University of Szeged

Doctoral School of Pharmaceutical Sciences

Ph.D. program:	Pharmaceutical Chemistry and Drugs Research
Programme Director:	Prof. Dr. István Szatmári
Institute:	Department of Medical Chemistry
Supervisors:	Prof. Dr. Tamás Martinek
	Dr. Zsófia Hegedüs

Vencel László Petrovicz

Ligand Development for the Inhibition of Transcription Factors

Final Examination Committee:

Chairman:	Prof. Dr. Dombi György
Members:	Dr. Fülöp Lívia
	Dr. Borics Attila

Reviewer Committee:

Chairman:	Prof. Dr. István Szatmári
Reviewers:	Dr. Kata Horvati
	Dr. Péter Kele
Secretary:	Dr. Gerda Szakonyi
Member:	Dr. Attila Borics

A. Introduction and aims

Inhibition of disease-relevant protein-protein interactions (PPIs) represents one of the most challenging frontiers in drug discovery. The free energy of binding in a PPI is distributed over a wide area (1200-3000 $Å^2$) forming hot-spot contacts on a typically flat surface with shallow binding clefts, which hinders the development of effective traditional small molecule ligands. In addition to the general challenges of inhibiting disease-relevant PPIs, transcriptional proteins involved in tumour development and other diseases present further obstacles for inhibitor development. These additional challenges arise from the overrepresentation of intrinsically disordered proteins (IDPs) in the transcriptional machinery, since the transactivation domains (TADs) of transcription factors (TFs) are predominantly disordered regions. This dynamic behaviour enables them to bind promiscuously to multiple targets, while they form often transient, but highly specific interactions, resulting in molecular switches that adapt to different external stimuli. In addition to promiscuity, nuclear localisation and allosteric regulation involving TF PPIs further complicate drug discovery approaches.

The scope of PPI inhibitor development efforts contains a diverse array of different structures ranging from peptide- to protein-based scaffolds. Up to today, one of the most utilised PPI inhibitor classes is monoclonal antibodies (mAbs) in medical practice, although they are facing serious limitations, including high production costs, stability issues, poor pharmacokinetic properties, and insufficient penetration through biological membranes. Peptidomimetic drugs can overcome these limitations; because of their relatively low molecular weight, peptidomimetics could potentially target PPIs inside the nucleus directly inhibiting TF activity. Design strategies include peptide-cyclisation and stapling, or the use of non-natural backbones as scaffolds, resulting in highly improved stability against proteases. The development of functional peptidomimetic PPI inhibitors can follow either a top-down (that is, rational modification of a chosen target's native ligand) or bottom-up (that is, high-throughput and fragment screening) approach. The former strategy heavily relies on detailed structural and mechanistic information about a given interaction, which can be more complex with IDP interaction partners. The alternative, fragment-based bottom-up ligand development could provide a way toward functional peptidomimetic inhibitors. However, using this approach would require a high-throughput screening strategy, which has limitations in relation to peptidomimetics, i.e., insufficient library size, synthetic and analytical difficulties. In vitro translation techniques can be adapted to incorporate non-natural side-chains but still have

limitations for monomers with artificial backbones that are often used in peptidomimetic design.

Our aim was to develop ways that address these challenges of targeting IDP-mediated interactions using peptidomimetic structures. The top-down design approach necessitates a detailed knowledge of a targeted PPI, because modifications in a disordered ligand could potentially influence not only its affinity, but also its folding or ligand cooperativity, which can be mechanistically important. Therefore, we first investigated the molecular details of the competition between HIF-1α and its negative feedback regulator CITED2 C-TADs for the partially shared binding site on the surface of p300/CBP.. This interaction initiates an adaptive response to hypoxic stress and is implicated as a potential target in solid tumours. The molecular mechanism of the competition between the two disordered ligands relies on negative cooperativity induced by allosteric communication. Our hypothesis was that $\alpha \rightarrow \beta^3$ amino acid replacement in CITED2 could shed light on the role of the different CITED2 binding motifs in the allosteric structural change, while additionally pave the way toward the top-down development of a functional peptidomimetic HIF-1α antagonist, which could lead to potential cancer therapeutics.

The majority of TAD interactions of TFs are mediated by a combination of short peptide motifs: short linear motifs (SLiMs) and molecular recognition features (MoRFs). These could potentially be mimicked by a library of artificial oligomers that fold into discrete and stable secondary structures, such as foldamers. Covalent linkage of these recognition motif mimetics could reproduce the multivalency of the interaction, which should be addressed at the screening stage. To overcome the limitations of high-throughput screening of peptidomimetics, and incorporate a multivalent ligand screening approach, we hypothesised that DNA-encoded libraries (DELs) could provide a solution. In DELs the library members are equipped with unique DNA-barcodes, which allows the selection of combinatorial libraries up to 10^6 members, since hit deconvolution is easily carried out by the use of high-throughput next generation DNA-sequencing (NGS). Our aim was a proof-of-concept synthesis of a DNA-encoded multivalent foldameric ligand that could be the foundation of bottom-up ligand development targeting PPIs that rely on multivalent interaction of recognition features.

B. Methods

Isothermal titration calorimetry

Titrations were performed in a pH 7.5 phosphate buffer using a Microcal VP ITC instrument. Direct titrations to TAZ1 were carried out using 5 μM protein in the cell and 60 μM ligand in the syringe at 35°C, using 5 μL injections with a 180 s spacing between the injections. For competition titrations the TAZ1/HIF-1$\alpha_{776-826}$ complex (1 to 1.2-1.4 molar ratio) was titrated with the different CITED2 variants using 60 μM concentration in the syringe, 5-10 μL injections with 180-240 s spacing in between. The raw data was integrated with NITPIC software. The thermograms of titrations were fitted with SEDPHAT to a model which takes into account ternary system formation.

Fluorescence anisotropy

Fluorescence anisotropy was measured using a Clariostar Plus microplate reader with excitation at 480 nm and emission at 535 nm at 35 °C. The buffer used for the experiment was 40 nM Na-phosphate, 100 mM NaCl, 1 mM DTT, 0.01% Triton-X, pH 7.4. Three repeats were performed for each titration with a control experiment in parallel. The competitor peptides (HIF-1$\alpha_{776-826}$, native CITED2 and variants) were diluted by mixing 40 μL peptide solution with 20 μL buffer in the first well, which was followed by a serial 2/3 dilution over 24 points in the well resulting in a 5 μM as the highest concentration in the first well. TAZ1 was added at a final concentration of 50 nM and tracer (Flu-HIF-1$\alpha_{786-826}$) at a final concentration of 25 nM to each well. Plates were incubated for 10 minutes at 35 °C before reading. The fluorescence polarisation signals were analysed in Origin Pro 9.5.

NMR measurements

Samples were prepared in 10 mM Tris pH 6.9, 1 mM DTT, 50 mM NaCl, 10% D_2O, 0.02% NaN_3 buffer. 90-100 μM ^{15}N-^{13}C-TAZ1$_{330-424}$ was mixed with 1.2-1.5 equivalent CITED2 peptide or equimolar amount of HIF-1α and CITED2 peptides. NMR experiments were carried out on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm CP-TCI triple-resonance cryoprobe. Measurements were carried out at 298 K, ^{1}H-^{15}N-HSQC, ^{1}H-^{13}C-HSQC including 3D HNCO, HNCA, HN(CO)CA, CBCA(CO)NH. Processing was carried out using Topsin 3.5, and data were analysed using NMRFAM-Sparky. Backbone and methyl resonance assignments were based on the previously published p300 TAZ1/CITED2 complex (BMRB 5788) and p300 TAZ1/HIF-1α complex (BMRB 5306).

C. Results and discussion

1. Site-directed allostery perturbation strategy to evaluate the role of CITED2 binding motifs in competition with HIF-1α

With the use of sequence truncation and binding motif specific α→β[3] amino acid replacements we got a clearer picture about the role of CITED2 binding motifs in the allosteric structural changes of TAZ1 during the competition with HIF-1α through a ternary complex, which is responsible for the unidirectional nature of the competition. This work revealed the essential role of CITED2's N-terminal residues in inducing the correct structural changes in TAZ1, and clarified the contribution of other motifs to the negative cooperativity between the competing ligands. Additionally, the replacement strategy could be a starting point in a top-down peptidomimetic development, by retaining the beneficial modifications, resulting in an allosteric, peptidomimetic inhibitor of HIF-1α.

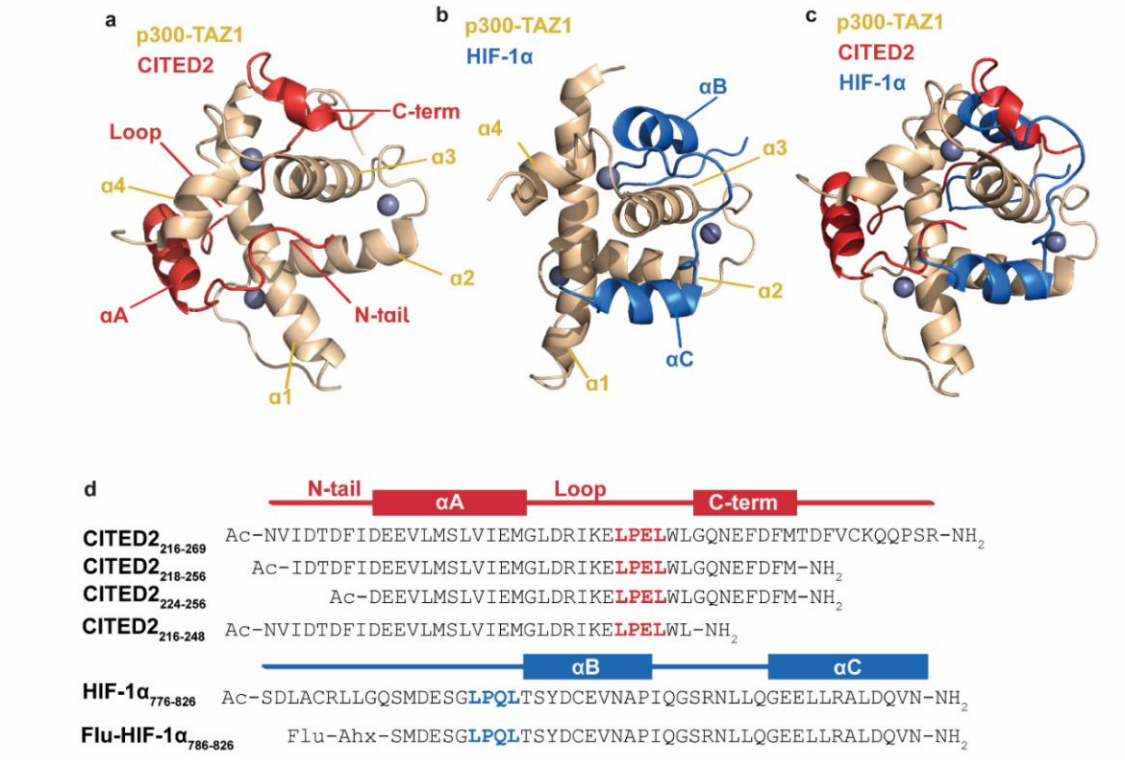

Figure 1. Structure of p300 TAZ1 bound to HIF-1α and CITED2. (a) Structure of CITED2 C-TAD (red) bound to TAZ1 domain (PDB:1P4Q). The helices of the TAZ1 domain are annotated α1–α4 from the N to C terminus, and the spheres represent Zn atoms. The binding motifs of CITED2 are annotated as N-tail, αA, loop, and C-term. (b) Structure of HIF-1α C-TAD (blue) bound to the TAZ1 domain (wheat) (PDB: 1L3E). HIF-1α helices are annotated αB–αC from the N to C terminus; αA was not observed for the p300-TAZ1 bound structure. (c) Overlayed structures of CITED2 and HIF-1α bound to TAZ1. TAZ1 is represented as the CITED2 bound conformation (d) Sequences of HIF-1α and CITED2 peptides used in this study, with their binding motifs indicated above. Flu: 5,6-carboxyfluorescein, Ahx: aminohexanecarboxylic acid.

1.1 Investigating the mechanistic role of CITED2 N- and C-terminal residues

The minimal length of CITED2 was determined first, which retains the binding properties and functionality of the full-length domain (CITED2$_{218-256}$) (Figure 1). Starting from this

sequence the truncation of CITED2 N-terminus (CITED2$_{218-223}$) slightly decreased the affinity; however, the cooperative aspect of competition was lost completely, indicating the loss of allosteric communication and unidirectionality. The truncation of the C-terminus (CITED2$_{249-256}$) significantly decreased the affinity for TAZ1, although the formation of the ternary complex remained favourable, indicating a retained allosteric communication between the ligands. NMR spectra showed that the N-terminus of CITED2 induces conformational changes in the mechanistically important α1 and α4 regions of TAZ1, which indicated that this region is responsible for the conformational lock that results in the unidirectional nature of competition (Figure 2).

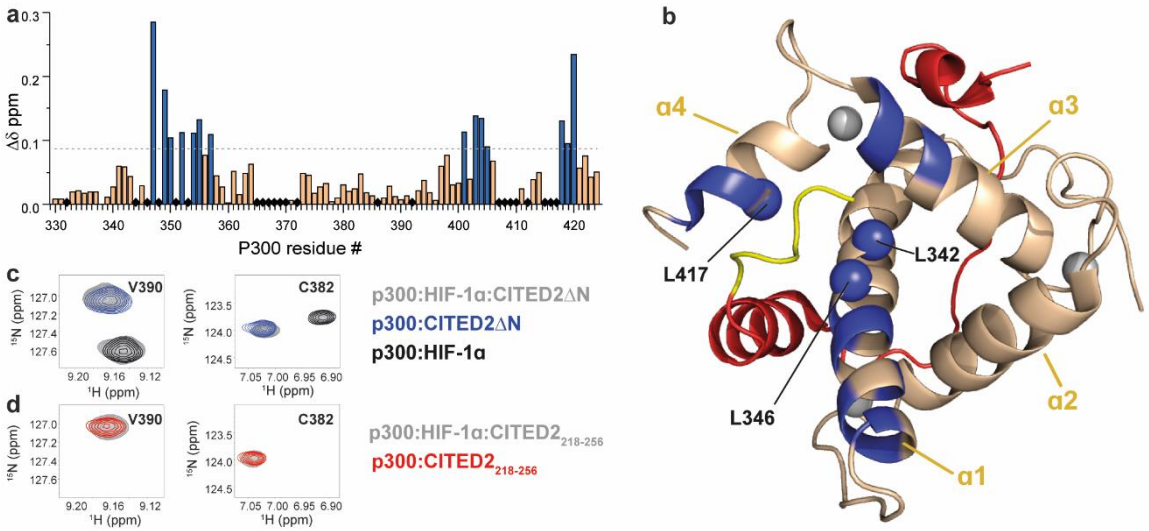

Figure 2. Structural comparison of CITED2$_{218–256}$ and CITED2ΔN bound to TAZ1 using NMR. (a) Weighted average ^{1}H, ^{15}N chemical shift difference between the TAZ1/CITED2ΔN and TAZ1/CITED2$_{218–256}$ complex. $\Delta\delta = [(\Delta\delta\ H)^2 + (\Delta\delta\ N/5)^2]^{1/2}$, residues above the significance level ($\Delta\delta > 0.9\ \Delta\delta$ average + σ) highlighted blue. Black diamonds indicate unassigned or proline residues. (b) Residues with significant chemical shift differences mapped onto the TAZ1/CITED2 structure and highlighted blue. Blue spheres represent core residues with significantly shifted CH_3 resonances. (PDB: 1P4Q, TAZ1 in wheat, CITED2 residues 224–259 in red, N terminal residues 218–223 represented in yellow). (c) Overlayed ^{1}H–^{15}N resonances of representative TAZ1 residues of samples containing TAZ1/HIF-1α/CITED2ΔN (grey), TAZ1/CITED2ΔN (blue) and TAZ1/HIF-1α (black). (d) Overlayed ^{1}H–^{15}N resonances for the same TAZ1 residues of samples containing TAZ1/HIF 1α/CITED2$_{218–256}$ (grey), TAZ1/CITED2$_{218–256}$ (red).

1.2 Backbone modification strategy to probe the contribution of CITED2 binding motifs to cooperativity and intramolecular interplay

Since sequence truncation and alanine scanning can misleadingly emphasise the importance of a given motif due to decreased TAZ1 affinity, a new strategy was developed using minimal site-directed perturbations that preserve the affinity of the parent peptide. The approach employs modular α→β^3 amino acid replacements in CITED2, targeting solvent exposed residues based on NMR structures and computational alanine scanning using BAlaS. Six CITED2 variants were synthesised: N-terminally modified (**1**), α-helix modified (**2**), three loop-modified variants (**3a-c**), and C-terminally modified (**4**). CD measurements confirmed that

all variants retained their disordered nature with a random coil conformation and slight helical content (Figure 3). The utilisation of this replacement strategy induces changes in competition efficiency (indicated by increased Δg) to reflect functionally relevant perturbations rather than reduced affinity, while NMR can monitor structural effects on TAZ1.

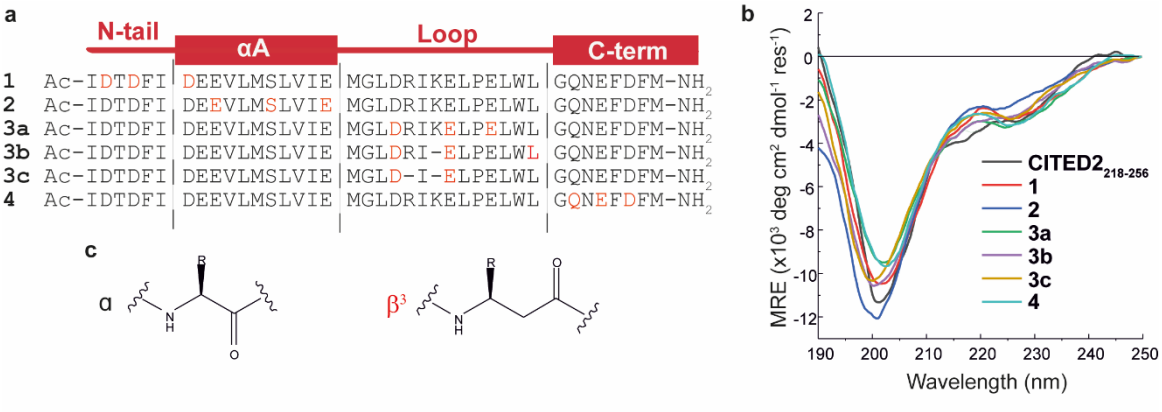

Figure 3. Sequences and overall structure of the β-amino acid modified CITED2 variants. (a) Sequences of modified CITED2 variants with binding motifs indicated above. α → β3 substitutions are highlighted in red, and deletions are represented by a line. (b) CD curves for CITED2 and variants in 20 mM Na-phosphate, 1 mM DTT, pH 7.4 at 20 mM concentration at room temperature. (c) Structure of α and homologous β^3-amino acid.

1.3 Impact of backbone modifications on the thermodynamic parameters

ITC measurements confirmed that variants **1**,**2** and **4** retained native-like binding affinity (K_D=10-19 nM) with slightly reduced enthalpic contribution and entropic cost. Loop modifications showed more pronounced effects: replacements in the conserved LPEL motif (**3a**) caused a 50-fold increase in K_D with altered thermodynamic profile, while preserving the conserved region and compensating the lengthening caused by β^3-amino acids with deletions (**3b-c**) restored native-like direct binding properties (K_D 98 and 29 nM, respectively). Cooperativity parameters for variants **2** and **4** remained similar to the native $CITED2_{218-256}$ (Δg=1.1-1.4 kcal mol^{-1}), indicating retained ternary complex formation and allosteric activity. N-terminal modification (**1**) reduced competition efficiency but less severely than truncation. Loop modifications (**3a-c**) showed the most dramatic effect with Δg>4 kcal mol-1 and Δh=0 kcal mol^{-1}, indicating a complete loss of ternary system formation and allosteric communication between the competing ligands. This tendency was also affirmed by fluorescence anisotropy measurements (Figure 4).

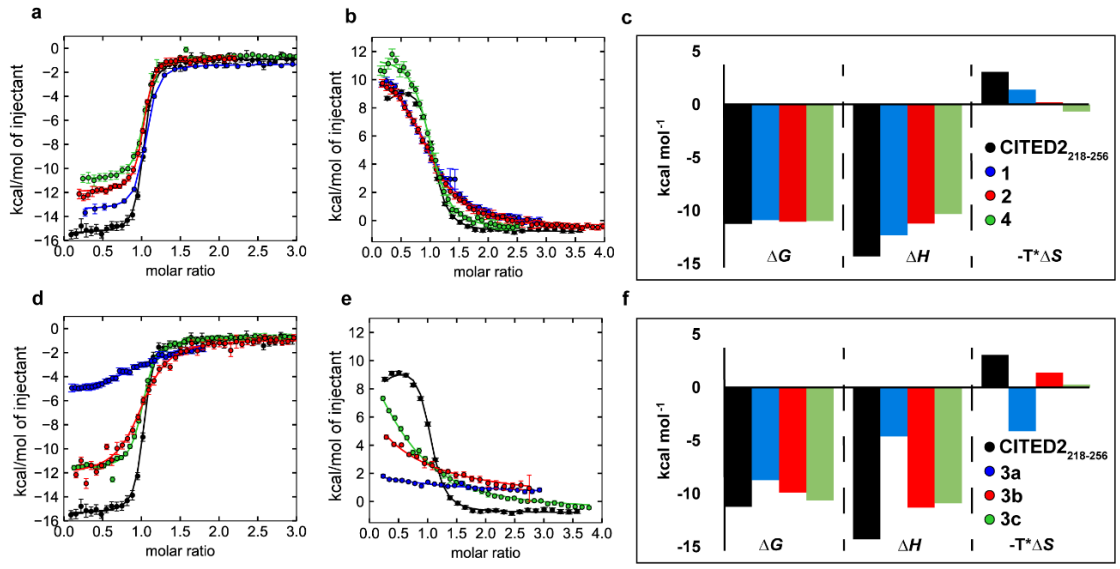

Figure 4. Binding and competition of different backbone modified CITED2 variants. (a) ITC thermogram for CITED2$_{218–256}$, **1**, **2** and **4** binding to TAZ1 (b) ITC thermogram for the titration of CITED2$_{218–256}$ **1** **2** and **4** to the preformed TAZ1–HIF-1α$_{776–826}$ complex. (c) Thermodynamic signatures of direct binding of CITED2$_{218–256}$, **1**, **2**, and **4** to TAZ1. (d) ITC thermogram for CITED2$_{218–256}$, **3a**, **3b**, and **3c** binding to TAZ1. (e) ITC thermogram for the titration of CITED2$_{218–256}$, **3a**, **3b**, and **3c** to the preformed TAZ1–HIF-1α$_{776–826}$ complex. (f) Thermodynamic signatures of direct binding of CITED2$_{218–256}$, **3a**, **3b** and **3c** to TAZ1.

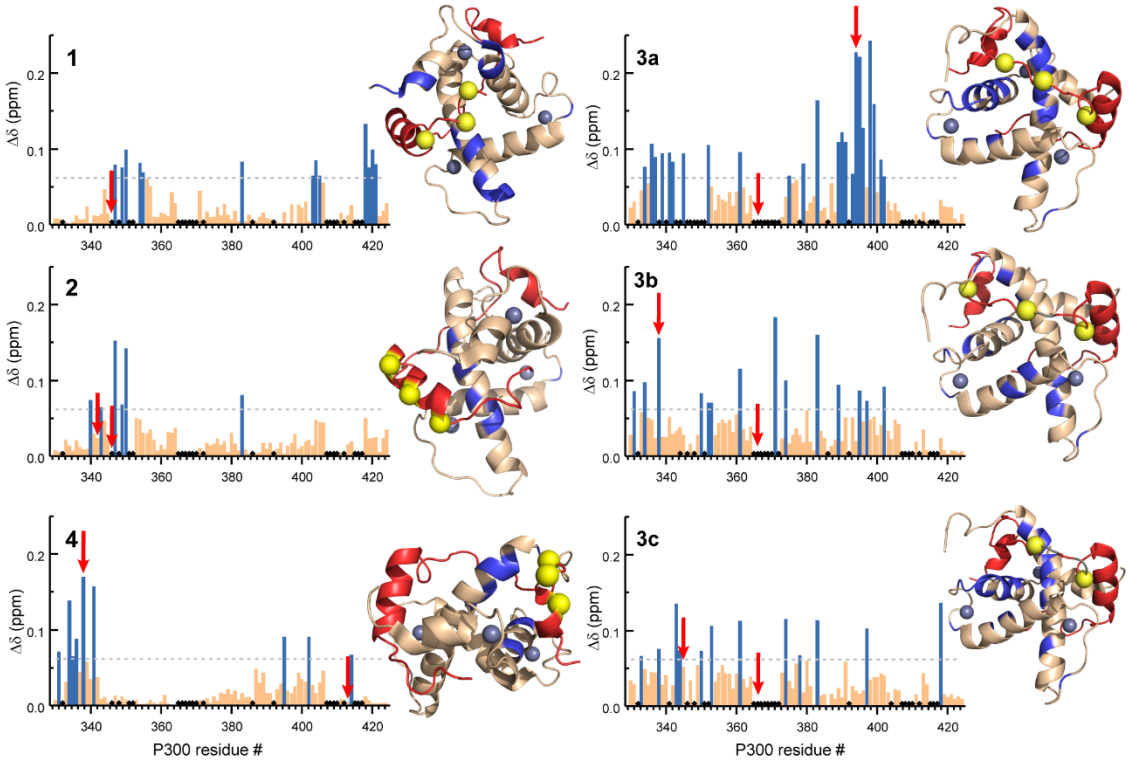

Figure 5. NMR chemical shift changes of TAZ1 upon binding to the β^3-amino acid modified CITED2 variants. Weighted average ^{1}H–^{15}N chemical shift differences between the TAZ1/CITED2 variants and TAZ1/CITED2$_{218–256}$ complexes. $\Delta\delta = [(\Delta\delta\ H)^2 + (\Delta\delta\ N/5)^2]^{1/2}$. The significance level was determined using the average Δδ and standard deviation for all the CITED2 variants, residues above the significance level (Δδ > 0.9 Δδ average + σ) highlighted blue in the bar graph and mapped onto the TAZ1-CITED2 structure, CITED2 residues in red, β^3-amino acid modifications sites represented as yellow spheres, removed amino acids highlighted grey. Black diamonds on the bar graphs indicate unassigned or proline residues. Red arrows indicate residues for which significant methyl proton chemical shift differences were detected.

1.4 Impact of the modifications on the conformation of TAZ1 in complex with the CITED2 variants

NMR spectra of CITED2 variants complexed with $^{13}C/^{15}N$-labelled TAZ1 revealed modification site dependent structural changes. Modifications in the N-terminus (**1**) caused conformational changes at the C-termini of α1, α3 and α4 helices resembling the results of the truncation experiments, further emphasising the essential role of the N-terminal tail of CITED2 in the competition. Variants **2** and **4** showed localised effects restricted to their binding sites, while loop modifications (**3a-c**) caused widespread TAZ1 structural perturbations distant from the binding site, indicating the loop's critical role in connecting recognition elements. Integrating ITC and NMR data showed that αA and C-terminal modifications were well-tolerated with minimal conformational and cooperativity impacts, suggesting a role in maintaining the high affinity nature of binding to TAZ1. N-terminal modifications revealed that this region is necessary for inducing the correct structural changes for the conformational lock. The drastic impact of loop modifications suggests its importance in maintaining intramolecular cooperativity between binding motifs during folding and binding (Figure 5.).

2. Proof of concept DNA-templated synthesis of an encoded multivalent foldamer ligand

In this work, we created a DNA-encoded trimeric foldamer construct that could mimic the recognition elements of disordered transcription factors and which is amplifiable by PCR as a proof-of-concept for future library synthesis. This DNA-templated synthesis contains the following steps: (i) conjugation of maleimide-modified foldamers to 5'-phosphorylated thiol modified codons; (ii) hybridisation of foldamer-codon conjugates and primers on a complementary template; (iii) ligation of codons and primers using T4 DNA ligase; (iv) strand separation using gel electrophoresis; (v) amplification of the ligated strand with PCR.

2.1 Thiol-maleimide conjugation of encoding oligonucleotides and H14 foldamers

For ligatable oligonucleotide conjugates, thiol modification was placed mid-sequence using a newly developed S-S-tBu protected thiol-containing thymidine phosphoramidite with improved metal-free synthesis, leaving the 5'-phosphate available for ligation. Conjugation was optimised using the transferrin receptor binding peptide HAIYPRH extended with glycine and N-terminal maleimido-propionic acid. As the first step, the oligonucleotide was deprotected by TCEP treatment. Following purification, conjugation proceeded overnight at room temperature

with a good yield (Figure 6.). The tert-butyl protecting group enables selective removal from CPG-bound oligonucleotides for solid-phase conjugation prior to cleavage, reducing purification steps. While alpha-peptides cannot withstand concentrated ammonia treatment due to racemisation, solid-phase conjugation with maleimido-propionic acid succeeded with lower yield, forming stable thioether bonds through thiosuccinimide hydrolysis. Despite the alkaline stability of foldamers, the experiments were continued with solution-phase conjugation for better yields.

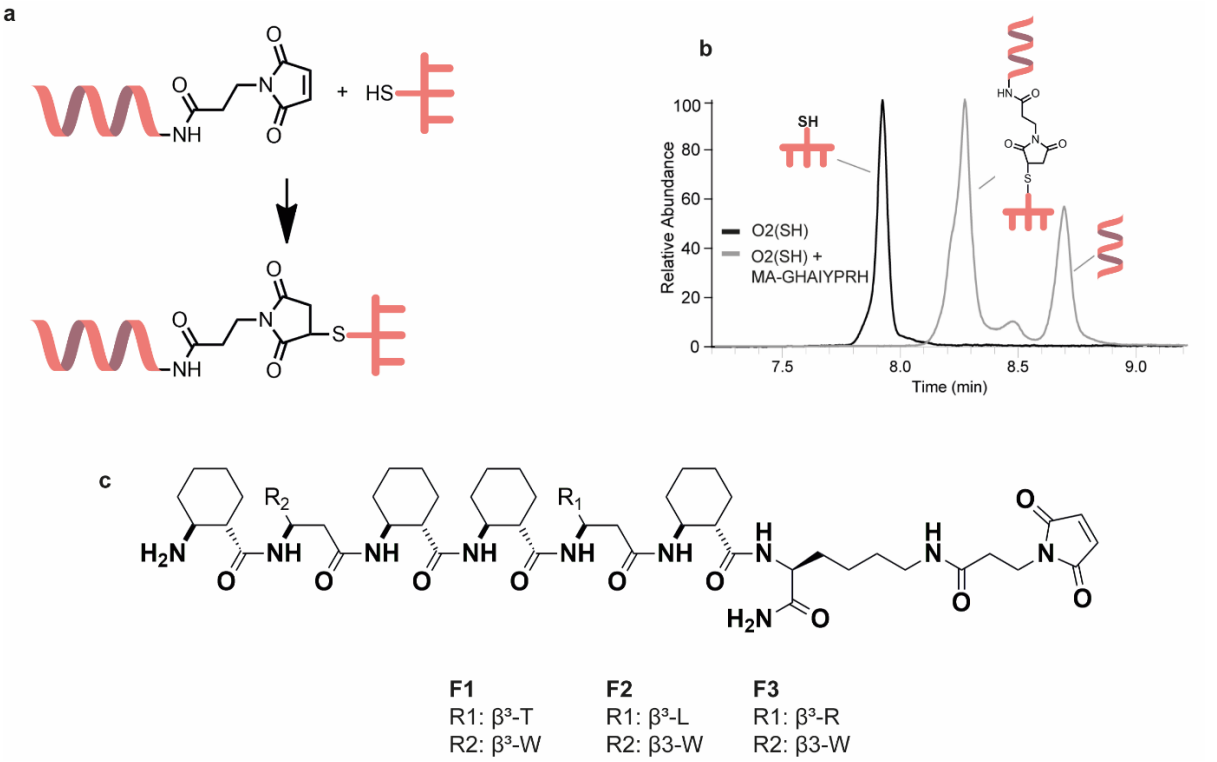

Figure 6. Thiol-maleimide conjugation in solution. (a) Reaction scheme for solution phase conjugation using 100 µM oligonucleotide, 200 µM TCEP, and 500 µM peptide, 10% DMSO, overnight room temperature. (b) HPLC-MS spectra of solution phase conjugation before and after the reaction (c) Primary structure of H14-foldamers used for conjugation

2.2 Ligation and amplification of the foldamer-oligonucleotide conjugates

The single-stranded DNA product presenting three foldameric ligands was prepared by templated ligation using T4 DNA ligase with a 60-mer template containing overhanging sequences for primer binding. The ligation was carried out using three-fold oligonucleotide excess and heat-cycles were introduced to inhibit the formation of truncated ligation products. Denaturing PAGE confirmed successful 60 bp strand formation (**LO1-3** and **LOSF1-3**) with appropriate truncated controls, proving that the presence of the foldamers did not influence ligation efficiency or yields. For PCR amplification, high-fidelity Phusion polymerase was used due to its ability to amplify C5-modified pyrimidine bases. Although amplification of **LOSF1-3** was not observed under the initial conditions, increasing the polymerase concentration to 0.15

U/µL resulted in the successful amplification of the foldamer modified ligated oligonucleotide (Figure 7).

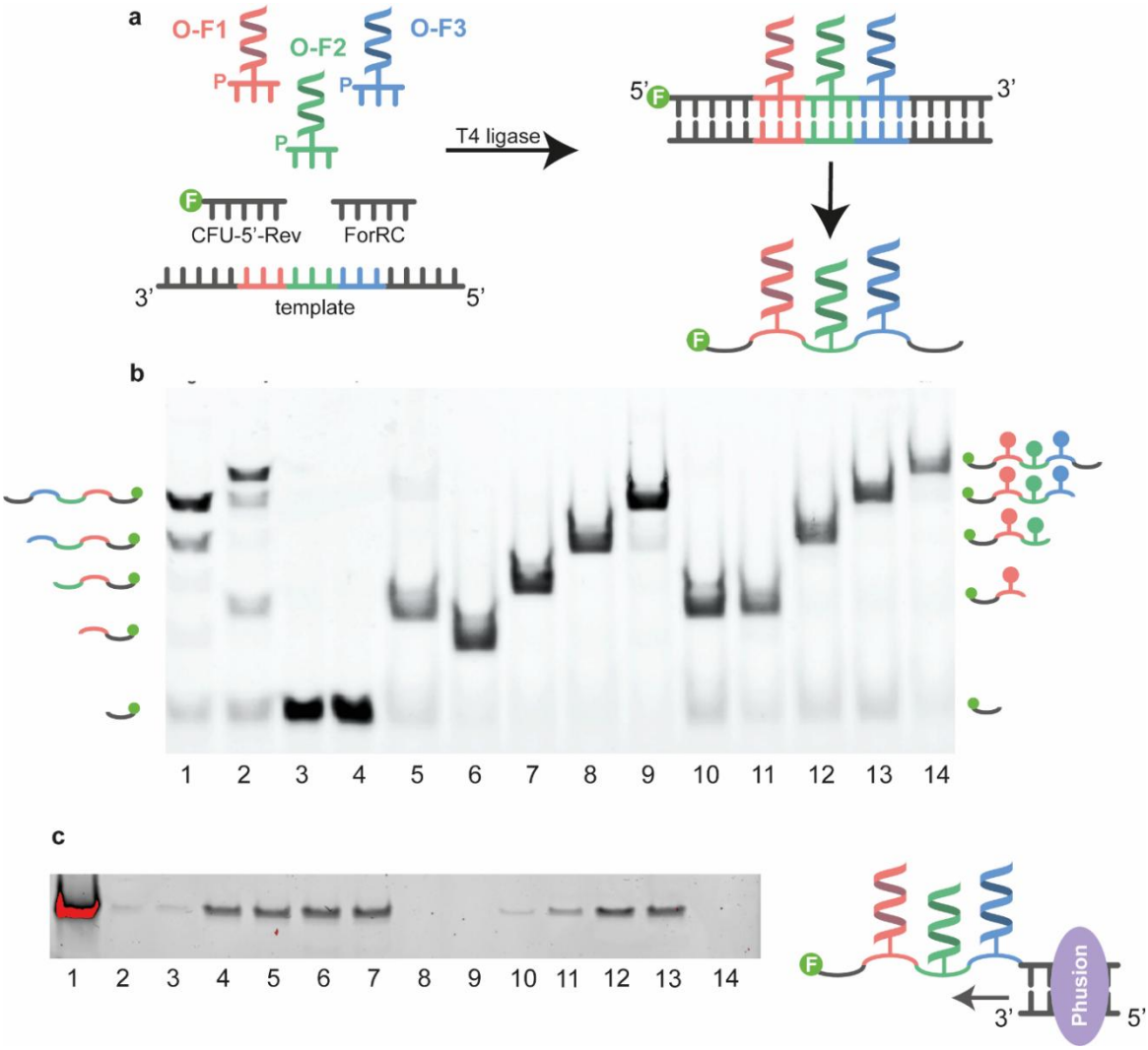

Figure 7. Templated ligation and amplification: (a) Scheme of the ligation process: to the template 5′ phosphorylated oligonucleotides (or foldamer–oligonucleotide conjugates) and extension sequences were added with fluorescent labelling on the 5′ end and ligated using T4 ligase. The ligation was carried out using 1 µM template and three equivalent excess oligonucleotides heated to 90 °C and cooled down to room temperature before adding T4 ligase. The heat–cool cycles and ligase additions were repeated three times. (b) Denaturing gel electrophoresis of the ligation mixtures with controls. Ligations performed without heat-cool cycles with mixtures containing non-functionalized oligonucleotides (Lane 1) or foldamer-oligonucleotide conjugates (Lanes 2–4). For control experiment template, (Lane 3) or T4 ligase (Lane 4) was omitted from reaction mixture. Lanes 5–9: Ligation performed using heat–cool cycles with a mixture containing template, T4 ligase, oligonucleotides **O1-3** and **F-5'-Rev**. The following reagents were omitted from the mixture: Lane 5: **O1**; Lane 6: **O2**; Lane 7: **O3**; Lane 8: **ForRC**; and Lane 9: All reagents present. Lanes 10–14: Ligation performed using heat–cool cycles with a mixture containing template, T4 ligase, oligonucleotide conjugates (**OSF1**, **OSF2** and **OSF3**), **ForRC**, and **F-5' Rev**. The following reagents were omitted from the mixture: Lane 10: **OSF1**; Lane 11: **OSF2**; Lane 12: **OSF3**; Lane 14: **ForRC**; Lane 15: All reagents present. Excluded reagents are indicated on the top of the gel and the full length and truncated products are shown on the sides of the gel. (c) Gel electrophoresis of the PCR products. Lane 1: ligated **LO1-3** hybridized to **template** amplified using 0.03 U/µL Phusion. Lane 2-7: amplification of the **template** directly from stock with 0.03 U/µL (Lane 2-3), 0.06 U/µL (Lane 4-5) and 0.15 U/µL (Lane 6-7). Lane 8 13: amplification of the ligated **LOSF1-3** extracted from denaturing PAGE with 0.03 U/µL (Lane 8-9), 0.06 U/µL (Lane 10-11) and 0.15 U/µL (Lane 12-13). Lane 14: negative control not containing any amplifiable oligonucleotide.

Full papers related to the thesis:

I. F. Hobor, Zs. Hegedüs, A. A. Ibarra, V. L. Petrovicz, G. Bartlett, R. B. Sessions, A. Wilson, T. A. Edwards (2022). Understanding p300-transcription factor interactions using sequence variation and hybridization. *RSC Chem. Biol.* **3**, 592-603

II. Z. Kupihár, Gy. Ferenc, V. L. Petrovicz, V. R. Fáy, L. Kovács, T. A. Martinek, Zs. Hegedüs (2023). Improved Metal-Free Approach for the Synthesis of Protected Thiol Containing Thymidine Nucleoside Phosphoramidite and Its Application for the Synthesis of Ligatable Oligonucleotide Conjugates. *Pharmaceutics* **15(1)**, 248

III. V. L. Petrovicz, I. Pasztuhov, T. A. Martinek, Zs. Hegedüs (2024). Site-directed allostery perturbation to probe the negative regulation of hypoxia inducible factor-1α. *RSC Chem. Biol.* **5**, 711-720

Scientific lectures related to the thesis

1. V. L. Petrovicz, Zs. Hegedüs, I. Pasztuhov, T. Martinek. A P300/HIF-1alfa kölcsönhatás allosztérikus szabályzásának feltérképezése a negatív regulátor módosításain keresztül.
   Zechmeister Előadói Nap 2022. Budapest – 2022.11.18.
2. V. L. Petrovicz, Zs. Hegedüs, I. Pasztuhov, T. Martinek. A P300/HIF-1alfa kölcsönhatás allosztérikus szabályzásának feltérképezése a negatív regulátor módosításain keresztül.
   26. Tavaszi Szél Konferencia. Miskolc – 2023.05.06., **3. helyezett**
3. V. L. Petrovicz, Zs. Hegedüs, I. Pasztuhov, T. Martinek. Selective allostery perturbation reveals the structural mechanism of the negative regulation for hypoxia inducible factor-1α.
   Zechmeister Előadói Nap 2023. Pécs – 2023.10.17.
4. V. L. Petrovicz, Zs. Hegedüs, I. Pasztuhov, T. Martinek. Kötőhely-specifikus allosztéria perturbáció a HIF-1α aktivitás szabályozásának feltérképezésére
   Zechmeister Előadói Nap. Budapest - 2024.10.15.
5. V. L. Petrovicz, Zs. Hegedüs, I. Pasztuhov, T. Martinek. Kötőhely-specifikus allosztéria perturbáció a HIF-1α aktivitás szabályozásának feltérképezésére
   Peptide Chemistry Symposium 2025. Balatonszárszó – 2025.05.20.