University of Szeged

Doctoral School of Pharmaceutical Sciences

Albert Szent-Györgyi Medical School

Department of Medical Chemistry

LIGAND DEVELOPMENT FOR THE INHIBITION OF TRANSCRIPTION FACTORS

Ph.D. thesis

Vencel László Petrovicz

Supervisors:

Prof. Tamás Martinek

Dr. Zsófia Hegedüs

2025

Table of Contents

1. Introduction

The inhibition of disease-relevant protein-protein interactions (PPIs) represent 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[1–3], 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 design. These additional challenges arise from the overrepresentation of intrinsic protein disorder in the transcriptional machinery, since the transactivation domains (TADs) of transcription factors (TFs) are predominantly disordered regions[4,5]. This dynamic behaviour enables them to bind promiscuously to multiple targets, while they form often transient, but highly specific interactions acting as molecular switches and adapting to different external stimuli[6]. Beside promiscuity, nuclear localisation and allosteric regulation involving TF PPIs further complicates 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 until today one of the most utilized PPI inhibitor class is monoclonal antibodies (mAbs) in medical practice[7–9], although a serious limitation of this drug class is the insufficient penetration through biological membranes[10–12]. Peptidomimetic drugs can serve as a potential alternative for PPI inhibitor development, which is proved by their wide range of applications in different clinical settings[13–15]. Thanks to 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 (i.e., rational modification of a chosen target's native ligand) or bottom-up (i.e., high throughput and fragment screening) approach. The former strategy heavily relies on the detailed structural and mechanistic information about a given interaction[16,17], 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[18,19]. However, using this approach would require a high-throughput screening strategy, which has limitations in relation with 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.

Based on the aforementioned development efforts and considerations, the foundation of my PhD research was to develop methods, which could support peptidomimetic inhibitor design against transcription factor interactions.

2. Aims

Our aim was to develop ways that addresses these challenges of targeting IDP mediated interactions using peptidomimetic structures. The top-down design approach necessitates a detailed knowledge about the targeted PPI, because modifications in a disordered ligand could potentially influence not only its affinity, but folding or ligand cooperativity, which can be mechanistically important. Therefore, first, we set out to investigate the molecular details of the competition between HIF-1α and its negative feedback regulator CITED2 C-TADs for the partially shared binding site on p300/CBP[20,21] surface. This interaction initiates an adaptive response to hypoxic stress, and is implicated as a potential target in solid tumours[22–26]. The molecular mechanism of the competition between the two disordered ligands relies on negative cooperativity induced by allosteric communication[27–29]. 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, that 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 folding into discrete and stable secondary structures, such as foldamers[30–32]. 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[33,34], 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)[35]. 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.

3. Literature background

3.1. Protein-protein interactions involving disordered ligands

The dominant concept in structural research of proteins in the 20th century was the 'sequence-structure paradigm' which states that the function of a given protein requires a folded structure[36,37]. However, emerging evidence revealed a novel protein class: intrinsically disordered proteins and regions (IDPs/IDRs). These lack stable secondary or tertiary structures, instead fluctuating between a wide range of conformational ensembles[6,38]. The importance of IDPs and IDRs is underscored in their prevalence in the human proteome, where 40% contain disordered sequences spanning 30 residues or more involved in diverse functions. (e.g.: entropic chains, assemblers, chaperones, post-translational modification sites)[39]. A highly important functional group is effectors, which modify the activity of their interaction partners and typically fold to distinct structures upon binding[40,41]. Due to their properties (e.g., transient binding, multiple interaction partners, posttranslational modifications, fast association and dissociation kinetics, dynamic complexes) these effectors have essential roles in the in cell signalling, transcription or translation[42,43].

3.1.1. Ensemble properties and functional features of IDPs

The sequence composition of intrinsically disordered regions is highly biased toward polar and charged residues, while bulky aromatic side-chains are significantly underrepresented compared to their folded counterparts[39,44,45]. The conformational space of a given IDR can be quantitatively described with its ensemble properties, such as the hydrodynamic radius, the end-to-end-distance or the radius of gyration, which function as probability distributions[46–48]. These properties are primarily encoded in the amino acid sequence, although other contributing factors, such as their biophysical (pH, salt concentration) or sequence context, proximity to biomolecules or post-translational modifications also have a significant contribution[49–52]. The polyelectrolyte class of IDRs has a high fraction of charged residues, and a high positive or negative net charge, which consequently makes them form swollen-coil ensembles.[6,53–56]. Additionally aromatic and aliphatic residues can form π-π or hydrophobic interactions that further influence ensemble properties[6,57].

The functional binding motifs participating in molecular recognition in IDRs can be loosely classified by the length of the disordered segments. Based on this classification system,

we can identify short linear motifs (SLiMs), molecular recognition features (MoRFs) or intrinsically disordered domains (IDDs).[6,39] Short linear motifs (or linear motifs-LM, MiniMotifs) are disordered peptide sequences ranging from 3 to 10 residues forming low-affinity interactions with structured or disordered interaction partners[58–60]. Although SLiMs enable specific binding with their interaction partners, they can possess substantial redundancy, meaning that only a fraction of the residues are conserved in a given SLiM (e.g.: NLxxxL motif for binding S-phase cyclins[61]). Additionally, a SLiM can bind to the binding pocket of multiple interaction partners, furthermore a given protein can bind multiple SLiMs at different binding sites[60]. The existence of multiple identical SLiMs in one sequence can also raise the affinity through allovalency[62]. By function, SLiMs can be grouped into two major families: functional binding motifs or sites for post-translational modifications. Importance of these motifs is clearly reflected in their abundance in the human proteome, with an estimated number in the hundred thousands[58,63].

Longer disordered segments spanning from 10 to 70 residues are called molecular recognition features (MoRFs) and participate in specific protein-protein interactions. Thanks to the plasticity of disordered ligands, a single MoRF can bind to different protein targets with distinct conformation. These sequences undergo disorder-to-order transitions upon binding, forming well-defined structures[63–65]. Based on the bound state, these structural elements can be divided into α-MoRFs forming alpha-helical (MLL binding to p300/CBP KIX domain)[66], β-MoRF forming beta-sheet (p21 disordered terminal inhibiting PCNA activity[67]) or *i*-MoRF forming irregular but rigid secondary structures[68]. Protein domains containing fully or mostly disordered regions that can include multiple SLiMs and MoRFs are termed intrinsically disordered domains (e.g.: N-TAD of p53)[69].

3.1.2. Molecular interactions of IDRs

The interactions of IDDs or IDRs containing long disordered segments are usually described by coupled folding and binding phenomena[41,70,71]. The mechanism of these events can be conceptualized with two extremities: conformational selection or induced fit[72]. Despite their disordered nature, the conformational ensemble of IDRs can be biased toward helical conformations that contain a low level of residual helicity. In conformational selection, a disordered segment samples a wide range of possible conformations enriched in helical states, where the target binding happens after the formation of the proper secondary structure[70,73]. In

an induced fit folding pathway, the interaction initiates by the formation of an encounter complex, which is stabilized by non-specific interactions, where the disordered component maintains its dynamic structure. The consecutive folding and binding of the disordered ligand, templated by the target happens after this state[74,75]. The entropic cost taken by the folding of the disordered ligand is compensated by the enthalpy contribution of the forming native contacts and the entropy-increasing effect of excluded surface water from the target.[76,77]

After folding and binding the disordered ligand can gain a completely rigid structure; however, the more common case is that the folded disordered ligand retains a considerable amount of disorder in its bound state, forming a 'fuzzy complex' with extensive protein-protein interfaces (PPIs)[63]. In a dynamic fuzzy complex, the bound structure of the IDR can be described with fast-exchanging conformational ensembles similar to its free state[78,79]. These dynamic complexes can be represented with the clamp or flanking models. In the clamp model, well-structured binding motifs are connected through a disordered linker (e.g.: HIF-1α-p300/CBP complex)[21], while in the flanking model, the opposite is true, where a structured binding motifs is flanked by disordered segments either on the amino- or carboxy-terminal or both (e.g.: p27 binding to cyclin-CDK complex)[80]. In extreme cases, the entire IDR can remain dynamically disordered after the binding event (random model). This is usually achieved by a repetitive SLiM pattern in the sequence of a given IDR, where at any given time only one or some of these SLiMs are in contact with the surface of the target, and they continuously replacing each other dynamically[81].

3.2. Hypoxic regulation through the HIF-1α/CITED2/TAZ1 interaction

A prime example for an interaction network involving disordered ligands is the HIF pathway, which plays a central role in the transcriptional response to hypoxic stress.

3.2.1. Role of HIF-1α in the hypoxic response

Normal oxygenation (8-10 kPa pO_2) is essential for cells to maintain optimal conditions for aerobic metabolism. To survive hypoxic stress (< 1.3 kPa pO_2), activation of the HIF-1 pathway allows cells to adapt to low oxygenation, by inducing responses that result in increased O_2 delivery and decreased O_2 consumption (e.g., vascularization, glycolysis)[82]. HIF-1 (Hypoxia-Inducible Factor-1) is a heterodimer consisting of an alpha subunit HIF-1α, and a constitutively expressed beta subunit (HIF-1β, also known as aryl hydrocarbon receptor nuclear

translocator, ARNT). The oxygen-labile subunit, HIF-1α similarly to its isoforms (HIF-2α, HIF-3α) is a member of the basic helix-loop-helix (bHLH) PAS superfamily[83]. HIF-1α has a nuclear localization signal (NLS), bHLH and two PAS (PAS A and PAS B) domains in the amino terminus, which are responsible for DNA binding and heterodimerization. The carboxy terminus contains an oxygen-dependent degradation domain (ODDD), a N-terminal, and a C-terminal transactivation domain (N-TAD and C-TAD) with an inhibitory domain (ID) in between, where another NLS is present[84,85] (Figure 1).

Under normoxia, the activity of constitutively expressed HIF-1α is regulated by multiple mechanisms. The most important among these is the degradation pathway, governed by oxygen-dependent prolyl hydroxylase domains (PHD1-3) acting as oxygen sensors. PHDs recognize two proline residues in the ODDD (P402 and P546), and consume 2-oxoglutarate as substrate (an intermediate of the Szentgyörgyi-Krebs cycle) for proline residue oxidation[84,86]. These post-translational modifications result in the recognition of hydroxylated prolines by the von Hippel-Lindau protein (pVHL), an E3-ubiquitin ligase. Following the polyubiquitination of HIF-1α, it is rapidly degraded by the 26S proteosome[87]. Inhibition of HIF-1 activity during normoxia is further regulated by the factor inhibiting HIF-1 (FIH-1), which hydroxylates an asparagine residue in the C-TAD (N803), disrupting the binding with its transcriptional coactivator p300/CBP, acting as a second oxygen sensor[88,89] (Figure 1).

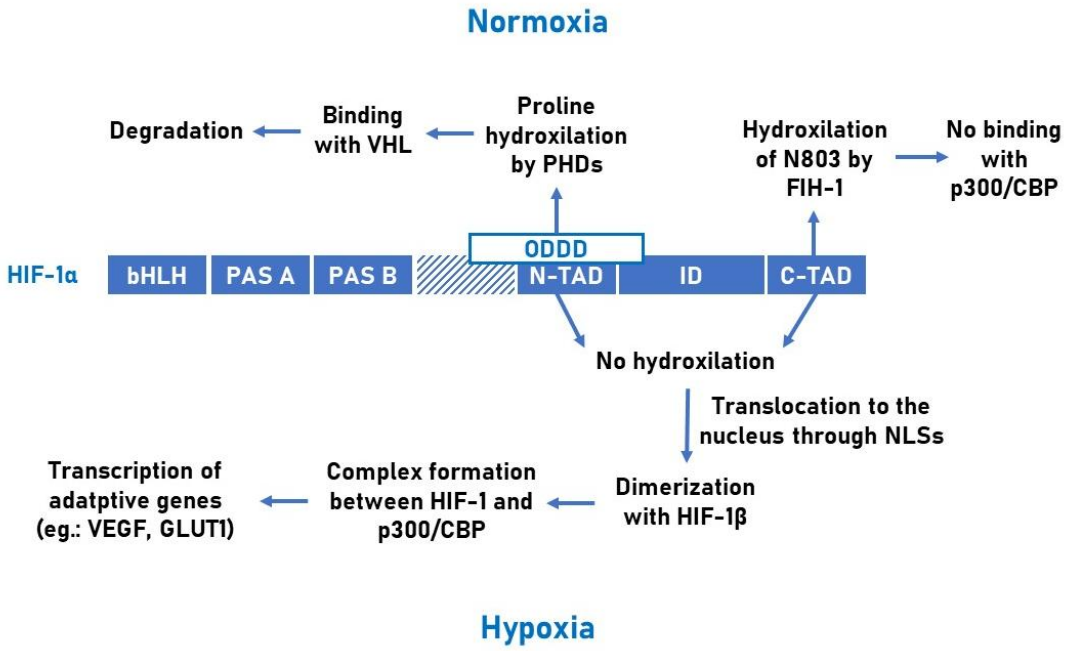


Figure 1. HIF-1α domains, regulation and activation during hypoxia

During hypoxic stress, low oxygenation inhibits the degradation pathway, resulting in HIF-1α accumulation in the cytosol[90]. The alpha subunit then translocates to the nucleus, where it dimerizes with HIF-1β[91]. The HIF-1 heterodimer targets genes that contain a hypoxia response element (HRE) in their regulatory region[82,92,93]. To become transcriptionally active, it must assemble with its co-activator p300/CBP through its disordered C-TAD[94]. P300/CBP histone acyltransferase is a crucial crossroads in transcriptional control, acting as a co-activator for multiple transcription factors (e.g.: p53, c-Myb, STAT1)[5,95,96]. Depending on cell type, the number of genes directly or indirectly induced by the HIF pathway ranges from 200 to 1-5% of all human genes. HIF-induced proteins help maintain the metabolic and survival needs of the cell, by either increasing O_2 supply or decreasing O_2 consumption[92,97]. The former is achieved by angiogenesis[98,99], modulation of iron transport[100], increasing the blood flow[101,102] or erythropoiesis[103]. The latter occurs via increased glycolytic rate with elevated expression of proteins involved in glycolysis[104,105] or glucose transport[106,107]. Additionally, the expression of pro- and anti-apoptotic proteins[108,109], secondary transcription factors[110,111] and self-regulatory proteins such as PHD3[112] or CITED2[24] are also upregulated[113].

Although HIF-1 activity is essential for the survival of hypoxic conditions and for the maintenance of stable O_2 homeostasis, it also serves a protective role in malignant solid tumours, since more than half of all malignancies have a median oxygen concentration of less than 10 mm Hg, rendering them hypoxic[114,115]. HIF-1 promotes tumour growth by increasing the levels of endothelial growth factors such as VEGF, stimulating the formation of new blood vessels to increase oxygen supply[116]. Simultaneously, HIF-1 is the main contributor of the Warburg effect, a characteristic metabolic shift for malignancies. This includes the upregulation of glucose transporters and glycolytic enzymes, enabling energy production through non-oxidative glucose metabolism[114,117]. Additionally, to allow vascularization, HIF-1 upregulates the expression of metalloproteases to degrade the extracellular matrix, and inhibits the production of E-cadherin resulting in metastasis, increasing morbidity[115,118,119]. These mechanisms render HIF-1 as a desirable target for cancer therapy. A potential node in the HIF-1 cascade that can be inhibited is the HIF-1α/p300 interaction, which would attenuate downstream induction of HIF-1 regulated genes[113].

3.2.2. Molecular insight into the competition between CITED2 and HIF-1α for p300/CBP

As mentioned above, HIF-1 induces the expression of self-regulatory proteins in addition to adaptive genes, which provide a negative feedback loop in the hypoxic response[92]. The most prominent self-inhibiting mechanism is carried out by CITED2 (p300/CBP interacting transactivator with Glu/Asp rich carboxy terminal domain 2; p35srj)[120]. CITED2 does not have a DNA binding domain itself; it acts as a transcriptional co-regulator by interacting with numerous transcription factors (e.g.: TFAP2, Myc)[121,122] and plays an essential role in embryonic development[123–125], stem cell survival, and renewal, among others[126,127]. Although ubiquitously and constitutively present at a baseline level, CITED2 expression is drastically increased by HIF-1 induction. CITED2 competes with HIF-1α C-TAD using its disordered carboxy-terminal transactivation domain (C-TAD) for the partially overlapping binding site on the TAZ1 domain of p300/CBP, implementing a highly efficient negative feedback mechanism in the HIF-1 cascade, acting as a molecular switch[27].

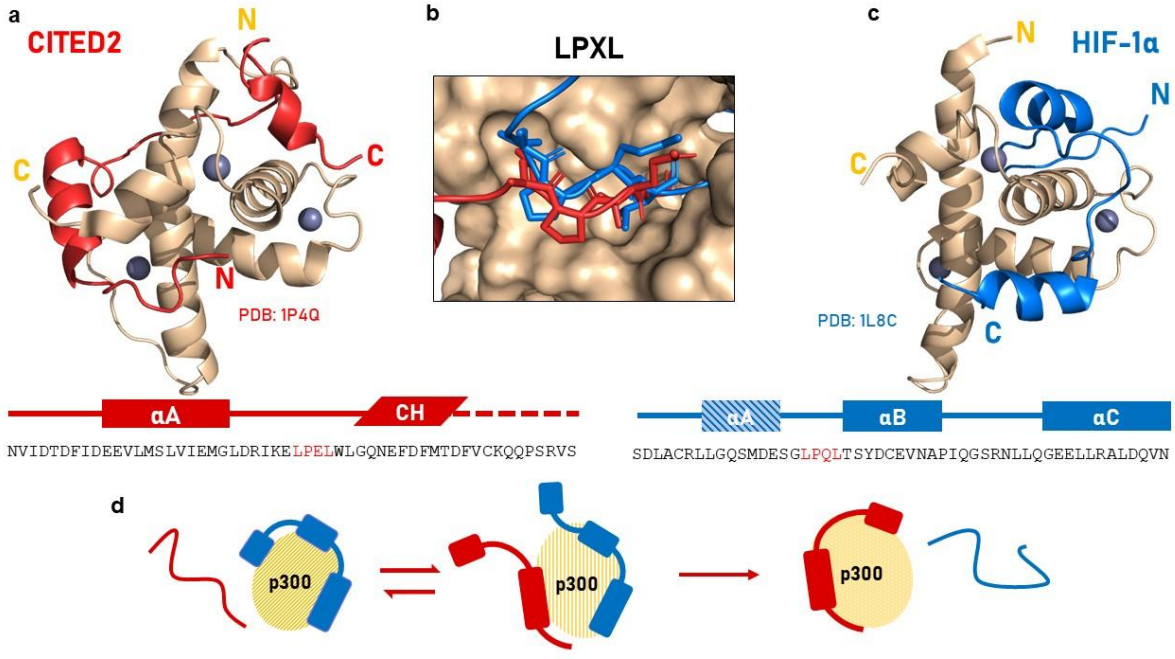


Figure 2. a) Structure of CITED2 in complex with p300. b) Overlapped structure of the LPEL and LPQL motif of CITED2 and HIF-1α, respectively in complex with p300. c) Structure of HIF-1α in complex with p300. d) Schematic representation of the competition mechanism between CITED2 and HIF-1α

TAZ1 is a globular domain consisting of four alpha helices designated as α1-4 and three highly conserved HCCC zinc-binding motifs[21]. The structure of TAZ1 has a considerable amount of conformational plasticity, enabling adaptation to multiple disordered ligands beyond

HIF-1α and CITED2 (e.g.: STAT2, NF-κB p65 subunit, p53)[21,128]. Binding of HIF-1α and CITED2 C-TAD domains proceeds through a transition state (encounter complex) governed by non-specific electrostatic interactions between the positively charged TAZ1 and the negatively charged, HIF-1α or CITED2 (net charge is -5 and -7 respectively)[129,130], where both C-TADs retain their disordered ensemble[131]. This state initiates the folding and binding of both IDPs forming specific contacts and distinct binding motifs on the TAZ1 surface. HIF-1α features three αMoRFs designated as αA, αB and αC with linker regions between them. CITED2 has an αMoRF at the amino terminus (αA), which is connected by a linker with a C-terminal hydrophobic amino acid rich structure. In their bound state they share a partially overlapping binding site and a conserved LPQ/EL binding motif targeting the same cleft on TAZ1. Considering the structure of TAZ1 in these two complexes, there are major differences in the conformation in its α1 and α4 helices[27,94,132]. Despite the similar low nanomolar affinities of these ligands accompanied by rapid association and slow dissociation kinetics, CITED2 displaces HIF-1α entirely in equimolar concentrations in a switch like, virtually irreversible manner[27,129,131,133]. The experimental and computational studies investigating the molecular mechanisms of the displacement agree on basic concepts; however, they provide alternative explanations or highlight different determinants for the unidirectional nature of the competition[27,29,129–132,134–137].

One aspect underlying the efficient displacement process is the difference in backbone dynamics between HIF-1α and CITED2 in their bound state. The loop region connecting αB and αC, and the N-terminus preceding the LPQL motif in HIF-1α retains a considerable amount of flexibility when bound to TAZ1, while CITED2 exhibits significantly less flexibility, resulting in a stronger thermodynamic coupling between its binding motifs[27,134,136]. The role of conformational plasticity of TAZ1 in the displacement cannot be ruled out either, as it becomes more rigid upon displacement of HIF-1α by CITED2[27,134,136,138].

Out of the several contributing factors one is that CITED2 has a more negative net charge than HIF-1α (-7 and -5 respectively), which is more pronounced in its N-terminal region, while in HIF-1α the acidic residues are more evenly distributed. This enables faster binding kinetics for CITED2, while the folding and binding of its N-terminal to the TAZ1/HIF-1α complex displaces the highly dynamic N-terminal region of HIF-1α, forming a transient ternary intermediate state[27,130]. This state is a crucial aspect of unidirectionality, since the binding of the αA helix of CITED2 induces allosteric conformational changes in the α4 helix of TAZ1

which destabilizes the binding of the high affinity αC helix of HIF-1α[27,137]. Another critical step of the competition involves the displacement of HIF-1α's conserved LPQL motif by CITED2's LPEL motif; mutations in this region significantly impair CITED2's competitive ability[27,134]. However, binding of αA helix and LPEL motif alone is insufficient for efficient inhibition. The absence of CITED2's C-terminal hydrophobic region reduces its competitive efficiency, indicating a key role in driving unidirectional displacement. Notably, deletion of the C-terminus drastically lowers CITED2's binding affinity also by approximately two orders of magnitude, potentially confounding competition measurements and obscure the C-terminal region's role in allosteric regulation[29]. Regardless of these aforementioned results the exact roles of CITED2's binding motifs in the allosteric structural changes and negative cooperativity remains elusive.

3.3. Ligand development for the disruption of protein-protein interfaces

Targeting therapeutically relevant PPIs is considered challenging due to the extensive interaction surfaces (1500-3500 $Å^2$), often shallow binding clefts or scattered hot-spot residues throughout the interface[3,139,140]. Additionally, the functional importance of disordered ligands in PPIs needs to be taken into account, which could further decrease the chance of finding traditional small molecule hits.[141]

3.3.1. Development of IDP-based PPI inhibitors with top-down design

A possible direction toward functional PPI inhibitors involves top-down development of peptidomimetics, where the starting point typically includes a fragment of the native ligand's sequence[142,143]. A prerequisite of this process is detailed structural information on the target protein in complex with its ligand of interest, supplemented with the molecular details or other factors contributing to the binding mechanism[144]. These native peptides that serve as starting points do not meet the requirements of drug candidates, mostly due to their poor proteolytic stability, beside their poor oral bioavailability or rapid renal clearence[145]. However, rational modification of the starting sequence can lead to ligands with improved pharmacokinetic properties. Modification strategies usually utilize either the rigidification of the peptide secondary structure, or the increase of the proteolytic stability by the implementation of non-natural amino acids in the sequence[17,144,146]. Peptide cyclization is an excellent strategy to develop drug-like ligands with improved proteolytic stability and in many cases increased membrane permeability, producing more than 50 therapeutics currently in use[147,148]. Another

approach is peptide stapling, where structure stabilization and proteolytic stability are achieved through the covalent linkage of amino acid side chains[149] or hydrogen bond surrogates (HBS), where the i, i+4 hydrogen bond is replaced with a hydrocarbon staple[150]. Such stapled peptides (Wilson at al.)[151] and HBS peptides (Arora at al.)[152] have been synthetized to mimic HIF-1α C-TAD, for the disruption of HIF-1α-p300/CBP interaction.

Another way to increase the proteolytic and structural stability of a ligand is through replacement of native residues by non-natural amino acids. One of the major and most studied possibility involves implementing β-amino acid residues in the sequence. These non-natural building blocks are the closest neighbours of their alpha counterparts, differing by the addition of an extra methylene group. These can be β^{2-}, β^{3-}, $\beta^{2,3}$- amino acids involving cyclic residues also[153–155]. While pure β-peptide foldamers can organize into diverse conformations depending on the residues (see in next section), chimeric α/β-peptides further widens the range of possible constructs[30]. Thanks to their close homology to α-amino acids, these residues can be incorporated in helices[156], sheets[157,158] or turns[159]. In helical regions α→β^3, α→cyclic $\beta^{2,3}$ replacements with rational design following a sequence based ααβ, αααβ or ααβααβ pattern typically yield more stable helical conformations due to the helix inducing effect of cyclic β-amino acids. Here, replaced residues occupy a strip on face of the helix. The reason behind this patterning is that replacements are much more tolerable in solvent exposed surfaces, than on the binding surface. This replacement strategy produced many chimeric α/β-peptidomimetics, resulting comparable affinity to the target with the native ligand[154] (e.g.: chimeric peptides mimicking HIV gp41[155] or chimeric BH3 targeting Bcl-x_L[160]).

3.3.2. Mimicry of Molecular recognition features (MoRFs) using foldamers

When structural data on therapeutic targets are unavailable, screening foldameric libraries offers an alternative approach to identify protein-protein interaction inhibitors. Foldamers are sequence-specific oligomers that mimic natural protein secondary structures by adopting defined three-dimensional conformations. Their non-natural frameworks encompass β-, γ- or δ-peptides, azapeptides, and aromatic backbones[30,32,161,162]. β-peptide and α/β-peptide foldamers represent the most extensively studied class, capable of self-organizing into various secondary structures including H10/12, H12, H14, and H14/15 helices. Cyclic β-amino acid residues like ACPC (2-aminocyclopentanecarboxylic acid) and ACHC (2-aminocyclohexanecarboxylic acid) effectively stabilize helical conformations, requiring as few

as 6 residues to form stable structures. (S,S)-ACHC units create stable H14 helices (3_{14}-helices) through hydrogen bonds between the *i* residue's amide proton and the *i*+2 position's carbonyl oxygen, while (S,S)-ACPC residues form H12-helices (2.5_{12}-helices) with hydrogen bonds between *i* and *i*+3 residues.[30]

A successfully utilized application of H14 β-peptide libraries bearing two proteinogenic side-chains is the screening against immobilized protein targets to develop peptidomimetic ligands in a bottom-up, fragment-based approach[18,19]. Despite the limited structural complexity of these ligands, they can identify either ortho- or non-orthosteric binding sites on protein surfaces, while the weakly binding foldameric fragments (K_D = 1μM - 1mM) can serve as starting points for de novo ligand development[163,164]. To ease the detection of weak binders, the H14 library members can be modified with photoreactive tags, which results in covalent crosslinking between the binding fragments and target protein after UV irradiation. This method is especially useful against targets with extremely shallow binding clefts[165]. To develop peptidomimetics with improved binding properties the covalent linkage of low affinity foldamer hits (binding to different surfaces of the target protein) proved to be successful by dynamic covalent ligation (DCL) approach. In these systems the weak binder foldamers are modified with thiol-groups, and incubated with the target protein. After the incubation the disulphide foldamer dimers with the highest affinity will be enriched due to the templating effect of the target protein[165,166].

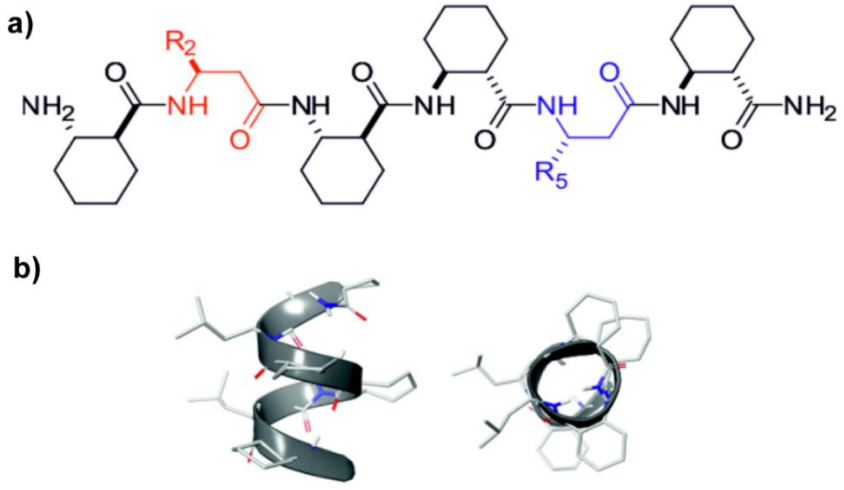


Figure 3. a) sequence and b) structure of H14 foldamers[166]

Although these screening approaches are able to identify binders with acceptable affinities, library size and low throughput are still a limitation. In vitro evolution methods that

can handle non-natural amino acids can be an alternative; however, not applicable to ligands that contain non-natural backbones exclusively, therefore alternative approaches are required.

3.3.3. DNA-encoded libraries for ligand development

DNA-encoded libraries (DEL) have proven to be a robust, cost-effective, and easily accessible alternative for screening vast compound libraries with sizes comparable to high-throughput screening technology, without requiring resource-intensive facilities. In a typical DEL, library members tagged with unique DNA-barcodes are selected against an immobilized target and the non-binders are washed away. The DNA-tags of eluted binding members are then amplified by PCR, and the amplified oligonucleotides are efficiently identified using second generation high-throughput sequencing[33,34,167,168]. Probably the most critical component of DEL is the library synthesis, which requires the use of DNA orthogonal chemistries[169]. The general library synthesis approach is the combinatorial split-and-pool synthesis, where the conjugation of building blocks and tagging with appropriate DNA-barcodes happens in (usually 2-3) iterative steps. This technique can easily produce library sizes in billions[33,34,170]. Since initial breakthroughs, this approach has been complemented with DNA-templated synthesis (DTS)[171], encoded self-assembling chemical libraries (ESAC)[172] or YoctoReactor technologies[173,174]. Although affinity selection remains most common, other cell-based (e.g.: selection against GPCRs) and in-solution selection methods emerged also (e.g.: one-bead-one-compound libraries, OBOC; DNA-programmed affinity labelling, DPAL)[35]. Utilization of DELs resulted in a number of drug candidates that entered clinical trials[175]. These include an sEH[176], a RIP1 kinase[177] and an autotaxin inhibitor[178].

The group of Liu and Hili developed a DTS-derived method termed ligase-catalysed oligonucleotide polymerization (LOOPER) for the selection of modified DNA-aptamer libraries[179]. LOOPER uses a template library that includes a set of a pentameric codons encoding the corresponding 256 different anticodons in which the adenine residue at the 5’ end is modifiable at its Hoogsteen face with various small molecules or polypeptides. The first LOOPER step hybridizes the modified anticodons and primers on the templates, which is followed by ligation with T4-DNA ligase[180–185]. After strand separation, the modified and ligated aptamer library is selected against an immobilized target. Similar to general DEL, the DNA code of the binders are amplified by PCR and sequenced for hit identification. A key advantage of LOOPER is the seamless integration with SELEX[186] (systematic evolution of

ligands by exponential enrichment). This means that the amplified oligonucleotides of binders are reused as templates after strand separation in iterative cycles of selection, enriching the highest affinity library members[187]. The LOOPER-SELEX process proved to be successful in selecting high affinity nanomolar ligands against multiple targets, with as few as six round of selection[179].

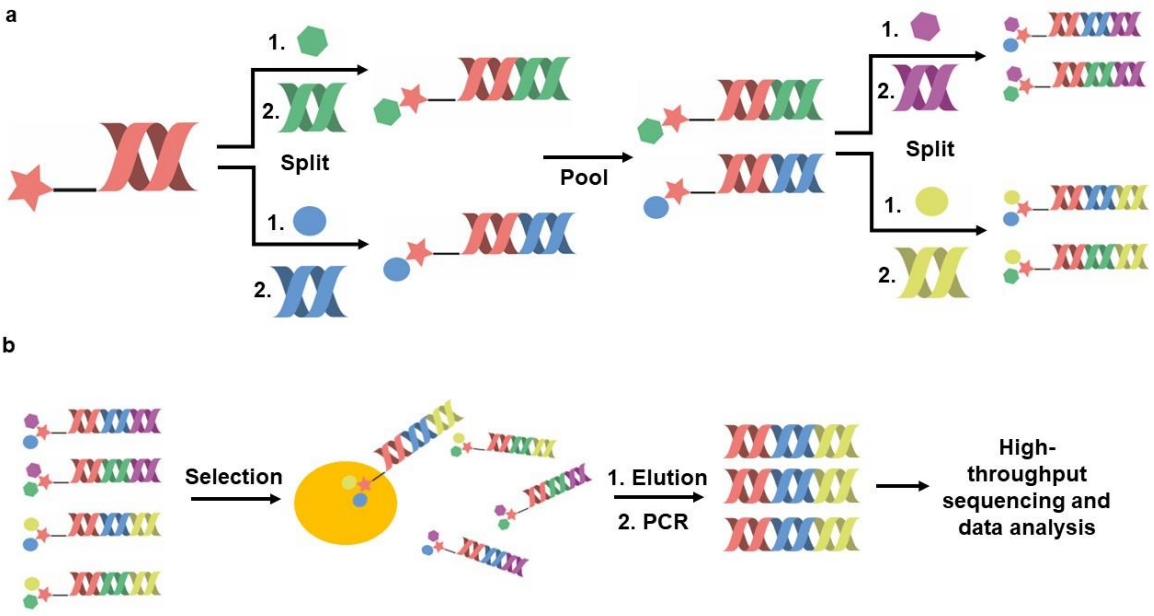


Figure 4. Schematic representation of a) split-and-pool library synthesis and b) DEL library selection and hit detection

4. Materials and methods

4.1. Peptide synthesis and purification

HIF-1α and CITED2 peptides were synthetized at 0.1 mmol scale on Tentagel R RAM resin (resin loading 0.19 mmol/g, Iris Biotech) using a microwave assisted automated peptide synthesizer, CEM Liberty Blue. Couplings were carried out using 5 equivalent amino acid excesses for α-amino acids, and 3 equivalent excesses for β-amino acids. Coupling reagents were DIC (Diisopropylcarbodiimide) and Oxyma (Ethyl cyano(hydroxiimino)acetate) dissolved in DMF (dimethylformamide). The deprotection solution contained 10% (w/v) piperazine and 5% (w/v) Oxyma dissolved in 10% absolute ethanol/NMP (N-methyl-2-pyrrolidon) mixture to efficiently reduce aspartimide formation[188]. Required amounts were calculated using the built-in reagent calculator of Liberty Blue. All amino α-amino acids were double coupled using standard high swelling (HS) Liberty Blue methods. For β-amino acids, a single coupling was used with increased coupling time (4 minutes).

The H14-foldamers were synthetized manually by SPPS in fritted SPE tubes at a 0.1 mmol scale on Tentagel R RAM resin (resin loading 0.19 mmol/g) using Fmoc chemistry. Couplings were carried out using 3 eq. β^3-amino acids 3 eq. HATU (1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate) and 6 eq. of DIPEA (N,N-Diisopropylethylamine) as coupling agents in DMF for 3 hours/coupling. Fmoc deprotection was done by a mixture containing 5% piperidine and 2% DBU (1,8 diazabicycloundec-7-ene) in DMF for 5 and 10 minutes, consecutively. The resin was washed using DMF and DCM after every step. The side-chain Mmt (4-methoxytrytil) protecting group lysin was cleaved using an acetic acid:trifluoroethanol:DCM 1:2:7 mixture. Coupling of maleimido propionic acid to the lysin side chain was carried out using standard coupling conditions described above, which was followed by the Fmoc deprotection of the lysin residue.

Acetylation: Peptides and foldamers were acetylated using 10 equivalents of acetic acid anhydride and DIPEA in DCM:DMF 1:1 in a fritted SPE tube for 20 minutes 2 times at room temperature.

Fluorescent labelling: Fmoc-aminohexane carboxylic acid (Ahx) was coupled manually to the N-terminus of HIF-$1\alpha_{786\text{-}826}$ using 5 equivalent amino acid excess, HATU as the activator, and DIPEA as the base for 3 hours at room temperature. Fmoc deprotection was carried out

using 2% DBU, 2% piperazine solution in DMF for 5 then 15 minutes. The resin was washed with DMF and DCM. 5(6)-carboxyfluorescein (Flu) was coupled using 3 equivalent excess, HATU, and DIPEA as coupling agents overnight at room temperature. The resin was washed with the deprotection solution before cleavage to remove non-specific carboxyfluorescein esters.

Methyl esterification of aspartimide products: The esterification of the apartimide side-products were carried out in MeOH solution containing 2% DIPEA[189].

Cleavage: Peptides were cleaved using TFA:DTT:TIS:Me_2S:H_2O (85:2.5:5:5:2.5) for 3 hours, after which TFA was evaporated, and the crude peptide was precipitated in ice-cold diethyl-ether. The precipitate was washed with ether, then redissolved in ACN:H_2O mixture and lyophilized.

Preparative RP-HPLC was carried out on a JASCO PU-4180 system equipped with a diode array detector (MD-4015) and an automatic fraction collector (Advantec, CHF122SC). Fraction collection was monitored and programmed using ChromNAV software. Peptides were dissolved in DMSO and purified on a Phenomenex Luna C18 (250 x 10 mm or 250 x 20 mm) column using appropriate gradients. The purity of the peptides was characterized by HPLC-UV and HPLC-MS measurements using Dionex Ultimate 3000 HPLC system interfaced with an LTQ XL (Thermo Scientific) ion trap mass spectrometer. For analytical measurements Aeris Widepore C18 (5mm, 100 Å, 250 x 4.6 mm) column and the following gradient: 0 min 5% B, 25 min 100% B, 1 ml/min flow rate. The following eluents were used for HPLC-UV methods: A: 0.1% TFA/H_2O; B: 0.1% TFA/ACN and for HPLC-MS: A 0.1% HCOOH/H_2O; B: 0.1% HCOOH/ACN.

4.2.Isothermal titration calorimetry

Titrations were performed in 40 mM Na-phosphate, pH 7.5 100 mM NaCl, 1 mM DTT, 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$\alpha_{776\text{-}826}$ complexes were prepared by titrating HIF-1$\alpha_{776\text{-}826}$ in 60 μM concentrations to TAZ1 in the cell using 5 μM concentration. These titrations were stopped when the amount of HIF-1$\alpha_{776\text{-}826}$ reached close to saturation, meaning when significant differences in injection heats were not observable. This resulted in a 1.2-1.5 excess of HIF-1$\alpha_{776\text{-}826}$ compared to TAZ1 in the cell. Following the

saturation of TAZ1, the excess volume was discarded from the cell, and the complex 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.

Data analysis: Integration of the raw thermograms were carried out using SVD analysis implemented in NITPIC[190] and were fitted to a competition model that includes cooperative interactions and ternary complex formation (Figure 1.)[191].

$$\mathrm{TAZ1 + HIF\text{-}1\alpha} \overset{K_{\mathrm{HIF}},\,\Delta H_{\mathrm{HIF}}}{\rightleftharpoons} \mathrm{[TAZ1{:}HIF\text{-}1\alpha]}$$

$$\mathrm{TAZ1 + CITED2} \overset{K_{\mathrm{CITED}},\,\Delta H_{\mathrm{CITED}}}{\rightleftharpoons} \mathrm{[TAZ1{:}CITED2]}$$

$$\mathrm{TAZ1 + HIF\text{-}1\alpha + CITED2} \overset{\alpha K_{\mathrm{CITED}},\,\Delta H_{\mathrm{terner}}}{\rightleftharpoons} \mathrm{[TAZ1{:}HIF\text{-}1\alpha{:}CITED2]}$$

$$\alpha = \frac{\mathrm{K_{[HIF:}\mathit{TAZ}1]/CITED}}{\mathrm{K_{CITED}}}$$

$$\Delta \mathrm{G_{[p300:HIF:CITED]}} = -\mathrm{RTln}(\alpha \mathrm{K_{CITED}}\,\mathrm{K_{HIF}}) = \Delta \mathrm{G_{HIF}} + \Delta \mathrm{G_{CITED}} + \Delta \mathrm{g}$$

$$\Delta \mathrm{H_{[p300:HIF:CITED]}} = \Delta \mathrm{H_{CITED}} + \Delta \mathrm{h}$$

Fitted parameters were as follows: K_{CITED}, $\Delta H_{\mathrm{CITED}}$, K_{HIF}, ΔH_{HIF}, Δh and the cooperative constant α ($K_{[\mathrm{TAZ1/HIF}]/\mathrm{CITED}}/K_{\mathrm{CITED}}$) including baseline and incompetent fraction fit for TAZ1. Two titrations were fitted simultaneously: CITED2 (or variant) to TAZ1 and CITED2 (or a variant) to TAZ1/HIF-1α complex where the concentration of HIF-1α was given as the molar ratio of $n_{\mathrm{HIF}}/n_{\mathrm{TAZ1}}$. Restraints were applied for K_{HIF} and ΔH_{HIF} based on average values of three direct HIF-1α to TAZ1 titrations. 68 % confidence intervals were calculated using the automatic confidence interval search built in SEDPHAT[190,192,193]. Figures were prepared using GUSSI[194].

4.3. Fluorescence anisotropy

Fluorescence anisotropy was measured in 384-well black plate 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 mM 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, where the fluorescently labelled Flu-HIF-1α was replaced with buffer in the reaction mixture, which was subtracted from the raw data. The competitor peptides (HIF-$1\alpha_{776\text{-}826}$, native CITED2 and the 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-1α$_{786-826}$) at a final concentration of 25 nM to each well. Plates were incubated for 10 minutes at 35 °C before reading. Fluorescence intensity and anisotropy were calculated according to the following equations:

$$I = (2PG) + S \quad (10)$$

$$r = (S - PG) \quad (11)$$

where I = total intensity, r = anisotropy, P = perpendicular intensity, S = parallel intensity, G = instrument factor set to 1. Apparent competition K_D ($K_{D,app}$) values for the competitor peptides were fitted using a method described previously, [195,196] see also Appendix I for fitting equations.

4.4. 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. ^{15}N-^{13}C-TAZ1$_{330-424}$ was mixed with 1.2-1.5 equivalent CITED2 peptide or equimolar amount of HIF-1a and CITED2 peptides. The complex was concentrated using 10K Amicon Ultra filters to its final volume resulting in 90-100 μM protein concentration. 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 when signal assignment was necessary, using pulse schemes implemented in Topspin 3.5 (Bruker), with excitation sculpting water suppression. Processing was carried out using Topsin 3.5, and data were analysed using NMRFAM-Sparky[197]. 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).

5.4. Preparation of foldamer-oligonucleotide conjugates in solution phase

The t-Bu-SS protection group was removed using TCEP in solution pH 7 and the oligonucleotide was purified using RP-HPLC. To prevent disulphide formation of the thiol functionalized oligonucleotides, 2 equivalents of TCEP were added to the stock solution and left for 2 h at room temperature. The maleimide thiol ligation was performed using 100 μM oligonucleotide and 5 equivalent foldamers (0.5 mM, diluted from 5 mM DMSO stock) in 50 mM Tris (pH 7) overnight at room temperature under nitrogen atmosphere. The reaction was followed by LCMS. The conjugates were purified using RP-HPLC.

4.5. Preparation of conjugates on solid phase

The CPG-bound oligonucleotides were treated with 100 mM TCEP pH 7 for 3 h and the TCEP solution was changed 3 times. The CPGs were washed with ACN and H2O and then reacted with 3-maleimidopropionic acid 10 mM in 50 mM Tris, pH 7.2 20% DMF, overnight. The cleavage was performed using concentrated ammonia, 24 h, at room temperature. The crude product was analysed using LCMS.

4.6. Templated ligation

The samples contained 1 µL 10× T4-ligase buffer, 1 µL 10 µM template, 0.6 µL 50 µM 5′ phosphorylated 10mer oligonucleotides or peptide oligonucleotide conjugates, 0.75 µL 40 µM extension sequences (ForRC and CFU-5'-REV), and 0.5 µL 5 U/µL ligase. The reaction mixtures were diluted with sterilized distilled water to 10 µL. The ligation was carried out in 4 consecutive cycles of 15 min ligation at room temperature, 1 min incubation at 90 °C followed by cooling down to room temperature, and a new aliquot of ligase addition. Before PAGE analysis, 10 µL 2× denaturing DNA dye (14 M Urea, 10× TBE, 0.01% bromophenol blue) was added to the mixtures.

4.7. Denaturing DNA gel electrophoresis

A 1 mm thick, 15% acrylamide/8 M urea/TBE gel was cast manually using 5% acrylamide/2.15 M urea/TBE as stacking gel. The gel was run in a 60 °C TBE buffer. To the 10 µL ligation mixtures 10 µL at 60 °C 2× denaturing DNA dye was added, of which 10 µL was loaded onto the gel, followed by running the gel at 110 V for 10 min then 150 V for the rest of the run.

4.8. PCR amplification

Each PCR tubes contained either 1 pmol **template** hybridised to **LO1-3**, 5 fmol **template** or 70 fmol **LOSF1-3,** 2.5 µL dNTP (2 mM), 0.5 µL **For primer** (10 µM) and **5'-CFU-Rev primer** (10 µM), 0.31 µL, 0.63 µL or 1.56 µL High-Fidelity Phusion polymerase (2,5 U/µL), 2.5 µL Sybr Green (1:1000) and 2.5 µL Pfu Buffer (10x) diluted with sterilised H_2O to 25 µL. The touch down PCR cycle was 5s 98°C denaturing step, 15s 55°C annealing step (gradually decreasing to 50 °C over the span of 20 cycles) and 45s 72°C extension step.

5. Results and discussion

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

CITED2 and HIF-1α both bind with their disordered C-terminal transactivation domain (C-TAD) to the TAZ1 domain of p300/CBP, and both form folded although dynamic structures (fuzzy complexes) on the TAZ1 surface[21,94,198,199] (Figure 5). Despite the similar high-affinity nature of their interaction with TAZ1, CITED2 is able to displace HIF-1α from its complex even at equimolar ratios, whereas the reverse process is virtually prohibited.

Among other contributing factors, the irreversibility of the competition is fundamentally based on the formation of a transient ternary intermediate, where the C-terminal helices of HIF-1α (α_B and α_C) are still in contact with TAZ1 when the α_A helix of CITED2 binds. This induces an allosteric structural change that locks the conformation of TAZ1 in a CITED2-bound state, rendering the interaction unidirectional.[27,29,130–132,134,137] The cooperative action of the CITED2 binding motifs drives the displacement of HIF-1α: the binding of the α_A helix of CITED2 destabilizes the binding of HIF-1α α_C, and the conserved LPEL motif of CITED2 replaces the LPQL sequence of HIF-1α, which is a key step in the competition[27,131,132]. The binding of α_A and LPEL is not sufficient for effective displacement, as the absence of the C-terminal diminishes the unidirectional nature of the competition[29]. However, the lack of the C-terminal hydrophobic region drastically decreases the affinity of CITED2 to TAZ1, which could mask the contribution to the allosteric structural change. Due to the fact that the transient ternary intermediate is key to the negative cooperativity between the ligands, it is essential to probe the interplay between the CITED2 binding motifs and to show how each contributes to the allosteric process.

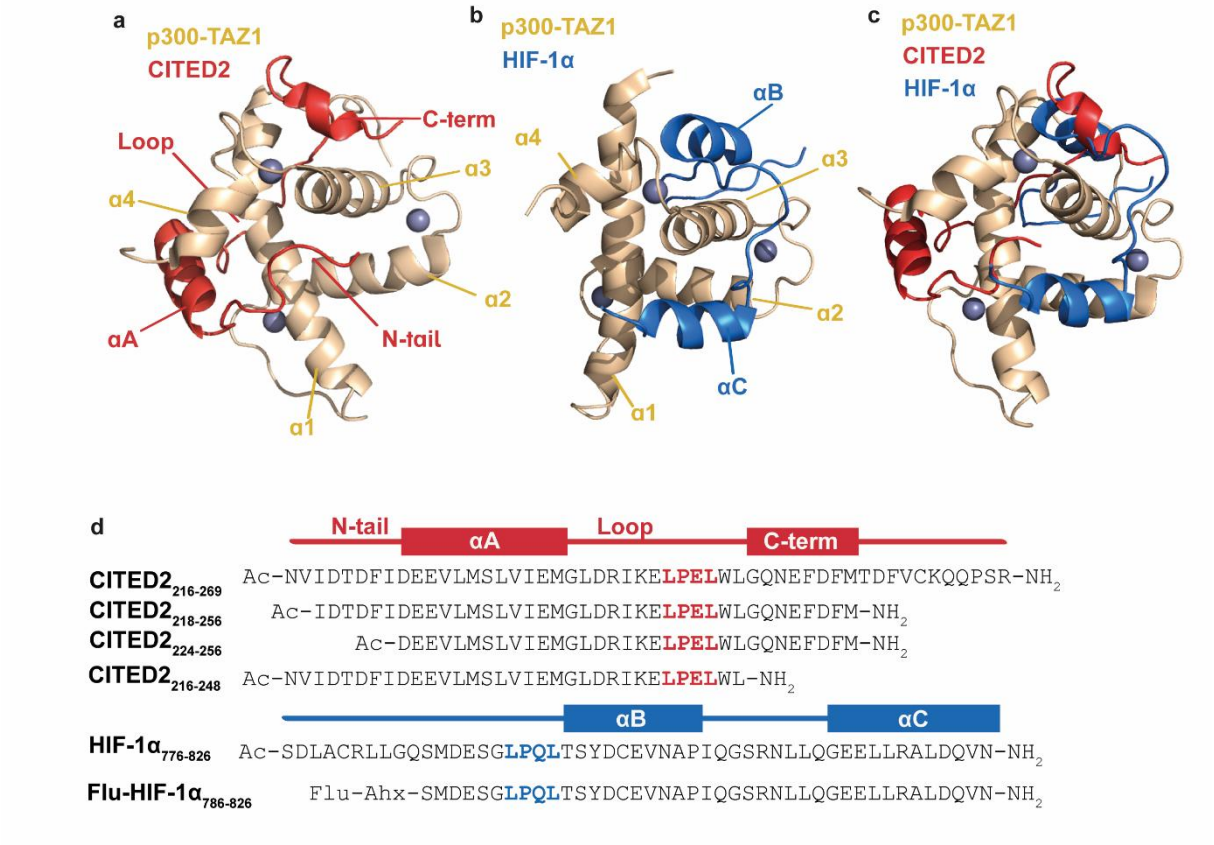


Figure 5. Structure of p300 TAZ1 bound to HIF-1α and CITED2. (a) Structure of CITED2 C-TAD (red) bound to TAZ1 domain (PDB:1P4Q). The helices of the TAZ1 domain are annotated α1–α4 from the N to C terminus, and the spheres represent Zn atoms. The binding motifs of CITED2 are annotated as N-tail, αA, 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.

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

Although previous studies have demonstrated which binding motifs of CITED2 are mechanistically important, the minimal length of human CITED2 that exhibits efficient native-like behaviour has not been determined. The full-length CITED2 C-TAD (CITED2$_{216-269}$) is the most commonly investigated in the literature and has disordered N- and C-terminal residues in its bound state. We compared this sequence variant with CITED2$_{218-256}$, which only contains residues that are reported to be in contact with TAZ1. This consideration is based on the highly homologous CBP and p300 domains, supported by existing NMR structures (Figure 5). To determine the thermodynamic parameters of direct binding and the cooperativity parameters of the competition, isothermal titration calorimetry (ITC) experiments were performed with the variants of CITED2 titrated to TAZ1 or to preformed TAZ1/HIF-1α complex. Thermograms

revealed that CITED2$_{218-256}$ had a slightly improved affinity for TAZ1 compared to the longer version (Table 1.) In general, a similar thermodynamic fingerprint was observed, characterised by favourable enthalpy and slightly unfavourable entropy. Competition thermograms were globally fitted to a model that allows the formation of the ternary system, from which the cooperativity parameters Δ*g* and Δ*h* were extracted. These parameters represent the additional Gibbs free energy and enthalpy for the ternary complex formation when CITED2 binds to the TAZ1/HIF-1α complex compared to the CITED2/TAZ1 binary interaction due to cooperative interactions[191,193,200]. In other words, the cooperativity parameters show the difference in Δ*G* and Δ*H* between the formation of CITED2/HIF-1α/TAZ1 and the CITED2/TAZ1 complexes. Competition titrations revealed a high degree of similarity between the two CITED2 ligands, with a small positive Δ*g* indicating that the formation of the CITED2/HIF-1α/TAZ1 complex is favourable in both cases and allosteric communication between the ligands is maintained, while the positive Δ*h* indicates entropy-driven stabilisation of the ternary complex. This suggests that the enthalpic cost of disrupting the high affinity contacts between TAZ1 and HIF-1α is compensated by the increasing disorder in the ternary complex due to the unfolding of HIF-1α. Taking into account these data, we continued our research with CITED2$_{218-256}$ as the parent peptide to investigate the individual action of its different binding motifs. CITED2$_{218-256}$ has the following four binding regions: CITED2 N-terminal region (N-tail, CITED2$_{218-223}$), an α-helix (CITED2$_{224-235}$) connected through a loop region (loop, CITED2$_{236-247}$) to an aromatic amino acid rich hydrophobic region at the C-terminal (C-term, CITED2$_{248-256}$).

To probe the contribution of the N-terminal residues and the C-terminal hydrophobic region to the ternary system formation, we compared the binding affinity and competition efficiency of the CITED2$_{218-256}$ minimal sequence, to a N-terminally truncated (CITED2$_{224-256;}$ CITED2ΔN) and a C-terminally truncated (CITED2$_{216-248;}$ CITED2ΔC) variant (Figure 6). To characterise the negative cooperativity between these ligands and HIF-1α, two sets of isothermal titration calorimetry (ITC) experiments were carried in a manner similar to those above. The direct binding parameters showed a fivefold increase in K_D for CITED2ΔN compared to CITED2$_{218-256}$ (K_D=50.5 nM and K_D=11 nM, respectively), although the reduced enthalpy contribution was somewhat compensated by a more favourable entropy (Figure 6b). The K_D of CITED2ΔC increased almost 30-fold, indicating the loss of many important contacts with TAZ1, and it showed an even higher degree of entropy compensation (Table 1) (Figure 6b). The reason for this enthalpy-entropy compensation phenomenon cannot be explained by

these data alone, but a possible cause may be that the CITED2/TAZ1 complex becomes more dynamic or fuzzy in the absence of the N- or C-terminus[201,202]. The results of the competition ITC titrations showed that the additional Gibbs energy for ternary system formation for the native ligand was Δg=1.1 kcal mol^{-1} (Δh=26.5 kcal mol^{-1}), indicating an entropy-driven negative cooperativity with HIF-1α through a transient ternary intermediate. In contrast, for CITED2ΔN we observed Δg >4 kcal mol^{-1} (Δh=0 kcal mol^{-1}), which means that the formation of the ternary intermediate is unfavourable, thus CITED2ΔN cannot compete through the allosteric pathway. In the case of CITED2ΔC the cooperativity parameters (Δg=1.27 kcal mol^{-1}; Δh=9.8 kcal mol^{-1}) indicated that although the unidirectional nature of the competition was not retained due to the decreased affinity of the ligand, the formation of the ternary intermediate remained favourable, indicating a retained allosteric communication between the ligands (Figure 6c).

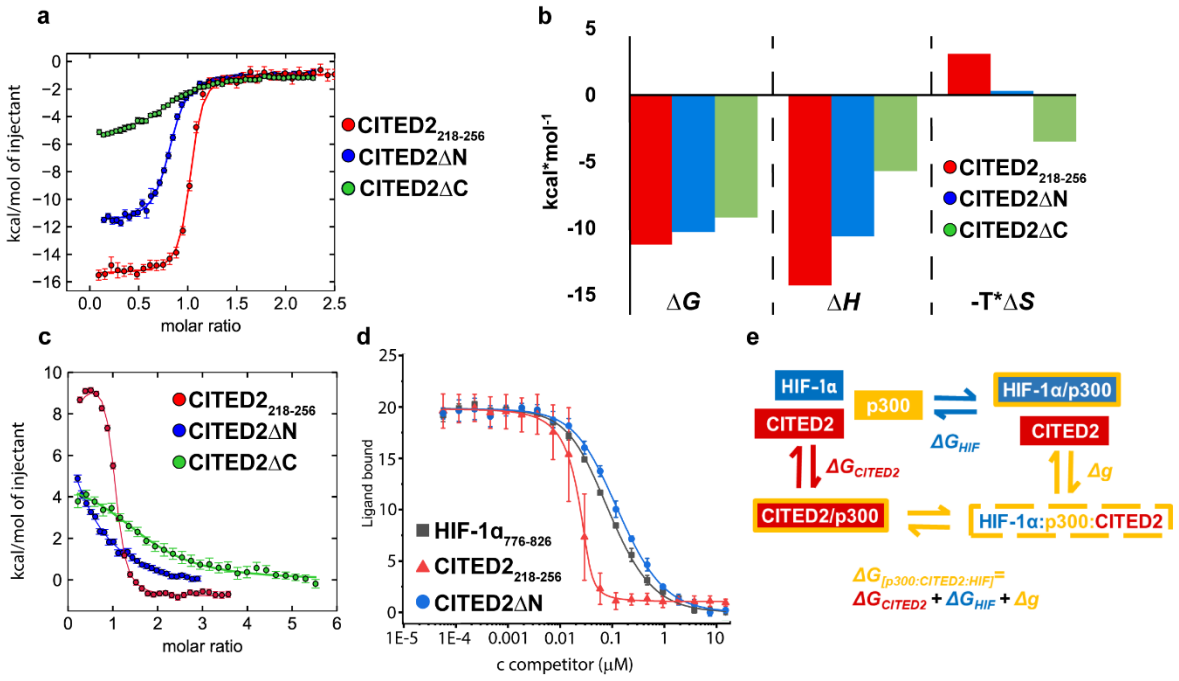


Figure 6. Binding and competition of different CITED2 sequences. (a) ITC thermograms and (b) thermodynamic signature for CITED2$_{218–256}$, CITED2ΔN and CITED2ΔC binding to TAZ1. (c) ITC thermograms for the titration of CITED2$_{218–256}$, CITED2ΔN and CITED2ΔC to the preformed TAZ1/HIF-1α$_{776–826}$ complex. ITC data were fitted globally to a model including ternary intermediate formation (Methods 5.2). Fitted parameters are listed in Table 1 (d) Fluorescence anisotropy (FA) competition of unlabelled CITED2$_{218–256}$, CITED2ΔN and HIF-1α$_{776–826}$ against fluorescein-labelled HIF-1α$_{786–826}$ (Flu-HIF-1α) in complex with TAZ1 using 50 nM protein and 25 nM tracer concentration, curves were fitted to a competition model using a fixed KD value for Flu-HIF-1α (KD = 64 nM) determined beforehand. (e) Schematic representation of the competition model which was used for the fitting of the competition thermograms, allowing ternary system formation.

To validate the results of the ITC experiments, fluorescence anisotropy (FA) titrations were carried out with fluorescently labelled HIF-1α (Flu-HIF-1α$_{786-826}$), from which the

apparent competition K_D ($K_{D,app}$) values were extracted. These values can be interpreted as the K_D that the examined peptide should have for binding to TAZ1, if a simple HIF-1α displacement is considered without the underlying allosteric effects. Therefore, a lower $K_{D,app}$ than the K_D determined by ITC indicates that the ligand is able to act through the transient ternary system formation and is a more effective competitor than expected from its direct binding affinity. These anisotropy data shows that the wild type CITED2$_{218-256}$ ($K_{D,app}$=2.0 ± 0.9 nM) is a much more efficient competitor than CITED2ΔN ($K_{D,app}$=86 ± 5 nM), with a similar $K_{D,app}$ value to HIF-1α displacing itself ($K_{D,app}$= 57 ± 4 nM) (Figure 6d).

Table 1. Thermodynamic parameters for direct binding and competition with TAZ1:HIF-1α. ITC data were fitted globally to a model that included ternary intermediate formation (Methods 5.2). Δg (and Δh) represents the additional Gibbs energy (and enthalpy) due to cooperative interactions when CITED2 is bound to the TAZ1/HIF-1α complex relative to CITED2 binding to free TAZ1. 68% confidence intervals of the fitting are included in brackets.

	Direct binding to TAZ1			**Cooperativity parameters, Binding to TAZ1:HIF-1α**		
	K_D (nM)	Δ*H* (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})	Δ*g* (kcal mol^{-1})	Δ*h* (kcal mol^{-1})	Δ*s* (cal mol^{-1} K^{-1})
CITED2$_{216-269}$	26.3 (20.8-33.1)	-13.2 (-12.9 to -13.4)	-10.0	0.31 (-0.34 to 0.9)	20.5 (19.9-21.2)	65.84
CITED2$_{218-256}$	11 (8-13)	-14.3 (-14.1 to -14.6)	-10.0	1.1 (0.9-1.3)	26.3 (25.9-27.0)	82.7
CITED2ΔN	51 (45-57)	-10.6 (-10.8 to -10.5)	-1.0	>4	0	n.a.
CITED2ΔC	303 (230-297)	-5.7 (-6.1 to -5.4)	11.3	1.27 (0.9-1.69)	9.8 (8.5-11.3)	27.7

To elucidate the structural change in TAZ1 due to the lack of the N-terminal residues of CITED2, solution NMR spectra were recorded with ^{15}N- and ^{13}C- labelled TAZ1 in complex with CITED2$_{218-256}$ and CITED2ΔN. The calculated chemical shift perturbations revealed mainly local conformational effects due to the truncation of the N-terminus, specifically in the α1 (residues 347-357), α3 (residues 401-407) and α4 (residues 418-420) helices, proving that these N-terminal residues are in contact with TAZ1. Since the core of TAZ1 is tightly packed with hydrophobic amino acids, CH_3 chemical shifts can shed light on the overall structural change. Significant differences between the two structures were observed for L342, L346 and L417 at the α1/α4 interface in TAZ1, indicating an overall change in orientation of these helices (Figure 7). Previous studies have demonstrated, that the conformational change of the α4 helix

is a key element in the displacement process[27,29,132]; however, here we showed that the effector of these structural changes is the N-terminal residues of CITED2.

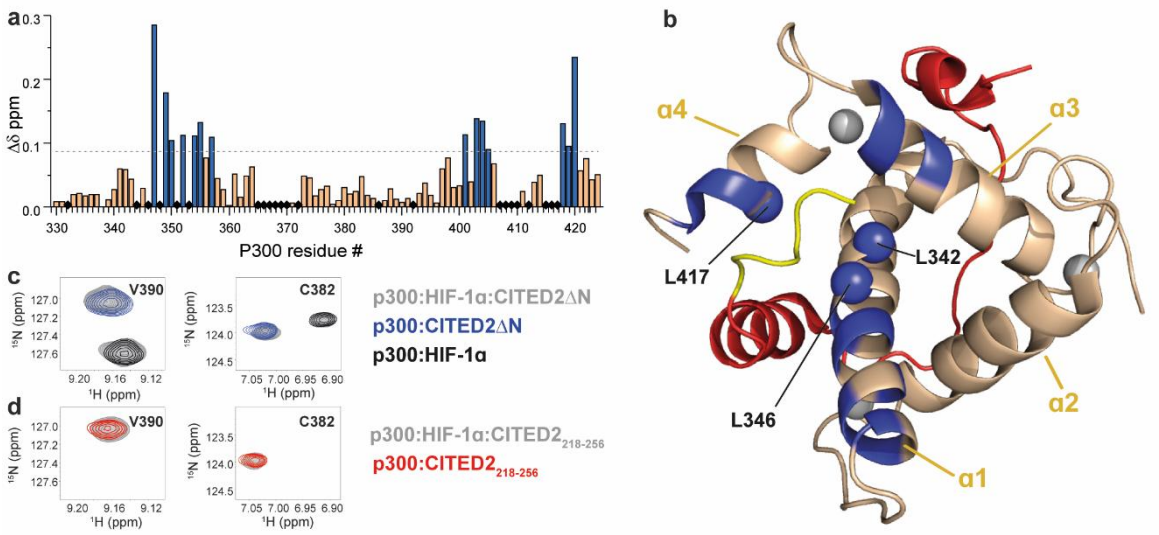


Figure 7. Structural comparison of CITED2$_{218–256}$ and CITED2ΔN bound to TAZ1 using NMR. (a) Weighted average ^{1}H, ^{15}N chemical shift difference between the TAZ1/CITED2ΔN and TAZ1/CITED2$_{218–256}$ complex. $\Delta\delta = [(\Delta\delta\ H)^2 + (\Delta\delta\ N/5)^2]^{1/2}$, residues above 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).

To investigate the reversibility of the interaction, samples of CITED2$_{218-256}$ and CITED2ΔN were prepared in equimolar ratios with HIF-1α and TAZ1. For the CITED2ΔN/HIF-1α/TAZ1 mixture chemical shifts were assigned both to the CITED2ΔN/TAZ1 and HIF-1α/TAZ1 complexes indicating the existence of both structures in solution. However, for the CITED2$_{218-256}$/HIF-1α/TAZ1 sample, the peaks showed correlation with only the CITED2$_{218-256}$/TAZ1 complex. These data strongly suggest that the presence of the CITED2 N-terminus is not only essential for the allosteric structural change in TAZ1 but also maintains the conformational lock needed to prohibit the reverse process (Figure 7c & d).

5.1.2. Backbone modification strategy to probe the contribution of CITED2 binding motifs to cooperativity and their intramolecular interplay

As we have seen above (CITED2ΔC), to obtain information on the contribution of CITED2 binding motifs to negative cooperativity with HIF-1α, the use of sequence truncation and alanine scanning techniques can be misleading, due to the potential drastic decrease in

affinity for TAZ1, which in turn diminishes competition efficiency. These effects can overemphasise the importance of a given motif in the induction of allosteric structural changes and unidirectionality. Based on this argument, a new strategy is needed, where the binding interface of a given motif is only minimally perturbed by targeted modifications and the modified CITED2 sequence resembles the properties of the parent peptide (e.g., thermodynamics of TAZ1 binding, disordered nature, side-chain functionality). The structure of TAZ1 is expected to adapt to these modifications due to its plasticity, maintaining the high-affinity nature of the CITED2/TAZ1 interaction. In this way, a change in competition efficiency with HIF-1α, which is indicated by the increase of the cooperativity parameter Δg, is most likely due to the perturbation of a functionally relevant region, and not to the decreased affinity of the ligand. The influence of the modifications on the structure of TAZ1 can be followed by NMR in parallel with the competition studies. These measurements can provide information on whether an allosterically relevant region of TAZ1 is affected or not by the given modification and give insight into the fine interplay between CITED2 binding motifs, and into the extent of cooperativity between the competing ligands.

Our hypothesis was that using a modular, motif-by-motif replacements of α→β^3 amino acids in CITED2 would fulfil our requirements for this strategy. These residues are often incorporated and tolerated in helical or loop regions[154,203–206]. To keep the level of perturbation to a minimum, the amino acids chosen for replacement were selected based on existing NMR-structures, where residues had the least direct contact with the surface of TAZ1. The selection was supported by computational alanine scanning using BAlaS[207,208], where residues that had a contribution of less than -4 kJ mol^{-1} to the binding free energy were selected for replacement.

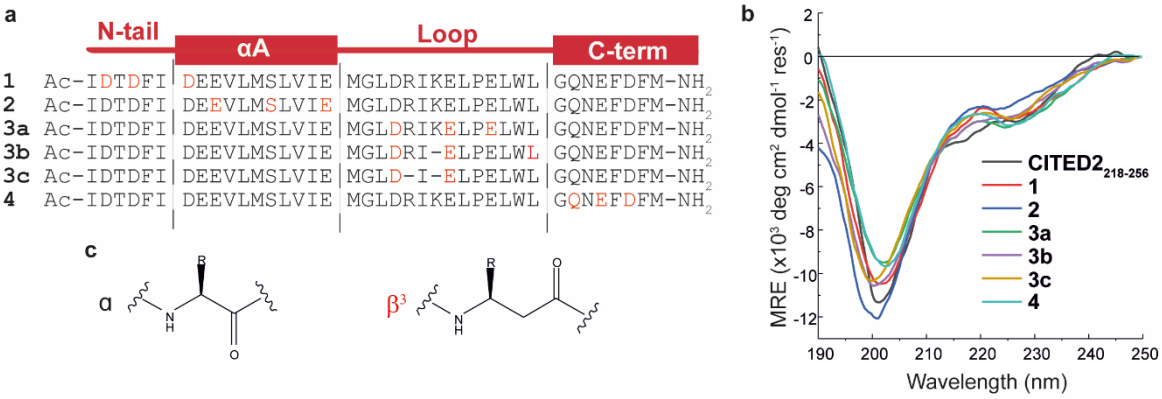


Figure 8. Sequences and overall structure of the β-amino acid modified CITED2 variants. (a) Sequences of modified CITED2 variants with binding motifs indicated above. α - β^3 substitutions are highlighted in red, and deletions are represented by a line. (b) CD curves for CITED2 and variants in 20 mM 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.

Based on the considerations above six CITED2 variants were synthetised: a N-terminally modified (**1**) an αA helix (**2**), 3 loop modified (**3a-c**), and a C-terminally modified (**4**). To ensure that the CITED2 variants retained their disordered nature, circular dichroism (CD) measurements were carried out. The CD spectra of **1-4** showed that all sequences had a random coil conformation with a slight helical content, as indicated by the small minimum at 220nm (Figure 8). Overall, it can be stated that the incorporation of β^3-amino acids did not induce folding of the free peptides.

5.1.3. Impact of backbone modifications on the thermodynamic parameters

The ITC method described above was used to prove that the variants retained native-like thermodynamic parameters of direct binding, and to evaluate how the site-specifically modified backbone affected the cooperativity parameters when competing with HIF-1α. Regarding the direct binding parameters **1, 2** and **4** showed similar K_D to that of the native interaction (ranging from 10 to 19 nM) with a slightly less enthalpic contribution and slightly less entropic cost (Figure 9) (Table 2). However, modification of the loop region showed a much more pronounced effect. Replacement in the conserved LPEL motif resulted in a 50-fold increase in the K_D, with a strikingly different thermodynamic profile, which could be contributed to the loss of important non-covalent intermolecular contacts. When the conserved region was left intact and the lengthening caused by the extra methylene groups of the beta amino acids was compensated for using amino acid deletions (**3c**), the affinity of the variants greatly improved and showed native-like thermodynamic behaviour (**3b-c** have 98 and 29 nM K_D, respectively). The observed enthalpy-entropy compensation phenomena in all variants (except **3a**) can be attributed to the loss or weakening of non-covalent contacts, which may be compensated for by more favourable entropy due to the extra, flexible methylene groups; however, other contributing factors cannot be ruled out based on these data alone (e.g., differences in solvation) (Figure 9) (Table 2).

The cooperative parameters of **2** and **4** remained similar to those of $CITED2_{218-256}$ (Δg=1.4 kcal mol^{-1} Δh=24.3 kcal mol^{-1} and Δg=1.1 kcal mol^{-1} Δh=23.2 kcal mol^{-1}, respectively) indicating that these variants retained their ability to form the ternary complex, and to act through allosteric changes. In the case of the N-terminal modification, the competition efficiency decreased, but not to as great an extent as in the truncation experiment. These data further emphasise the essential role of this region in maintaining the negative cooperative nature

of the competition (Figure 9) (Table 2). The most pronounced effect was observed when the residue replacements were carried out in the loop region, in parallel to the tendency of direct binding parameters. The additional Gibbs energy needed for the formation of the ternary system was, in all cases (**3a-c**), greater than 4 kcal mol^{-1} (and Δh = 0 kcal mol^{-1}) meaning that these ligands lost the ability to form the ternary complex with HIF-1α, which in turn diminished their ability to initiate the correct allosteric changes in the structure of TAZ1, thus indicating a central role for this region in governing the competition (Figure 9) (Table 2).

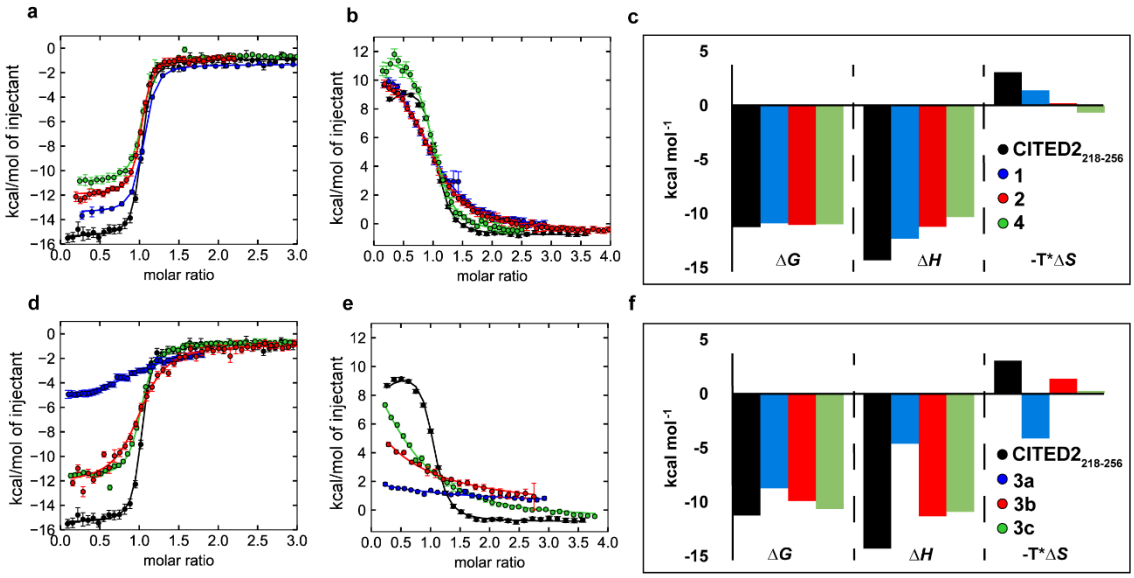


Figure 9. Binding and competition of different backbone modified CITED2 variants. (a) ITC thermogram for CITED2$_{218–256}$, **1**, **2** and **4** binding to TAZ1 (b) ITC thermogram for the titration of CITED2$_{218–256}$ **1** **2** and **4** to the preformed TAZ1/HIF-1α$_{776–826}$ complex. (c) Thermodynamic 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. Fitted parameters are listed in Table 1. ITC data were fitted globally to a model including ternary intermediate formation (Methods 5.2)

Similarly to the truncation experiments, fluorescence anisotropy measurements were carried out with the β^3-amino acid modified variants to validate the ITC results (Figure 10). These measurements reproduced the same tendency as the ITC titrations with lower, native-like $K_{D,app}$ values for **2** and **4** indicating cooperativity between the ligands (Table 2.) , and higher ones for **1** and **3a-c** which resemble their corresponding K_D for the binary interaction with TAZ1, indicating the loss of ability to induce allosteric changes in TAZ1 that locks the conformation into the CITED2 bound state.

Table 2. Thermodynamic parameters for direct binding and competition with TAZ1/HIF-1α. ITC data were fitted globally to a model that included ternary intermediate formation (Methods 5.2). Δg (and Δh) represents the additional Gibbs energy (and enthalpy) due to cooperative interactions when CITED2 is bound to the TAZ1/HIF-1α complex relative to CITED2 binding to free TAZ1. 68% confidence intervals of the fitting are included in brackets.

	Direct binding to TAZ1			Cooperativity parameters, Binding to TAZ1:HIF-1α			KD, app (FA)
	K_D (nM)	Δ*H* (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})	Δ*g* (kcal mol^{-1})	Δ*h* (kcal mol^{-1})	Δ*s* (cal mol^{-1} K^{-1})	(nM)
CITED2$_{218-256}$	11 (8-13)	-14.3 (-14.1 to -14.6)	-10.0	1.1 (0.9-1.3)	26.3 (25.9-27.0)	82.7	2.0 ± 0.9
1	19 (16-23)	-12.3 (-12.5 to -12.1)	-4.6	2.6 (2.3-2.9)	26.2 (24.4-29.0)	76.7	9.6 ± 2.4
2	16 (13-20)	-11.1 (-10.9 to -11.3)	-0.4	1.4 (1.2-1.6)	24.3 (23.4-26.1)	74.1	1.7 ± 0.7
3a	675 (595-814)	-4.6 (-4.8 to -4.4)	13.3	>4	0	n.a.	> 600
3b	98 (83-108)	-11.3 (-11.6 to -11.1)	-4.6	>4	0	n.a.	112 ± 13
3c	29 (26-33)	-10.9 (-11.0 to -10.8)	-8.9	>4	0	n.a.	76 ± 6
4	17 (12-23)	-10.3 (-10.6 to -10.1)	2.1	1.1 (0.9-1.4)	23.2 (22.1-24.1)	71.6	4.9 ± 0.8

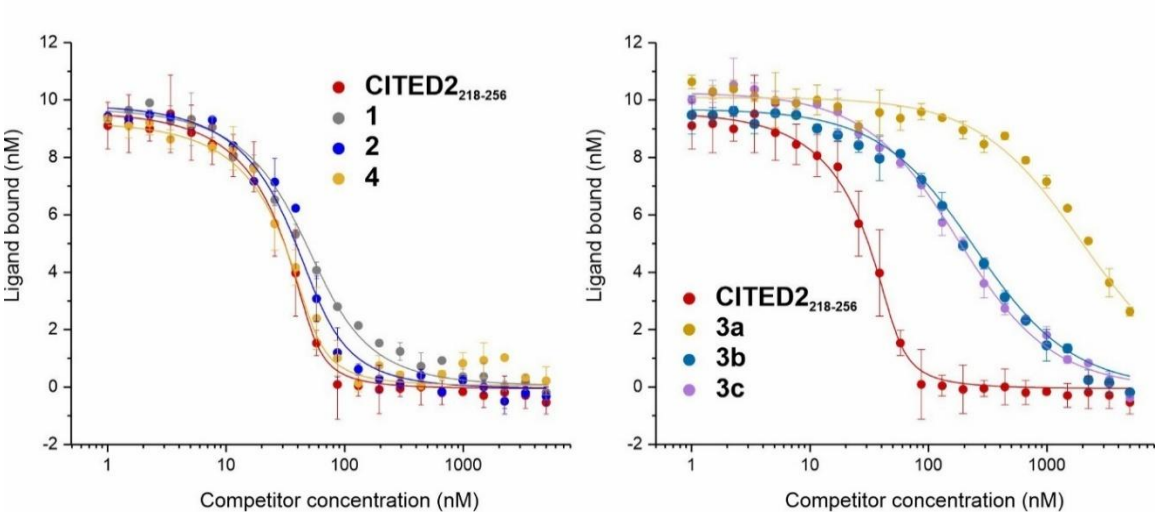


Figure 10. Fluorescence anisotropy competition titrations of 1-4: Measurements were carried out in 40 mM Na-phosphate, 100 mM NaCl, 1 mM DTT, 0.01% Triton-X, pH 7.4 buffer. CITED2 and variants were serially diluted on the plate using 2/3 dilution series, with 5 µM as the highest concentration to which 50 nM p300$_{330-420}$ and 25 nM Flu-HIF-1α were added. Plates were read after a 10-minute incubation at 35°C.

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

To shed light on the connection between the changes in competition efficiency and the conformational differences in TAZ1 structure, ^{1}H-^{15}N-NMR spectra were recorded of the CITED2 variants in complex with ^{13}C- and ^{15}N-labelled TAZ,1 and the weighted chemical shift differences were calculated relative to the CITED2/TAZ1 complex. Although the chemical shift

differences were generally low (Δδ<0.2 ppm), the observed changes were highly dependent on the modification site. In the case of **1** the conformational effects on TAZ1 structure were most pronounced at the C-terminal of α1, α3 and α4 helices, which resemble the results of the truncation experiments where the N-terminus was deleted, giving further support that the N-terminus of CITED2 has a direct role in the formation of the conformational lock. Similar localised effects were seen in the case **2** and **4**, where the significant structural differences were restricted to the interaction site of the modified binding motifs. β^3-amino acid replacements in the loop region had a widespread impact on the TAZ1 structure; differences were observed distant to the loop binding site. These results indicate that the loop region is critical for connecting the recognition elements.

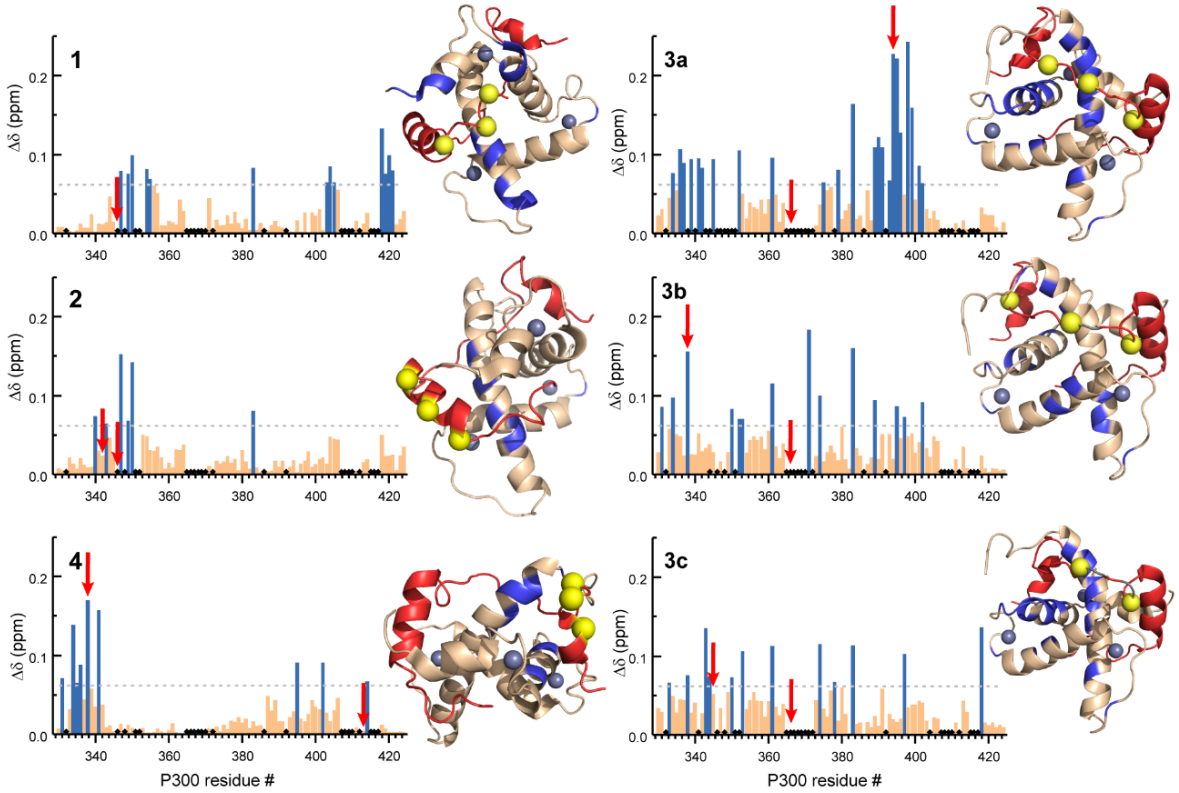


Figure 11. NMR chemical shift changes of TAZ1 upon binding to the β^3-amino acid modified CITED2 variants. Weighted average $^1H–^{15}N$ chemical shift differences between the TAZ1:CITED2 variants and TAZ1:CITED2$_{218–256}$ complexes. $\Delta\delta = [(\Delta\delta\ H)^2 + (\Delta\delta\ N/5)^2]^{1/2}$. 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.

To connect the results of ITC measurements and NMR data we, can conclude that modifications of the αA and C-terminal helix of CITED2 were tolerated well, because the

conformational perturbation and impact on cooperativity parameters were minimal. Although the N-terminal variants retained their ability to act through the allosteric pathway, the free energy gate of ternary system formation increased significantly, which, taken together with the structural impact, strongly suggests that the unidirectionality of the competition is contributed by this region. The detrimental effect of loop modifications on the competition efficiency taken together with the structural data indicates a role in maintaining the intramolecular cooperativity between the binding motifs during the folding and binding process (Figure 11).

According to these results β^3-amino acid replacements could be an efficient technique for the investigation of other interaction involving disordered TFs or disordered ligands in general, at a mechanistic, molecular level. An upside of this approach is that the replacements tolerated by the investigated IDP can be a starting point for top-down drug development. Building on our findings, our future work will combine the tolerated modifications within the CITED2 sequence with additional optimisation strategies including the incorporation of other non-natural building blocks (e.g., D-amino acids, N-methyl and fluorinated residues) to develop a functional, protease-resistant hypoxic response inhibitor, that preserves the negative cooperative competition mechanism with HIF-1α.

5.2. Proof of concept synthesis of a DNA-templated encoded foldamers

Protein-protein interfaces (PPI) generally cannot be targeted with small molecule drugs due to the extensive surface of these interaction, where peptidic or foldameric fragments have proven to be useful[3,139,140,146]. Interactions mediated by IDPs often utilise multivalency, which if addressed at the screening stage, would provide a more efficient strategy for ligand development. However, screening of peptidomimetic compounds by conventional high-throughput screening (HTS) or fragment-based methods faces serious limitations due to problems involving the synthesis or hit deconvolution. An alternative approach is the use of display technologies, but implementation of residues with non-natural backbones is inefficient due to the limitations of in vitro translational machinery[209]. DNA-encoded library (DEL) techniques, however, potentially offer a good alternative due to their accessibility, versatility, efficiency and cost-effectiveness[34].

In a general DEL system, the library members are covalently tagged with a DNA barcode, where all the members can be pooled together, and can be selected against an

immobilized target in a single tube. After washing away the non-binders, the hits can be identified after PCR amplification with the use of next generation sequencing (NGS)[33,170,210–213]. Although the literature describes DELs where cyclic or linear peptides are tagged with the DNA barcodes, these are either monomeric or not directly amplifiable by PCR and are readily degradable by proteases[185,214,215], while the utilization of foldamers in relation with this technique is not mentioned.

As a proof of concept, we set out to devise a complete workflow to create a trimeric, and amplifiable DEL member that presents beta-peptidic H14 foldamers as ligands. These structures consist of (S,S)-2-aminocyclohexanecarboxylic acid and β^3-amino acid residues (Figure 14) presenting two proteinogenic side chains on the surface of a stable H14 helix, making them promising candidates to probe PPI interfaces with extensive surface area[30]. The devised method 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 (Figure 12b.).

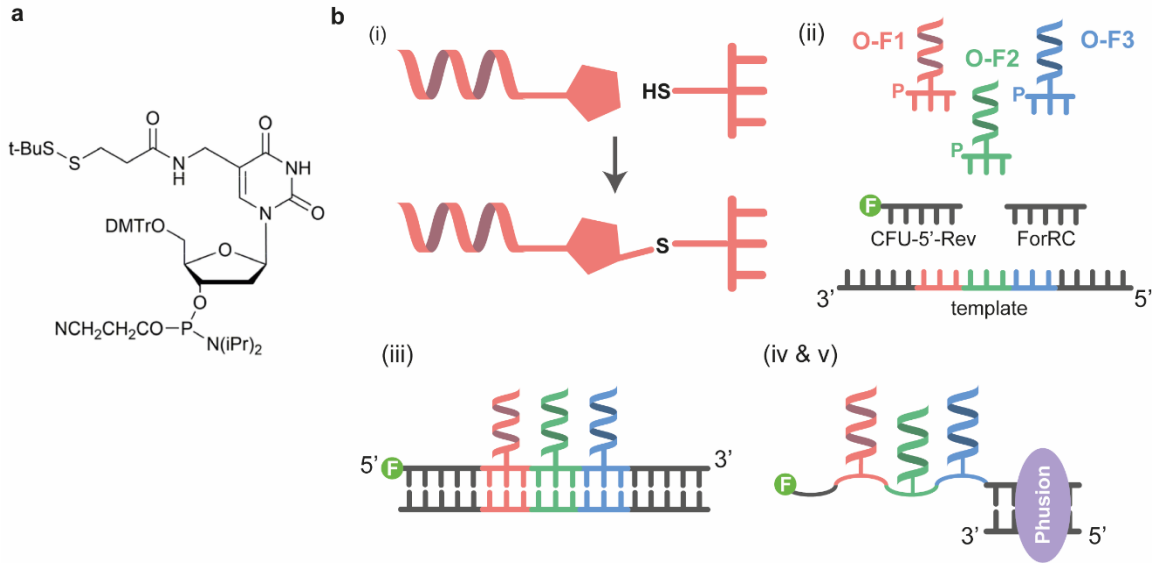


Figure 12. (a) Structure of the tBuS-protected thiol-modified thymidine nucleoside phosohoramidite (b) Schematic representation of the workflow of the library member synthesis: (i) conjugation of oligonucleotides and H14-foldamers via thiosuccinimide linkage (ii) hybridization of the conjugates and primers on the encoding template (iii) template guided ligation of the conjugates and primers (iv & v) separation of the strands with denaturing PAGE, isolation of the ligated product, and amplification via Phusion polymerase. (Thiol modified nucleotide synthesis was developed by Zoltán Kupihár, University of Szeged, and oligonucleotide synthesis was carried out by Györgyi Ferenc, BRC Szeged)

5.2.1. Thiol-maleimide conjugation of encoding oligonucleotides and H14 foldamers

To produce ligatable oligonucleotide conjugates, the thiol modification must be placed in the middle of the oligonucleotide sequence, bearing a thiol protection group that is orthogonal with the synthesis process. This was achieved by using a newly developed synthesis of a S-S-tBu protected thiol-containing thymidine nucleoside phosphoramidite[216] (Figure 12a.). The phosphoramidite could be easily incorporated into oligonucleotide sequences with general CPG method leaving the 5'-phosphate group available for ligation. Prior to the conjugation of foldamers, the reaction was optimized with an alpha-peptide sequence, the transferrin receptor binding **HAIYPRH** which is implemented in tumour-targeting delivery systems[217]. This sequence was extended with a glycine and a maleimido-propionic acid moiety on the N-terminus. For the conjugation reaction, the thiol-modified **O2-S-S-tBu** needed to be deprotected first to remove the tert-butyl protecting group, which was carried out by TCEP treatment. Following the deprotection and purification of **O2-SH**, the conjugation mixture was prepared with the addition of TCEP to prevent disulfide bond formation and the addition of the maleimide modified peptide. The reaction proceeded to give good yield overnight at room temperature; however, the production of a maleimide-TCEP conjugate ylene byproduct was observed[218]. To circumvent this problem, the deprotected **O2-SH** was instantly frozen and freeze-dried, and the conjugation reaction was performed under argon atmosphere to prevent disulfide formation (Figure 13a-b).

Table 3. Oligonucleotide sequences used for conjugation and ligation[1]

Name	Sequence (5'-3')
O1	p-TGTCTGAACC
O2	p-TCACTCTTGC
O3	p-ACTTTCGCAC
O1-SH	p-TGTCT(SH)GAACC
O2-SH	p-TCACT(SH)CTTGC
O3-SH	p-ACTTT(SH)CGCAC
5'-CFU-Rev primer	CFU-AGAATGCTGGGCAAT
ForRC	p-CTTTTATCACGGCCC
For primer	p-GGGCCGTGATAAAAG
Template	GGGCCGTGATAAAAGGGTTCAGACAGCAAGAGTGAGTG CGAAAGTATTGCCCAGCATTCT

[1]T(SH): thiol modified thymidine monomer; p: 5′ phosphate; CFU: 5′-carboxyfluorescein.

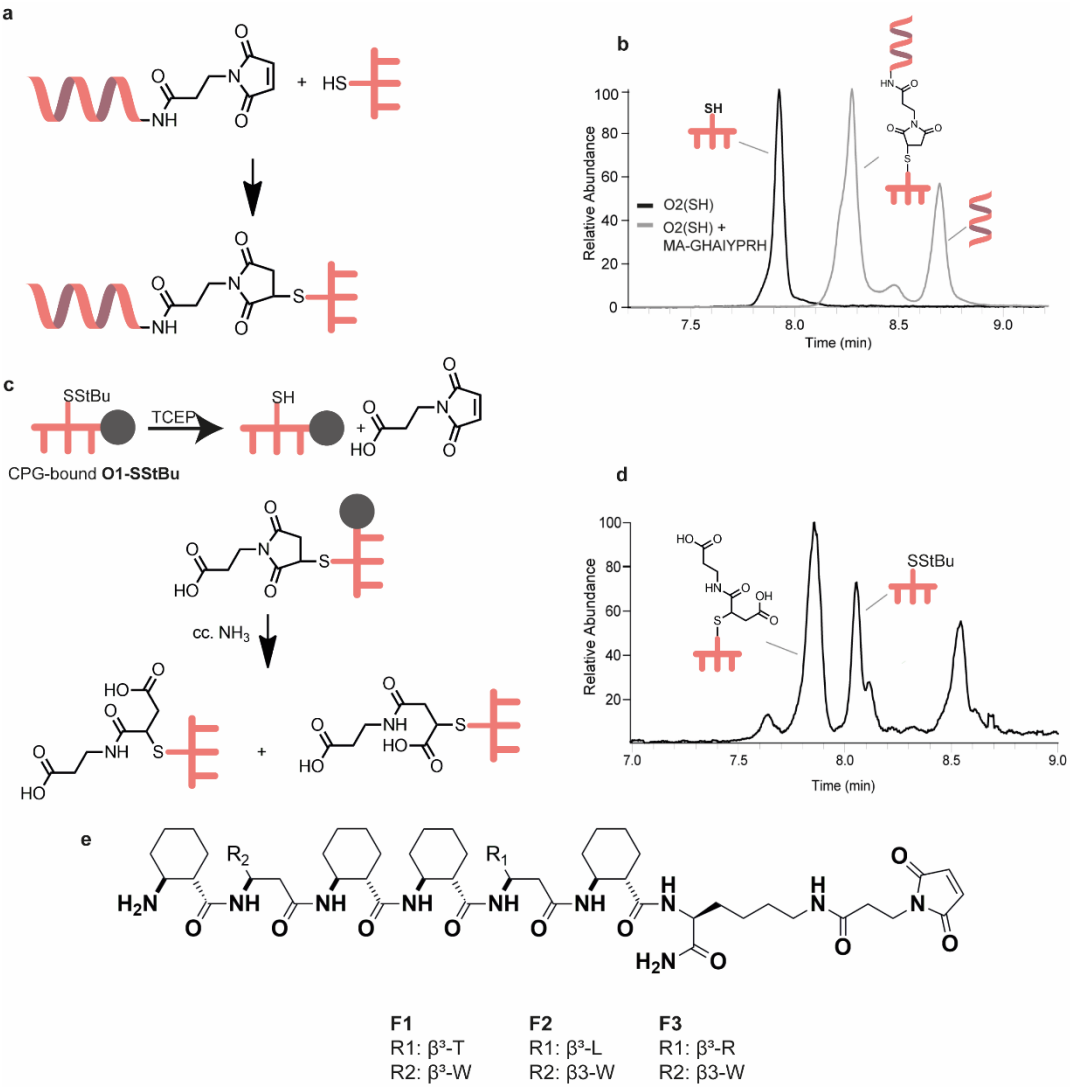


Figure 13. Thiol-maleimide conjugation in solution and on solid support. (a) Reaction scheme for solution phase conjugation using 100 µM oligonucleotide, 200 µM TCEP, and 500 µM peptide, 10% DMSO, overnight room temperature. (c) Reaction scheme for solid phase conjugation. TCEP treated CPG-bound oligonucleotide was reacted with 3-maleimidopropionic acid (10 mM in 50mM Tris, pH 7.2, 20% DMF) overnight, followed by cleavage using concentrated NH3, 24 h. (e) Primary structure of H14-foldamers used for conjugation

An advantage of the tert-butyl protected thiol modified residue, is that the tert-butyl protecting group can be selectively removed from the CPG-bound oligonucleotide, and the chosen ligands can be conjugated on solid phase, prior to cleavage. This would allow a reduction in the number of purification steps. Since the alpha-peptides cannot withstand the concentrated ammonia treatment during oligonucleotide cleavage due to racemization of the alpha carbon atoms, the feasibility of this approach was tested with maleimido-propionic acid. The results showed that the conjugation was successful with a lower yield than in solution phase, and the thiosuccinimde moiety was hydrolysed to form two constitutional isomers with a more stable thioether bond[219]. With a carefully chosen linker length this should not influence

the rotation of the attached ligands. Although the foldamers can withstand alkaline conditions due to the achiral alpha-carbon of β^3-amino acids and the rigid cyclized alpha-carbon of the aminocyclohexane-carboxylic acid residues, the preparation of the foldamer-oligonucleotide conjugates was carried out utilizing the solution phase process, because of the better yield (Figure 13).

5.2.2. Ligation and amplification of the foldamer-oligonucleotide conjugates

With our foldamer-oligonucleotide conjugates in hand, we set out to prepare a single-stranded DNA presenting three foldameric ligand by templated ligation. For this purpose, T4-DNA ligase was used, which is able to ligate short oligonucleotide sequences if a complementary DNA strand is present. To govern the assembly of our conjugates a 60-mer template DNA was synthetised, containing two overhanging 15'mer sequence for the hybridisation of the CFU-labelled reverse primer (CFU-Rev) and the complementary sequence of the forward primer (ForRC). The 5' CFU labelling of the reverse primer was served to ease the detection of the ligated products after gel electrophoresis.

The ligation was carried out with a three-fold oligonucleotide excess relative to the template concentration at room temperature, with or without the foldamers attached. Although the initial approaches yielded the desired 60 bp product in both cases, the presence of truncated side products rendered the method for further optimization. To overcome incomplete ligation, presumed to arise from the hindered hybridisation of the oligonucleotides, heat cycles were introduced to the ligation process. This modified method contained a room temperature ligation step followed by a 90 °C heating step to dissociate the partially ligated intermediates; this cycle was repeated three times, with fresh T4 DNA ligase added after each cycle when the mixture cooled to room temperature due to the heat inactivation of the enzyme. A possible improvement in the future could be the replacement of T4 ligase with thermostable ligases to reduce cost and to automatise the ligation step[220]. Denaturing PAGE analysis revealed that the applied heat-cycle protocol led to the desired 60 bp DNA-strand when all the necessary components were present (termed as **LO1-3** and **LOSF1-3**), and produced the adequate truncated product in the control experiment, from which one of the components was omitted. An exception occurred when **O1** or **O1-SH** was excluded from the reaction mixture, forming a non-specific 30 bp product; however, this did not interfere with the ligation process when all the components were

present. The presence of the attached foldamers did not influence the efficiency of the ligation process, the yields or the distribution of the products compared to the unmodified oligonucleotides, with the exception of the product size, which was slightly higher for the foldamer-functionalized lanes, although this was expected due to the increase in mass (Figure 14a-b). With the optimized ligation method, we set out to prepare a ligation in preparative scale and isolate the full-length foldamer-functionalised single-stranded product. Using denaturing PAGE, the desired product was extracted and subjected to PCR-amplification. PCR amplification is an essential step in a DEL selection to produce sufficient number of copies from the encoding DNA-barcodes of the binders, enabling them to be subjected to NGS for identification. The polymerase of choice was the high fidelity Phusion polymerase, which is a class-B polymerase that can amplify C5 modified pyrimidine bases, since the modifications reside in the major groove of the duplex DNA, hence it does not interfere with duplex formation and enzymatic polymerisation with DNA[221,222]. Initial experiments with the manufacturer's standard protocol, carried out with a touch-down PCR method and analysed by denaturing PAGE, showed a high intensity band for the amplified product of the **LO1-3/template** duplex as control, and a faint band for **LO1-3** alone but the foldamer functionalized **LOSF1-3** did not produce copies in a detectable amount. During parameter optimisation, raising the polymerase concentration five-fold to 0.15 U/μL proved to be beneficial, with both **LO1-3** and **LOSF1-3** forming high intensity bands (Figure 14c). The benefit of this is that in future applications, the amplified product of the selected hits can be reused for ligation in multiple iterative steps for the enrichment of the tightest binders in the DEL.

As the results above indicate ligatable and amplifieable DEL members presenting three foldameric ligands can be created. Since the foldamers are known binders of calmodulin protein[166], we tested their interaction as the next step of the in vitro selection cycle, but we did not observe binding. A possible reason behind this could be the electrostatic repulsion between the protein and the encoding DNA-sequence. The solutions for this problem in the future could be the lenghtening of the linker regions of the oligo-foldamer conjugates or redesigning the arrangement of the constuct to reduce the interference of the target protein.

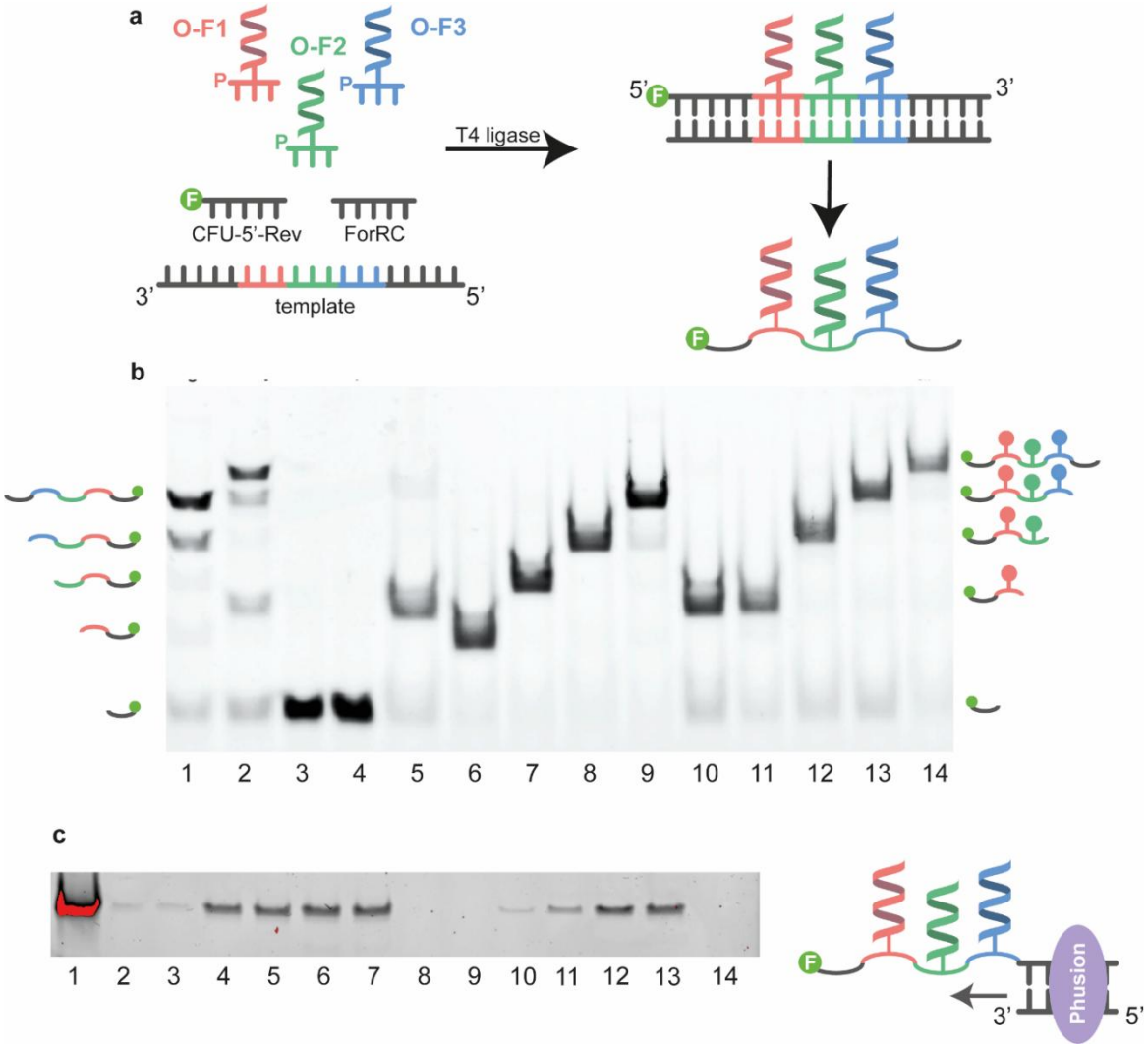


Figure 14. Templated ligation and amplification: (a) Scheme of the ligation process: to the template 5′ phosphorylated oligonucleotides (or foldamer–oligonucleotide conjugates) and extension sequences were added with fluorescent labelling on the 5′ end and ligated using T4 ligase. The ligation was carried 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.

6. Conclusion

Targeting the PPIs of intrinsically disordered transcriptions factors is a highly desired direction of drug development efforts, due to their involvement in various cancerous conditions, among others. Peptidomimetic drugs could successfully overcome the additional challenges surrounding IDP PPI interactions, since their size allow them to be transported to the nucleus, while they could mimic the multivalent binding of IDPs.

Top-down development of peptidomimetics requires detailed structural and mechanistic information about a given IDP PPI interaction, to allow the rational design of a drug-like ligand based on a natural interaction partner. To gain insight into the allosteric competition mechanism between HIF-1α and CITED2 C-TADs for the TAZ1 domain of p300/CBP, we developed a strategy, whereby the modular modification of CITED2 binding motifs provides information about their contribution to the negative cooperativity between the competing ligands. The site-directed conformational perturbation of TAZ1 were carried out using a rational $\alpha \rightarrow \beta^3$-amino acid replacements throughout the CITED2 binding motifs in a way, that retained the native-like binding parameters, but induced site-specific conformational changes in TAZ1. The combination of NMR measurements and competition data revealed which CITED2 binding motif induces the conformational change necessary for the unidirectional competition. This strategy could be useful to investigate other systems involving IDPs if cooperativity or allosteric structural changes are present. An advantage of this approach is that the combination of beneficial modifications serves as a starting point for peptidomimetic inhibitor development, while giving clues about the molecular mechanism of a given competition system.

Since many of the TF interactions are multivalent and mediated by short SLiM and MoRF motifs, mimicking these with foldameric ligands could be a possible direction toward functional peptidomimetic inhibitors. DELs provide an efficient way to create a combinatorial library of compounds, to select them against an immobilised target, and to easily analyse the hits thanks to the encoding DNA barcodes. To combine the benefits of both, we developed a novel DEL system, where an amplifiable single-stranded oligonucleotide presents three beta-peptide H14 ligands, potentially mimicking the multivalent binding of IDPs.

In summary, both strategies developed here can overcome the limitations of addressing the challenges targeting IDP mediated interactions, and result in starting points for further drug development.

7. Summary

1. Using a site-directed allostery perturbation strategy we evaluated the role of CITED2 binding motifs in competition with HIF-1α.

1.1. Sequence truncation strategy

1.1.1. The truncation of CITED2 N-terminus ($CITED2_{218-223}$) decreased the affinity slightly, however the cooperative aspect of competition was lost completely indicating the loss of allosteric communication and unidirectionality.

1.1.2. Truncation of the C-terminus ($CITED2_{249-269}$) decreased the affinity to TAZ1 significantly although the formation of ternary complex remained favourable indicating a retained allosteric communication between the ligands.

1.1.3. 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 giving the unidirectional nature of competition.

1.2. Backbone modification strategy: A backbone modification strategy was introduced to investigate the rest of the binding motifs and their interplay, where solvent exposed amino acids were replaced with β^3-analogues in a motif-by-motif fashion, with the aim to retain a native like high affinity interaction with TAZ1, while inducing minimal perturbations at the binding surface. The structural change could be linked to the decrease in competition efficiency which in turn shed light on the role of a given motif in allostery.

1.2.1. Several β^3-amino acid modified CITED2 variants were synthetized and characterized producing peptides with native like behaviour, regarding the high affinity to TAZ1, the intrinsically disordered conformation and side-chain configuration.

1.2.2. Replacements in CITED2 αA ($CITED2_{224-235}$) and C-terminal ($CITED2_{249-256}$) helices did not affect the ability of these ligands to act through the allosteric pathway and the modifications caused only local conformational perturbations indicating a role in maintaining the high affinity tether for the other motifs to bind.

1.2.3. Modification in the N-terminus of CITED2 resulted in a significantly higher free energy cost for ternary system formation which in turn decreased the effectiveness of

the competition, and the resulted conformational change was similar to the N-terminally truncated product further emphasizing the importance of this region to create the conformational lock in competition.

1.2.4. Modifications in CITED2 loop region ($CITED2_{236-248}$) completely diminished the negative cooperativity between the ligands, furthermore the conformational perturbations were distributed on the entirety of TAZ1 surface, which indicates a role where this region is responsible for the correct placement of the other binding motifs and allows the cooperative binding between them, which is in line with literature data.

2. Proof of concept synthesis of a foldameric DNA-templated library member: The synthesis of an PCR amplifiable DNA-templated library member with 3 foldameric ligand was carried out using a 10-mer oligonucleotide with a thiol modified thymidine phosphoramidite nucleotide incorporated in the 5^{th} position, leaving the 5' end free for ligation.
   2.1. Conjugation of foldamers in aqueous phase: the β-peptidic H14 foldamers functionalized with a maleimide moiety was conjugated to the appropriate thiol modified nucleotide forming via a thiosuccinimide linkage.
   2.2. Conjugation on solid support: To test the applicability of solid phase conjugation, maleimido propionic acid was coupled to the deprotected oligonucleotide, and the conjugate was successfully cleaved resulting in the more stable hydrolysed thioeter linked small molecule.
   2.3. Templated ligation of the foldamer modified oligonucleotides and extension sequences were successfully carried out, with the introduction of a heat-cool cyle.
   2.4. The PCR amplification of the foldamer modified strand was carried out using Phusion polymerase in concentration raised 5-fold compared to standard protocol.

8. References

1 Scott, DE et al. (2016). *Nat. Rev. Drug Discov.* 15**,** 533–550 //
2 Guo, W et al. (2014). *Bioorg. Med. Chem. Lett.* 24**,** 2546–2554 //
3 Yan, C et al. (2008). *Protein J.* 27**,** 59–70 //
4 Liu, J et al. (2006). *Biochemistry* 45**,** 6873–6888 //
5 Brodsky, S et al. (2021). *Curr. Opin. Struct. Biol.* 71**,** 110–115 //
6 Holehouse, AS & Kragelund, BB. (2024). *Nat. Rev. Mol. Cell Biol.* 25**,** 187–211 //
7 Weiner, LM et al. (2010). *Nat. Rev. Immunol.* 10**,** 317–327 //
8 Kothari, M et al. (2024). *Cureus* 16**,** //
9 Crescioli, S et al. (2024). *mAbs* 16**,** 1–35 //
10 Keizer, RJ et al. (2010). *Clin. Pharmacokinet.* 49**,** 493–507 //
11 Chen, C et al. (2025). *mAbs* 17**,** //
12 Yu, YB et al. (2021). *Pharm. Res.* 38**,** 3–7 //
13 Qvit, N et al. (2017). *Drug Discov. Today* 22**,** 454–462 //
14 Nada, H et al. (2024). *Signal Transduct. Target. Ther.* 9**,** 1–32 //
15 Al Musaimi, O et al. (2025). *Pharmaceuticals* 18**,** 291 //
16 van Wier, SP & Beekman, AM. *Chem. Soc. Rev.* 54**,** 1684–1698 //
17 Wang, X et al. (2021). *Front. Chem.* 9**,** //
18 Erlanson, DA et al. (2016). *Nat. Rev. Drug Discov.* 15**,** 605–619 //
19 Bon, M et al. (2022). *Mol. Oncol.* 16**,** 3761–3777 //
20 Semenza, GL. (2002). *Proc. Natl. Acad. Sci.* 99**,** 11570–11572 //
21 Freedman, SJ et al. (2002). *Proc. Natl. Acad. Sci. U. S. A.* 99**,** 5367–5372 //
22 Nguyen, LK et al. (2013). *J. Cell Sci.* 126**,** 1454–1463 //
23 Yfantis, A et al. (2023). *Cells* 12**,** 798 //
24 Fernandes, MT et al. (2020). *World J. Clin. Oncol.* 11**,** 260–274 //
25 Cairns, RA et al. (2007). *Proc. Natl. Acad. Sci.* 104**,** 9445–9450 //
26 Burroughs, SK et al. (2013). *Future Med. Chem.* 5**,** 553–572 //
27 Berlow, RB et al. (2017). *Nature* 543**,** 447–451 //
28 Petrovicz, VL et al. (2024). *RSC Chem. Biol.* 5**,** 711–720 //
29 Berlow, RB et al. (2022). *Proc. Natl. Acad. Sci.* 119**,** //
30 Martinek, TA & Fülöp, F. (2012). *Chem. Soc. Rev.* 41**,** 687–702 //
31 Pelay-Gimeno, M et al. (2015). *Angew. Chem. Int. Ed.* 54**,** 8896–8927 //
32 Goodman, CM et al. (2007). *Nat. Chem. Biol.* 3**,** 252–262 //
33 Flood, DT et al. (2020). *Isr. J. Chem.* 60**,** 268–280 //
34 Mannocci, L et al. (2011). *Chem. Commun.* 47**,** 12747–12753 //
35 Song, Y & Li, X. (2021). *Acc. Chem. Res.* 54**,** 3491–3503 //
36 Redfern, O et al. (2008). *Curr. Opin. Struct. Biol.* 18**,** 394–402 //
37 Dunker, AK et al. (2001). *J. Mol. Graph. Model.* 19**,** 26–59 //
38 Forman-Kay, JD & Mittag, T. (2013). *Structure* 21**,** 1492–1499 //
39 van der Lee, R et al. (2014). *Chem. Rev.* 114**,** 6589–6631 //
40 Wright, PE & Dyson, HJ. (2009). *Curr. Opin. Struct. Biol.* 19**,** 31–38 //
41 Yang, J et al. (2019). *Protein Sci.* 28**,** 1952–1965 //
42 Wright, PE & Dyson, HJ. (2015). *Nat. Rev. Mol. Cell Biol.* 16**,** 18–29 //
43 Miao, J & Chong, S. (2025). *Curr. Opin. Genet. Dev.* 90**,** 102285 //
44 Dyson, HJ. (2016). *Biophys. J.* 110**,** 1013–1016 //
45 Zhao, B & Kurgan, L. (2022). *Biomolecules* 12**,** 888 //

46 Ahmed, MC et al. (2020). *Methods Mol. Biol. Clifton NJ* 2141**,** 429–445 //
47 Waszkiewicz, R et al. (2024). *J. Phys. Chem. Lett.* 15**,** 5024–5033 //
48 Emenecker, RJ et al. (2023). *bioRxiv* doi:10.1101/2023.10.29.564547 //
49 Bah, A & Forman-Kay, JD. (2016). *J. Biol. Chem.* 291**,** 6696–6705 //
50 Usher, ET et al. (2024). *Biophys. J.* 123**,** 4082–4096 //
51 Owen, I & Shewmaker, F. (2019). *Int. J. Mol. Sci.* 20**,** 5501 //
52 Wang, Q et al. (2022). *Proc. Natl. Acad. Sci.* 119**,** //
53 Kudaibergenov, SE & Nuraje, N. (2018). *Polymers* 10**,** 1146 //
54 Bugge, K et al. (2025). *Nat. Commun.* 16**,** 3242 //
55 Das, RK & Pappu, RV. (2013). *Proc. Natl. Acad. Sci.* 110**,** 13392–13397 //
56 Jeon, J & Dobrynin, AV. (2003). *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 67**,** //
57 Espinoza-Fonseca, LM. (2011). *Mol. Biosyst.* 8**,** 237–246 //
58 Tompa, P et al. (2014). *Mol. Cell* 55**,** 161–169 //
59 Maiti, S & De, S. (2022). *J. Magn. Reson. Open* 10–11**,** 100029 //
60 Bhowmick, P et al. (2015). *Adv. Exp. Med. Biol.* 870**,** 291–318 //
61 Faustova, I et al. (2021). *EMBO J.* 40**,** //
62 Bugge, K et al. (2020). *Front. Mol. Biosci.* 7**,** //
63 Morris, OM et al. *Open Biol.* 11**,** 210222 //
64 Vacic, V et al. (2007). *J. Proteome Res.* 6**,** 2351–2366 //
65 Trivedi, R & Nagarajaram, HA. (2022). *Int. J. Mol. Sci.* 23**,** 14050 //
66 Thakur, JK et al. (2014). *Nucleic Acids Res.* 42**,** 2112–2125 //
67 Gulbis, JM et al. (1996). *Cell* 87**,** 297–306 //
68 Brett, TJ et al. (2002). *Structure* 10**,** 797–809 //
69 Krois, AS et al. (2018). *Proc. Natl. Acad. Sci.* 115**,** 11302–11310 //
70 Arai, M et al. (2015). *Proc. Natl. Acad. Sci.* 112**,** 9614–9619 //
71 Dogan, J et al. (2014). *Phys. Chem. Chem. Phys.* 16**,** 6323–6331 //
72 Csermely, P et al. (2010). *Trends Biochem. Sci.* 35**,** 539–546 //
73 Kjaergaard, M et al. (2010). *Proc. Natl. Acad. Sci. U. S. A.* 107**,** 12535–12540 //
74 Lacy, ER et al. (2005). *J. Mol. Biol.* 349**,** 764–773 //
75 Dahal, L et al. (2017). *Biophys. J.* 113**,** 2713–2722 //
76 Skriver, K et al. (2023). *Curr. Opin. Struct. Biol.* 83**,** 102697 //
77 Reid, KM et al. (2023). *J. Phys. Chem. B* 127**,** 7839–7847 //
78 Fuxreiter, M & Tompa, P. (2012). *Adv. Exp. Med. Biol.* 725**,** 1–14 //
79 Gianni, S et al. (2021). *Acc. Chem. Res.* 54**,** 1251–1259 //
80 Lacy, ER et al. (2004). *Nat. Struct. Mol. Biol.* 11**,** 358–364 //
81 Mittag, T et al. (2008). *Proc. Natl. Acad. Sci.* 105**,** 17772–17777 //
82 Soni, S & and Padwad, YS. (2017). *Acta Oncol.* 56**,** 503–515 //
83 Wang, GL et al. (1995). *Proc. Natl. Acad. Sci.* 92**,** 5510–5514 //
84 Yang, C et al. (2021). *Chin. J. Nat. Med.* 19**,** 521–527 //
85 Jiang, B-H et al. (1996). *J. Biol. Chem.* 271**,** 17771–17778 //
86 Chowdhury, R et al. (2009). *Structure* 17**,** 981–989 //
87 Maxwell, PH et al. (1999). *Nature* 399**,** 271–275 //
88 Lando, D et al. (2002). *Genes Dev.* 16**,** 1466–1471 //
89 McNeill, LA et al. (2002). *Biochem. J.* 367**,** 571–575 //
90 Berra, E et al. (2001). *EMBO Rep.* 2**,** 615–620 //
91 Chilov, D et al. (1999). *J. Cell Sci.* 112 (Pt 8)**,** 1203–1212 //
92 Kaluz, S et al. (2008). *Clin. Chim. Acta Int. J. Clin. Chem.* 395**,** 6–13 //
93 Semenza, GL. (2002). *Biochem. Pharmacol.* 64**,** 993–998 //

94 Dames, SA et al. (2002). *Proc. Natl. Acad. Sci. U. S. A.* 99**,** 5271–5276 //
95 Dyson, HJ & Wright, PE. (2016). *J. Biol. Chem.* 291**,** 6714–6722 //
96 Ferrie, JJ et al. (2024). *Mol. Cell* 84**,** 234–243 //
97 Schödel, J et al. (2011). *Blood* 117**,** 207–217 //
98 Liu, Y et al. (1995). *Circ. Res.* 77**,** 638–643 //
99 Ramakrishnan, S et al. (2014). *J. Neuroimmune Pharmacol. Off. J. Soc. NeuroImmune Pharmacol.* 9**,** 142–160 //
100 Yang, L et al. (2018). *Biochem. Biophys. Res. Commun.* 507**,** 128–135 //
101 Kietzmann, T et al. (1999). *Blood* 94**,** 4177–4185 //
102 Fink, T et al. (2002). *Blood* 99**,** 2077–2083 //
103 Jiang, B-H et al. (1997). *J. Biol. Chem.* 272**,** 19253–19260 //
104 Cui, X et al. (2017). *Oncotarget* 8**,** 24840–24852 //
105 Semenza, GL et al. (1996). *J. Biol. Chem.* 271**,** 32529–32537 //
106 Hayashi, M et al. (2004). *J. Endocrinol.* 183**,** 145–154 //
107 Sadlecki, P et al. (2014). *BioMed Res. Int.* 2014**,** //
108 Bruick, RK. (2000). *Proc. Natl. Acad. Sci.* 97**,** 9082–9087 //
109 Piret, J-P et al. (2005). *J. Biol. Chem.* 280**,** 9336–9344 //
110 Miyazaki, K et al. (2002). *J. Biol. Chem.* 277**,** 47014–47021 //
111 Salnikow, K et al. (2008). *Carcinogenesis* 29**,** 1493–1499 //
112 Minamishima, YA et al. (2009). *Mol. Cell. Biol.* 29**,** 5729–5741 //
113 Wei, J et al. *Mini-Rev. Med. Chem.* 18**,** 296–309 //
114 Balamurugan, K. (2016). *Int. J. Cancer J. Int. Cancer* 138**,** 1058–1066 //
115 Semenza, GL. (2004). *Cancer Res.* 64**,** 1312–1313 //
116 Shi, Y-H & Fang, W-G. (2004). *World J. Gastroenterol.* 10**,** 1082–1087 //
117 Sharma, A et al. (2022). *Front. Genet.* 13**,** //
118 Zhang, W et al. (2015). *PLoS ONE* 10**,** //
119 Zhao, J et al. (2024). *Heliyon* 10**,** //
120 Bhattacharya, S et al. (1999). *Genes Dev.* 13**,** 64–75 //
121 Chou, Y-T et al. (2012). *Cell Death Differ.* 19**,** 2015–2028 //
122 Bragança, J et al. (2003). *J. Biol. Chem.* 278**,** 16021–16029 //
123 Kuna, M & Soares, MJ. (2024). *BioEssays News Rev. Mol. Cell. Dev. Biol.* 46**,** //
124 Weninger, WJ et al. (2005). *Dev. Camb. Engl.* 132**,** 1337–1348 //
125 Fame, RM et al. (2016). *J. Neurosci.* 36**,** 6403–6419 //
126 Lawson, H et al. (2021). *Stem Cell Rep.* 16**,** 2784–2797 //
127 Li, Q et al. (2012). *J. Biol. Chem.* 287**,** 29088–29100 //
128 De Guzman, RN et al. (2005). *Biochemistry* 44**,** 490–497 //
129 Lindström, I et al. (2018). *Sci. Rep.* 8**,** 7872 //
130 Wang, Y & Brooks III, CL. (2020). *J. Phys. Chem. Lett.* 11**,** 864–868 //
131 Chu, W-T et al. (2020). *Proc. Natl. Acad. Sci. U. S. A.* 117**,** 5595–5603 //
132 Appling, FD et al. (2021). *Struct. Lond. Engl. 1993* 29**,** 1327–1338 //
133 Sancho, DD & Best, RB. (2011). *Mol. Biosyst.* 8**,** 256–267 //
134 Ruiz-Ortiz, I & Sancho, DD. (2020). *Phys. Chem. Chem. Phys.* 22**,** 8118–8127 //
135 Gao, M et al. (2019). *Biophys. J.* 117**,** 1301–1310 //
136 Berlow, RB et al. (2019). *Biochemistry* 58**,** 1354–1362 //
137 Wen, B et al. (2022). *J. Phys. Chem. Lett.* 13**,** 9201–9209 //
138 Nyqvist, I et al. (2019). *J. Phys. Chem. B* 123**,** 2882–2888 //
139 Smith, MC & Gestwicki, JE. (2012). *Expert Rev. Mol. Med.* 14**,** //
140 Chen, J et al. (2013). *Protein Sci. Publ. Protein Soc.* 22**,** 510–515 //

141 Liu, Z & Huang, Y. (2014). *Protein Sci. Publ. Protein Soc.* 23**,** 539–550 //
142 Swenson, CS et al. (2024). *Chem. Rev.* 124**,** 13020–13093 //
143 Wilson, AJ. (2009). *Chem. Soc. Rev.* 38**,** 3289–3300 //
144 van Wier, SP & Beekman, AM. *Chem. Soc. Rev.* 54**,** 1684–1698 //
145 Lenci, E & Trabocchi, A. (2020). *Chem. Soc. Rev.* 49**,** 3262–3277 //
146 Wójcik, P & Berlicki, Ł. (2016). *Bioorg. Med. Chem. Lett.* 26**,** 707–713 //
147 Ji, X et al. (2024). *Angew. Chem. Int. Ed.* 63**,** //
148 Costa, L et al. (2023). *Pharmaceuticals* 16**,** 996 //
149 Moiola, M et al. (2019). *Molecules* 24**,** 3654 //
150 Jedhe, GS & Arora, PS. (2021) Academic Press in *Methods in Enzymology* (ed. Petersson, E. J.) 656**,** 1–25 //
151 Hetherington, K et al. (2020). *Chem. Weinh. Bergstr. Ger.* 26**,** 7638–7646 //
152 Henchey, LK et al. (2010). *J. Am. Chem. Soc.* 132**,** 941–943 //
153 Werner, HM & Horne, WS. (2015). *Curr. Opin. Chem. Biol.* 28**,** 75–82 //
154 Cabrele, C et al. (2014). *J. Med. Chem.* 57**,** 9718–9739 //
155 Horne, WS et al. (2009). *Proc. Natl. Acad. Sci.* 106**,** 14751–14756 //
156 Johnson, LM & Gellman, SH. (2013). *Methods Enzymol.* 523**,** 407–429 //
157 Olajos, G et al. (2015). *Chem. Weinh. Bergstr. Ger.* 21**,** 6173–6180 //
158 Hegedüs, Z et al. (2013). *J. Am. Chem. Soc.* 135**,** 16578–16584 //
159 Sharma, GVM et al. (2008). *Chem. Asian J.* 3**,** 969–983 //
160 Horne, WS et al. (2008). *Angew. Chem. Int. Ed.* 47**,** 2853–2856 //
161 Nizami, B et al. (2020). *Nucleic Acids Res.* 48**,** 1122–1128 //
162 Vagner, J et al. (2008). *Curr. Opin. Chem. Biol.* 12**,** 292–296 //
163 Tököli, A et al. (2020). *Chem. Sci.* 11**,** 10390–10398 //
164 Simon, MA et al. (2022). *Sci. Rep.* 12**,** 5904 //
165 Wéber, E et al. (2025). *Angew. Chem. Int. Ed.* 64**,** //
166 Bartus, É et al. (2017). *ChemistryOpen* 6**,** 236–241 //
167 Lanne, A et al. (2023). *Drug Discov. Today* 28**,** 103670 //
168 Satam, H et al. (2023). *Biology* 12**,** 997 //
169 Shi, Y et al. (2021). *RSC Adv.* 11**,** 2359–2376 //
170 Satz, AL et al. (2022). *Nat. Rev. Methods Primer* 2**,** 1–17 //
171 Gartner, ZJ et al. (2004). *Science* 305**,** 1601 //
172 Melkko, S et al. (2004). *Nat. Biotechnol.* 22**,** 568–574 //
173 Hansen, MH et al. (2009). *J. Am. Chem. Soc.* 131**,** 1322–1327 //
174 Blakskjaer, P et al. (2015). *Curr. Opin. Chem. Biol.* 26**,** 62–71 //
175 Gironda-Martínez, A et al. (2021). *ACS Pharmacol. Transl. Sci.* 4**,** 1265–1279 //
176 Ding, Y et al. (2021). *Bioorg. Med. Chem.* 41**,** 116216 //
177 Harris, PA et al. (2017). *J. Med. Chem.* 60**,** 1247–1261 //
178 Martin, RE et al. (2025). *Bioorg. Med. Chem. Lett.* 123**,** //
179 Movahedi, M et al. (2021). *Aptamers* 5**,** 22–30 //
180 Guo, C et al. (2015). *J. Am. Chem. Soc.* 137**,** 11191–11196 //
181 Kong, D et al. (2016). *Angew. Chem. Int. Ed Engl.* 55**,** 13164–13168 //
182 Lei, Y & Hili, R. (2017). *Org. Biomol. Chem.* 15**,** 2349–2352 //
183 Lei, Y et al. (2019). *Org. Biomol. Chem.* 17**,** 1962–1965 //
184 Guo, C et al. (2019). *ChemBioChem* 20**,** 793–799 //
185 Guo, C & Hili, R. (2017). *Bioconjug. Chem.* 28**,** 314–318 //
186 Xu, Y et al. (2021). *Front. Bioeng. Biotechnol.* 9**,** //
187 Kong, D et al. (2020). *ACS Synth. Biol.* 9**,** 43–52 //

188 Michels, T et al. (2012). *Org. Lett.* 14**,** 5218–5221 //
189 Stathopoulos, P et al. (2005). *J. Pept. Sci.* 11**,** 658–664 //
190 Zhao, H et al. (2015). *Methods San Diego Calif* 76**,** 137–148 //
191 Houtman, JCD et al. (2007). *Protein Sci.* 16**,** 30–42 //
192 Krainer, G et al. (2012). *Anal. Chem.* 84**,** 10715–10722 //
193 Sigurskjold, BW. (2000). *Anal. Biochem.* 277**,** 260–266 //
194 Brautigam, CA. (2015) Academic Press in *Methods in Enzymology* (ed. Cole, J. L.) 562**,** 109–133 //
195 Lee, CW et al. (2010). *Proc. Natl. Acad. Sci.* 107**,** 19290–19295 //
196 Roehrl, MHA et al. (2004). *Biochemistry* 43**,** 16056–16066 //
197 Lee, W et al. (2015). *Bioinformatics* 31**,** 1325–1327 //
198 Freedman, SJ et al. (2003). *Nat. Struct. Biol.* 10**,** 504–512 //
199 Guzman, RND et al. (2004). *J. Biol. Chem.* 279**,** 3042–3049 //
200 Brautigam, CA et al. (2016). *Nat. Protoc.* 11**,** 882–894 //
201 Tompa, P & Fuxreiter, M. (2008). *Trends Biochem. Sci.* 33**,** 2–8 //
202 Graham, TA et al. (2001). *Nat. Struct. Biol.* 8**,** 1048–1052 //
203 Reinert, ZE & Horne, WS. (2014). *Org. Biomol. Chem.* 12**,** 8796–8802 //
204 Mortenson, DE et al. (2018). *ChemBioChem* 19**,** 604–612 //
205 Haase, HS et al. (2012). *J. Am. Chem. Soc.* 134**,** 7652–7655 //
206 Checco, JW & Gellman, SH. (2016). *Curr. Opin. Struct. Biol.* 39**,** 96–105 //
207 Wood, CW et al. (2020). *Bioinformatics* 36**,** 2917–2919 //
208 Ibarra, AA et al. (2019). *ACS Chem. Biol.* 14**,** 2252–2263 //
209 Fujino, T et al. (2016). *J. Am. Chem. Soc.* 138**,** 1962–1969 //
210 Madsen, D et al. (2020). *Prog. Med. Chem.* 59**,** 181–249 //
211 Cao, C et al. (2014). *Chem. Commun.* 50**,** 10997–10999 //
212 Chan, AI et al. (2015). *Curr. Opin. Chem. Biol.* 26**,** 55–61 //
213 O'Reilly, RK et al. (2017). *Acc. Chem. Res.* 50**,** 2496–2509 //
214 Onda, Y et al. (2021). *Chem. – Eur. J.* 27**,** 7160–7167 //
215 Zhao, F et al. (2024). *J. Am. Chem. Soc.* 146**,** 1946–1956 //
216 Kupihár, Z et al. (2023). *Pharmaceutics* 15**,** 248 //
217 Li, S et al. (2021). *Mol. Pharm.* 18**,** 1480–1485 //
218 Kantner, T & Watts, AG. (2016). *Bioconjug. Chem.* 27**,** 2400–2406 //
219 Fontaine, SD et al. (2015). *Bioconjug. Chem.* 26**,** 145–152 //
220 Shi, J et al. (2023). *Front. Microbiol.* 14**,** //
221 Zasedateleva, OA et al. (2023). *Int. J. Mol. Sci.* 24**,** 13643 //
222 Hottin, A & Marx, A. (2016). *Acc. Chem. Res.* 49**,** 418–427 //

9. Acknowledgements

First and foremost, I want to express my gratitude to my Ph.D. supervisors. As the head of the department, Prof. Tamás Martinek provided all the necessary financial support and created a truly innovative research environment where I could fully indulge in creativity and scientific freedom. I also owe my heartfelt gratitude to my co-supervisor, Dr. Zsófia Hegedüs, whose encouragement, patience and altruistic help accompanied me throughout my entire Ph.D. journey, making it feel like a smooth sailing. I consider myself very lucky to be her first Ph.D. student. Once again, I cannot be grateful enough for the support and mentorship of both my supervisors, and I am thankful that I can consider them as friends.

I am also grateful to Prof. Enikő Forró, who allowed me to start my academic journey at the Institute of Pharmaceutical Chemistry as an undergraduate researcher and cemented my desire to pursue a carrier in the scientific field.

I would like to thank István Pasztuhov and Márk Tresztián, who supported our efforts with their hard work as undergraduate students. I wish them good luck on their scientific journey, as they hold great potential. I also owe gratitude to Dr. Zoltán Kupihár for his synthetic work on the thymidine-phosphoramidite, and also to Dr. Györgyi Ferenc who provided us the oligonucleotides necessary for our work.

I would like to thank all my colleagues as well; they created a professionally and emotionally supporting environment at the department. I am glad to say that my time here was free of arguments and I will leave with great memories of everyone. I am especially thankful to have gained lifelong friends throughout my Ph.D. journey, namely Dr. Attila Tököli, Dr. Zita Ibolya Papp, Dr. Éva Kovács-Bartus, Arijit Shakar, Dr. Kaushik Bhaumik, Csanád Vidéki, Mihály Gyémánt, Bence Nagymihály and L. Szabolcs Ferenczi. It was also heartwarming, that my friendship with Máté Héthelyi and Martin Cseh, which began during my undergraduate years, continued as we became colleagues as well as friends. I need to mention Dr. Lívia Fülöp and Prof. György Dombi, who stand as role models for me with their knowledge, expertise and scientific attitude. I hope that one day I can live up to them.

I am also grateful to my university friends Péter Szijártó, Gábor Rektenwald, János Böcz, Tamás Szilasi, Tamás Pócsik and Ottó Rácz for not forgetting me, while I was immersed in the lab. I want to thank Máté Fekete, who became my best friend throughout my Ph.D. years, as

well as Cintia Felleg and Zoltán Csendes, who were always there for me to let off some steam together.

My biggest gratitude goes to the most important people in my life —my family —, firstly to my grandma, who struggled through the elementary school with me as a hard-headed child; my sister who always listened when I felt down, and was always proud of me. This whole journey of mine would not have been possible without my parents. Their unconditional love, financial and emotional support allowed me to fully address myself to scientific work, which I have wanted to pursue since my teenage years. They should get at least one letter from my Ph.D., so thanks again, Mom and Dad.

Once again, I extend my heartfelt thanks to everyone who has contributed to my academic and personal growth.

Appendix

ITC fitting and data analysis

$$[\text{TAZ1/HIF}] = K_{\text{HIF}}\,[\text{TAZ1}]\,[\text{HIF}] \quad (1)$$

$$[\text{TAZ1/CITED}] = K_{\text{CITED}}\,[\text{TAZ1}]\,[\text{CITED}] \quad (2)$$

$$[\text{TAZ1/HIF/CITED}] = K_{[\text{TAZ1/HIF}]/\text{CITED}}\,[\text{TAZ1/HIF}]\,[\text{CITED}] \quad (3)$$

$$[\text{TAZ1/HIF/CITED}] = K_{[\text{TAZ1/CITED}]/\text{HIF}}\,[\text{TAZ1/CITED}]\,[\text{HIF}] \quad (4)$$

Cooperativity constant α for CITED2 binding to the preformed TAZ1/HIF is defined by:

$$K_{[\text{TAZ1/HIF}]/\text{CITED}} = \alpha\, K_{\text{CITED}} \quad (5)$$

$$\alpha = K_{[\text{TAZ1/HIF}]/\text{CITED}} / K_{\text{CITED}} \quad (6)$$

from which the cooperative enhancement of binding CITED2 to the preformed TAZ/HIF-1α complex was calculated as follows:

$$\Delta G_{[\text{TAZ1/HIF/CITED}]} = -RTln(\alpha\, K_{\text{CITED}}\, K_{\text{HIF}}) = \Delta G_{\text{CITED}} + \Delta G_{\text{HIF}} + \Delta g \quad (7)$$

$$\Delta H_{[\text{TAZ1/HIF/CITED}]} = \Delta H_{\text{CITED}} + \Delta H_{\text{HIF}} + \Delta h \quad (8)$$

$$-T\Delta s = \Delta g - \Delta h \quad (9)$$

Equations used for fitting model and data analysis of the fluorescence anisotropy competition titrations

Fluorescence intensity and anisotropy were calculated according to the following equations as previously described:

$$I = (2PG) + S \quad (10)$$

$$r = (S - PG) \quad (11)$$

where I = total intensity, r = anisotropy, P = perpendicular intensity, S = parallel intensity, G = instrument factor set to 1.

Apparent competition K_D ($K_{D,app}$) values for the competitor peptides were fitted using a method described previously[195,196], using the following equations in Origin Pro:

$$r = \frac{(\gamma r_{TAZ:HIF} + \gamma r_{HIF})}{1+\gamma} \quad (12)$$

where r = fluorescence anisotropy, *TAZ* is $\text{TAZ1}_{330\text{-}420}$, *HIF* is Flu-HIF-$1\alpha_{786\text{-}826}$, $r_{\text{TAZ:HIF}}$ is the anisotropy of TAZ1 bound Flu-HIF-$1\alpha_{786\text{-}826}$, while r_{HIF} is the anisotropy of unbound HIF-$1\alpha_{786\text{-}826}$. γ is given by the following equation:

$$\gamma = \frac{[TAZ:HIF]}{[HIF]_t - [TAZ:HIF]} \quad (13)$$

where *[TAZ:HIF]* is the concentration of TAZ1 bound Flu-HIF-1α$_{786\text{-}826}$, *[HIF]*$_t$ is the total concentration of Flu-HIF-1α$_{786\text{-}826}$. *[TAZ:HIF]* is given by:

$$[TAZ{:}HIF] = \frac{[HIF]_t(2\sqrt{a^2-3b}\,cos\frac{\theta}{3}-a)}{3K_{HIF}+(2\sqrt{a^2-3b}\,cos\frac{\theta}{3}-a)} \quad (14)$$

where

$$\theta = arccos\frac{-2a^3+9ab-27c}{2\sqrt{(a^2-3b)^3}} \quad (15)$$

$$a = K_{HIF} + K_{CIT} + [HIF]_t + [CIT]_t - [TAZ]_t \quad (16)$$

$$b = K_{HIF}K_{CIT} + K_{HIF}([CIT]_t - [TAZ]_t) + K_{CIT}([HIF]_t - [TAZ]_t) \quad (17)$$

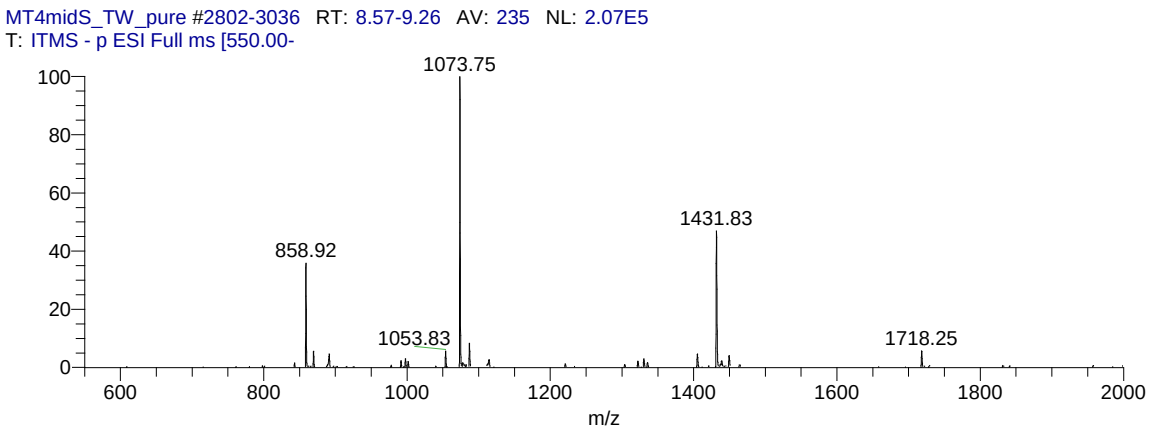
$$c = -[TAZ]_t K_{HIF} K_{CIT} \quad (18)$$

K_{HIF} is the dissociation constant for Flu-HIF-1α$_{786\text{-}826}$, K_{CIT} is the dissociation constant for CITED2 or a given variant, while [*TAZ*] is the total concentration of TAZ1$_{330\text{-}420}$.

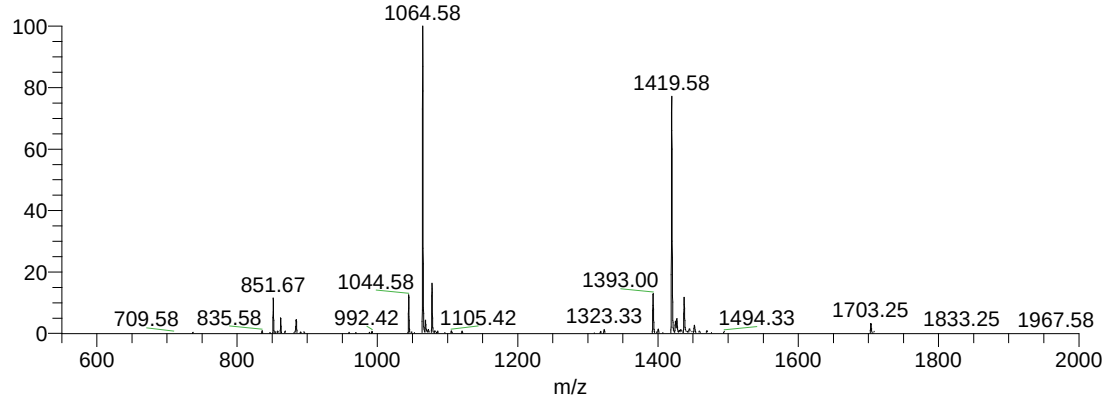
Fitted parameters were K_{HIF}, K_{CIT}, $r_{TAZ:HIF}$, r_{HIF}, $[HIF]_t$, $[TAZ]_t$, where K_{HIF}, $r_{TAZ:HIF}$, r_{HIF}, $[HIF]_t$ were kept constant during fitting. K_{HIF} was determined in a separate experiment in which HIF-1α$_{776\text{-}826}$ was the competitor (Figure 3.). $[HIF]_t$ was determined using fluorescence intensity measurements relative to a fluorescein calibration curve. $[TAZ]_t$ was kept as a floating parameter to account for incompetent protein fraction. Anisotropy values were normalized and represented as Ligand-bound (LB) values as follows:

$$L_B = \frac{r-r_{HIF}}{r_{TAZ:HIF}-r_{HIF}}[HIF] \quad (19)$$

MS spectra of the foldamer oligonucleotide conjugates



MT5midS_LWpure #2992-3111 RT: 9.14-9.48 AV: 120 NL: 2.16E5
T: ITMS - p ESI Full ms [550.00-



MT16_RW_conj_220721 #2885-2912 RT: 8.81-8.90 AV: 28 NL: 1.26E4
T: ITMS - p ESI Full ms [550.00-