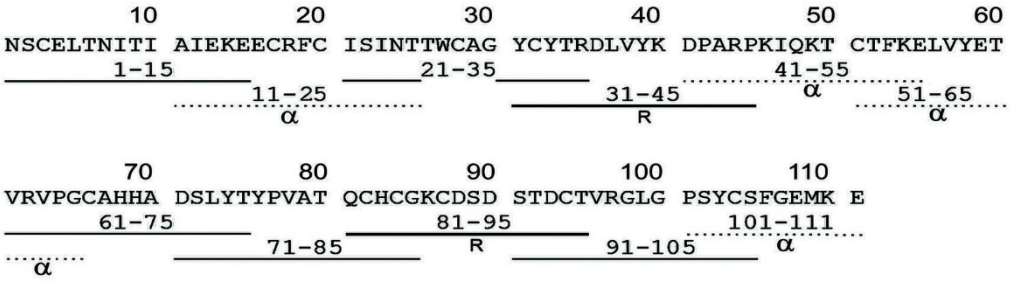

Figure 2

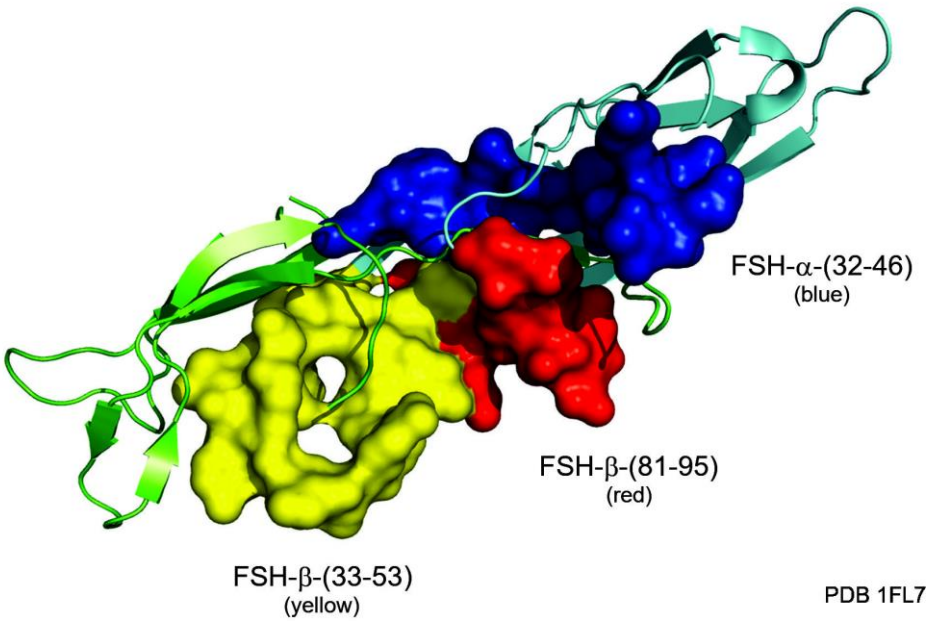

Figure 3

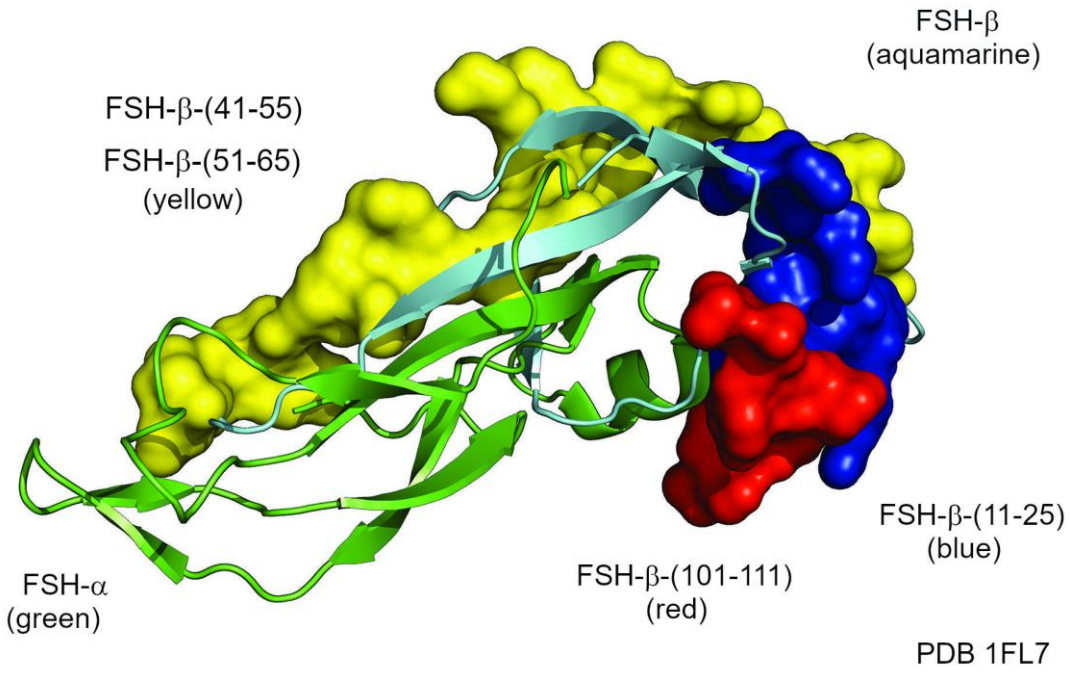
FSH-β
(aquamarine)
FSH-β-(41-55)
FSH-β-(51-65)
(yellow)
FSH-β-(11-25)
(blue)
FSH-α
(green)
FSH-β-(101-111)
(red)
PDB 1FL7

Figure 4

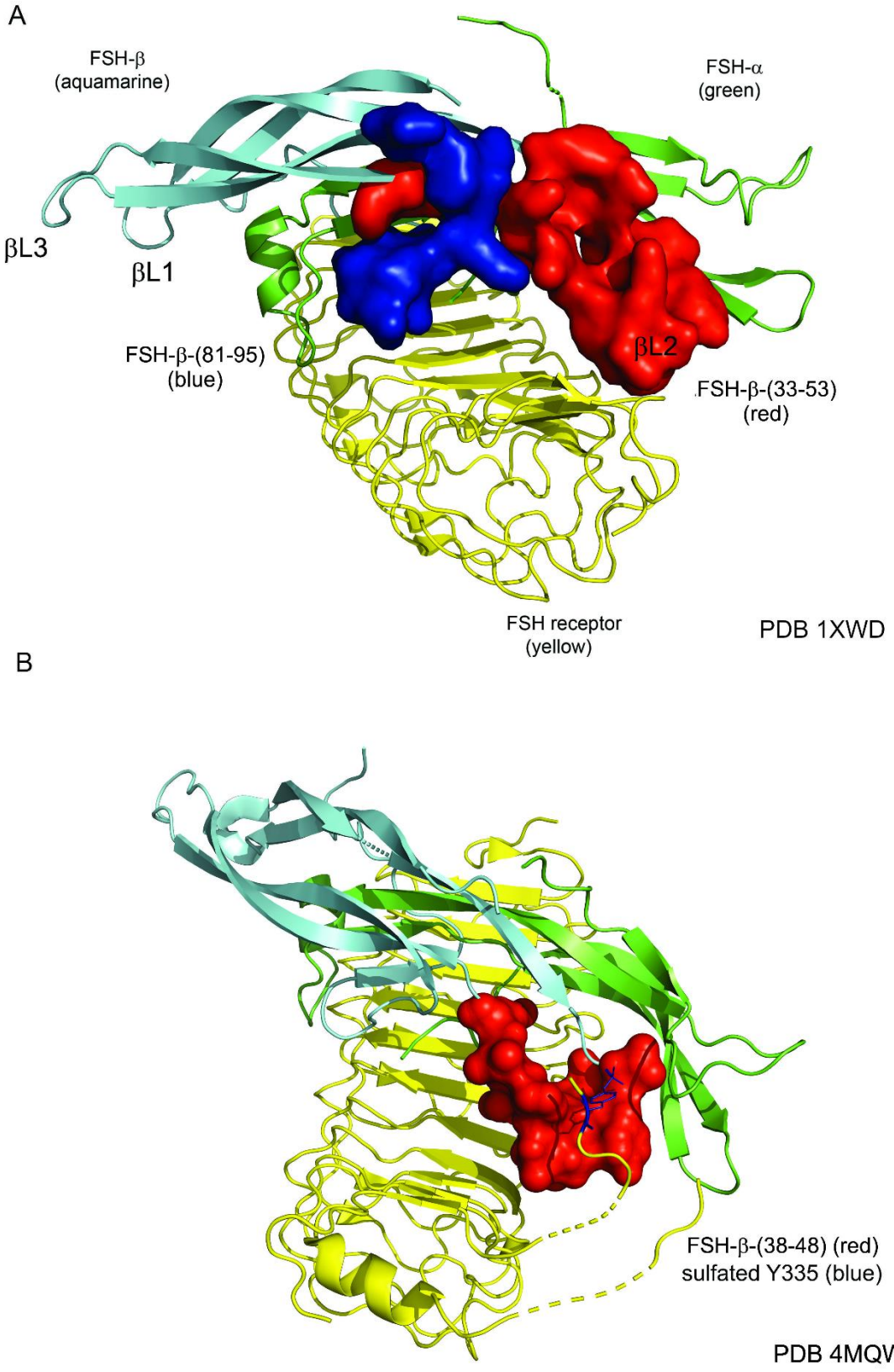
A
FSH-β
(aquamarine)
FSH-α
(green)
βL3
βL1
βL2
FSH-β-(81-95)
(blue)
FSH-β-(33-53)
(red)
FSH receptor
(yellow)
PDB 1XWD
B
FSH-β-(38-48) (red)
sulfated Y335 (blue)
PDB 4MQW

Figure 5

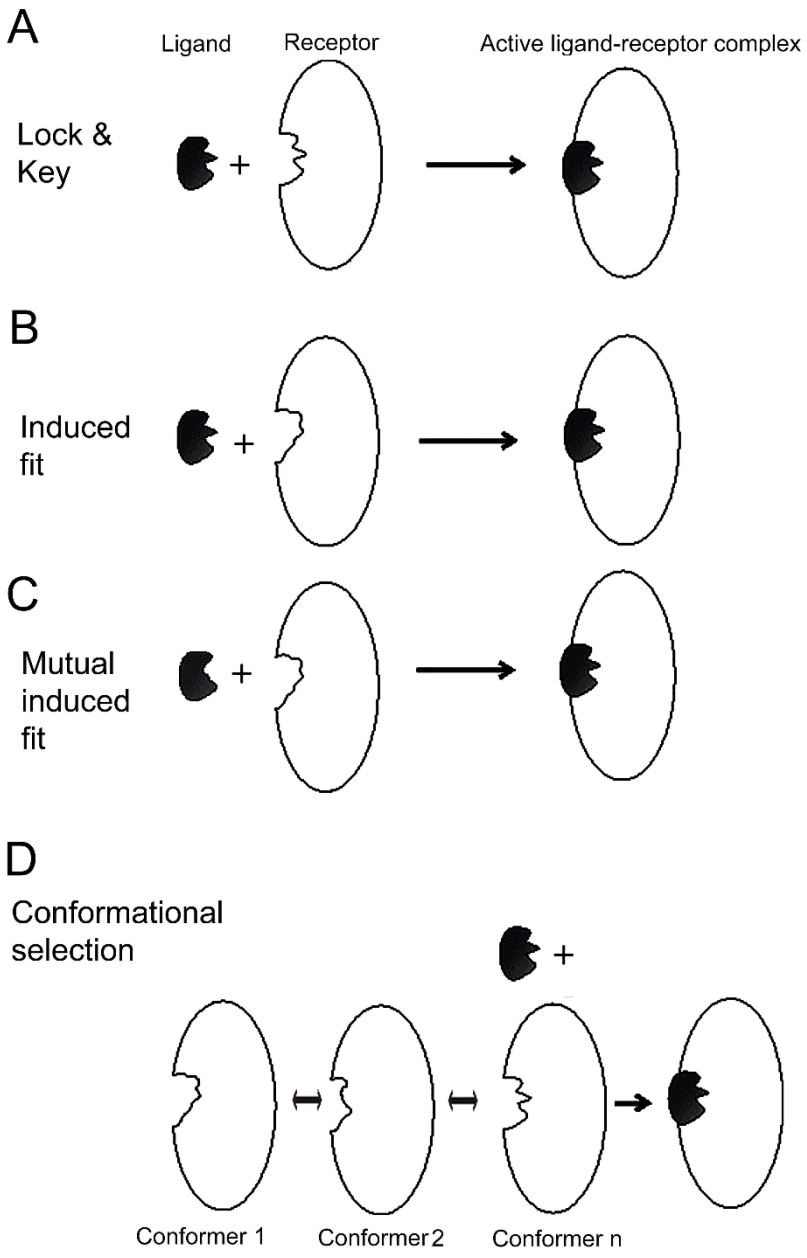
A
Ligand
Receptor
Active ligand-receptor complex
Lock &
Key
B
Induced
fit
C
Mutual
induced
fit
D
Conformational
selection
Conformer 1
Conformer 2
Conformer n

Figure 6

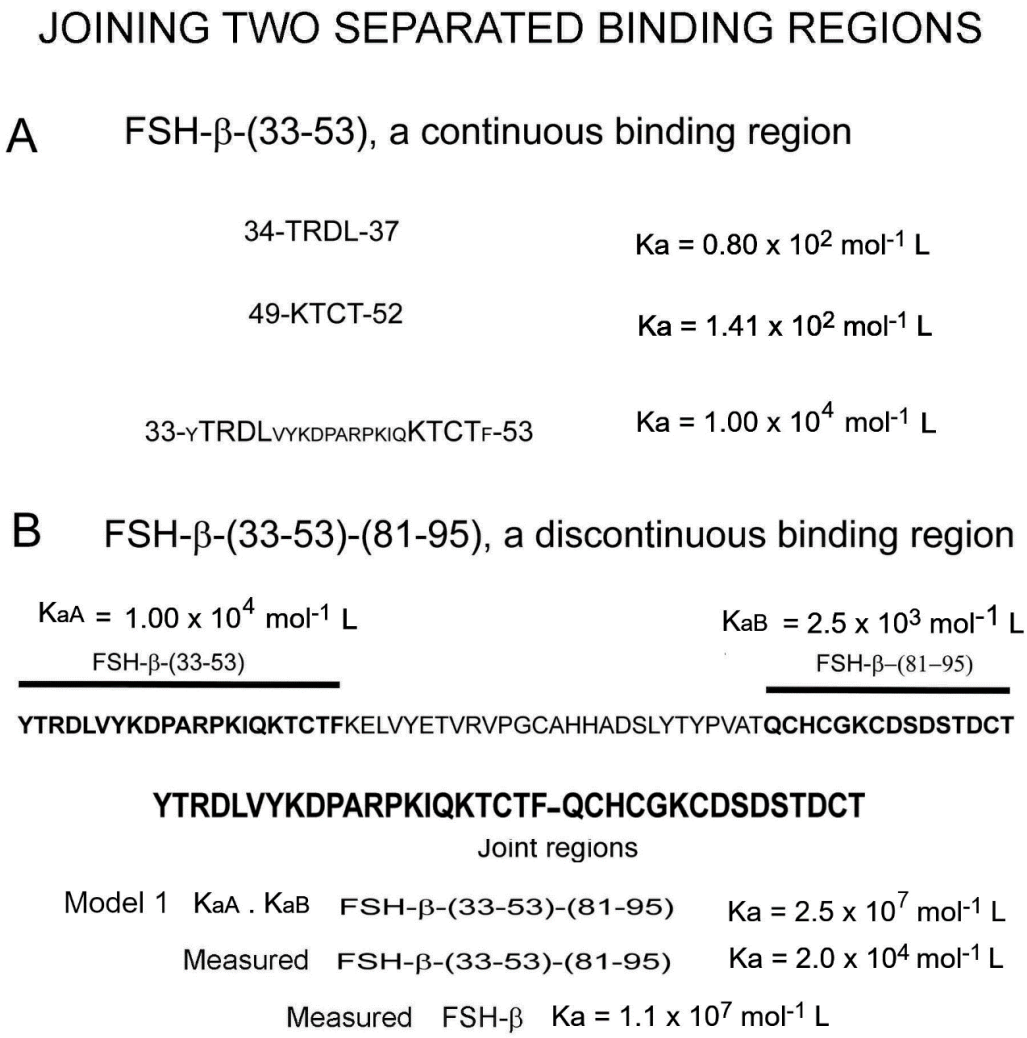
JOINING TWO SEPARATED BINDING REGIONS
A FSH-β-(33-53), a continuous binding region
34-TRDL-37
Ka = 0.80 x 10² mol⁻¹ L
49-KTCT-52
Ka = 1.41 x 10² mol⁻¹ L
33-YTRDLVYKDPARPKIQKTCTF-53
Ka = 1.00 x 10⁴ mol⁻¹ L
B FSH-β-(33-53)-(81-95), a discontinuous binding region
KaA = 1.00 x 10⁴ mol⁻¹ L
FSH-β-(33-53)
KaB = 2.5 x 10³ mol⁻¹ L
FSH-β-(81-95)
YTRDLVYKDPARPKIQKTCTFKELVYETVRVPGCAHHADSLYTYPVATQCHCGKCDSDSTDCT
YTRDLVYKDPARPKIQKTCTF-QCHCGKCDSDSTDCT
Joint regions
Model 1 KaA . KaB FSH-β-(33-53)-(81-95) Ka = 2.5 x 10⁷ mol⁻¹ L
Measured FSH-β-(33-53)-(81-95) Ka = 2.0 x 10⁴ mol⁻¹ L
Measured FSH-β Ka = 1.1 x 10⁷ mol⁻¹ L

Figure 7

Principle of $\Delta G°$ additivity

$$\Delta G° = \left(\sum_i \Delta G°_i\right) + \Delta G°_{int}$$

If $\Delta G°_{int} = 0$

$$\Delta G° = \sum_i \Delta G°_i$$

and

$$Kd\ (total) = \prod_i Kd_i$$

This represents the Kd expected for a synthetic peptide containing 'i' independent binding regions

If interactions are included as a "c" correction factor,

$$\Delta G° = \left(\sum_i \Delta G°_i\right) + \Delta G°_{int} = c \sum_i \Delta G°_i$$

If $\Delta G°_i$ is a linear function of x,

$$c \sum_i \Delta G°_i (x) = c\ F(x)$$

If F(x) = BSA(x)
and c = γ'

$$\Delta G° = \gamma' BSA(x)$$

Valid for proteins or small ligands

w/o conformational changes upon binding

If nonadditive terms are included,

$$\Delta G° = \gamma'' BSA(x) + \Delta G°_{na}$$

Valid for fragments

conformational changes upon binding

γ'' = slope, $\Delta G°_{na}$ = ordinate, BSA (x) = BSA of each binding region

Figure 8

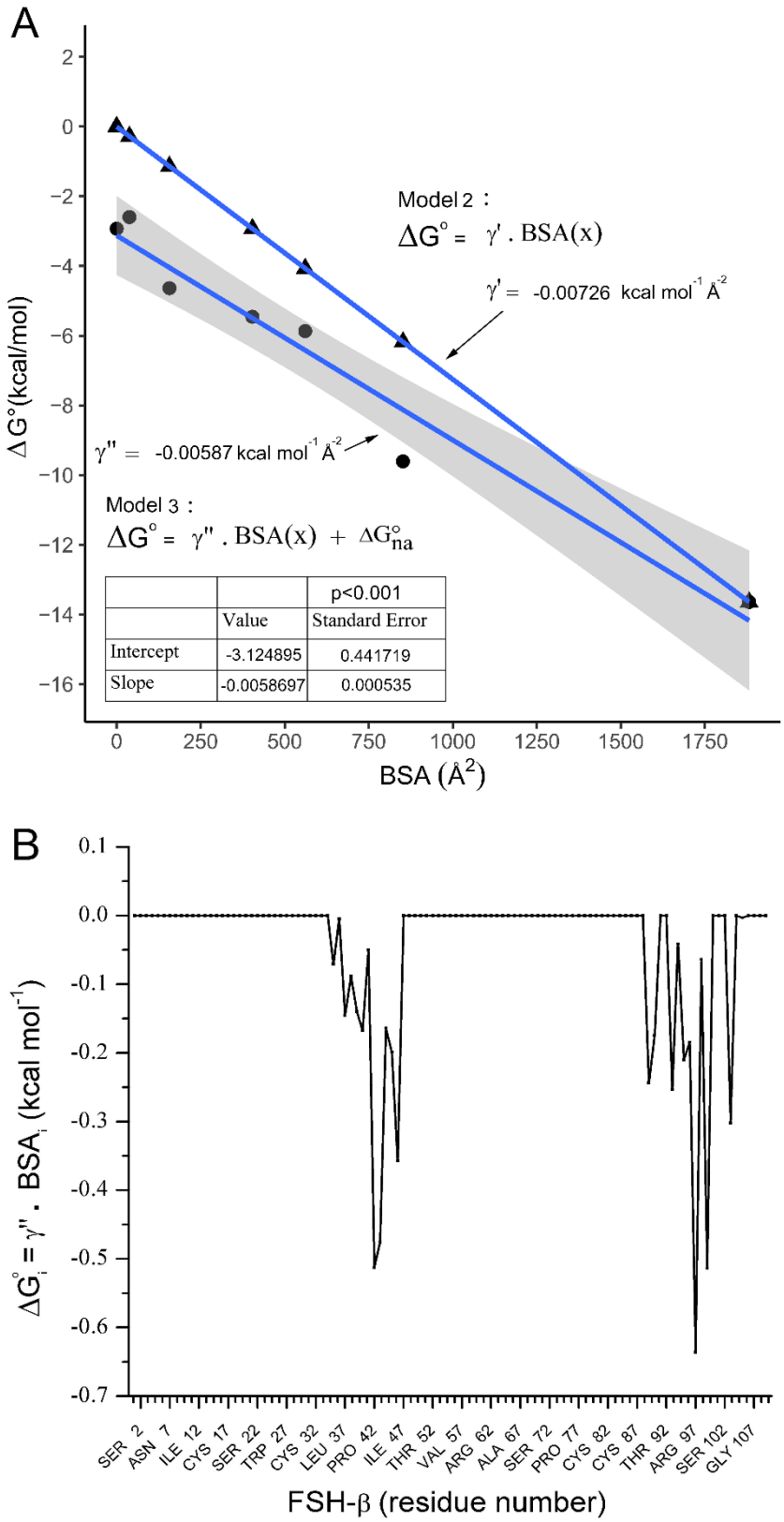
A
Model 2 :
$\Delta G^{\circ} = \gamma' . BSA(x)$
$\gamma' = -0.00726$ kcal mol^{-1} Å^{-2}
$\gamma'' = -0.00587$ kcal mol^{-1} Å^{-2}
Model 3 :
$\Delta G^{\circ} = \gamma'' . BSA(x) + \Delta G^{\circ}_{na}$
p<0.001
Value
Standard Error
Intercept -3.124895 0.441719
Slope -0.0058697 0.000535
ΔG°(kcal/mol)
BSA (Å^2)
B
$\Delta G^{\circ}_i = \gamma'' . BSA_i$ (kcal mol^{-1})
SER 2 ASN 7 ILE 12 CYS 17 SER 22 TRP 27 CYS 32 LEU 37 PRO 42 ILE 47 THR 52 VAL 57 ARG 62 ALA 67 SER 72 PRO 77 CYS 82 CYS 87 THR 92 ARG 97 SER 102 GLY 107
FSH-β (residue number)

Figure 9

For Table of Contents Use Only (Graphical Abstract)

Overlapping synthetic peptides as a tool to map protein-protein interactions − FSH as a model system of nonadditive interactions.

Tomás A. Santa-Coloma

Overlapping synthetic peptides as a tool to map protein-protein interactions − FSH as a model system of nonadditive interactions

Tomás A. Santa-Coloma