

---

# **DESIGN AND SYNTHESIS OF LIGANDS TO MODULATE IMPORTANT CELLULAR PATHWAYS**

---

**IVAN de PAOLA**

Dottorato in Scienze Biotecnologiche – XXIV ciclo  
Indirizzo Biotecnologie Industriale e Molecolare  
Università di Napoli Federico II





Dottorato in Scienze Biotecnologiche – XXIV ciclo  
Indirizzo Biotecnologie Industriale e Molecolare  
Università di Napoli Federico II



---

# **DESIGN AND SYNTHESIS OF LIGANDS TO MODULATE IMPORTANT CELLULAR PATHWAYS**

---

**Ivan de Paola**

Dottorando: Ivan de Paola  
Relatore: Prof. Ettore Benedetti  
Correlatore: Dr.ssa Laura Zaccaro  
Coordinatore: Prof. Giovanni Sannia







## INDEX

<b>ABBREVIATIONS</b>	pag. 1
<b>SUMMARY</b>	pag. 3
<b>RIASSUNTO</b>	pag. 5
<b>1. INTRODUCTION</b>	pag. 13
1.1 <i>Protein-Protein interactions: Structural Basis</i>	pag. 13
1.2 <i>Helical Structure in Protein Interface</i>	pag. 15
1.3 <i>Helical Structure mimetics: Peptides and Peptidomimetics</i>	pag. 17
1.4 <i>p53 Protein</i>	pag. 21
1.4.1 <i>p53: a cancer suppressor</i>	pag. 21
1.4.2 <i>p53-TFIIH interaction</i>	pag. 22
1.5 <i>KCTD Protein</i>	pag. 25
1.5.1 <i>POZ/BTB domain</i>	pag. 25
1.5.2 <i>Potassium Channel Tetramerization Domain (KCTD) Proteins</i>	pag. 26
1.5.3 <i>E3 ubiquitin-protein ligases</i>	pag. 26
1.6 <i>Aim of the work</i>	pag. 30
<b>2. MATERIALS AND METHODS</b>	pag. 32
2.1 <i>Materials</i>	pag. 32
2.2 <i>p53 Mimetic Peptides</i>	pag. 32
2.2.1 <i>Peptides synthesis</i>	pag. 32
2.2.2 <i>Circular Dichroism Studies</i>	pag. 33
2.2.3 <i>Isothermal Titration Calorimetry Studies</i>	pag. 33
2.2.4 <i>Cloning, Expression and Purification of Recombinant Proteins</i>	pag. 33
2.2.5 <i>Media, Plasmids and Strains</i>	pag. 34
2.2.6 <i><math>\beta</math>-Galactosidase Activation Assay</i>	pag. 34
2.2.7 <i>NMR Samples</i>	pag. 34
2.2.8 <i>NMR Spectroscopy Experiments</i>	pag. 35
2.3 <i>Cul3 Mimetic Peptides</i>	pag. 35
2.3.1 <i>Peptides Synthesis</i>	pag. 35
2.3.2 <i>Circular Dichroism Studies</i>	pag. 36
2.3.3 <i>Expression and Purification of POZ/BTB domains of KCTD11</i>	pag. 36
2.3.4 <i>ELISA assays for Cul3 peptides</i>	pag. 37

<b>3. RESULTS</b>	pag. 38
3.1 <i>p53 Mimetic Peptides</i>	pag. 38
3.1.1 <i>First Generation: Design of p53-13, NC15 and NC17 Peptides</i>	pag. 38
3.1.2 <i>Synthesis of p53-13, NC15 and NC17 Peptides</i>	pag. 39
3.1.3 <i>CD studies</i>	pag. 41
3.1.4 <i>ITC Experiments</i>	pag. 42
3.1.5 <i>NMR Experiments</i>	pag. 42
3.1.6 <i>Second Generation: Design of N-Cap series peptides</i>	pag. 43
3.1.7 <i>Synthesis of A-Cap, W-Cap and E-Cap peptides</i>	pag. 44
3.1.8 <i>CD studies</i>	pag. 45
3.1.9 <i>ITC Experiments</i>	pag. 46
3.1.10 <i>NMR Experiments</i>	pag. 46
3.1.11 <i>E-Cap in vivo experiment</i>	pag. 48
3.2 <i>Cul3 Mimetics</i>	pag. 50
3.2.1 <i>Design of Cul3-wt, Cul3-KK and Cul3-Y(P)</i>	pag. 50
3.2.2 <i>Synthesis of Cul3-wt, Cul3-KK and Cul3-Y(P)</i>	pag. 51
3.2.3 <i>CD Studies</i>	pag. 55
3.2.4 <i>ELISA Assay</i>	pag. 56
3.2.5 <i>Design of Cul3-stapled</i>	pag. 57
3.2.6 <i>Synthesis of Cul3-stapled</i>	pag. 57
3.2.7 <i>CD Studies</i>	pag. 60
3.2.8 <i>ELISA Assay</i>	pag. 61
<b>4. CONCLUSION and FUTURE PERSPECTIVES</b>	pag. 62
<b>5. BIBLIOGRAPHY</b>	pag. 64
<b>SUPPLEMENTARY INFORMATION</b>	pag. 68

## ABBREVIATIONS

<b>ATF</b>	Artificial Transcription Factor
<b>BOC</b>	Butossicarbonyl
<b>BTB</b>	Bric-à-brac, Tramtrack and Broad
<b>CUL</b>	Cullin
<b>DBD</b>	DNA Binding Domain
<b>DCM</b>	Dicloromethan
<b>DIPEA</b>	Diisopropilethylammina
<b>DMF</b>	Dimethylformammide
<b>EDT</b>	Ethanedithiol
<b>ELISA</b>	Enzyme-linked immunosorbent assay
<b>ESI</b>	Electrospray Ionization-Mass Spectrometry
<b>Fmoc</b>	9-Fluorenilmetossicarbonyl
<b>HATU</b>	2-(1H-9-Azobenzotriazole-1-yl)-1,1,3,3,- tetramethyluronium
<b>HBTU</b>	2-(1H-Benzotriazole-1-ile)-1,1,3,3- tetramethyluronium esafluoro phosphate
<b>Hh</b>	Hedgehog
<b>HOBT</b>	Hydroxybenzotriazole
<b>ITC</b>	Isothermal Titration Calorimetry
<b>KCTD</b>	K <sup>+</sup> Channel Tetramerization Domain
<b>HSQC</b>	Heteronuclear Single Quantum Coherence
<b>LC-MS</b>	Liquid chromatography-mass spectrometry
<b>MeOH</b>	Methanol
<b>NMP</b>	N-Methyl-2-pyrrolidone
<b>NMR</b>	Nuclear magnetic resonance spectroscopy
<b>NOE</b>	Nuclear Overhauser Effect Spectroscopy
<b>OtBu</b>	O-t-Butyl
<b>POZ</b>	Poxvirus Zinc Finger
<b>PyBOP</b>	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
<b>RP-HPLC</b>	Reverse Phase High Performance Liquid Chromatography
<b>TAD</b>	Trans Activaction Domain
<b>tBu</b>	t-Butyl
<b>TEV</b>	Tobacco Etch Virus
<b>TFA</b>	Trifluoroacetic acid
<b>TIS</b>	TriisopropylSilane
<b>TRIS</b>	Tris (hidroxy methyl) amino methane
<b>Trx</b>	Thioredoxin A
<b>UV-Vis</b>	Ultraviolet-Visible Spectroscopy
<b>VEGF</b>	Vascular Endothelial Growth factor



## SUMMARY

The study of the protein-protein interactions (PPI) is fundamental to correlate structure and function in biological systems. The understanding of molecular features that regulate these interactions is the starting point for the design of specific ligands able to modulate main cellular activities allowing the development of new systems for biotechnology applications.

Recent studies by Arora and co-workers highlighted that roughly 60% of the protein complexes found in the PDB display helical structure at interface such as  $\alpha$ - and  $3_{10}$ -helices. The average length of helical domains in interacting interfaces is rather small and spans two to three helical turns (or eight to twelve residues). These complexes suggest that it may be possible to develop small molecules that mimic such structure domains and display functionality in a similar fashion.

Peptides are molecules able to mimic structural motifs enabling the detailed analysis of the interaction at the level of individual amino acid residues. Moreover several progresses in chemical synthesis allowed the introduction of different building blocks and modification at peptide backbone, helpful to stabilize defined secondary structures. This can be used to design peptides and peptidomimetics, mimicking sequences involved in PPI as molecules of biotechnological interest.

Malfunctions in transcriptional regulation are associated with many human diseases and there is considerable interest in biomedical and biotechnological field in developing artificial transcription factors (ATFs) that must minimally contain a DNA binding domain (DBD) and a transactivation domain (TAD).

Designing artificial TADs is proven difficult because these systems interact with multiple target proteins as part of their normal function and, moreover, there are only a limited number of high-resolution structures of TADs in complex with target proteins.

The p53 protein is a transcriptional activator factor with a fundamental role in cellular survival. Recent structural studies by Omichinski and co-workers gave insights into interaction between p53 TAD2 and ttf1/p62 subunit of the transcription factor TFIIH and contributed to the understanding of the way p53 protein works in transcriptional activation.

On the basis of these information new molecules, able to interfere with the binding between p53 TAD2 and ttf1 of TFIIH, were designed. In our hypothesis peptide analogues of p53 TAD2 with enhanced helical propensity could yield a more potent artificial TAD. Therefore, several peptides mimicking the helical fragment 47-55 of p53 TAD2 were designed by using different molecular tools, such as N- and C-capping boxes, and synthesized.

A structural analysis of all obtained peptides by CD experiments highlighted an increased helical content relative to native p53-13 peptide and one of these peptides, E-Cap, showed a higher binding affinity than other peptides in ITC experiments.

Furthermore NMR experiments were performed to elucidate how E-Cap peptide interacts with ttf1. NMR titrations spectra displayed that E-Cap peptide binds along the same interface of p53<sub>25-75</sub>, forming a 9-residue  $\alpha$ -helix when in complex with ttf1. NMR studies of E-Cap peptide in complex demonstrated the presence of both the  $i, i+3$  and the  $i, i+4$  interactions for Leu6 with Leu9 and Trp10 residues, respectively.

The potential ability of E-Cap peptide to activate transcription *in vivo* was investigated in yeast cells by activation assay of *lacZ* with LexA fusion peptide. E-Cap peptide

showed to be a more potent transcriptional activator than positive control Gal4, p53-13 native sequence and AH and VP2 transcriptional activators.

Besides mutation studies for E-Cap and p53-13 peptides provided evidence that both *i, i+3* interaction between Leu6-Leu9, and *i, i+4* interaction between Leu6-Trp10 have an important role for the *in vivo* activity.

In conclusion, a new peptide with high content of helical conformation and able to activate transcription *in vitro* was designed using only proteinogenic amino acids. This result will be used to develop a new interesting biotechnological system such as "programmable ATF", linking obtained TAD to new synthetic DBD.

BTB/POZ domain-containing proteins are identified in diverse cellular locations and participate in different processes ranging from transcriptional repression/activation to cytoskeletal organization. In particular, KCTD are a class of protein containing BTB/POZ domains with different biological function as KCTD11 that inhibits tumor growth in medulloblastoma. POZ/BTB domain of KCTD proteins mediates interaction with Cullin, a family of proteins ubiquitin ligases that induce proteolysis of target proteins by proteasome. Recently, different studies highlighted that several members of KCTD family protein bind to Cullin3 (Cul3).

The understanding of the structural features required for the interaction of the Cul3 with KCTDs is an intriguing challenge and is the basis for the future design of molecules with agonist/antagonist function for biotechnological applications.

To achieve this goal the attention was focused on KCTD11 that mediates interaction with Cul3 by its POZ/BTB domain at N-terminus.

Starting from the proposed model of KCTD11-Cul3 interaction, peptide molecules mimicking the helical fragment of Cul3 49-68, were designed. To evaluate the role of some aromatic residues of Cul3 in the interaction with KCTD11 two mutant peptides were developed (Cul3-KK and Cul3-Y(P)). Structural analysis by CD experiments highlighted that these peptide and Cul3-wt native sequence peptide are unfolded. ELISA assays showed that Cul3-KK and Cul3-Y(P) have a lower affinity than Cul3-wt for KCTD11.

Subsequently, in order to test if the enhance of the helical propensity match with a gain in binding affinity, opportunely constrained sequences were designed and synthesized. Cul3 stapled peptide, with a hydrocarbon bridge in central position of sequence, showed indeed a higher helical structure than the native Cul3-wt peptide. Preliminary ELISA assays indicate that both Cul3-wt and Cul3 stapled peptides were able to interact with KCTD11.

Future studies by ITC and NMR will allow the evaluation of the  $K_d$  and the molecular features governing the interaction with KCTD11 required for the future perspectives in biotechnology field.

## RIASSUNTO

### PREMESSE SCIENTIFICHE

#### Interazioni Proteina-Proteina e Peptidomimetici

Le interazioni tra le proteine svolgono un ruolo chiave nella regolazione dell'omeostasi cellulare. La comprensione della relazione tra la struttura e la funzione è fondamentale per comprendere le basi molecolari di tali interazioni.

Lo studio delle superfici d'interazione tra proteine, la cui struttura è stata risolta mediante tecniche NMR o raggi X e depositata nel PDB, ha evidenziato che in circa il 60% dei casi sono presenti motivi elicoidali come  $\alpha$ -eliche ed eliche  $3_{10}$ .

In particolare, è stato osservato che la lunghezza media delle eliche coinvolte nelle interazioni proteina-proteina è relativamente piccola e comprende in genere due o tre *turn*, cioè 8/12 residui. All'interno di queste brevi sequenze elicoidali sono contenuti i residui essenziali per l'interazione in quanto ognuno di questi contribuisce, interagendo con i residui dell'altra proteina, all'energia libera di legame.

I peptidi sono potenzialmente molecole capaci di mimare le superfici di legame tra proteine e quindi di essere utilizzati per modulare importanti processi cellulari. Purtroppo, estrapolati dalle proteine, essi raramente conservano la conformazione nativa mentre adottano un insieme di strutture diverse, perdendo la capacità di legarsi specificamente ai loro bersagli. Inoltre, l'instabilità proteolitica dei peptidi è un altro fattore che ne limita l'utilizzo. Per questi motivi negli ultimi anni sono state sviluppate diverse metodologie per stabilizzare la struttura elicoidale di peptidi e per aumentare la loro resistenza alle proteasi.

Queste strategie sono state utilizzate per ottenere molecole, derivate da domini coinvolti in interazioni tra proteine, quali ligandi specifici importanti per lo sviluppo di nuovi sistemi da utilizzare in ambito biotecnologico.

#### p53: Un potente attivatore della trascrizione

L'alterata regolazione della trascrizione è una delle principali cause di malattie umane. Per questo motivo c'è un notevole interesse nello sviluppo di fattori di trascrizione "artificiali" (ATF) che possano attivare o reprimere un gene specifico [6]. Affinché una biomolecola possa funzionare come ATF deve contenere un dominio di transattivazione della trascrizione (TAD) in grado di legare un fattore di trascrizione e un dominio di legame specifico al DNA (DBD).

Mentre negli ultimi anni sono stati sviluppati diversi tipi di domini artificiali di legame al DNA come poliammidi pirrolo-imidazolo (PIP), PNA e proteine a dita di zinco, più complicato si è rivelata la progettazione di domini di trascrizione artificiali a causa delle scarse informazioni strutturali.

Recentemente il gruppo del professor Omichinski dell'Università di Montreal in Canada ha risolto, mediante tecniche NMR, la struttura del complesso molecolare costituita dal dominio TAD della proteina p53 e il fattore di trascrizione generale TFIIH.

La proteina p53 è un potente attivatore della trascrizione che agisce in seguito a danni alla cellula. I geni bersaglio di p53 sono diversi e possono mediare risposte cellulari differenti: arresto del ciclo cellulare nel punto di regolazione G1/S,

attivazione del sistema di riparazione del DNA e inizio della cascata di segnali che inducono l'apoptosi nel caso in cui il danno al DNA sia irreparabile.

La struttura NMR del complesso tra il dominio tfb1 di TFIIH del lievito (p62 nell'uomo) e la sequenza di p53<sub>20-73</sub> contenente il dominio TAD2, ha evidenziato che p53 contiene una  $\alpha$ -elica anfipatica tra la Pro47 e la Thr55 che interagisce con i foglietti  $\beta$ 5- $\beta$ 7 ed il *loop* fra  $\beta$ 5 e  $\beta$ 6 del dominio PH della subunità tfb1 di TFIIH. Studi di mutagenesi sito diretta hanno dimostrato che i tre residui idrofobici Phe54, Trp53 e Ile50 del dominio TAD2 di p53 sono fondamentali per l'interazione con tfb1/p62. Tali residui occupano una faccia dell'elica anfipatica e stabiliscono delle interazioni idrofobiche con TFIIH, mentre il residuo Glu51 stabilisce interazioni elettrostatiche fondamentali per la stabilizzazione del complesso. Queste informazioni dimostrano che l'interazione tra p53 e tfb1 è basata su un riconoscimento di superficie, in cui è necessaria la presenza della struttura elicoidale del frammento 47-55 di p53 per un corretto orientamento dei residui coinvolti nel legame.

### Il dominio POZ/BTB e la famiglia di proteine KCTD

Il dominio POZ/BTB (*Drosophila melanogaster* *bric-à-brac*, *tramtrack* and *broad*) è stato ritrovato in molte proteine con differenti funzioni: repressione della trascrizione, regolazione del citoscheletro, canali ionici e proteine coinvolte nel processo di ubiquitinazione e degradazione. In tutti i domini POZ/BTB è possibile identificare un motivo strutturale comune caratterizzato da tre foglietti beta allineati e cinque alfa eliche, coinvolto in numerose interazioni proteina-proteina.

Questo dominio è altamente conservato nella regione N-terminale di una nuova famiglia di proteine KCTD ( $K^+$  Channel Tetramerization Domain), i cui membri sembrano essere coinvolti in numerosi processi biologici.

Differenti studi hanno dimostrato che la maggior parte delle proteine KCTD interagisce con le culline, una famiglia di 7 proteine conosciute come ubiquitina-proteina ligasi, che legano molecole di ubiquitina alla proteina bersaglio, che viene così riconosciuta dal proteasoma e degradata.

La comprensione delle basi molecolari del riconoscimento tra il dominio POZ/BTB e le culline è importante per la progettazione di nuove molecole utili per applicazioni biotecnologiche.

KCTD11 è un oncosoppressore antagonista della rete di segnali denominati Hedgehog (Hh) ed è coinvolto nella deacetilazione e attivazione di Gli1 e Gli2 [12], due fattori di trascrizione chiave nella via metabolica di Hh. KCTD11 forma un complesso multiproteico con la Cullina3 (Cul3) e HDAC mediando la deacetilazione di Gli1 nella rete di segnali di Hh.

Recentemente il gruppo della dott.ssa Pedone dell'IBB-CNR di Napoli, ha proposto un modello di interazione tra KCTD11 e Cul3 derivato dalla struttura cristallografica di Skp1/Cul1, omologhi rispettivamente di KCTD11 e Cul3. In questo modello ogni subunità del dominio POZ/BTB del tetramero di KCTD11 interagisce con una molecola di Cul3. In particolare, l' $\alpha$ -elica della regione 49-68 di Cul3 si posiziona all'interfaccia tra le varie subunità del dominio POZ/BTB. Inoltre, è stato proposto che i residui Tyr58, Tyr62 e Tyr125 di Cul3 e il residuo Phe102 di KCTD11 sono fondamentali per l'interazione (Figura 2).

## RISULTATI OTTENUTI

Sulla base delle informazioni strutturali a disposizione, un approccio simile è stato utilizzato per entrambe i sistemi studiati. Sono stati infatti progettati e sintetizzati peptidi che, conservando i residui fondamentali per l'interazione, hanno una buona propensione ad assumere una conformazione elicoidale, necessaria per il riconoscimento dei bersagli proteici.

### Progettazione, Sintesi, Caratterizzazione strutturale e Attività in vitro ed in vivo di peptidi analoghi di p53 TAD2

Uno degli obiettivi della presente tesi è stato quello di ottenere un potenziale dominio di transattivazione artificiale, partendo dai dati strutturali relativi al complesso tra p53 TAD2 e tfb1 di TFIIH.

L'analisi di tale struttura ha evidenziato il ruolo chiave per l'interazione dei residui Ile50, Glu51, Trp53 e Phe54 situati nella regione elicoidale 45-57 di p53 TAD2.

Sono state progettate due generazioni di peptidi, entrambe contenenti la sequenza 45-57 di p53.

Nella prima generazione di peptidi mimetici di p53 TAD2 sono stati conservati i residui fondamentali per il riconoscimento di TFIIH e sono state aggiunte opportune sequenze denominate N- e C-*capping*. Queste sequenze si ritrovano all'inizio e alla fine di strutture elicoidali, dove formano legami a idrogeno e interazioni idrofobiche specifiche che aumentano la propensione alla formazione di un'  $\alpha$ -elica della sequenza proteica. Inoltre, nei peptidi le estremità N- e C-terminali sono state rispettivamente acetilate e ammidate per eliminare le cariche presenti che possono essere cause di destabilizzazione dell'elica.

In particolare, il peptide p53-13 contiene la sequenza naturale 45-57 della proteina p53 con le due estremità acetilate e ammidate.

Il peptide NC17 è formato da 17 amminoacidi, conserva i residui fondamentali per il riconoscimento di TFIIH e contiene un residuo di Glu al posto della Gln52 in modo da formare un ponte salino  $i, i+4$  con la Lys all'estremità C-terminale. I residui dei motivi N-*capping* e di Schellman al C-terminale inseriti nella sequenza sono stati scelti in base alla loro frequenza statistica in quelle posizioni.

Il peptide NC15 contiene 15 residui, lo stesso motivo N-*capping* di NC17 e in questo caso i residui di Trp53 e Phe54 fanno parte del motivo di Schellman per la stabilizzazione della conformazione elicoidale.

Tutti i peptidi sono stati sintetizzati in fase solida utilizzando la chimica Fmoc. I prodotti ottenuti sono stati purificati mediante RP-HPLC preparativo. La purezza e l'identità dei peptidi sono state verificate mediante LC-MS.

Peptidi	Sequenze
p53-13	Ac- <sup>45</sup> Leu-Ser-Pro-Asp-Asp- <sup>50</sup> Ile-Glu-Gln-Trp-Phe- <sup>55</sup> Thr-Glu-Asp-NH <sub>2</sub>
NC17	Ac-Leu-Thr-Pro-Asp-Glu-Phe-Ile-Glu-Glu-Trp-Phe-Leu-Lys-Asp-His-Gly-Ile-NH <sub>2</sub>
NC15	Ac-Leu-Thr-Pro-Asp-Glu-Phe-Ile-Glu-Gln-Trp-Phe-Lys-His-Gly-Ile-NH <sub>2</sub>

**Tabella 1. Sequenze amminoacidiche dei peptidi mimetici di p53 TAD2 della prima generazione. In rosso sono evidenziati i residui fondamentali per l'interazione, in blu i residui appartenenti al motivo N-capping mentre in verde i residui del motivo di Schellman.**

Al fine di studiare le proprietà conformazionali dei peptidi sintetizzati sono state eseguite misure di dicroismo circolare. Lo spettro CD di p53-13 è caratterizzato da un profondo minimo a 200nm, caratteristico dell'assenza di struttura secondaria definita.

Lo spettro CD del peptide NC15 presenta due minimi a 206 e 220 nm, un crossover ( $\lambda_0$ ) a 199nm ed un massimo intenso a 193 nm, indice di un buon contenuto di struttura elicoidale. Lo spettro CD del peptide NC17 presenta invece due minimi a 202 e 221 nm, non presenta crossover, indicando uno scarso contenuto di struttura elicoidale.

Per valutare l'affinità di legame per il dominio PH di tfb1 dei peptidi sintetizzati sono stati effettuati esperimenti ITC in collaborazione con il gruppo del professore Omichinski dell'Università di Montréal in Canada. I risultati ottenuti hanno mostrato che il peptide p53-13 ha una costante di dissociazione apparente ( $K_d$ ) di  $1,6 \pm 0,3 \mu\text{M}$  per tfb1, mentre NC15 e NC17 non sono stati in grado di legare tfb1.

Questi dati sono stati confermati anche da esperimenti di etero-correlazione <sup>1</sup>H-<sup>15</sup>N NMR. La titolazione di tfb1 con il peptide p53-13 ha evidenziato che quest'ultimo compete con lo stesso sito di legame di p53 TAD2.

Studi analoghi condotti sui peptidi NC15 e NC17 non hanno evidenziato significativi spostamenti dei segnali di tfb1.

La seconda generazione di peptidi è stata progettata sulla base dei dati riportati in letteratura sul peptide mimetico del fattore VEGF (QK). In queste sequenze sono stati inseriti un motivo N-capping e due residui di Leu in posizione  $i, i+3$ .

I peptidi differiscono tra loro per alcune modifiche apportate all'interno del motivo N-capping. In particolare, nella sequenza di E-Cap il Glu in posizione 3 è stato scelto in base a dati statistici relativi alla probabilità di trovare il residuo in tale posizione; in W-Cap il Trp in posizione 3 è stato scelto per omologia con il peptide QK; mentre in A-Cap l'Ala in posizione 3 è stata scelta per analogia con VP16C, un fattore di trascrizione, capace anch'esso di legarsi a tfb1.

I residui di leucina sono stati inseriti nelle posizioni 6 e 9, rispettivamente in  $i$  e  $i+3$  dell'elica, in quanto questi residui contribuiscono alla stabilizzazione della conformazione elicoidale attraverso interazioni idrofobiche e si trovano sul lato dell'elica opposto a quello dell'interazione con tfb1.

I peptidi sono stati sintetizzati in fase solida utilizzando la chimica Fmoc ed i prodotti ottenuti sono stati purificati mediante RP-HPLC preparativo. La purezza e l'identità dei peptidi sono state verificate mediante LC-MS.

Peptidi	Sequenze
E-Cap	Ac-Leu-Thr-Glu-Glu-Glu-Leu-Ile-Glu-Leu-Trp-Phe-Thr-Glu-NH <sub>2</sub>
A-Cap	Ac-Leu-Thr-Ala-Glu-Glu-Leu-Ile-Glu-Leu-Trp-Phe-Thr-Glu-NH <sub>2</sub>
W-Cap	Ac-Leu-Thr-Trp-Glu-Glu-Leu-Ile-Glu-Leu-Trp-Phe-Thr-Glu-NH <sub>2</sub>

**Tabella 2. Sequenze amminoacidiche dei peptidi mimetici di p53 TAD2 della seconda generazione.**

**In rosso sono evidenziati i residui fondamentali per l'interazione, in verde i due residui di Leu in posizione i, i+3, in blu le modifiche apportate al motivo N-capping**

Le sequenze ottenute sono state analizzate mediante tecniche di dicroismo circolare per valutarne il contenuto elicoidale. Gli spettri CD presentano due minimi a 222nm e 207nm, un crossover tra 193 e 197 nm e un massimo a ~190nm, indice di un contenuto elicoidale simile, maggiore rispetto al peptide p53-13 ma inferiore a quella di NC15.

Analisi di interazione mediante ITC hanno mostrato che E-Cap, W-Cap e A-Cap sono in grado di legarsi a tfb1, con una K<sub>d</sub> rispettivamente di 0.24 ± 0.03 μM, 1.3 ± 0.1 μM e 1.9 ± 0.3 μM.

Esperimenti di etero-correlazione HSQC <sup>1</sup>H-<sup>15</sup>N NMR hanno evidenziato che tutti i peptidi della nuova generazione provocano un cambio dei segnali di tfb1 nella regione localizzata tra i foglietti β5, β6 e β7 e il *loop* tra β5 e β6. Inoltre sono stati condotti degli studi su E-Cap per comprendere come la presenza delle due leucine influisce sulla capacità di legame.

La struttura di E-Cap in complesso con tfb1 conferma che i residui Leu6 e Leu9 formano un ponte idrofobico, sul lato opposto dell'elica rispetto a quello dell'interazione. Inoltre, la Leu6 stabilizza ulteriormente la struttura elicoidale di E-Cap formando dei contatti con l'anello aromatico del Trp10 in posizione *i+4*.

La potenziale funzione di E-Cap come attivatore della trascrizione *in vivo* è stata analizzata esprimendo in cellule di lievito E-Cap legato a LexA, un dominio di legame al DNA specifico per il gene reporter *lacZ*. Oltre al controllo positivo Gal4, la capacità di attivatore della trascrizione di E-Cap è stata paragonata a quella di altri attivatori AH e VP2 e alla sequenza p53-13 anch'essi co-espressi con LexA.

E-Cap ha mostrato una capacità di attivazione della trascrizione maggiore rispetto alle altre sequenze, confermando che stabilizzando la conformazione elicoidale di piccoli analoghi di p53 è possibile creare dei potenti domini di transattivazione artificiali.

Tali peptidi saranno ulteriormente ottimizzati per lo sviluppo di attivatori artificiali di trascrizione con elevate potenzialità applicative nel campo delle biotecnologie.

### Progettazione, Sintesi, Caratterizzazione strutturale e Saggi ELISA di peptidi analoghi di Cul3

Lo sviluppo di mimetici di Cul3 specifici per l'interazione con il dominio POZ/BTB della regione N-terminale di KCTD11 sono stati condotti in collaborazione con il gruppo della dr.ssa Pedone e della dr.ssa Di Gaetano presso l'IBB-CNR di Napoli.

Precedenti studi di mutagenesi e modellistica molecolare hanno evidenziato che la regione che media il riconoscimento tra Cul3 e KCTD11 presenta dei residui

aromatici fondamentali per l'interazione (Tyr58, Tyr62 e Tyr125 di Cul3 e Phe102 e Tyr102 di KCTD11) [13].

Sulla base di questi studi un altro obiettivo del presente lavoro di ricerca è stato quello di progettare analoghi di Cul3 in grado di legare il dominio POZ/BTB di KCTD11, allo scopo di contribuire alla comprensione dei determinanti strutturali fondamentali per l'interazione tra Cul3 e la famiglia di proteine KCTD.

Inizialmente, è stato progettato il peptide Cul3-wt corrispondente alla regione 49-68 di Cul3. Successivamente sono stati progettati due analoghi, Cul3-KK e Cul3-Y(P), in cui la Tyr58 e la Tyr62 sono state sostituite rispettivamente con due residui di Lys e un residuo di Tyr fosforilata, in modo da verificare l'importanza nell'interazione di questi residui.

Peptidi	Sequenze
Cul3-wt	Ac- <sup>49</sup> Asn-Ser-Gly-Leu-Ser-Phe- <sup>55</sup> Glu-Glu-Leu-Tyr-Arg- <sup>60</sup> Asn-Ala-Tyr-Thr-Met- <sup>65</sup> Val-Leu-His-Lys-NH <sub>2</sub>
Cul3-KK	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-Leu-Lys-Arg-Asn-Ala-Lys-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>
Cul3-Y(P)	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-Leu-Tyr(P)-Arg-Asn-Ala-Tyr-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>

**Tabella 3. Sequenze amminoacidiche dei peptidi Cul3-wt, Cul3-KK e Cul3-Y(P). In rosso sono evidenziate le Tyr fondamentali per l'interazione, in verde le Lys che sostituiscono le Tyr e in blu la Tyr fosforilata.**

Tutti i peptidi sono stati acetilati al N-terminale e amidati al C-terminale in modo da eliminare le cariche presenti che possono destabilizzare il dipolo dell'elica. Inoltre sono stati sintetizzati anche i derivati biotinilati da utilizzare per gli studi di interazione *in vitro*, mediante saggi ELISA.

I peptidi sono stati sintetizzati in fase solida utilizzando la chimica Fmoc. I prodotti ottenuti sono stati purificati mediante RP-HPLC preparativo. La purezza e l'identità dei composti sono state verificate mediante LC-MS.

L'analisi conformazionale mediante dicroismo circolare ha evidenziato un basso contenuto di struttura elicoidale sia per Cul3-wt che per Cul3-KK. Gli spettri CD di entrambi i peptidi mostrano, infatti, un minimo intorno a 222nm e un profondo minimo a 202-203 nm, un cross-over a 192nm e un massimo a 190nm.

Studi preliminari di interazione con KCTD11 sono effettuati mediante saggi ELISA per confermare la capacità di Cul3wt di interagire e valutare se Cul3-KK e Cul3-Y(P), si legano con minore affinità. L'andamento delle curve risultanti dagli esperimenti indica che Cul3-wt si lega al dominio POZ/BTB di KCTD11 con una discreta affinità, mentre i due mutanti si legano con un'affinità inferiore.

Come nel caso degli analoghi di p53 TDA2, è stato progettato un peptide che presenta un maggiore contenuto elicoidale per valutare l'effetto sull'affinità per KCTD11. In particolare alla sequenza nativa di Cul3 sono stati sostituiti i residui Leu57 e Ala61 con due  $\alpha,\alpha$ -amminoacidi,  $\alpha$ -(4') pentenil- alanina (S5). Questi residui permettono di creare un ponte idrocarburico mediante reazione di metatesi (RCM) tra i residui  $i, i+4$  della sequenza, stabilizzando il peptide in conformazione elicoidale.

Peptidi	Sequenze
Cul3-stapline	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-S5-Tyr-Arg-Asn-S5-Tyr-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>
Cul3-stapled	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-S5-Tyr-Arg-Asn-S5-Tyr-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>

**Tabella 4. Sequenze amminoacidiche dei peptidi Cul3-stapline e Cul3-stapled. In rosso sono evidenziati i residui necessari per RCM**

I peptidi sono stati acetilati al N-terminale e amidati al C-terminale in modo da eliminare le cariche presenti su di esse che possono essere la causa di destabilizzazione dell'elica. Anche in questo caso dei peptidi sono stati sintetizzati anche i derivati biotinilati da utilizzare per gli studi di interazione *in vitro*, mediante saggi ELISA.

I peptidi sono stati sintetizzati in fase solida utilizzando la chimica Fmoc. Anche la reazione di metatesi è stata condotta in fase solida e i prodotti ottenuti sono stati purificati mediante RP-HPLC preparativo. La purezza e l'identità dei composti sono stati verificati mediante LC-MS.

Lo studio conformazionale mediante CD ha evidenziato che Cul3-stapled ed il suo analogo lineare Cul3-stapline hanno un buon contenuto elicoidale, simile tra loro e sicuramente maggiore rispetto a Cul3-wt.

L'analisi preliminare dei risultati ottenuti dai saggi ELISA ha confermato che entrambi i peptidi sono in grado di legarsi al dominio POZ/BTB di KCTD11 con un'affinità maggiore rispetto a Cul3-wt.

Studi futuri mediante tecniche ITC ed NMR saranno effettuati per determinare la costante di affinità e per comprendere le basi molecolari dell'interazione con il dominio POZ/BTB di KCTD11 e di altri membri della famiglia di proteine KCTD. Recenti dati di letteratura indicano infatti che molti domini POZ/BTB di questa classe di proteine mediano l'interazione con Cul3. Pertanto, lo sviluppo di molecole in grado di legare questi sistemi proteici coinvolti in numerose patologie umane può essere un punto di partenza per lo sviluppo di interessanti agenti terapeutici.



## 1. INTRODUCTION

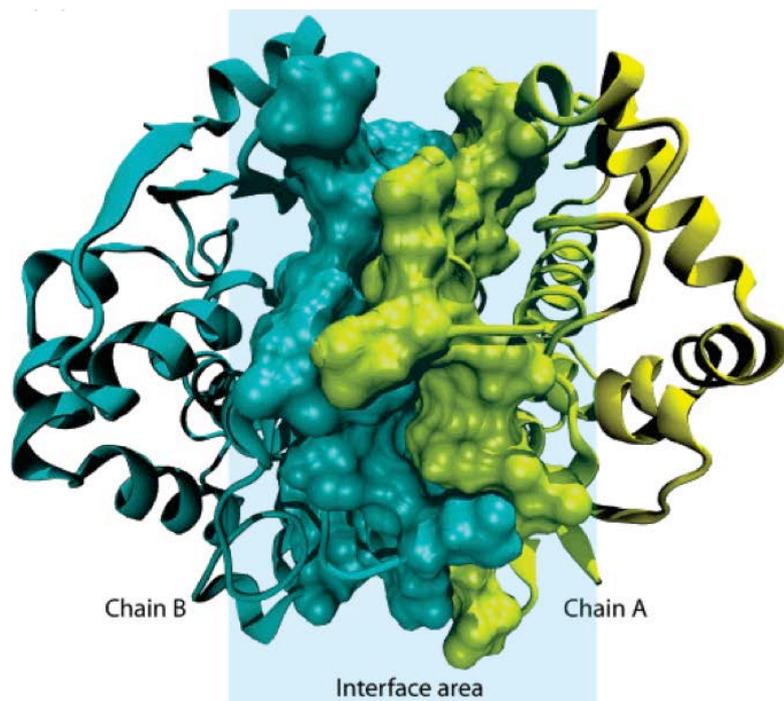
Sequencing of the human genome has greatly expanded the knowledge of the primary structures of proteins.

Proteins play a central role in key biological signaling processes and these events require protein interaction networks.

The study of the protein-protein interactions (PPI) is fundamental to correlate structure and function in biomolecular systems. The understanding of molecular features that regulate these interactions is the starting point for the design of specific ligands able to modulate main cellular activities allowing the development of new systems for biotechnology applications.

### 1.1 Protein-Protein interactions: Structural Basis

In PPI independently folded domains of proteins form an interface in which residues located on different protein chains take place, giving rise to interaction surfaces (Figure 1) [1].



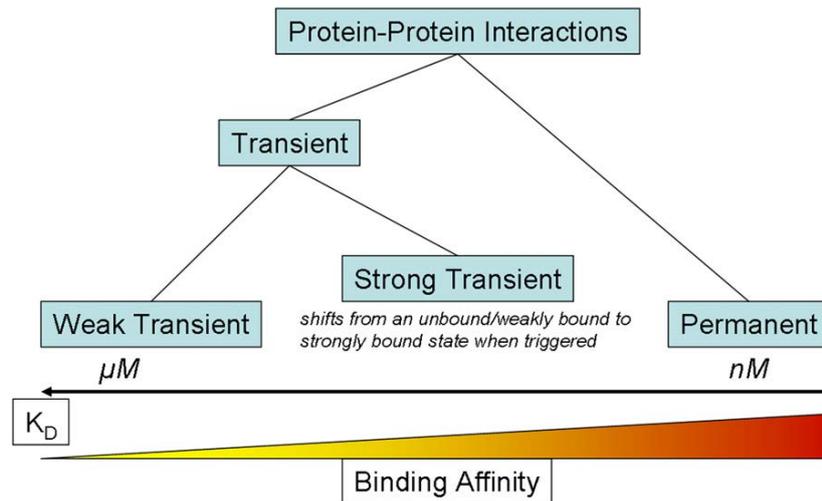
**Figure 1.** The figure represents two interacting proteins (human glutathione S-transferase, PDB ID: 10GS). [Keskin, O., et al., J Mol Biol, 2005]

Some proteins exert their functions just in complexed form, so that PPI in this case can be defined as "obligate" or "permanent" [2]. Generally these kinds of protein complexes represent important cellular machinery such as proteasome, ribosome, DNA replication and transcription, ATP synthase and many others [3]. The interactions between different subunits of these complexes are very strong showing binding affinity ( $K_d$ ) values in nM range [2].

Many others proteins interact "transiently" since the interaction is broken and formed continuously and the complex readily undergoes changes in the oligomeric state. In

this case the binding affinity between different proteins is weak with  $K_d$  values typically in  $\mu M$  range.

Thornton and co-workers [4] [5] have highlighted that it is possible to identify different protein transient associations: weak transient complexes that show a dynamic mixture of different oligomeric states and strong transient complexes where proteins interact only after an activation signal. However, a *continuum* between transient and obligate interaction occurs and the stability of all complexes depends on the physiological conditions and cellular environment.



**Figure 2. Different Types of Protein-Protein Interactions on the Basis of Their Binding Affinities [Perkins, J.R., et al., Structure, 2010]**

The analysis of interacting interfaces that determine the stability of the complex such as size, shape and surface complementarity is required for the knowledge of molecular properties that regulate PPI [4], [5].

Ramos and co-workers [6] have described that the standard-size of PPI interfaces is  $1200\text{--}2000 \text{ \AA}^2$ , but we can distinguish different size if binding affinity is more or less strong. Large size interfaces in the range  $2000\text{--}4660 \text{ \AA}^2$  occur mostly between proteases and a particular class of inhibitors, G-proteins and other components of the signal transduction system [6].

Inside the surface area different kinds of forces contribute to the binding free energy of PPI.

Hydrophobicity is a leading force in PPI and hydrophobic interactions in proteins occur between amino acid residues of nonpolar regions through van der Waals contacts [5] [6].

Electrostatic force is a further force involved in PPI and electrostatic complementarity of interacting protein surfaces promotes complex formation and defines the lifetime of complexes [6].

Hydrogen bonds are formed between protein contact surfaces and the surrounding water molecules [5] and their average number is proportional to the area of subunit surfaces.

Summing the free energy gain produced by all the individual van der Waals, electrostatics interactions and by desolvation process, the total gain produces a high stabilization of the protein–protein complex.

It is observed that there is a higher conservation of residues in interfaces than in the rest of protein and some of these contribute dominantly to the binding free energy.

The average percentage of aromatic residues in interfaces demonstrates their importance in PPI.

Tryptophan is a hydrogen-bonding donor and contributes with its indole ring to aromatic  $\pi$ -interactions. Moreover, it has a large hydrophobic surface and can protect fragile hydrogen bonds from water. Its mutation to an alanine generates a large cavity, due to the significant difference in sizes, which can create a highly complex destabilization.

Tyrosine has both the phenolic ring that can participate to aromatic  $\pi$ -interactions and the 4-hydroxyl group that make hydrogen bonds.

Other amino acids are most frequently found in interfaces of PPI. Arginine is capable of multiple types of favorable interactions by its guanidinium group such as hydrogen bond and salt-bridge. Analysis of various complexes shows that aspartate and asparagine are favored over glutamate and glutamine and isoleucine is more frequent than leucine [6].

Systematic substitutions of these residues in the interfaces by alanine (Ala scanning) and the measurement of the difference in binding free energy between the wild type protein and the mutants, have highlighted a substantial decrease. The sites where alanine mutations cause a significant drop in the binding free energy of at least 2.0 kcal/mol are defined as "hot spots" [7].

Generally all residues can contribute indirectly to PPI favoring the orientation of "hot spots" and/or contributing by electrostatic contacts [6].

"Hot spots" tend to organize packing tightly and to form networks of interactions that are called "hot regions" [9]. The hot regions of protein-protein interfaces are often organized as preferred secondary structures: helices,  $\beta$ -sheets and loops.

## **1.2 Helical structure in Protein Interface**

The helical structures represent one of most frequent secondary motifs found in proteins and play crucial roles in many PPI [10].

Most helices located within the protein core are vital for the overall stability of protein tertiary structure and some helices exposed on the surface of a protein are bioactive regions for the recognition of proteins, DNA and RNA [11]. Recent studies by Arora and co-workers highlighted that roughly 60% of the protein complexes found in the Protein Data Bank (PDB) display helical interface [12].

The  $\alpha$ -helix is characterized by a circular rotation of 3.6 residues *per* complete turn that places side-chains in  $i$ ,  $i+3$ ,  $i$ ,  $i+4$  and  $i$ ,  $i+7$  positions close in the space [10]. The carbonyl groups and amide protons in these positions, orientated in opposite directions, create hydrogen bonds that run in the same direction along the surface of the helix, thus stabilizing the structure. The residues spaced  $i$ ,  $i+2$ ,  $i$ ,  $i+5$  and  $i$ ,  $i+6$  place the side-chain pairs on opposite faces of the helix, avoiding any interaction in a monomeric helix.

Other kinds of helical structures are occasionally found in proteins and in naturally occurring peptides such as  $3_{10}$ - and  $\pi$ -helices.

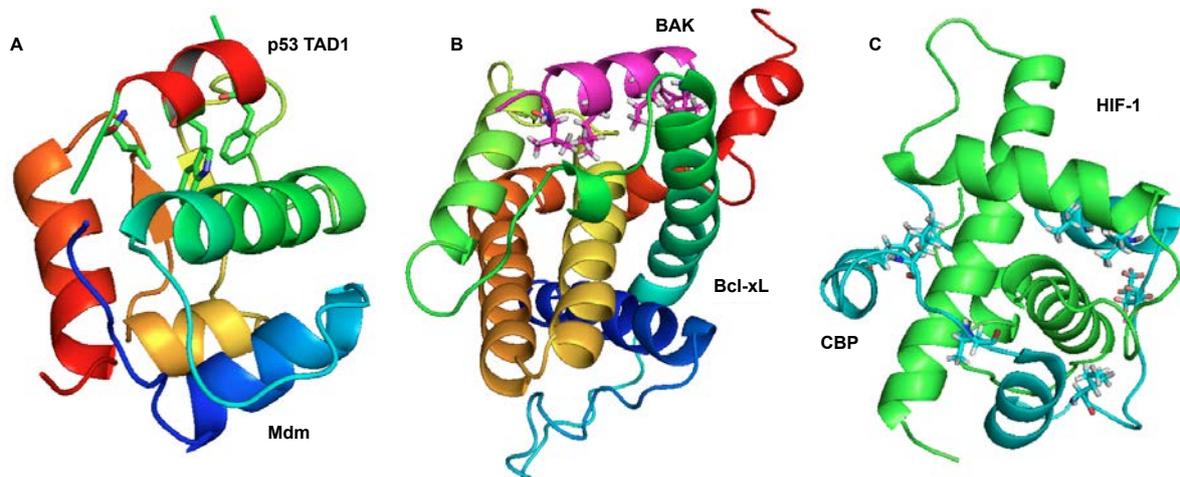
The  $3_{10}$ -helix occurs in about 3-4% of residues in protein crystal structures, usually at

the N- and C-terminal of  $\alpha$ -helices. It is characterized by more tightly wound than the  $\alpha$ -helix with hydrogen bonds between  $i, i+3$  residues. The resulting steric clash renders this conformation less favorable than the  $\alpha$ -helix. There is no energy barrier between  $3_{10}$ -helical and  $\alpha$ -helical structures, and the two forms can therefore easily interconvert, so that the  $3_{10}$ -helix is proposed to be an intermediate in the  $\alpha$ -helices formation [10].

The  $\pi$ -helix is extremely rare and energetically unfavorable respect with  $\alpha$ -helix. Its geometrical features enforce hydrogen bonds between  $i, i+5$  residues. Evidence for  $\pi$ -helices show that are more likely to be associated with protein function than structure of proteins [10], [11].

Several examples of helical surface that plays a major role in mediating a PPI are reported in the literature.

The interaction between trans activation domain (TAD) of p53 and Mdm2 is one of most studied case of PPI involving an  $\alpha$ -helix (Figure 3A) [3]. The anti-apoptotic proteins Bcl-x<sub>L</sub> and Mcl-1 that exert their effect sequestering pro-apoptotic proteins (such as BAK) through the interaction with key residues exposed from an  $\alpha$ -helix surface [3] (Figure 3B) and the hypoxia inducible transcription factor (HIF-1) that interacts through its three distinct  $\alpha$ -helical motifs with coactivator CBP (Figure 3C), represent other examples [3].



**Figure 3. A.** In p53/Mdm2 interaction three residues from a helix (in red) in the p53 activation domain reside in a Mdm2 deep hydrophobic groove (PDB code: 1YCR) **B.** The interaction between Bcl-x<sub>L</sub> and BAK involves four key hydrophobic residues of BAK (purple) binding within a hydrophobic cleft on Bcl-xL (PDB code: 1BXL) **C.** HIF-1 (green) binds to CBP (cyan) through three distinct  $\alpha$ -helical motifs (PDB code: 1L8C).

### **1.3 Helical structure mimetics: Peptides and Peptidomimetics**

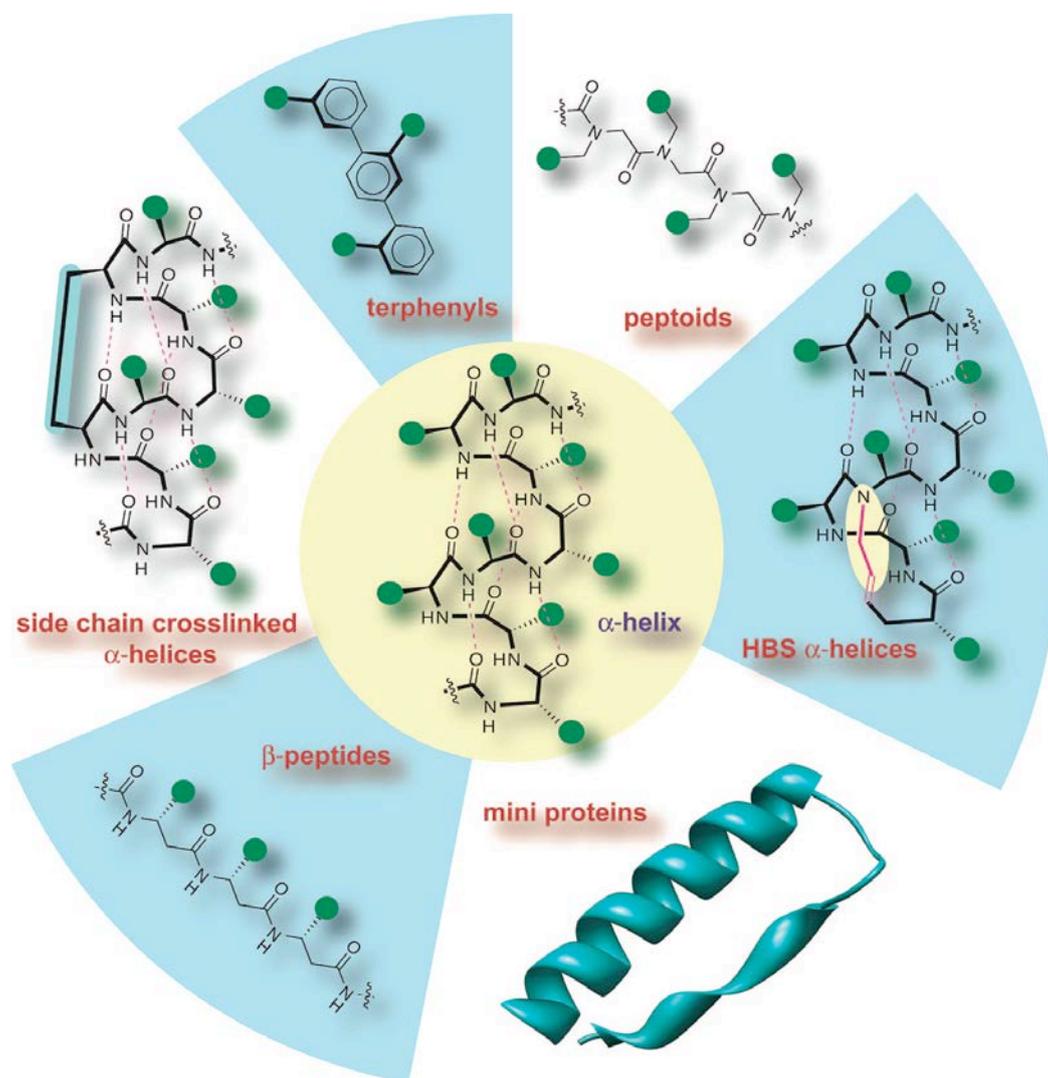
The inhibition of molecular interactions as PPI is a promising strategy for the controlled interference of biological functions. This approach requires the availability of molecules or library of molecules capable of conformationally mimicking defined surface of binding preserving the right arrangement of functional sites.

It is possible to recognize two structural types of protein binding sites: *sequentially continuous binding sites* that are located in single, defined stretches of the protein sequence and *discontinuous binding sites* where the amino acids are not located in continuous stretches of the sequence but are placed into spatial proximity by protein folding [13].

As previously reported,  $\alpha$ -helices constitute the largest class of protein secondary structures and play a considerable role in mediating protein–protein interactions. Recent analysis by Arora and co-workers of the PDB highlight that 62% of the resolved protein complexes are featured by helical interfaces and that the average length of helical domains in interacting interfaces is rather small and spans two to three helical turns (or eight to twelve residues) [14]. These complexes suggest that it may be possible to develop small molecules that potentially interfere with the interactions between biomolecules. Peptides are adequate molecules to mimic protein-binding site, enabling the detailed analysis of the interaction at the level of individual amino acid residues [15]. They are able to reproduce protein secondary structures of selected interface with high affinity and specificity. Moreover, since recent progresses in chemical synthesis methods it is possible to introduce different building blocks and modification at peptide backbone. Unfortunately, peptides composed of less than fifteen amino acid residues rarely retain their conformation once excised from the protein and the ability to bind their specific targets [16, 17].

Stabilization of peptides in the helical structure should reduce their conformational heterogeneity and increase their resistance to proteases as these enzymes typically bind their substrates in the extended conformation. The proteolytic stability of the helix should thus be directly proportional to its conformational stability [17].

The scientific community focuses much of its attention on studying different approaches to either stabilize the  $\alpha$ -helical conformation in peptides and mimic this motif with natural and non-natural scaffolds. These approaches can be summarized into three general categories: helix stabilization, helical foldamers and helical surface mimetics.



**Figure 4. Strategies that stabilize the  $\alpha$ -helical conformation in peptides or mimic this domain with non-natural scaffolds [Patgiri, A., et al., Acc. Chem. Res., 2008]**

Since the discovery of  $\alpha$ -helix structure in proteins, several studies have provided a set of rules for helix stabilization introducing several molecular tools such as amino acids residues with large helix forming propensities, aliphatic and aromatic residues that allow hydrophobic van der Waals interaction, polar and charged amino acids that promote stabilizing hydrogen bonds, amino acid residues able to stabilize aromatic-charge or aromatic-sulfur interactions, N- and C-capping groups that prevent unfavorable charge interactions with the helix dipole [18].

The spatial proximity of residues placed at positions  $i$ ,  $i+4$  in an  $\alpha$ -helix is further exploited to stabilize helices by the incorporation of metal ligands at these positions introducing natural or diacetic-based unnatural amino acids [18].

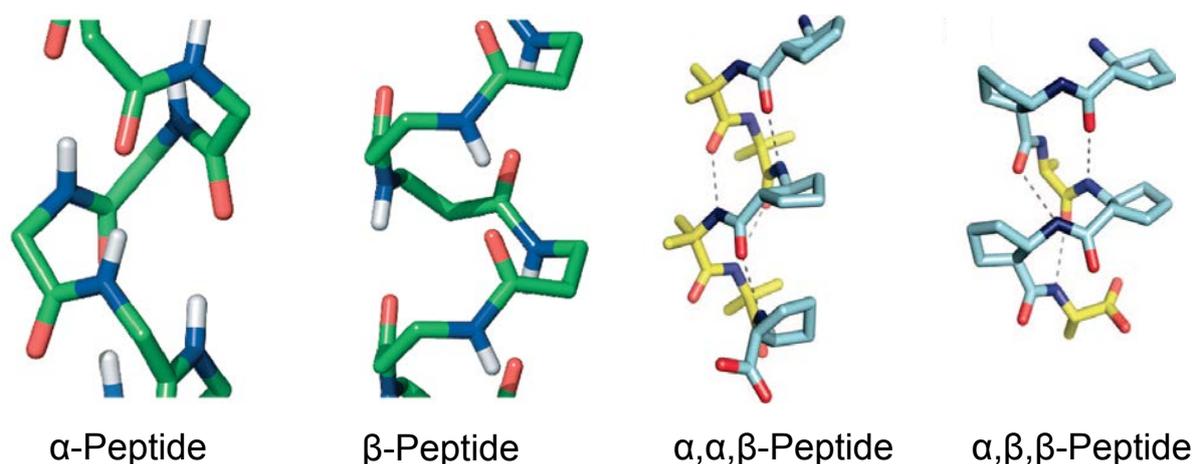
Along with the helix-coil transition theory in peptides, another approach for helix-stabilization consists of introducing constraints or non natural amino acids to form covalent bonds between the  $i$ ,  $i+4$  and  $i$ ,  $i+7$  side chain groups. Lactam [19] and disulfide bridges [20] have been the earliest side chain crosslinks used.

Recently, further covalent helix stabilization is achieved by Grubbs and co-workers using ruthenium-catalyzed ring-closing metathesis (RCM) of O-allyl serine residues

to introduce olefinic crosslinks between  $i$ ,  $i+4$  and  $i$ ,  $i+7$  positions in a helical sequence [21]. Likewise, Schafmeister and Verdine have described hydrocarbon bridged helices obtained by olefin metathesis reaction [22]. The hydrocarbon linker is employed because it is expected to be chemically more stable than linkers built from amide or disulfide bonds. These hydrocarbon-stapled peptides also show an increased tendency to penetrate the cell membrane likely for the lipophilic nature of the linker [23].

In the last years click chemistry is emerged as a good synthetic tool to insert 1,4-disubstituted triazole linkage as isostere of the amide bond in peptide structure [24]. Recently, Dawson and co workers have described an efficient method for side-chain cyclization of peptides by click chemistry, introducing cross-link between  $i$ ,  $i+3$  side-chain groups and favoring the formation of an  $\alpha$ -helix structure [25].

Helical foldamers, such as  $\beta$ -peptides and peptoids (Figure 4), are composed of amino acid analogues and are capable of adopting conformations similar to those found in natural proteins [17]. Foldamer design, based on heterogeneous backbones, is more advantageous relative to homogenous backbones (Figure 5) [26]. The number of candidate backbones is indeed larger if heterogeneous backbones for a given set of monomer classes are included. For example if only  $\alpha$ - and  $\beta$ -amino acids as building blocks are considered, the homogeneous approach limits the use to  $\alpha$ -peptides or  $\beta$ -peptides while the heterogeneous approach allows many different combinations (such as  $\alpha$ - $\beta$ - $\alpha$ - $\beta$ - $\alpha$ - $\beta$ -,  $\alpha$ - $\alpha$ - $\beta$ - $\alpha$ - $\alpha$ - $\beta$ -,  $\alpha$ - $\beta$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$ -,  $\alpha$ - $\alpha$ - $\beta$ - $\beta$ -, to name just a few). Each of these heterogeneous backbones offers a potentially distinctive way to project sets of side chains in the space. By analogy to proteins, a specific three-dimensional arrangement of these side-chains allows more distinct shapes that can be generated with foldamers.



**Figure 5. Helical Foldamers: The heterogeneous approach [Horne, W.S. and S.H. Gellman, Acc. Chem. Res., 2008]**

Many helix structures consisting of  $\beta$ -peptides are also reported. These conformation types are distinguished by the number of atoms in the hydrogen-bonded ring that is formed in solution: 8-helix, 10-helix, 12-helix, 14-helix, and 10/12-helix. Generally  $\beta$ -peptides form a more stable helix than  $\alpha$ -peptides [27].

Finally, the helical surface mimetics utilize conformationally restricted scaffolds with attached functional groups that resemble the  $i$ ,  $i+4$  and  $i$ ,  $i+7$  pattern of side chain positioning along the face of an  $\alpha$ -helix [12]. These mimetics generally derive from

small building blocks. For example Hamilton and coworkers have introduced 3,2',2'-functionalized terphenyl derivatives as structural and functional mimetics of  $\alpha$ -helices [28] acting as potent antagonists of calmodulin and inhibitors of the assembly of the hexameric HIV fusion protein gp41 [29]. More recently they have designed a new  $\alpha$ -helix scaffold that mimic the  $\alpha$ -helical BH3 domain of Bak. This moiety consists of a trispyridylamide scaffold that exhibits a preferred conformation with all O-substituents located on the same side of the molecule mimicking the  $i$ ,  $i+4$  and  $i$ ,  $i+7$  residues in the  $\alpha$ -helix.

## 1.4 *p53* Protein

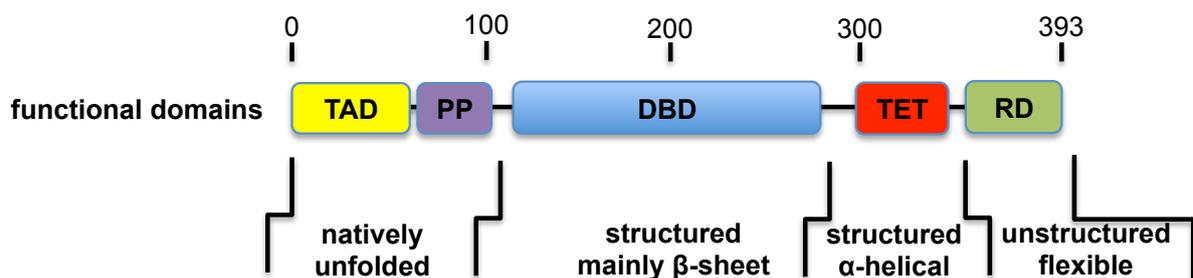
### 1.4.1 *p53*: a cancer suppressor

The *p53* gene was the first tumor-suppressor gene to have been identified. It expresses for a nuclear protein of 393 amino acids, a key regulator protein for cell cycle regulation, apoptosis and DNA repair [30].

Normally, *p53* protein cellular levels are very low because it has a short half-life. Its localization and degradation are mediated by *Mdm2* that binds *p53* protein blocking its transcriptional action and tagging for ubiquitin-dependent proteolysis degradation. *Mdm2* is also the main transcriptional target for *p53* protein, directly activating expression of its own negative regulator and producing a potent negative feedback regulatory loop. It is worth to note that the regulation of *p53* occurs post-translationally, since such kind of regulation would appear to be optimal in a stress response. *p53* protein levels have been shown to dramatically rise in response to different stresses, including ionizing and UV radiations, growth factor deprivation, treatment with chemotherapeutic drugs, oncogene activation, heat shock or virus infections. *p53* protein expression has also been shown to be induced following hypoxia in the early stages of tumorigenesis, explaining what *p53* loss may be selected for. The primary phenotypic consequences of *p53* induction are cell cycle arrest or apoptotic cell death. Alterations in *p53* protein expression or in its regulated signals cascade are the main causes of ~50% human tumors, therefore this protein and its network have been subject of many studies to design and synthesize ligands that can modulate its functions [31].

Active form of *p53* protein is a phosphorylated tetramer polypeptide that works as a potent transcriptional activator [32], [33].

In the structure of protein it is possible to recognize 5 different functional domains (Figure 6):



**Figure 6. Domain organization of *p53*:** The amino-terminal transactivation domain (TAD) is natively unfolded and interacts with components of the transcriptional machinery and includes the *Mdm2* and TFIID binding site. The central DNA binding “core” domain (DBD) holds 95.1 % of the oncogenetic mutations. The carboxy-terminal region of *p53* consists of a tetramerization domain (TET) and a regulatory domain (RD). [Romer L., et al, *Angew Chem Int Ed*, 2006]

(i) the transactivation acidic domain (TAD) at N-terminal region that encodes the transactivation function. This region is relatively acid and has been shown to interact with components of the transcriptional machinery, such as TATA-binding protein (TBP), TBP-associated factors (TAFs) and the co-activator proteins p300/CBP. The *Mdm2* oncoprotein has been shown also to bind *p53* in the N-terminal region where it negatively regulates *p53* transactivation function;

(ii) a proline-rich domain located adjacent to the TAD domain that is functionally important for p53 tumor suppressor activity;

(iii) a central DNA binding domain (DBD). The main p53 mutations are found within this region and disrupt DNA binding by p53;

(iv-v) the C-terminal region of p53 has a relatively basic charge and contains sequence-nonspecific DNA binding activity. A discrete region within this C-terminal domain regulates oligomerization of p53. The C-terminal domain also contains several nuclear localization signals as well as a nuclear export signal that is thought to regulate p53 subcellular localization.

#### **1.4.2 p53-TFIIH interaction**

The N-terminal region TAD of p53 protein has been extensively studied because it is associated to the transcriptional activation function of the protein [34]. This region is rich of acidic residues and can be divided in two subdomains, TAD1 (residues 1-40), and TAD2 (residues 40-83) that show independent transcriptional activation functions [35].

TAD1 and TAD2 functionally interact with a variety of transcriptional regulatory proteins. TAD1 specifically recognizes Mdm2, the TBP and CBP/p300, whereas TAD2 interacts with the replication protein A (RPA70) and others.

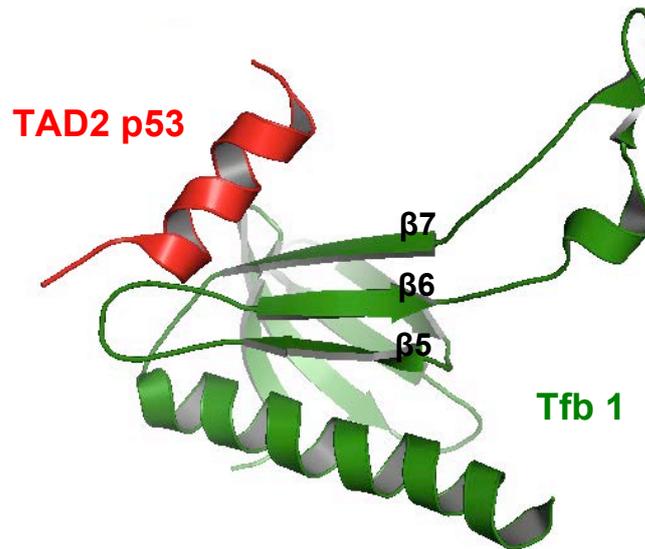
The transcription factor II H (TFIIH) is one of general transcription factors (GTF) known to be a target of TAD2 of p53. TFIIH is composed of ten subunits that can be divided into two groups, the core (XPB [Ssl2], p62 [Tfb1], p52 [Tfb2], p44 [Ssl1], p34 [Tfb4], and TTDA [Tfb5] in human and yeast respectively) and the CAK complex (cdk7 [Kin28], cyclin H [Ccl1], and MAT1 [Tfb3]) that are linked by XPD [Rad3] subunit. TFIIH is the only GTF to possess enzymatic activity and its helicase activities (XPB and XPD) are essential to the formation of the open complex during transcription initiation, whereas its kinase activity (cdk7) is required for the phosphorylation of the C-terminal domain of the RNA Pol II largest subunit. This phosphorylation enables RNA Pol II to progress from the initiation phase to the elongation phase of transcription [36].

The interaction between p62 subunit of TFIIH core (tfb1 in yeast) and TAD of p53 protein has been shown to be important for the recruitment of p53 to the TFIIH complex [35]. This interaction is directly correlated with the ability of the p53 TAD to stimulate transcriptional elongation. In addition, TFIIH plays a role in regulating the transcriptional activity of p53. The phosphorylation of the TAD of p53 at Ser33 by the cdk7 is required for p53 to induce apoptosis in response to UV and ionizing radiation damage.

In the recent years Omichinski and co-workers have been interested in the study of interaction of TAD of different transcription factors with p62/tfb1 subunit of TFIIH [35], [37], [38].

The structure of interaction between p53<sub>20-73</sub> TAD2 with tfb1 subunit of the general transcription factor TFIIH has been solved by NMR. Tfb1 subunit of TFIIH has been chosen because of longer stability respect with p62; moreover the authors have demonstrated that tfb1 and p62 interact with TAD2 in similar way.

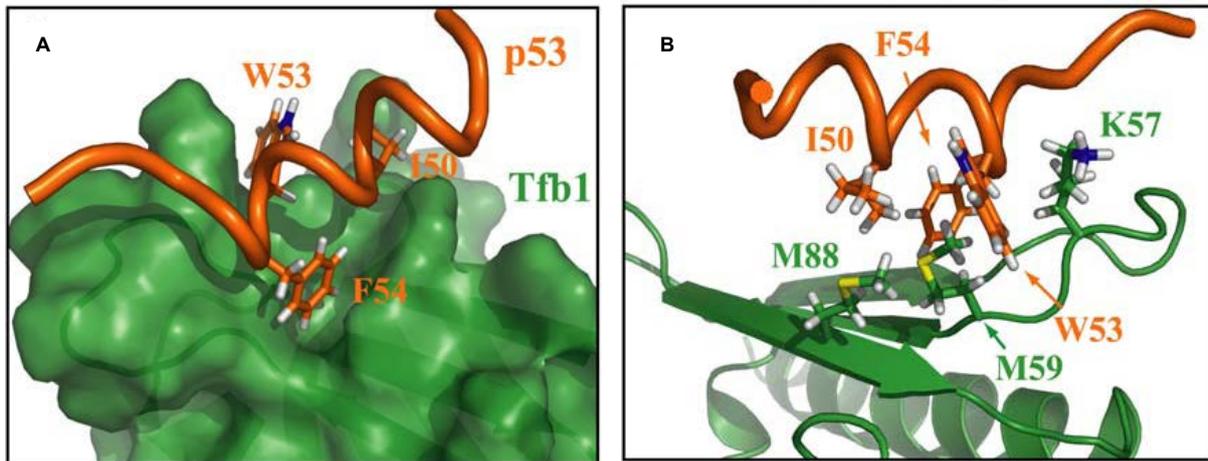
The structure of Tfb1 in the complex presents a typical PH fold (Figure 7), which consists of two perpendicular antiparallel  $\beta$ -sheets arranged in a  $\beta$ -sandwich and flanked on one side by a long  $\alpha$ -helix.



**Figure 7. Ribbon model for the lowest-energy conformer of the complex Tfb1 (in green)/p53 (in red) (PDB code: 2GS0). [Di Lello, P., et al. Mol Cell, 2006]**

The p53 TAD2 in the complex is unstructured except for a short amphipathic  $\alpha$ -helix involving residues 47–55. This short helix interacts with the surface of Tfb1 formed by  $\beta$ -strands  $\beta$ 5,  $\beta$ 6, and  $\beta$ 7 and the loop connecting  $\beta$ 5 to  $\beta$ 6 (Figure 7). The surface of this region is characterized by the presence of some positively charged residues, Lys47, Arg61 and Arg86 that help to position the negatively charged residues of p53 TAD2.

The hydrophobic side of TAD2 helix interacts with tfb1 by three fundamental hydrophobic residues (Phe54, Trp53 and Ile50) establishing a series of close contacts (Figure 8). The aromatic side chain of Phe54 fills a shallow pocket formed by Gln49, Ala50, Thr51, Met59, Leu60, and Arg61 of tfb1. In this location, the aromatic ring of Phe54 is positioned to participate in an amino-aromatic interaction with Gln49 and in a cation- $\pi$  interaction with Arg61 of tfb1. The indole ring of Trp53 is inserted into a narrow cleft formed by the side chains of Met59 and Lys57 of tfb1. In this orientation, the side chain of Trp53 is positioned to make a cation- $\pi$  interaction with Lys57 on one side of the indole ring and a sulfur- $\pi$  interaction with Met59 on the other side. The third hydrophobic residue of p53, Ile50 makes van der Waals contacts to Met 59 and Met88 of tfb1.



**Figure 8. (A)** The 3D structure of Tfb1 is shown as a ribbon within the transparent molecular surface (in green), whereas the backbone trace of p53 is represented as a tube (in orange). **(B)** The side chains of the three hydrophobic residues (Phe54, Trp53, and Ile50) on the amphipathic side of the p53 helix are also shown. The aromatic ring of Phe54 from p53 is positioned in a shallow pocket on the surface of Tfb1. [Di Lello, P., et al. Mol Cell, 2006]

Other acidic residues of p53 TAD2 helix are positioned to mediate complex formation with tfb1: Glu56 residue forms a salt bridge with Lys57 of tfb1, contributing to stabilize the complex. These four residues are conserved in p53 helix mediating the interaction also between p53 TAD2 and tfb1 subunit of TFIIH.

Further electrostatic interactions involve Glu51 residue of p53 TAD2 and Arg61 or Arg86 residues of tfb1. Moreover, two important phosphorylation sites are located on TAD2, Ser46 that is phosphorylated by p38 mitogen activated protein kinase (MAPK) and Thr55 that is phosphorylated by TATA-associated factor 1 (TAF1). The phosphorylation of these two sites has a significant effect on formation of tfb1-TAD2 complex. Phosphorylation at Ser46 by MAPK and HIPK has been postulated to play an important role in p53-induced apoptosis favoring the folding of amphipathic helix in 47-55 region, whereas the phosphorylation at Thr55 by TAF1 has been associated with an increase in Mdm2-mediated degradation of p53 favoring the folding of TAD1 helix region that interacts with Mdm2.

## 1.5 *KCTD Protein*

### 1.5.1 *POZ/BTB Domain*

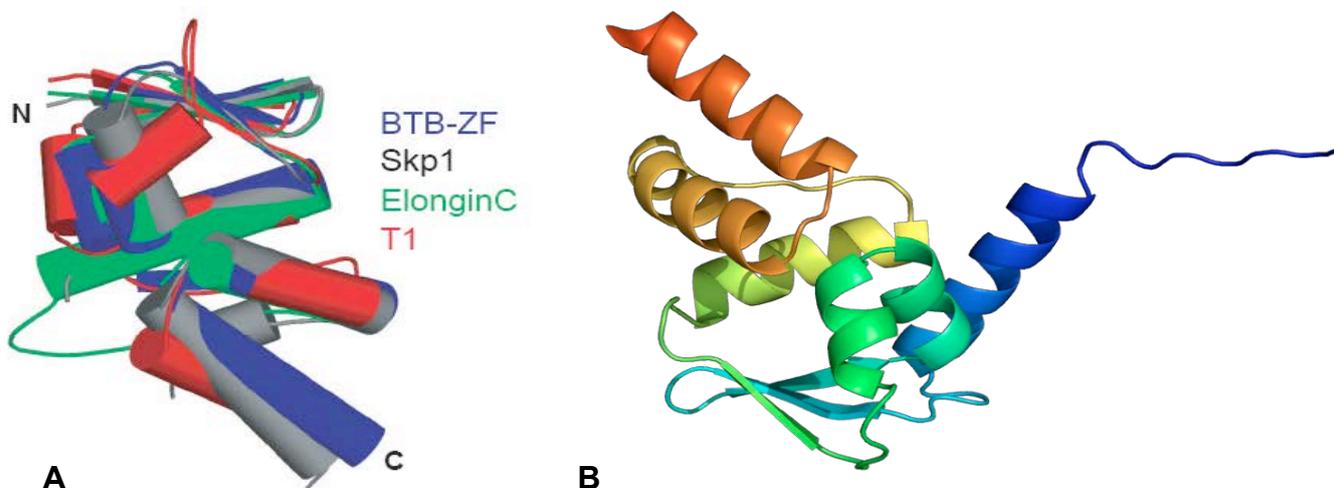
The BTB domain (also known as the POZ domain) was originally identified as a conserved motif present in the *Drosophila melanogaster* *bric-à-brac*, *tramtrack* and *broad* (BTB) complex transcription regulators and in many *poxvirus* *zinc finger* (POZ) proteins [39].

BTB/POZ domain is typically found in a single copy per protein, but in some proteins it is combined with a variety of additional protein-protein interaction domains [40].

POZ (T1) domains are contained at the N-terminus of all the mammal potassium channels and are combined with numerous other domains forming a variety of distinct domain architectures. Given the extremely wide representation and functional significance of BTB/POZ domains, the structural basis of these interactions is of the high interest.

The biological properties of BTB/POZ-containing proteins seem to be contingent on the protein-protein interaction properties of this domain. In particular, BTB/POZ domains appear to participate in a variety of processes ranging from transcriptional repression/activation [41] to cytoskeletal organization [42]. In several proteins, BTB/POZ domains mediate homodimerization or multimerization [43].

Several BTB/POZ structures were determined by X-ray crystallography, establishing the structural similarity between different examples of the fold [44], [45]. A comparison of the solved structures from the PDB of BTB domains such as BTB-ZF proteins, Skp1, ElonginC and T1 domains (Figure 9), shows a common region of approximately 95 amino acids consisting of a cluster of 5  $\alpha$ -helices constituted in part of two  $\alpha$ -helical hairpins (A1/A2 and A4/A5), and capped at one end by a short solvent-exposed three stranded  $\beta$ -sheet (B1/B2/B3). An additional hairpin-like motif consisting of A3 and an extended region links the B1/B2/A1/A2/B3 and A4/A5 segments of the fold.



**Figure 9. (A) Superposition of the BTB core fold from currently known BTB structures and (B) ribbon model of POZ/BTB domain. The BTB core fold (approximately 95 residues) is retained across four sequence families: BTB-ZF, Skp1, ElonginC and T1 families (PDB code: 1BUO, 1FQV, 1VCB, 1T1D). [Stogios PJ, et al., Genome Biology, 2005]**

### **1.5.2 Potassium Channel Tetramerization Domain (KCTD) Proteins**

Recently, it was highlighted the presence of POZ/BTB domain in a new class of human proteins that share a significant sequence identity (~30%) with the tetramerization domain of K<sup>+</sup> voltage-gated channel T1 proteins, denoted as KCTD (K<sup>+</sup> Channel Tetramerization Domain).

The human genome has 21 *kctd* genes that encode potassium channel tetramerization domain containing proteins. KCTD proteins are soluble with N-terminal BTB/POZ domains and variable C-terminus [46].

Although scarce, the information available about KCTD proteins suggest that these proteins might be involved in development and cellular differentiation. High expression of transcripts in fetal tissues and low levels in adults suggest that they play a role during development. Thus, KCTD12 is implicated in the maturation of inner ear neurons by linkage to a progressive dominant auditory neuropathy. A point mutation in KCTD7 is associated with neurodegeneration and progressive myoclonic epilepsy [47]. KCTD5 is reported to interact with human Golgi reassembly stacking protein 55 (GRASP55) the first known Golgi target for mitotic ERK signaling [48] and phosphorylation for G2 phase Golgi fragmentation [49]. KCTD11 is proposed to suppress Hedgehog signaling and to inhibit tumor growth in medulloblastoma [50].

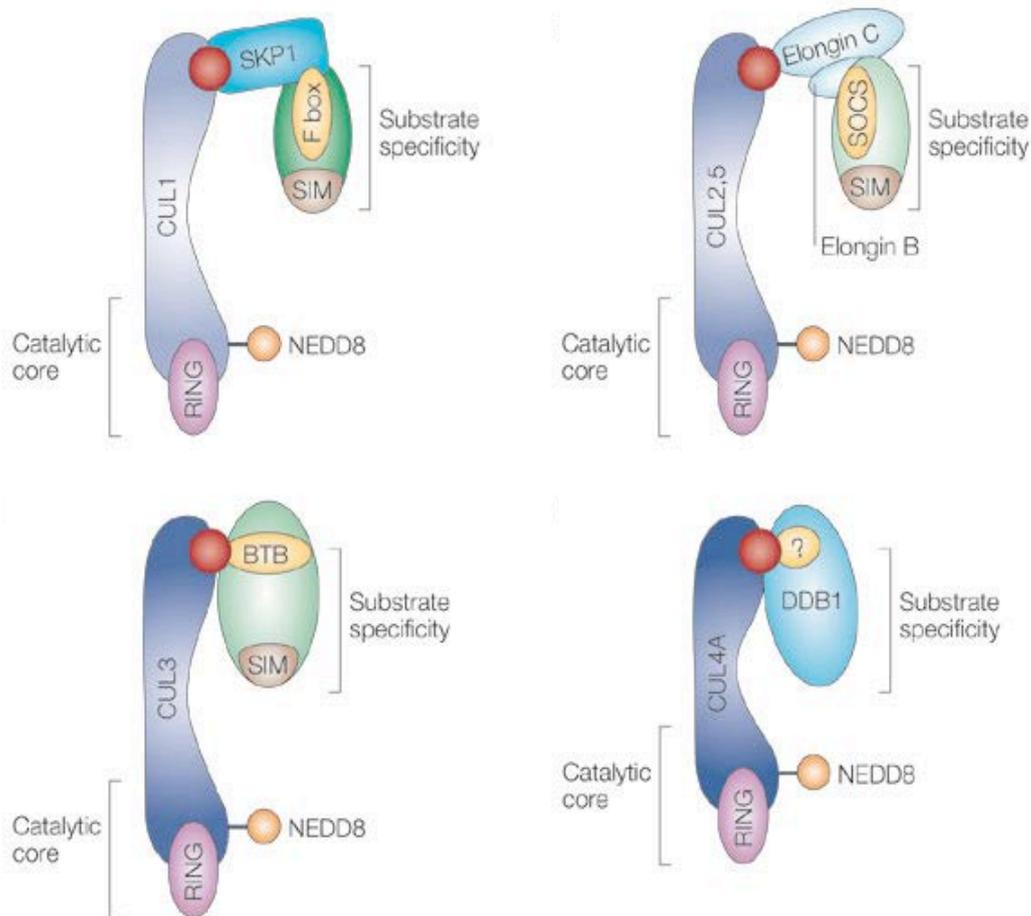
### **1.5.3 E3 ubiquitin-protein ligases**

Ubiquitin-dependent proteolysis regulates protein abundance and serves as a central regulatory function in many biological processes.

Ubiquitination is involved in a wide range of cellular functions, such as cell proliferation, differentiation, and apoptosis, mainly by targeting proteins for degradation by the proteasome, but it is also involved in protein transport and signaling through additional mechanisms.

The ubiquitination of a target protein is mediated by the ubiquitin–protein ligases, which represent a diverse super-family of proteins and protein complexes.

The largest known class of ubiquitin ligases is cullin- RING ubiquitin ligase such as Cul3, one of the seven cullins found in the human genome (Cul1, Cul2, Cul3, Cul4A, Cul4B, Cul5 and Cul7) (Figure 10). Most of them bind to adaptors through their BTB domains, which, in turn, bind to additional proteins that work as substrate-specific adaptors. In particular it was demonstrated that Cul3 interacts specifically with different KCTD proteins [51].



**Figure 10. The common catalytic core of cullin–RING ligases, which consists of a RING protein and a cullin-family member, defines this modular class of ubiquitin ligase [Petroski, M.D., et al., Nat Rev Mol Cell Biol, 2005]**

Recent analyses have shown that the C-terminal region of KCTD5 is indispensable for interaction with Cul3, while the BTB/POZ domain interacts with additional five aminoacids on the N-terminal region. The Cul3 region involved in this interaction is the N-terminus, as described for other BTB/POZ proteins.

More recently it has been demonstrated by co-immunoprecipitation experiments that also KCTD7 is able to homodimerize and to bind Cul3 like KCTD5, suggesting a similar mechanism of interaction [52].

KCTD11 also forms a multiproteins complex with Cul3 and HDAC mediating Gli deacetylation in Hedgehog (Hh) pathway [50]. Hedgehog is a master regulator of tissue development and its pathway deregulation is responsible for the onset of several tumors, including medulloblastoma.

The relevant interest for KCTD11-Cul3 complex led to the study of molecular features governing the interaction.

On the basis of molecular modeling studies and mutagenesis experiments [53], Pedone and co-workers proposed a model of KCTD11-Cul3 interaction derived from the structure of the known complex Skp1/Cul1 (Figure 11) [54], homologous of KCTD11 and Cul3, respectively.

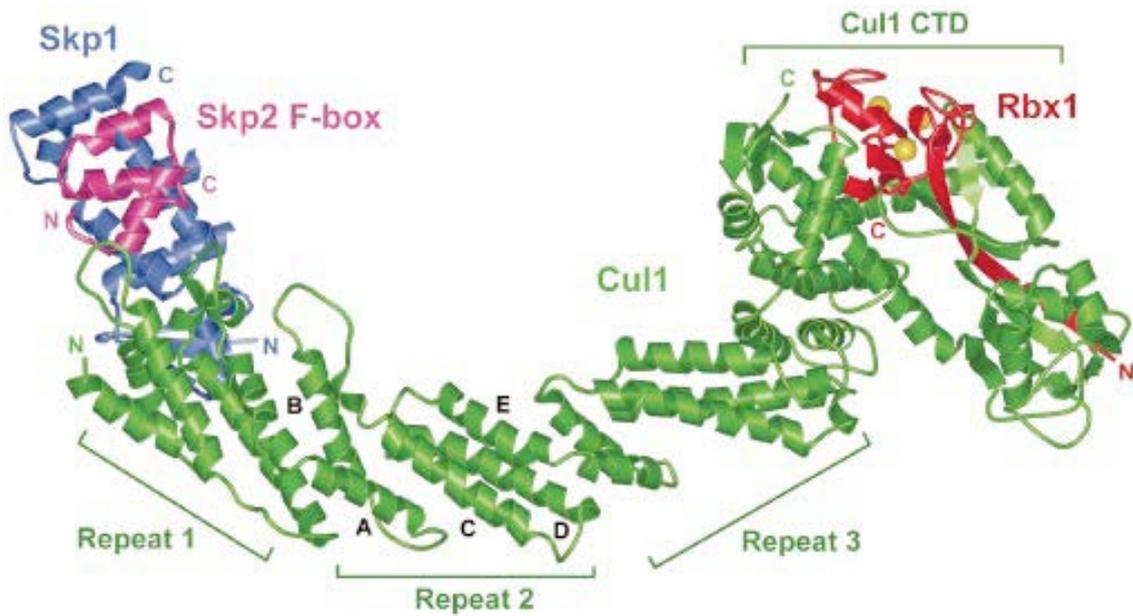


Figure 11. Overall structure of the Cul1–Rbx1–Skp1–F box Skp2 quaternary complex. [Zheng, N., et al. Nature, 2002]

In this model, each subunit of the tetramer of POZ/BTB domain interacts with a Cul3 molecule, forming a complex of huge size which interacting with other fusion partners. Interestingly, the  $\alpha$ -helix of the region 49-68 of Cul3 docks into a cavity formed at the inter-subunit interface within the tetramer. The analysis of the model also shows that, in this complex, a larger portion of the POZ/BTB domain is involved in the complex interface if compared to the 1:1 model. Indeed, each Cul3 molecule interacts with two distinct BTB subunits (Figure 12A). The residue Phe102 of KCTD11 directly interacts with a cluster of aromatic side chains of Cul3 made of Tyr58, Tyr62 and Tyr125 (Figure 12B).

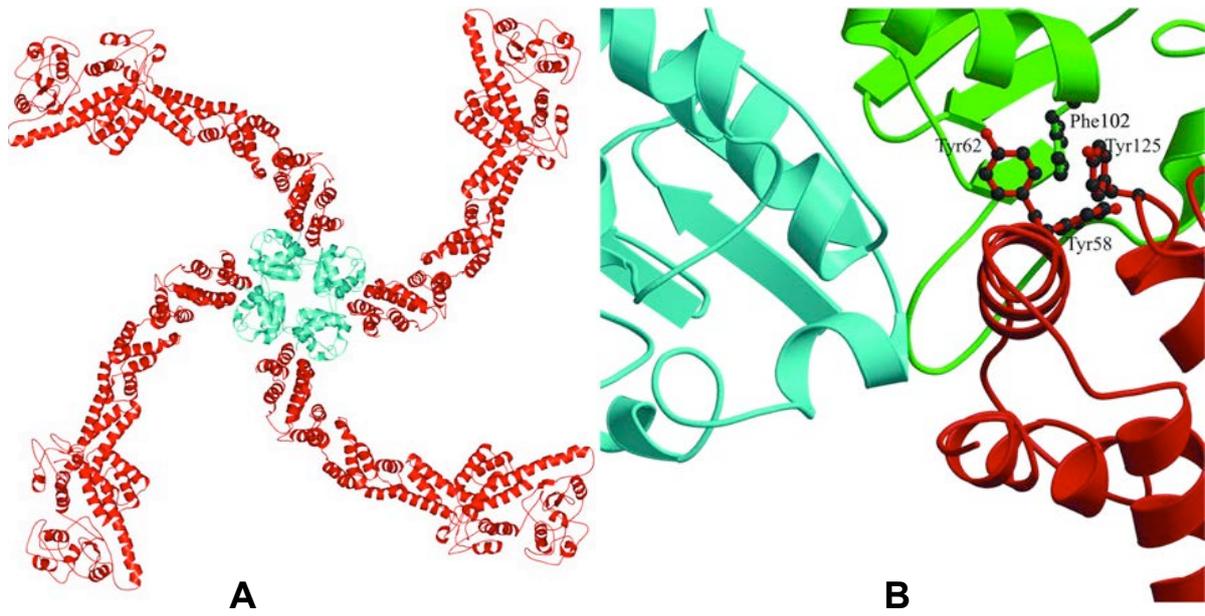


Figure 12. (A) A view of the complex between the KTCD11 POZ/BTB tetramer with four molecules of Cul3 (red) and (B) a snapshot of the KTCD11 BTB-Cul3 interface [Correale, S., et al. *Biochimie*, 2011]

## 1.6 Aim of the work

In the recent years the scientific community focused the attention on the design and synthesis of conformationally constrained molecules useful for development of new systems for biotechnological applications.

The study of  $\alpha$ -helical motifs occurring in PPI of some important molecular pathways was the starting point of this PhD thesis. The analysis of molecular features of the selected PPI allowed the design of peptide molecules able to mimic the structural motifs involved in these interactions.

Malfunctions in transcriptional regulation are associated with many human diseases and there is considerable interest in developing artificial transcription factors (ATFs) that must minimally contain a DBD and a TAD.

The ability to regulate the expression of any targeted gene using a 'programmable' ATF offers a powerful tool for functional genomics and for development of intriguing biotechnological systems.

The p53 protein has a fundamental role in cellular survival due to its transcriptional activator function.

Recently, structural studies obtained by NMR spectroscopy gave insights into interaction between p53 TAD2 and ttf1/p62 subunit of TFIIH and contributed to the understanding of the way p53 protein works in transcriptional activation.

On the basis of this structure, the design and synthesis of molecules that mimic the helical fragment of p53 TAD2 and are able to bind ttf1 and to function as TADs were performed.

It was hypothesized that peptides with increased helical propensity, which retain four key residues (IEXWF) in the  $\alpha$ -helix of p53 TAD2 (residues 47 to 55 in p53) mimic p53 TAD2 and are stronger *in vivo* activators.

Several peptides were designed by using different molecular tools such as N- and C-capping boxes to enhance helical propensity. The peptides were synthesized on solid phase method and were analyzed by CD experiments to evaluate the helical content. ITC and NMR experiments allowed to demonstrate the interaction with ttf1 subunit of TFIIH. Subsequently, studies in a yeast model system were performed to investigate the TAD properties of the obtained molecules.

BTB/POZ domain-containing proteins are identified in diverse cellular locations and participate in different processes, ranging from transcriptional repression/activation to cytoskeletal organization.

KCTD are a class of protein containing BTB/POZ domains with different biological functions, most of them unknown so far.

KCTD11 was proposed as a suppressor of Hedgehog signaling and to inhibit tumor growth in medulloblastoma. POZ/BTB domain of this protein mediates interaction with Cul3, a protein that is associated to the class of ubiquitin ligases that induce proteolysis of target proteins by proteasome.

The study of the molecular features governing the interaction of the Cul3 with KCTD11 could aid to understand the recognizing mechanism between KCTD family and proteasome machinery. These results could be used for the development of new entities able to work as agonist or antagonist of KCTD proteins involved in de-regulated pathways.

Recently, studies of molecular modeling and mutagenesis provided a model of

KCTD11-Cul3 interaction. On the basis of this information, the design and synthesis of peptides encompassing the Cul3 49-68 helical region involved in the interaction with KCTD11 were firstly performed. Subsequently, an analogue containing hydrocarbon-bridge as molecular tool to stabilize helical conformation was synthesized.

The conformation structure of all peptides was analyzed to estimate the helical content by CD experiments. The interaction with KCTD proteins were investigated by ELISA assays.

## 2. MATERIALS and METHODS

### 2.1 *Materials*

All amino acids, N- $\alpha$ -Fmoc-O-phospho-L-tyrosine, Rink-amide 4-methylbenzhydrylamine (MBHA) and Novasyn TGR resin, HBTU, HOBt and HATU were purchased from Novabiochem. DIPEA, TFA, Piperidine, N-(+)-biotinyl-6-aminocaproic acid, acetic anhydride and scavengers were supplied from Fluka. (S)-N-Fmoc-2-(4'-pentenyl) alanine for stapled peptides synthesis was supplied from Anaspec. Grubb's catalyst for ring close metathesis (RCM) was supplied from Sigma-Aldrich. The solvents used in the synthesis, purification and characterization of the peptides were purchased from Romil.

Polypropylene reaction vessels and sintered polyethylene frits were supplied by Alltech Italia.

All materials used for gene amplification were supplied by Stratagene Cloning Systems. All synthetic oligonucleotides were purchased from Sigma (Italy). All aqueous solutions were made using water purified by a Milli-Q water system (Millipore). All the chemicals used were of the highest grade available.

Expression vector pETM11 and pETM20 were from Novagen.

*E. coli* DH5 $\alpha$ , *E. coli* BL21(DE3) codon plus RIL were purchased from Stratagene. BL21(DE3), BL21 (DE3)STAR strains were supplied by Invitrogen. The identity of the inserts in the resulting recombinant plasmids was confirmed by DNA sequencing (MWG-Biotech).

### 2.2 *p53 MIMETIC PEPTIDES*

This work has been carried out in collaboration with Prof. Omichinski at University of Montreal. In particular, expression and purification of protein and labeled peptides, NMR experiments and  $\beta$ -Galactosidase Activation Assay were performed at Biochemistry Department of University of Montréal.

#### 2.2.1 *Peptides synthesis*

All peptides were obtained by Fmoc solid-phase strategy. The syntheses were carried out with Rink Amide MBHA resin (loading 0.63 mmol g<sup>-1</sup>), using all standard amino acids. The resin was swollen in DMF for 30 minutes prior to synthesis.

Fmoc protecting group was removed by treatment with 30% piperidine in DMF (2 x 10 min).

Active ester-coupling reactions were performed under a 10-fold excess of amino acid respect with the resin and HBTU (9.8 equiv)/HOBt (9.8 equiv)/DIPEA (20 equiv) in DMF for 1 hour. The coupling efficiency was assessed by Kaiser test. Each step was followed by resin washings (3 x 5 min).

Before cleavage from the resin, the peptides were acetylated by two treatments with a solution of acetic anhydride (0.5M) /DIPEA (0.015M)/HOBt (0.125M) in DMF for 10 minutes. NC15, NC17 and p53-13 peptides were cleaved off the solid support by a solution of TFA/H<sub>2</sub>O/EDT/TIS (94.5:2.5:2.5:1 v/v/v/v) for 2 hours whereas A-Cap, E-Cap, W-Cap peptides were cleaved off from the resin by treatment with TFA/H<sub>2</sub>O/EDT (90:5:5 v/v/v) for 2 hours. The resins were then filtered and the

peptides were precipitated using cold anhydrous diethyl ether.

NC15, NC17 and p53-13 crude products were purified by preparative RP-HPLC at flow rate of 20 mL/min on a Shimadzu LC-8A system, equipped with a diode array detector SPD-10A, using a Phenomenex C18 column (21 x 250 mm; 15  $\mu$ m; 300 Å). The binary solvent gradients chosen are different for each peptide: NC15 was purified with a linear gradient of H<sub>2</sub>O (0.1% TFA)/CH<sub>3</sub>CN (0.1% TFA) from 20 to 80% of CH<sub>3</sub>CN (0.1% TFA) in 30 min; NC17 with a linear gradient from 30% to 70% of CH<sub>3</sub>CN (0.1% TFA) in 30 minutes and p53-13 with a linear gradient from 30% to 80% of CH<sub>3</sub>CN (0.1% TFA) in 20 minutes.

A-Cap, W-Cap and E-Cap peptides were purified using Ammonium Formate buffer 0.1M (pH 7.0) and Ammonium Formate buffer/ CH<sub>3</sub>CN (pH 7.0) (1:1 v/v) with a linear gradient from 45% to 100% in 25 minutes for A-Cap, from 40% to 100% in 25 minutes for W-Cap and from 50% to 95% of ammonium formate buffer/ CH<sub>3</sub>CN (pH 7.0) in 30 minutes for E-Cap.

The identity and purity of the peptides were assessed by LCQ Deca XP Max LC/MS system equipped with a diode array detector combined with an electrospray ion source and ion trap mass analyzer (Thermo Finnigan) using a Phenomenex C18 column (250 x 2 mm; 5  $\mu$ m; 300 Å) at flow rate of 200  $\mu$ L/min.

### **2.2.2 Circular Dichroism Studies**

Circular dichroism analysis were performed on a Jasco J-810 spectropolarimeter at 25°C in 10 mM phosphate buffer at pH=7.1 using a quartz cell with a 0.1 cm path length in the 190–260 nm range. All peptide concentrations were determined by UV spectroscopy using A<sub>280</sub>. The results are reported as mean residue ellipticity [ $\theta$ ].

### **2.2.3 Isothermal Titration Calorimetry Studies**

The ITC experiments were performed in 20 mM TRIS (pH=7.5) for tfb1. The protein and peptides concentrations were determined from A<sub>280</sub>. All titrations were done at least in duplicates and were fit to a single binding site mechanism with 1:1 stoichiometry.

### **2.2.4 Cloning, Expression and Purification of peptides and recombinant Proteins**

The bacterial expressed p53 peptide analogs, E-Cap, p53-13, p53-13(LL) and mutants were constructed by inserting the BamHI-EcoRI-digested DNA (IDT) into pGEX-2T plasmid. Tfb1 was cloned as described in [35]. Mutants of tfb1 were prepared using the QuickChange II site-directed mutagenesis kit (Stratagene).

The p53 analogs and tfb1 were expressed as GST-fusion proteins in *E. coli* host strain TOPP2 and bound to GSH-resin (GE Healthcare) [35]. The resin bound protein was incubated overnight with thrombin (Calbiochem). After cleavage, the supernatant was purified by FPLC over a SP-Sepharose High Performance column. Uniformly (>98%) <sup>15</sup>N-labeled and <sup>15</sup>N/<sup>13</sup>C-labeled proteins were prepared in minimal media containing <sup>15</sup>NH<sub>4</sub>Cl with or without <sup>13</sup>C<sub>6</sub>-glucose as the sole nitrogen and carbon source.

### **2.2.5 Media, Plasmids and Strains**

All yeast strains were grown in synthetic complete media (SC; 0,67% yeast nitrogen base w/o amino acids, 2% glucose and amino acids drop-out mix) lacking uracil and histidine. The EGY48 (Mat alpha leu2-3 his3-11,15 trp1-1 ura3-1 6lexAops-LEU2) strain was transformed with the LexA operator-Lac-Z fusion plasmid pSH18-34 combined with either pEG202NLS (pEG202 derivative with SV40 nuclear localization sequence between LexA and polylinker) as a negative control, pSH17-4 (GAL4-activation domain cloned into pEG202 backbone) as a positive control or pEG202NLS with LexA fused to the activation domains to be tested.

### **2.2.6 $\beta$ -Galactosidase Activation Assay**

Yeast strains were transformed with the LexA operator-*lacZ* fusion plasmid pSH18-34 (eight LexA binding sites) and various LexA-fusion proteins to determine the relative abilities of the fusion proteins to activate transcription in the yeast model system. In all cases, cells were grown to mid-logarithmic phase in growth media lacking Ura and His. The number of cells used for the various LexA fusion proteins was adjusted to obtain reliable readings of optical density at 420 nm. For each measurement,  $\beta$ -galactosidase activity was determined from three independent cultures and average values are given. One unit of  $\beta$ -galactosidase is defined as the amount that hydrolyzes 1  $\mu$ mol of ONPG to *o*-nitrophenol and galactose per minute per cell. Results are presented as the mean of the percentages of the  $\beta$ -galactosidase units of the tested peptides on the  $\beta$ -galactosidase units of the GAL4 positive control  $\pm$  standard error of the mean (SEM). Western blot analysis was performed with an anti-LexA antibody to verify expression of all LexA-fused peptides.

### **2.2.7 NMR Samples**

For the NMR chemical shift mapping studies with labeled Tfb1PH, the samples consisted of 0.5 mM  $^{15}\text{N}$ -Tfb1PH in 20 mM sodium phosphate (pH=6.5), 1 mM EDTA and 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$ ; unlabeled p53 analogs (NC17, NC15, E-Cap, W-Cap or A-Cap) were added to a final ratio of 1:2. For the chemical shift mapping studies with labeled E-Cap peptide, the sample consisted of 0.5 mM of  $^{15}\text{N}$ -E-Cap in 20 mM sodium phosphate (pH=6.5), 1 mM EDTA and 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  to which unlabeled tfb1 was added to a final ratio of 1:2. For the competition experiment, an HSQC was first collected with a sample containing 0.8 mM of  $^{15}\text{N}$ -p53<sub>40-73</sub> (p53TAD2) in 20 mM sodium phosphate (pH=6.5), 1.0 mM EDTA and 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$ . Then 1 mM of unlabeled tfb1 was added and a second HSQC was collected.

Finally, 0.8 mM of unlabeled E-Cap peptide was added and a third HSQC spectrum was recorded. The structural studies of the E-Cap peptide in complex with tfb1 were performed on two samples. The first contained 0.5 mM of  $^{15}\text{N}$ -E-Cap and 0.5 mM unlabeled tfb1 in 20 mM sodium phosphate (pH=6.5), 1 mM EDTA and 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$ . The second sample contained 0.5 mM  $^{15}\text{N}/^{13}\text{C}$ -Tfb1PH and 0.5 mM unlabeled E-Cap in 20 mM sodium phosphate (pH=6.5), 1 mM EDTA. For studies in  $\text{D}_2\text{O}$ , the sample was dissolved in 99.996%  $\text{D}_2\text{O}$ .

### **2.2.8 NMR Spectroscopy Experiments**

The NMR experiments were carried out at 295 K on Varian Unity Inova 500, 600 and 800 MHz spectrometers. For the chemical shift mapping studies, 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC were recorded. Intramolecular nuclear Overhauser effects (NOEs) for E-Cap were obtained from 3D  $^{15}\text{N}$ -edited NOESY-HSQC ( $\tau_m=140$  ms) and 2D  $^{13}\text{C}/^{15}\text{N}$ -{F1/F2}-filtered  $^1\text{H}$ - $^1\text{H}$  NOESY ( $\tau_m=40$  and 100 ms). The NMR data were processed with NMRPipe/NMRDraw and analyzed with CcpNmr.

The NOE-derived distance restraints were divided into four classes defined as strong (1.8-2.8 Å), medium (1.8-3.4 Å), weak (1.8-5.0 Å) and very weak (3.3-6.0 Å). Backbone dihedral angles were derived with the program TALOS. The structure of E-Cap was calculated using the program CNS, with a combination of torsion angle and Cartesian dynamics and starting from an extended structure with standard geometry. The quality of structures was assessed using PROCHECK-NMR and MOLMOL. The figures were generated with the program PyMol (<http://www.pymol.org>).

## **2.3 CuI3 MIMETIC PEPTIDES**

The expression and purification of KCTD11 protein has been carried out in collaboration with Dr. Emilia Pedone and Dr. Sonia Di Gaetano at IBB-CNR of Naples.

### **2.3.1 Peptide syntheses**

Cul3wt, Cul3Y(P) and stapled Cul3 peptides encompassing the region 49-68 of Cullin3 sequence were obtained by Fmoc solid-phase strategy [Table xx]. To mimic the charge status of the fragment within the parent protein, the N-terminus and the C-terminus of the peptides were acetylated and amidated, respectively. The syntheses were carried out with Novasyn TGR resin (substitution  $0.24 \text{ mmol g}^{-1}$ ), using standard amino acids except N- $\alpha$ -Fmoc-O-phospho-L-tyrosine for Cul3Y(P) and (S)-N-Fmoc-2-(4'-pentenyl) alanine for stapled peptides synthesis.

The resin was swollen in DMF for 30 min prior to synthesis. Coupling reactions for standard amino acids were performed by using 10 equivalents of Fmoc protected amino acids respect with the resin activated in situ with HBTU (9.8 equiv)/HOBt (9.8 equiv)/DIPEA (20 equiv) in DMF for 1 h.

Coupling reactions of N- $\alpha$ -Fmoc-O-phospho-L-tyrosine were instead performed by using 5 equivalents of amino acid activated in situ with PyBop (5 equiv)/DIPEA (10 equiv) in DMF overnight.

Coupling reactions of (S)-N-Fmoc-2-(4'-pentenyl) alanine were performed by using 2 equivalents of amino acid activated in situ with HATU (1.9 equiv)/DIPEA (4 equiv) in DMF overnight.

The coupling efficiency was assessed by Kaiser test. Each step was followed by resin washings (3 x 5 min).

Fmoc protecting group was removed by treatment with 30% piperidine in DMF two times for 10 minutes. Before the cleavage from the resin, all the peptides were acetylated or biotinylated at N-terminus to obtain the corresponding derivatives. The

acetylation reaction was carried out two times for 10 minutes using a solution of acetic anhydride (0.5 M)/DIPEA (0.015 M)/HOBt (0.125 M) in DMF.

Biotinylated peptides were obtained using a solution of N-(+)-biotinyl-6-aminocaproic acid (2 equiv)/HATU (1.9 equiv)/DIPEA (4 equiv) in DMF overnight.

The metathesis reactions were performed on aliquotes of peptidyl-resins containing acetylated or biotinylated peptides by treatment with a solution of Grubb's catalyst 0.1M (1st generation) in DCM, 2 times for 2 hours.

Cul3-wt and Cul3-Y(P) peptides were cleaved off the resin by treatment with a mixture of TFA/H<sub>2</sub>O/EDT/TIS (94: 2.5: 2.5: 1 v/v/v/v) for 2 hours at room temperature, while Cul3 stapled peptides were cleaved off the resin by treatment with a mixture of TFA/H<sub>2</sub>O/EDT/Thioanisole/Phenol (82,5: 5: 2.5: 5: 5 v/v/v/v).

The resins were filtered and the crude peptides were precipitated with diethyl ether, dissolved in a H<sub>2</sub>O/CH<sub>3</sub>CN (1: 1) solution and lyophilized. The peptides were purified by preparative RP-HPLC on the Shimadzu system equipped with a UV-Vis detector SPD10A using a Phenomenex Jupiter Proteo C12 column (21.2 x 250 mm; 4 μm; 90) and a linear gradient of H<sub>2</sub>O (0.1% TFA)/CH<sub>3</sub>CN (0.1% TFA) from 20 to 70% of CH<sub>3</sub>CN (0.1% TFA) in 20 min at flow rate of 5 ml/min. The collected fractions containing the peptides were lyophilized.

The identity and purity of the peptides were assessed by an ESI-LC-MS instrument (ThermoFinnigan, NY, USA) equipped with a diode array detector combined with an electrospray ion source and ion trap mass analyzer using a Phenomenex Jupiter Proteo C12 column (150 x 2 mm; 4 μm; 90 Å) and a linear gradient of H<sub>2</sub>O (0.1% TFA)/CH<sub>3</sub>CN (0.1% TFA) from 20 to 80% of CH<sub>3</sub>CN (0.1% TFA) in 20 minutes at flow rate of 200 μl/min.

### **2.3.2 Circular Dichroism Studies**

Circular dichroism analysis were performed on a Jasco J-810 spectropolarimeter at 25°C in 10 mM phosphate buffer at pH=7.1 using a quartz cell with a 0.1 cm path length in the 190–260 nm range. All peptide concentrations were determined by UV spectroscopy using A<sub>280</sub>. The results are reported as mean residue ellipticity [θ].

### **2.3.3 Expression and Purification of POZ/BTB domains of KCTD11**

After a first screening of small-scale expression cultures, performed using different strains, temperatures, IPTG concentration and induction length, recombinant constructs were transferred into *E. coli* strain that assured the best expression level in soluble phase.

Cell pellets from 500 ml cultures were re-suspended in 10 ml of 20 mM Tris/HCl pH 7.5 (buffer A), supplemented with a protease inhibitor cocktail (Complete EDTA-FREE, Roche). Crude extracts were prepared by disrupting the cells with 10 min pulses at 20 Hz (Misonix Sonicator 3000) and centrifugating lysates at 16 000 rpm for 30 min. The crude extracts were applied on a HisTrap HP (GE Healthcare) equilibrated with buffer A containing 0.3 M NaCl and 10 mM imidazole. Proteins were eluted with the same buffer A supplemented with 300 mM imidazole.

For KCTD11 the active fractions were pooled, extensively dialyzed against 50 mM Tris/HCl pH 7.5, 0.5 mM EDTA and successively digested with the protease TEV (Tobacco Etch Virus). The digested samples were applied onto a HisTrap performed

as above described. The flow-through samples were applied onto a 1.6 cm \_ 60 cm column (HiLoad Superdex 200, GE Healthcare) connected to an AKTA system (GE Healthcare) and eluted with Buffer A containing 0.15 M NaCl. at a flow rate of 1 ml/min.

#### **2.3.4 ELISA Assays for Cul3 peptides**

For ELISA assays, 5 µg/ml streptavidin in phosphate/citrate buffer pH 5.0 was incubated overnight at 37°C for coating. First, binding was executed with 0.8 µM biotinylated Cul3 peptides in Phosphate Buffered Saline (PBS) 1X for 1 h at room temperature. Second, binding was performed with different concentrations of His-TrxA/KCTD11 BTB (0.8, 1.5, 3.8, 7.6 and 15.6 µM) in PBS 1X. His-TrxA was used as negative control in the same concentrations. As blocking solution 1% BSA, 0.05% Tween-20 in PBS 1X was used for 1 h at room temperature. To reveal the occurred interaction mouse anti-His monoclonal antibody was incubated in 1: 1000 dilution at room temperature for 2 h; then, horseradish peroxidase-conjugated anti-mouse antibody (Pierce) was diluted 1:10 000 in PBS 1X and incubated at room temperature for 1 h. The colorimetric reaction has been carried out with SIGMAFAST OPD reagent (Sigma Aldrich), according to the manufacturer's instructions. Finally, a Model680 Microplate Reader (Bio-Rad, Hercules, CA-USA) has been used for readings at 490 nm; data were processed by a Microplate Manager 5.2 program. The reported data are mean values of triplicate experiments.

## 3. RESULTS

### 3.1 *p53 Mimetic Peptides*

The design of peptides that mimic the properties of p53 TAD2 as transcriptional activator was based on information obtained by NMR structure of ttf1/TAD2 complex. This study highlighted that the helical region of p53 TAD2 between residues 47-55 contains the structural features for binding with ttf1/p62 subunit of TFIID. Since the residues Ile51, Glu52, Trp53 and Phe54 in the  $\alpha$ -helix of TAD2 are fundamental for the interaction, the design strategy consisted of keeping fixed the three dimensional conformation of these residues and of introducing molecular tools to stabilize the helical secondary structure.

#### **3.1.1 Design of the first generation of peptides: p53-13, NC15 and NC17**

Three peptide sequences named p53-13, NC17 and NC15 were designed. The helix conformation of these peptides was stabilized by using N- and C-capping sequences. These motifs are defined as specific patterns found at or near the ends of helices that, involving hydrogen bonding and hydrophobic interactions, increase the helical propensity of a peptide sequence.

The residues at the N-terminal of an  $\alpha$ -helix are indicated as N'-N<sub>cap</sub>-N1-N2-N3-N4. N<sub>cap</sub> is the residue with non helical  $\varphi$  and  $\psi$  angles immediately preceding the N-terminus and N1 is the first residue with helical  $\varphi$  and  $\psi$  angles. Similarly, the C-terminal residues are defined as C4-C3-C2-C1-C<sub>cap</sub>-C'-C'' [10], [55].

N-terminal *capping box* is a motif where the side chain of the N<sub>cap</sub> residue forms a hydrogen bond with the backbone of N3 and the side chain of N3 forms a hydrogen bond with the backbone of N<sub>cap</sub> [10], [55].

The *Schellman motif* has two backbone-backbone hydrogen bonds between the amide NH at C'' and the carbonyl group at C3 and between amide group at C' and the carbonyl group at C2. The associated hydrophobic interaction is between C3 and C'' while polar residues are highly favored at C1 position. The C' residue is typically a glycine [10], [55].

p53-13 peptide encompasses the 45-57 sequence of p53 protein without N- and C-capping motifs (Table 1). This sequence was chosen as peptide reference to evaluate the effect of the helical stabilization on the biological activity.

NC17 peptide consists of 17 amino acids and preserves the 50-54 sequence of p53 except for the Gln52 that was replaced with Glu to favour the stabilization of the helix by the *i, i+4* salt bridge with Lys at C-terminal region (Table 1). NC17 contains the N-capping box and Schellman motif at C-terminal region [55] whose residues were chosen according to the frequency of occurrence in the corresponding box positions. In particular in the capping motif it was preferred to retain Asp48 present in the native sequence instead of inserting Glu as reported.

NC15 peptide contains 15 residues encompassing the 50-54 sequence of p53 and the N-capping box of NC17 (Table 1). In this case Schellman motif at C-terminal region includes Trp54 and Phe53 residues directly involved in the ttf1 interaction and the stabilization of the helix by the *i, i+4* salt bridge involves Glu51 and Lys at C-terminal region.

All peptides were acetylated and amidated at the N- and C-termini respectively, to avoid the interference of terminal charges with helix dipole and their sequences are reported in Table 1.

**Table 1**

Peptide	Sequence
p53-13	Ac- <sup>45</sup> Leu-Ser-Pro-Asp-Asp- <sup>50</sup> Ile-Glu-Gln-Trp-Phe- <sup>55</sup> Thr-Glu-Asp-NH <sub>2</sub>
NC17	Ac-Leu-Thr-Pro-Asp-Glu-Phe-Ile-Glu-Glu-Trp-Phe-Leu-Lys-Asp-His-Gly-Ile-NH <sub>2</sub>
NC15	Ac-Leu-Thr-Pro-Asp-Glu-Phe-Ile-Glu-Gln-Trp-Phe-Lys-His-Gly-Ile-NH <sub>2</sub>

**Red Line** Key residues for interaction with Tfb1

**Blu Line** N-Capping motif residues

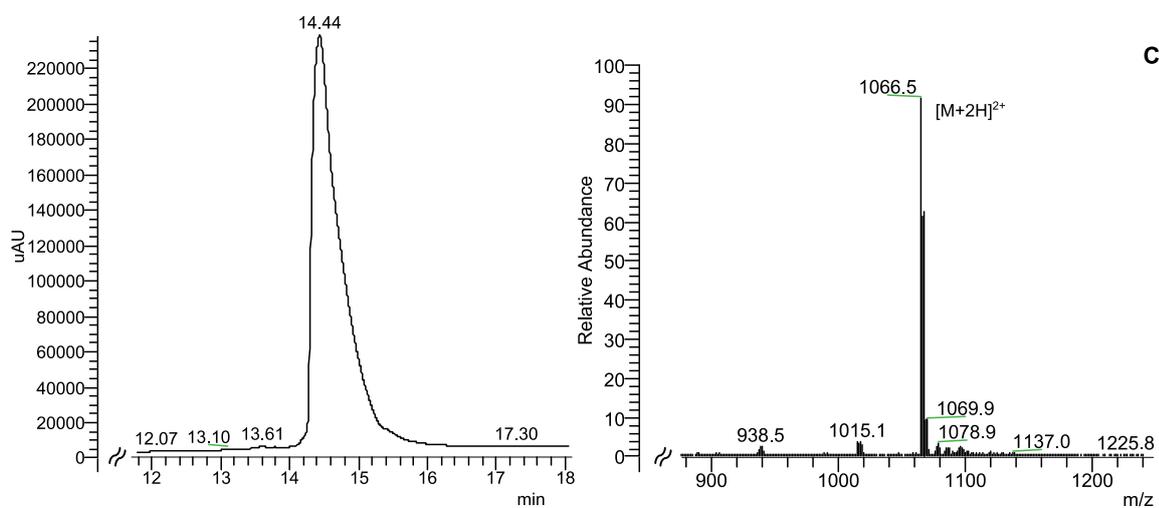
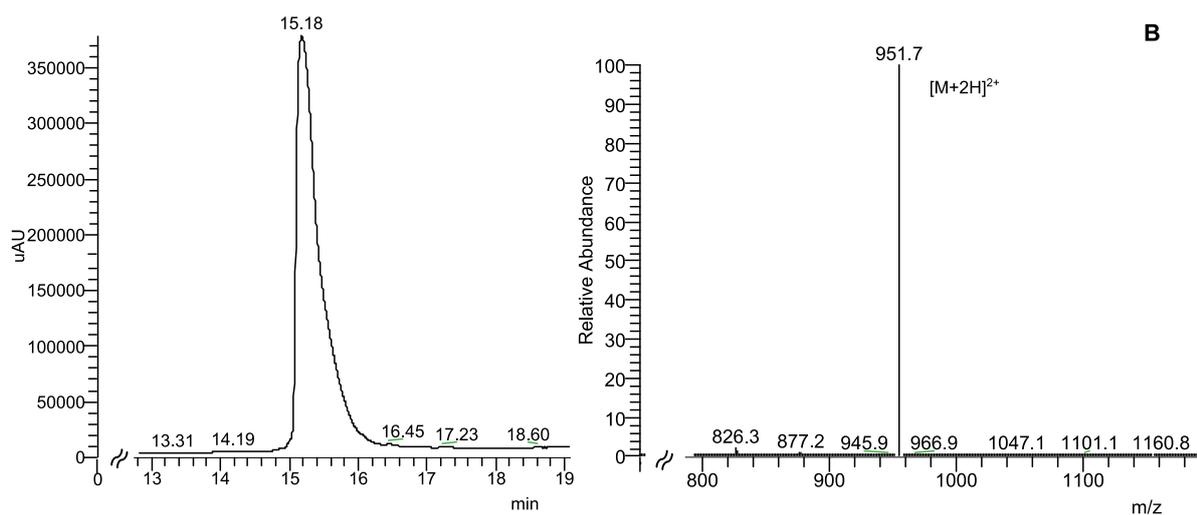
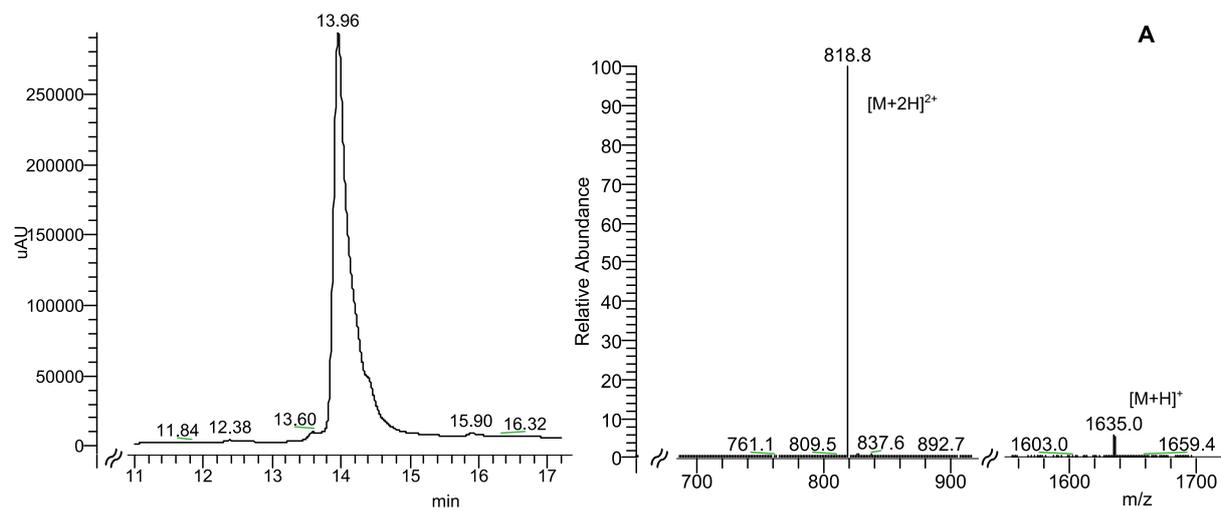
**Green Line** Schellman motif residues

### **3.1.2 Synthesis of p53-13, NC15 and NC17 Peptides**

The peptides were synthesized by Fmoc solid phase strategy. The identity and purity of the obtained peptides were assessed by LC-MS (Figure 18, 19, 20). The synthesis yields theoretical masses and molecular ion peaks for peptides were reported in Table 2.

**Table 2**

Peptide	Yield (%)	Theoretical Mass (Da)	Molecular ion peak
p53-13	27,5	1635,7	1635,0
NC15	13,5	1901,2	1901,4
NC17	50	2130,4	2131,0



**Figure 13. (A) LC-MS of p53-13, (B) of NC15 and (C) NC17.**

### 3.1.3 CD studies

The conformational behavior of synthesized peptides was investigated by circular dichroism experiments (Figure 14).

CD spectrum of p53-13 peptide is characterized by a deep minimum at 200 nm peculiar of a peptide without a secondary structure conformation.

CD spectrum of NC17 peptide is characterized by two minima at 202 and 221 nm without crossover corresponding to a poor content of helical conformation while the NC15 spectrum shows two minima at 220 and 206 nm and a crossover at 199 nm typical of a peptide with some helical content.

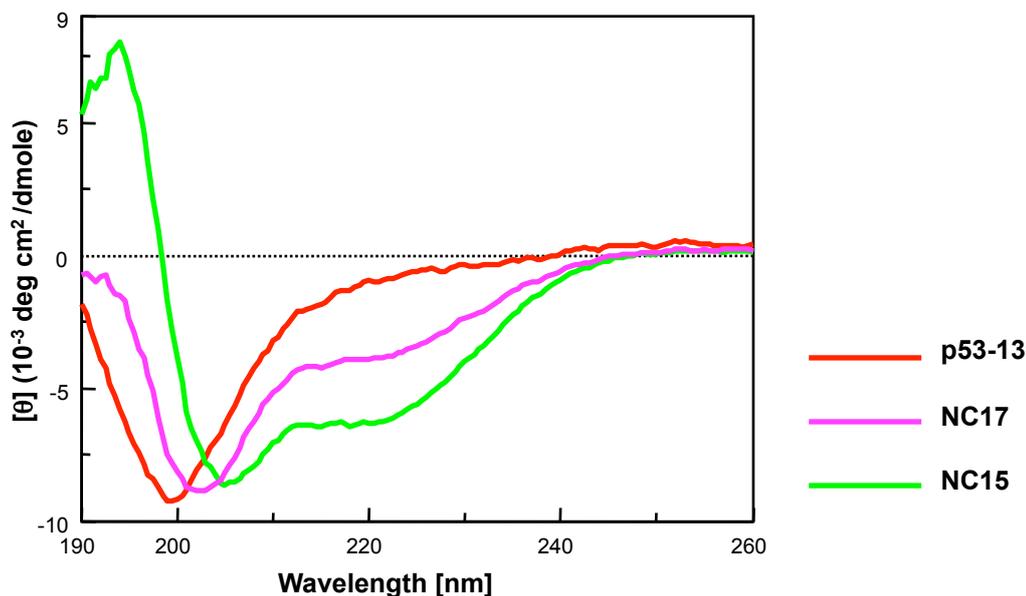


Figure 14. CD Spectra of p53-13 (red line), NC15 (green line) and NC17 (pink line) peptides.

The analysis of data obtained by CD experiments (Table 3) highlighted that the ratios between the CD signal at 222 nm and the signal of the other minimum at lower wavelength, which is often used as a mean to confirm the presence of helical content [56] are 0,24, 0,73 and 0,58 for p53-13, NC15 and NC17 peptides, respectively.

These values indicate that the introduction of N-capping box and Schellman motif in the peptides improve the  $\alpha$ -helix content respect with p53-13.

Table 3

Peptide	$\theta_{222}$	$\theta_{min}$	$\theta_{222}/\theta_{min}$	$\theta_{190}$	cross
p53-13	-2486	-10422	0,239	-3405	187
NC15	-7008	-9538	0,735	5979	198
NC17	-5522	-9421	0,586	-46	189

### 3.1.4 ITC Experiments

The ability of obtained peptides to bind tfb1 was tested by ITC experiments. p53-13 peptide displayed an apparent dissociation constant ( $K_d$ ) of  $1,6 \pm 0,3 \mu\text{M}$  for the interaction with tfb1. On the contrary, ITC experiments indicated that neither NC15 nor NC17 are able to bind tfb1 under the same conditions.

### 3.1.5 NMR Experiments

2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled tfb1<sub>1-115</sub> in the free form and in the presence of peptides respectively were performed. Several residues of p53-13 undergoing dramatic changes in  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts are identical to those observed with p53 (20-73) protein in complex with tfb1<sub>1-115</sub> (Figure 15). The residues that showed the most significant changes in  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts upon formation of the complex are located in the region of strands  $\beta_5$ ,  $\beta_6$ ,  $\beta_7$  and loop between  $\beta_5/\beta_6$  of tfb1 structure. Unfortunately NC15 and NC17 have not shown a good  $K_D$  while p53-13 has shown a higher binding affinity in comparison with its cognates.

Although the addition of a N-capping box and a C-capping box increases the helical propensity, both NC15 and NC17 possess significantly lower affinity for tfb1 than p53-13.

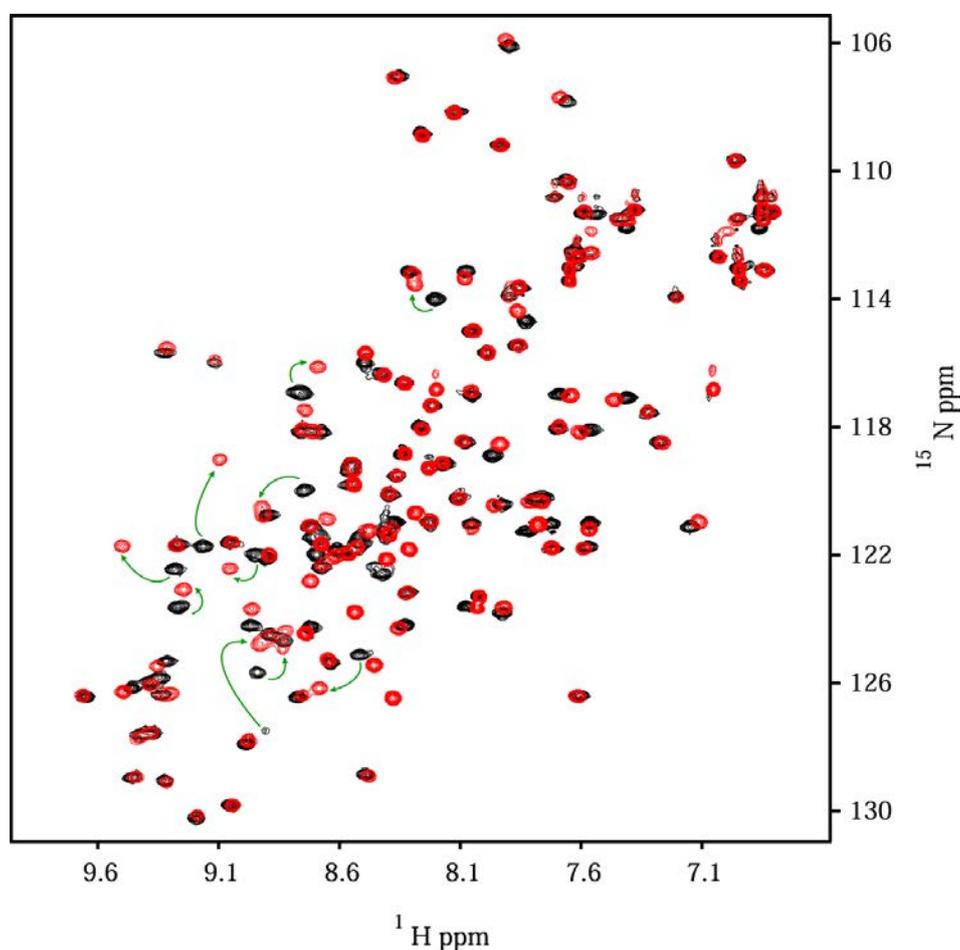


Figure 15. Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra of  $^{15}\text{N}$ -labeled Tfb1 in the free form (black) and in the presence of p53-13 (red).

### 3.1.6 Design of the second generation of peptides: N-Cap series

A second generation of molecules referred as E-Cap, W-Cap and A-Cap was designed in which the C-capping motif was not included.

In these peptides a N-capping motif was combined with a hydrophobic bridge that involve the side chains of two leucines. These residues were inserted in positions 6 and 9 with  $i, i+3$  spacing so that the hydrophobic interaction would form in the center of the peptide, but on the opposite face of the  $\alpha$ -helix relative to the ttf1-binding interface. This approach is similar to VEGF mimetic peptide (QK) design [57].

The three peptides are all 13 residues long and differ only for N-capping motif composition. The sequences are reported in Table 4.

In E-Cap peptide Glu at N1 was selected for its high frequency of occurrence in the capping box [55] and for its acidic features mimicking the phosphorylation at Ser 46 in the native protein.

In W-Cap Trp was chosen for homology with the QK-peptide and in A-Cap Ala was selected since it is present in this position in VP16, the activator transcriptional factor of Herpes Simplex virus that share the same binding interface of p53 TAD2 on ttf1 [37].

Table 4

Peptide	Sequence
E-Cap	Ac-Leu-Thr-Glu-Glu-Glu-Leu-Ile-Glu-Leu-Trp-Phe-Thr-Glu-NH <sub>2</sub>
A-Cap	Ac-Leu-Thr-Ala-Glu-Glu-Leu-Ile-Glu-Leu-Trp-Phe-Thr-Glu-NH <sub>2</sub>
W-Cap	Ac-Leu-Thr-Trp-Glu-Glu-Leu-Ile-Glu-Leu-Trp-Phe-Thr-Glu-NH <sub>2</sub>

**Red Line** Key residues for interaction with Ttf1

**Blue Line** N-Capping mutant residues

**Green Line** Leucine for  $i, i+3$  hydrophobic bridge residues

### 3.1.7 Synthesis of A-Cap, W-Cap and E-Cap peptides

The peptides were synthesized by Fmoc solid phase strategy. The purification of N-Cap peptides required different conditions compared to the standard methods because of the high content of acidic residues and poorly solubility in acid solution.

The purified peptides (99%) were analyzed by LC-MS (Figure 23,24,25). The yields, theoretical masses and ion molecular peaks are reported in Table 5.

Table 5

Peptide	Yield (%)	Theoretical Mass (Da)	Molecular ion peak
E-Cap	16,7	1692,9	1693,4
A-Cap	20,8	1634,8	1635,5
W-Cap	18,4	1750,0	1751,4

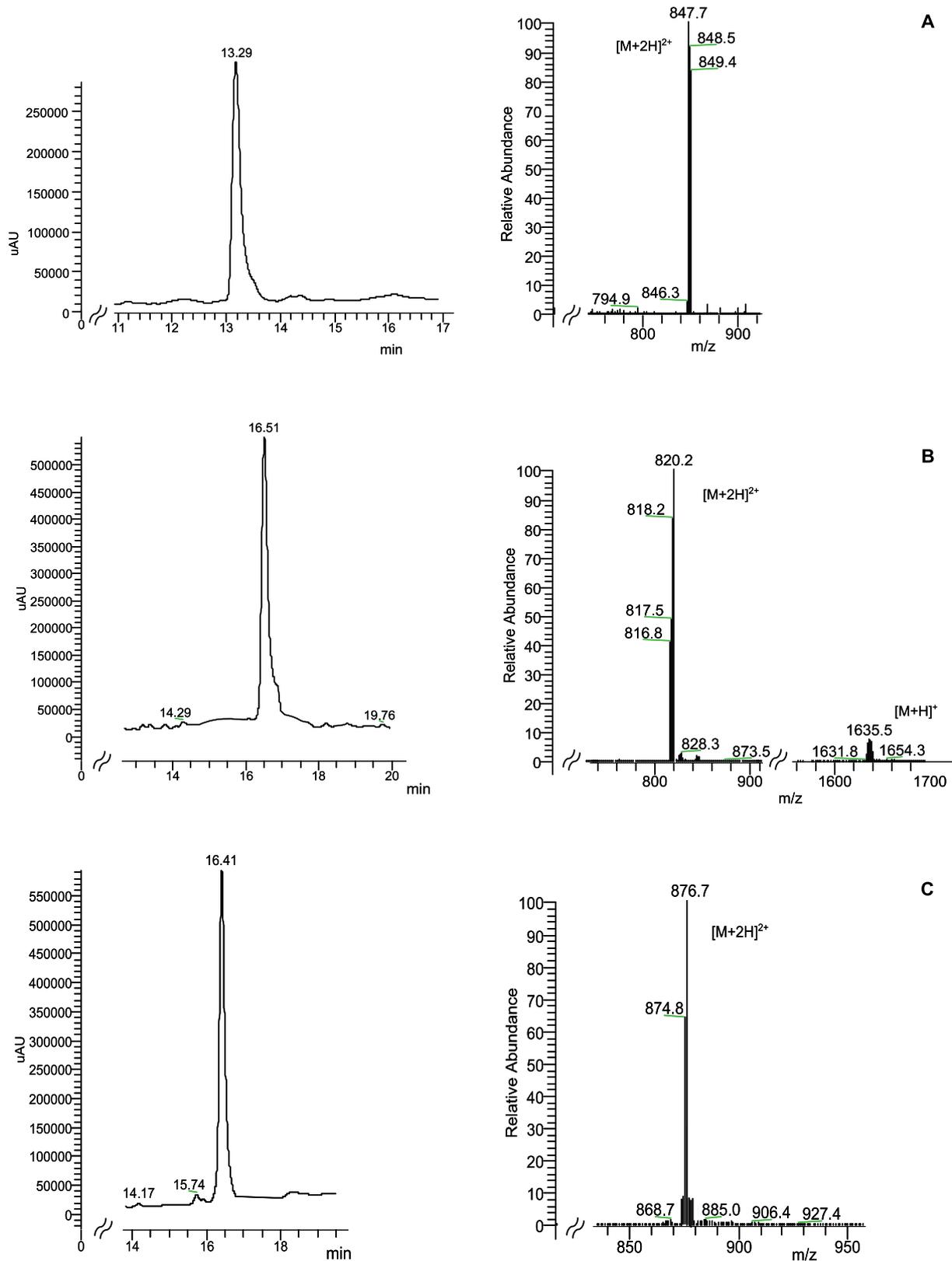


Figure 16. LC-MS (A) of E-Cap, (B) of A-Cap and (C) W-Cap

### 3.1.8 CD studies

The structural properties of E-Cap, A-Cap and W-cap peptides were analyzed by CD experiments (Figure 17).

The shape of the obtained CD spectra is very like, highlighting similar content of  $\alpha$ -helix structure.

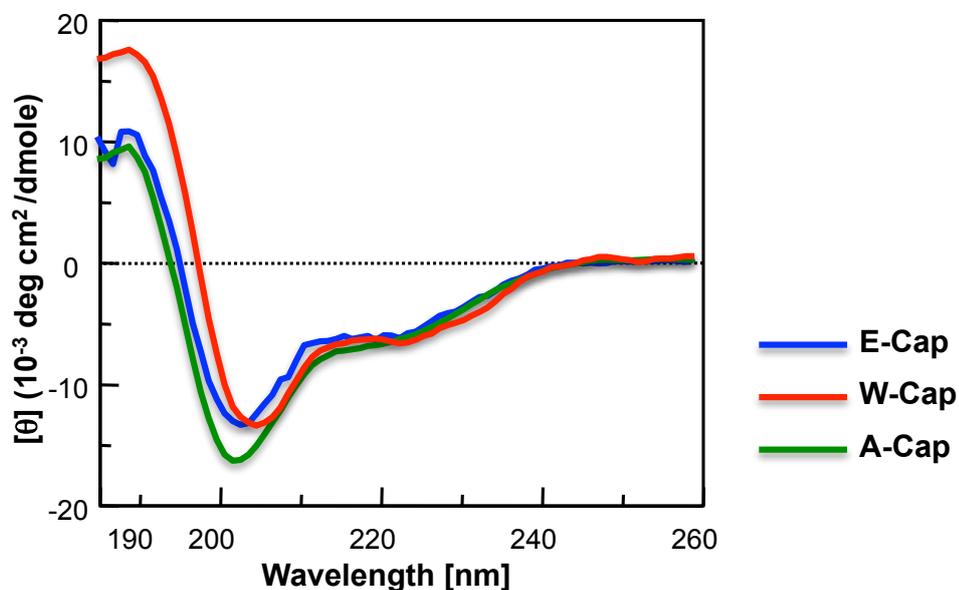


Figure 17. CD Spectra of E-Cap (blue line), W-Cap (red line) and A-Cap (green line)

The ratios between the CD signal at 222 nm and the signal of the other minimum at lower wavelength are 0,52, 0,54 and 0,52 for E-Cap, A-Cap and W-Cap, respectively. A comparison with the CD spectra of first generation peptides clearly indicates an increase of the secondary structure compared to p53-13 but a reduction of helical content compared to NC15.

Table 6

Peptide	$\theta_{222}$	$\theta_{\min}$	$\theta_{222}/\theta_{\min}$	$\theta_{190}$	cross
E-Cap	-6983	-13359	0,523	4217	193
A-Cap	-7135	-13112	0,544	7669	194
W-Cap	-7227	-13790	0,524	17409	197

### 3.1.9 ITC Experiments

All the peptides were able to bind tfb1 as verified by ITC experiments. In particular E-Cap peptide has an apparent  $K_d$  of  $0.24 \pm 0.03 \mu\text{M}$  while A-Cap and W-Cap have an apparent  $K_d$  values of  $1.3 \pm 0.1 \mu\text{M}$  and  $1.9 \pm 0.3 \mu\text{M}$  respectively, similar to p53-13.

### 3.1.10 NMR Experiments

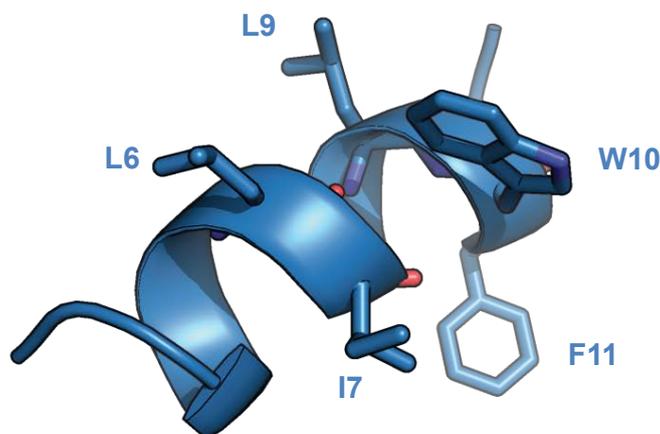
NMR titration experiments were performed to determine the binding site of the N-Cap peptides on tfb1.

Addition of E-Cap to  $^{15}\text{N}$ -labeled tfb1 produced changes in both the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts for several signals in the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of tfb1 (Figure 19). As p53 TAD2, a mapping on the structure of tfb1 have highlighted that most significant changes of signals of the residues are located within the strands  $\beta 5$ ,  $\beta 6$  and  $\beta 7$  and in the loop between  $\beta 5$  and  $\beta 6$ .

Similar changes in chemical shifts are also observed in titrations of tfb1 with both A-Cap (Figure 20) and W-Cap (Figure 21) but with lower intensity than E-Cap.

So the introduction of N-capping box and dileucine hydrophobic bridge increases the  $\alpha$ -helix content of all three peptides but improves the affinity mostly of E-Cap for tfb1. Although CD studies indicate that E-Cap possesses a higher helical content than p53-13, NMR experiments with E-Cap did not show the presence of NOE signals characteristic of an  $\alpha$ -helical conformation in the free form. Additional NMR studies of E-Cap in complex with tfb1 show that like p53 TAD2, it transitions to form a nine-residue  $\alpha$ -helix from Glu3 to Phe11 (Figure 18). The structure of E-Cap in complex with tfb1 confirms that the side-chains of Leu6 and Leu9 are in close proximity to each other, and on the opposite side of the helix relative to the binding interface with tfb1.

In addition, the side-chain of Leu6 is in position to further stabilize the helix through contacts with the aromatic ring of Trp10 in the  $i+4$  position.



**Figure 18. Ribbon model of the average structure of the E-Cap peptide from the complex with Tfb1PH. The side chains of the three hydrophobic residues (I7, W10 and F11) and the leucines (L6 and L9) are highlighted to show that they are on opposite faces of the helix**

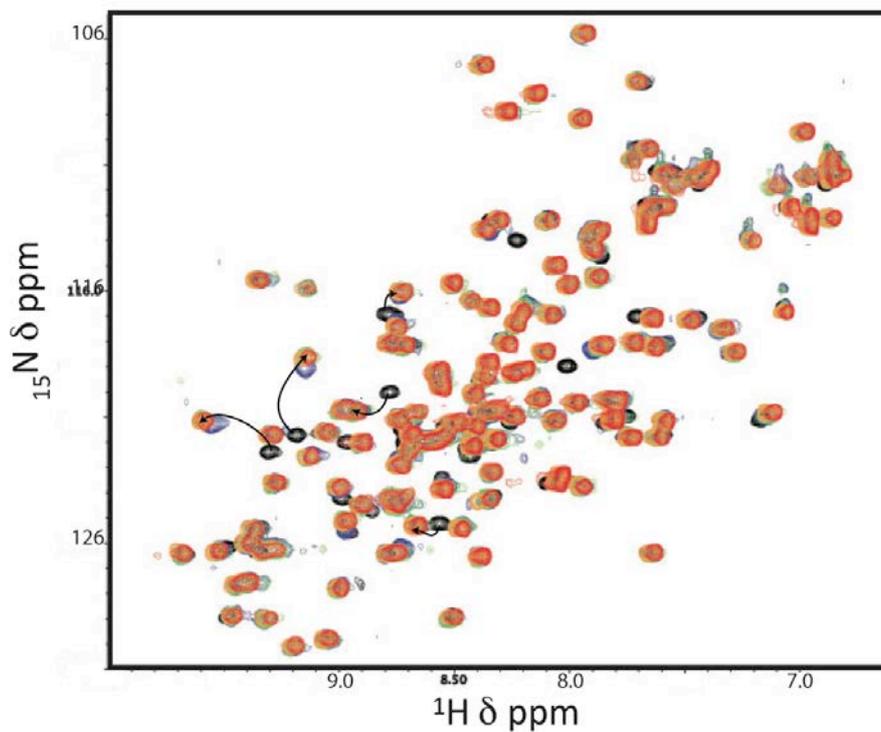


Figure 19. Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra of  $^{15}\text{N}$ -labeled tfb1 in the free form (black) and in the presence of 0.5 (blue), 1.0 (green) and 1.25 (red) equivalents of E-Cap

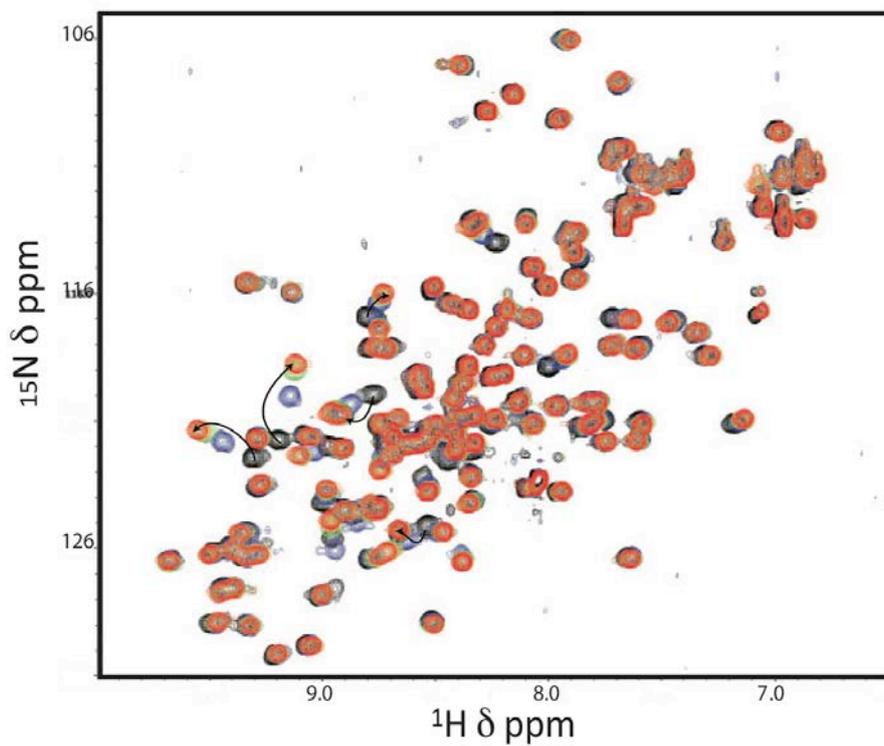
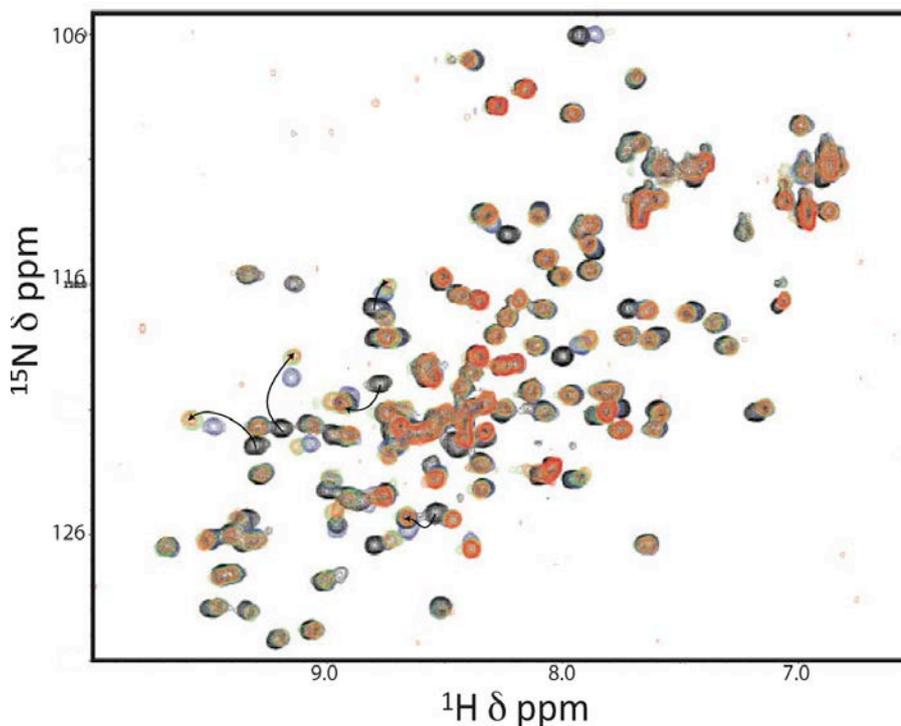


Figure 20. Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra of  $^{15}\text{N}$ -labeled Tfb1 in the free form (black) and in the presence of 0.5 (blue), 1.0 (green) and 1.25 (red) equivalents of A-Cap



**Figure 21. Overlay of the 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra of  $^{15}\text{N}$ -labeled tfb1 in the free form (black) and in the presence of 0.5 (blue), 1.0 (green) and 1.25 (red) equivalents of W-Cap**

### **3.1.11 E-Cap in vivo experiment**

The potential ability of E-Cap to activate transcription *in vivo* was investigated in yeast cells. E-Cap sequence was expressed fused to the DNA-binding domain (DBD) of LexA and its activity was measured for the *lacZ* reporter gene relative to a positive control (Gal4-TAD/LexA-DBD) whose activity is established as 100 %.

In this system, the E-Cap-LexA-DBD fusion protein activates transcription at  $161 \pm 9$  % of the positive control. In comparison, the native p53-13 peptide fused to the LexA-DBD activates transcription at only  $8 \pm 3$  %. Thus, E-Cap is ~20 fold more potent than p53-13 as a transcriptional activator in this *in vivo* system.

To compare the *in vivo* activity of the E-Cap relative to other known artificial TADs, its activity relative to two models artificial TADs, the AH 47 and VP2 45 peptides was compared. The AH-LexA-DBD and VP2-LexA-DBD activates transcription at approximately 5% of the positive control (Figure 22) which is similar to p53-13, but over 30-fold less than E-Cap. These results strongly support the idea that stabilizing the helical character of short analogs of p53 TAD2 can lead to a significant enhancement of their *in vivo* transcriptional activity.

To investigate the role of two leucine residues in *i*, *i*+3 positions to activate the transcriptional *in vivo*, the leucine residue at position 9 was mutated to native glutamine residue found in p53-13 to generate E-Cap-(LQ). This change lowers the *in vivo* activity to  $67 \pm 8$  % of the positive control, and thus E-Cap-(LQ) is ~50% less active than E-Cap.

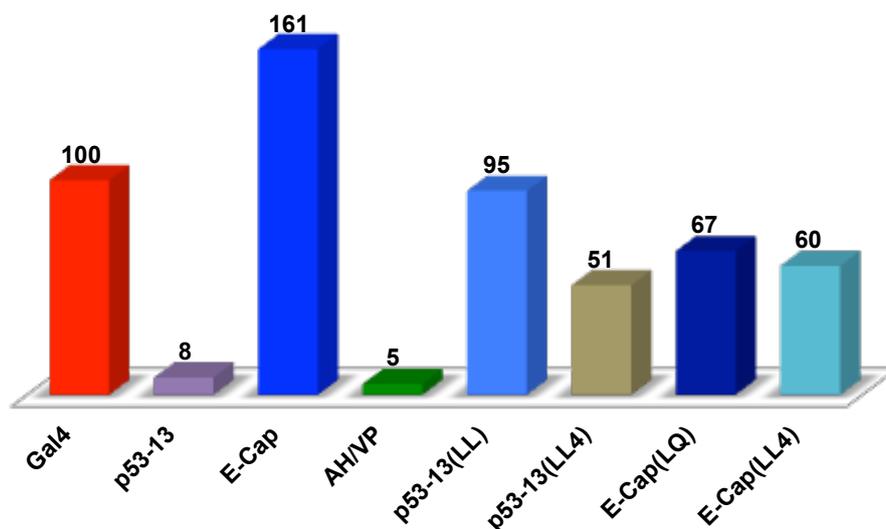
Next, two leucine residues were inserted at equivalent positions of p53-13 to generate p53-13-(LL). The *in vivo* transcriptional activation of p53-13-(LL) is 10-fold higher than p53-13, but 50% less active than E-Cap. These results support the idea

that both the N-Cap motif, and the dileucine-bridge contribute significantly to the *in vivo* transcriptional activity of E-Cap.

Experimental and theoretical studies indicate that two hydrophobic amino acids separated by either three (*i, i+3* spacing pattern) or four residues (*i, i+4* spacing pattern) enhance the helical propensity of peptides through side-chain interactions. Therefore since the insertion of the two leucines with an *i, i+3* spacing pattern enhances the *in vivo* activation of E-Cap and p53-13-(LL), the role of leucines with an *i, i+4* spacing pattern on *in vivo* activity was also tested.

Leucine residues were introduced at positions 4 and 8 in the p53-13 peptide generating p53-13-(LL4) and at positions 5 and 9 in the E-Cap peptide E-Cap-(LL4). These positions are again chosen in an attempt to place the leucine bridge in the center of the peptide, but on the backside of the helix relative to the Tfb1PH binding interface. In the yeast activation assay, LexA-p53-13-(LL4) displays  $51 \pm 8$  % of the activity of GAL4-LexA. This corresponds to a 6-7-fold increase in activity compared to the native p53-13 ( $8 \pm 3$ %), but only about half of the activity that we observe when the leucines are in the *i, i+3* spacing pattern in p53-13-(LL) ( $95 \pm 6$  %). Similarly, E-Cap-(LL4) displays  $60 \pm 2$  % of the activity of GAL4-LexA, but this corresponds to roughly 40% of the activity compared to E-Cap-(LL) ( $161 \pm 9$  %). The increased *in vivo* activity observed when two leucines were inserted in either the *i, i+3* or *i, i+4* spacing pattern of p53-13 is consistent with *in vitro* studies with model peptides showing that both spacing patterns are able to increase their helical propensity/stability.

In conclusion, these results strongly support the idea that stabilizing the helical character of short analogs of p53 TAD2 can lead to a significant enhancement of their *in vivo* transcriptional activity.



**Figure 22.** LexA-peptide fusion proteins were co-transformed into yeast with the reporter LexA operator-Lac-Z fusion plasmid pSH18-34. Results are presented as the mean of the percentages of the  $\beta$ -galactosidase units of the tested fusion proteins on the  $\beta$ -galactosidase units of the LexA-GAL4TAD positive control

## 3.2 Cul3 Mimetics

### 3.2.1 Design of Cul3-wt, Cul3-KK & Cul3-Y(P)

The studies of complex between Cul3 and BTB/POZ domain of KCTD11 by molecular modeling and mutagenesis studies [50], [53] highlighted that the interaction is stabilized by a cluster of aromatic residues Tyr58, Ty62 and Tyr125 located in the helix 55-65 of the Cul3 and Phe102 and Tyr103 of KCTD11. It was shown that mutations of these residues inhibit the formation of the complex. Although other Cul3 residues are involved in KCTD11 recognition, the attention was focused on the role played by the sequence mentioned above.

A Cul3-based peptide as a potential KCTD11 binder was therefore designed by considering the peptide corresponding to the 49–68 region in Cul3 sequence (named Cul3-wt hereafter). The location of this fragment within Cul3 three-dimensional structure is shown in Table 7.

To check the role of aromatic side chains in Cul3 – KCTD11 complex model, a Cul3-wt peptide analogue (Cul3-KK) in which Tyr58 and Tyr62 were substituted with Lys residues was also designed.

Cul3 is a target for tyrosine kinase and recently it was highlighted that Tyr58 is a site of phosphorylation in some cancer cell lines [58]. To verify the effect of this phosphorylation on binding with KCTD11, Cul3-Y(P) peptide was designed. This peptide preserves the same sequence of Cul3-wt peptide with Tyr(PO<sub>3</sub>H<sub>2</sub>) in place of Tyr58.

Table 7

Peptide	Sequence
Cul3-wt	Ac- <sup>49</sup> Asn-Ser-Gly-Leu-Ser-Phe- <sup>55</sup> Glu-Glu-Leu-Tyr-Arg- <sup>60</sup> Asn-Ala-Tyr-Thr-Met- <sup>65</sup> Val-Leu-His-Lys-NH <sub>2</sub>
Cul3-KK	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-Leu-Lys-Arg-Asn-Ala-Lys-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>
Cul3-Y(P)	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-Leu-Tyr(P)-Arg-Asn-Ala-Tyr-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>

**Red Line** Tyrosine key residues for POZ/BTB-Cul3 interaction

**Green Line** Lysine mutant residues

**Blue Line** Tyr phosphorylated residue

### 3.2.2 Synthesis of Cul3-wt, Cul3-KK and Cul3-Y(P)

The peptides were synthesized by Fmoc solid phase strategy. All peptides were also obtained as biotinylated analogues for the following ELISA assays.

The identities and the purities (99%) of peptides were assessed by LC-MS. The yields, the theoretical masses and ion molecular peaks are reported in Table 8.

**Table 8**

<b>Peptide</b>	<b>Yield (%)</b>	<b>Theoretical Mass (Da)</b>	<b>Molecular ion peak [M + 2H]<sup>2+</sup></b>
AcCul3-wt	21	2413,7	2411,6
BiotinylCul3-wt	20,4	2709,3	2708,0
AcCul3-Y(P)	32,4	2492,2	2492,0
BiotinylCul3-Y(P)	29,7	2789,3	2786,6
AcCul3-KK	17,5	2343,7	2345,6
BiotinylCul3-KK	13,2	2639,4	2640,9

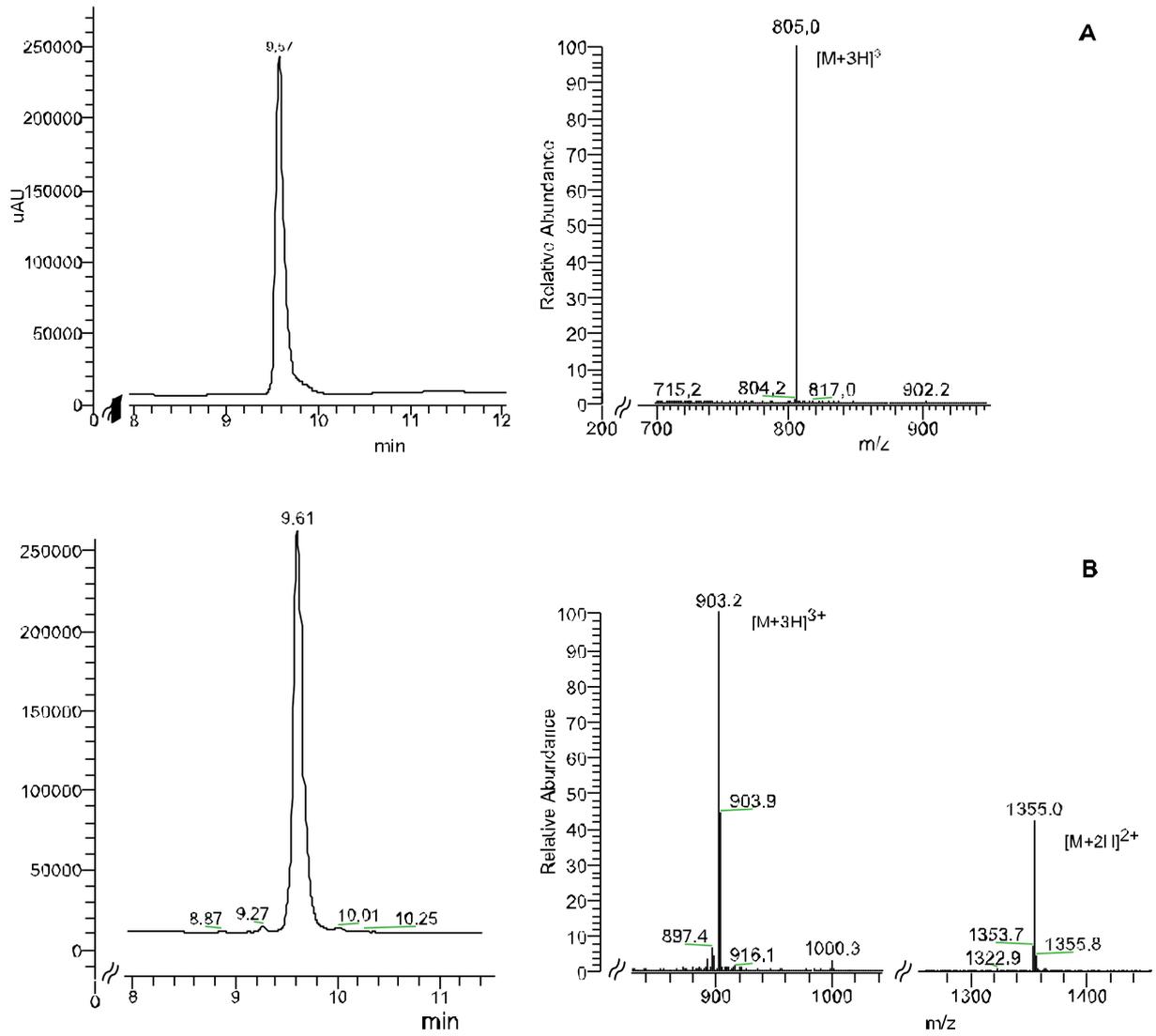


Figure 23. LC-MS (A) of Cul3-wt and (B) of BiotCul3-wt

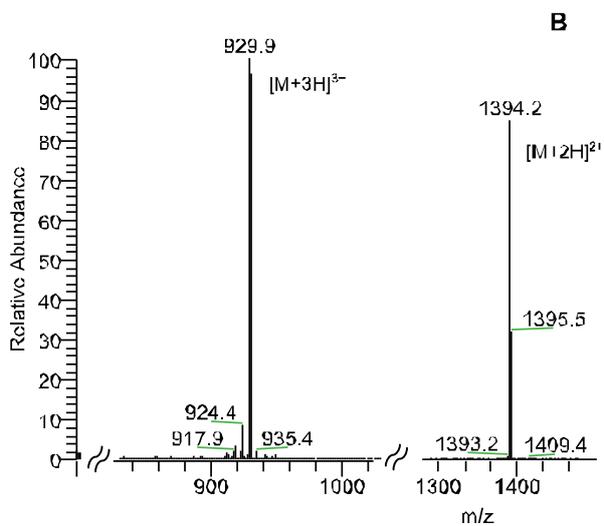
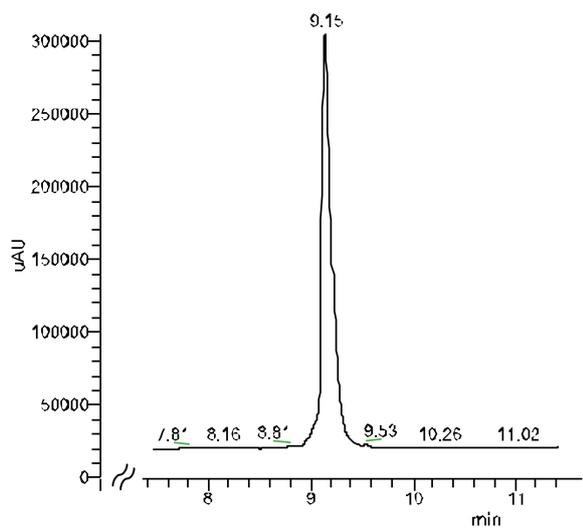
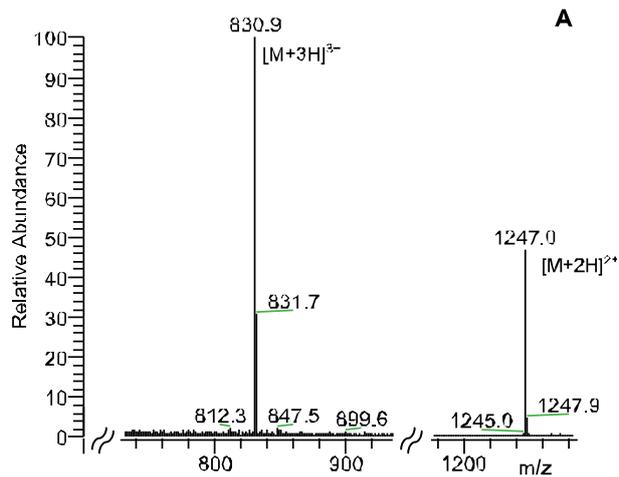
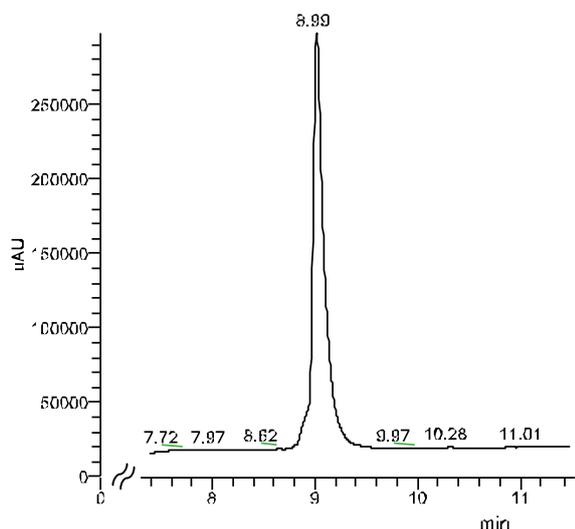
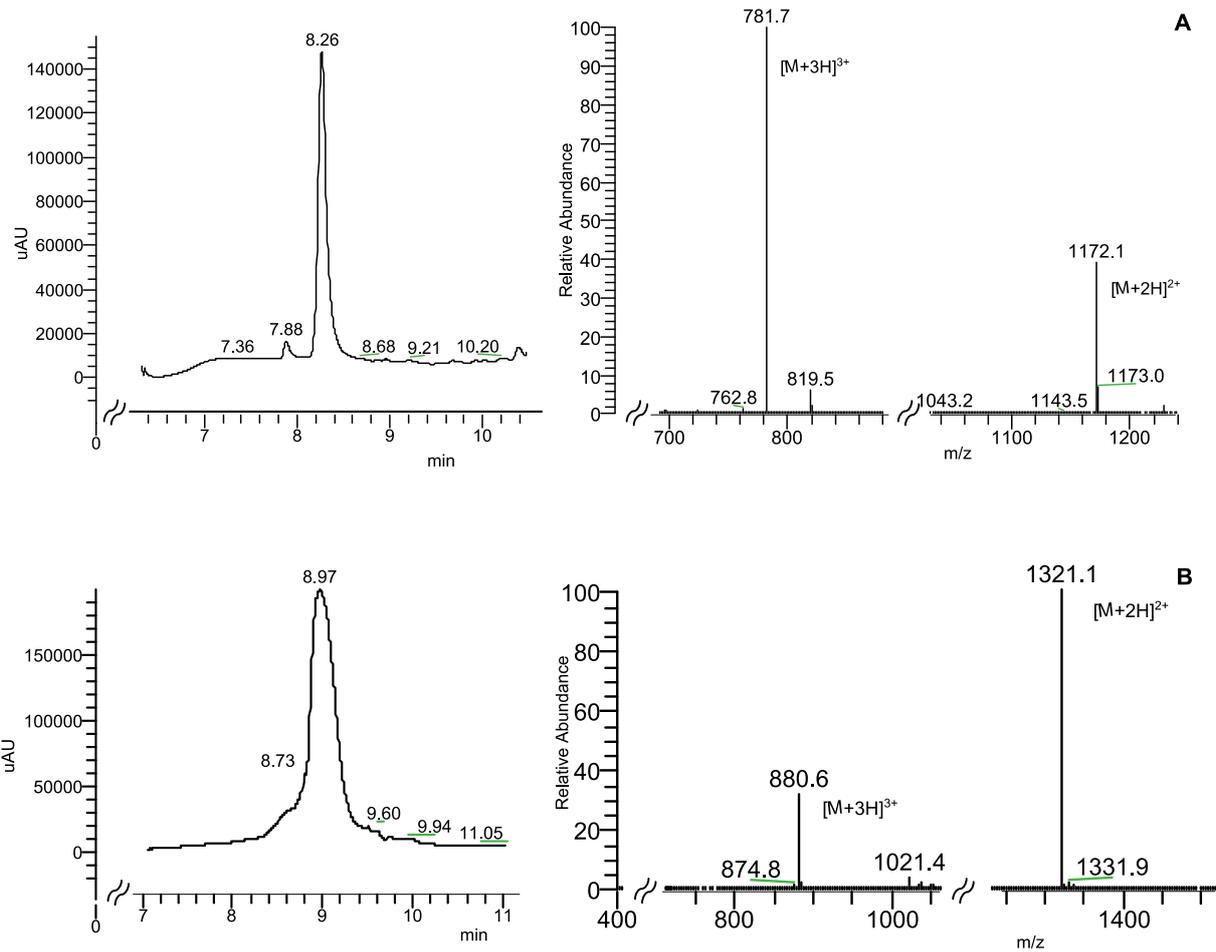


Figure 24. LC-MS (A) of Cul3-Y(P) and (B) of BiotCul3-Y(P)



**Figure 25. LC-MS (A) of Cul3-KK (B) of BiotCul3-KK**

### 3.2.3 CD Studies

The CD spectra of Cul3-wt and Cul3-KK are similar and show a deep minimum at ~202–203 nm, a second minimum at ~222 nm, a cross over at ~192–193 nm, and a maximum at a wavelength lower than 190 nm (Figure 26). The general features of the both spectra are indicative of a low helical content [59].

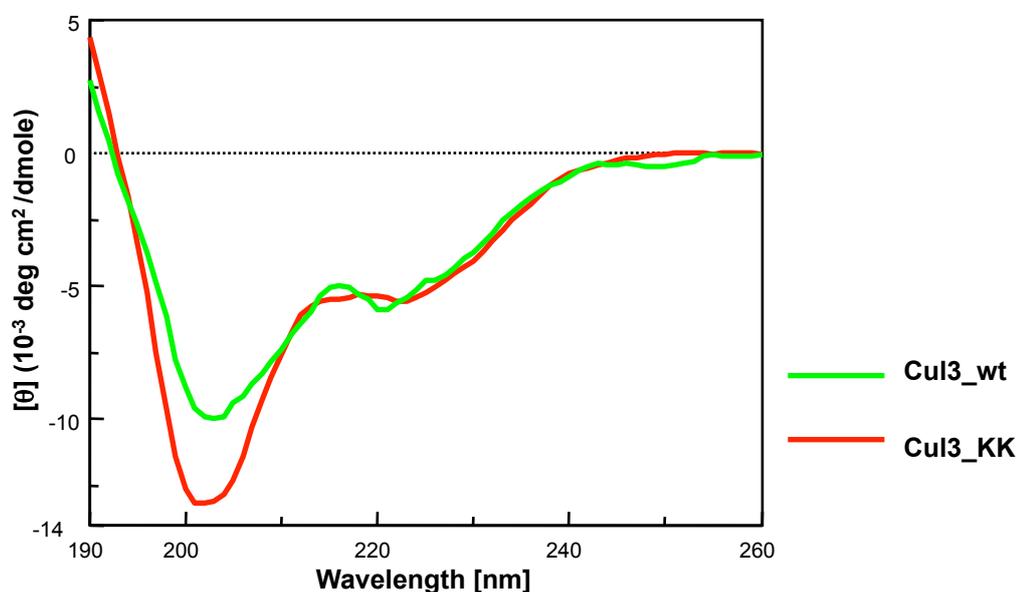


Figure 26. CD spectra of Cul3-wt (green line) and Cul3-KK (red line)

The analysis of data obtained by CD experiments (Table 9) highlighted that the ratios between the CD signal at 222 nm and the signal of the other minimum at lower wavelength are 0.49 and 0.42 for Cul3-wt peptide and Cul3-KK peptide, respectively. These values confirm that these peptides possess low helical contents.

Table 9

Peptide	$\theta_{222}$	$\theta_{\min}$	$\theta_{222}/\theta_{\min}$	$\theta_{190}$	cross
Cul3-wt	-4599	-9283	0,495	1146	192
Cul3-KK	-5560	-13190	0,422	4338	193

### 3.2.4 ELISA Assay

The ELISA assays were performed with biotinylated Cul3-wt, Cul3-KK and Cul3-Y(P) and KCTD11-Trx fused protein. Trx protein used as negative control.

The shape of the binding curve (Figure 27) shows that Cul3-wt peptide is able to bind to KCTD11 with a higher affinity than Cul3-KK and Cul3-Y(P) peptides [59].

These results could confirm the importance of Tyr58 and Tyr62 residues in the interaction in accordance with molecular modeling studies.

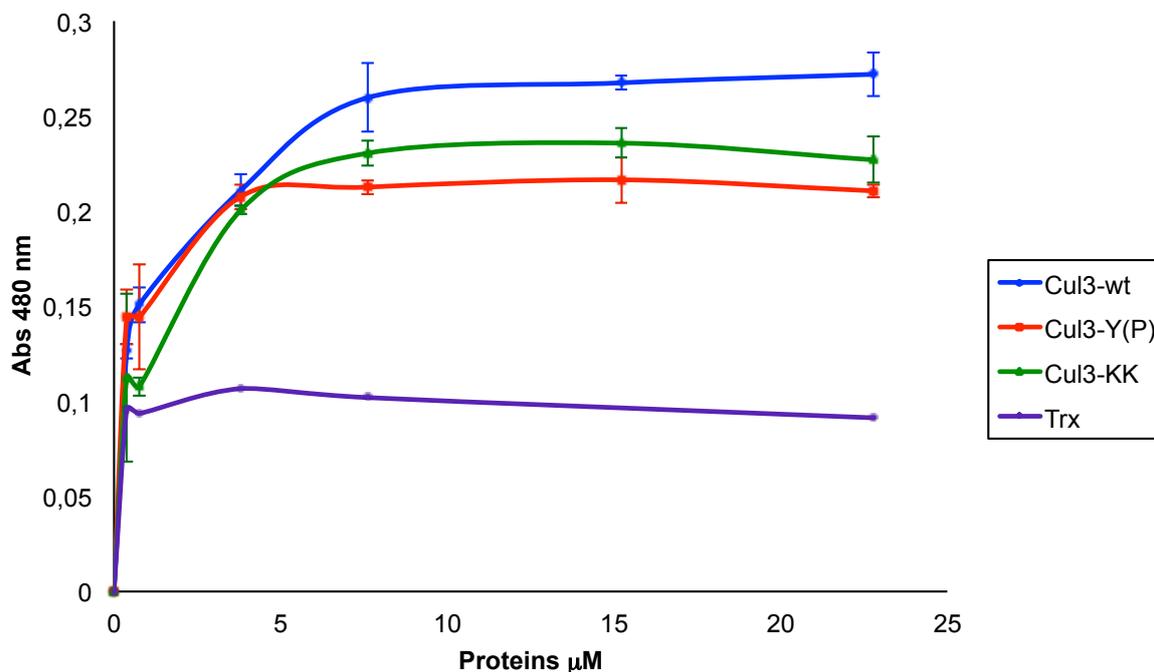


Figure 27. Binding curves obtained from ELISA assays on the proteins KCTD11 using the peptides BiotinCul3-wt (blue line), BiotinCul3-KK (yellow line) and BiotinCul3-Y(P) (pink line)

### 3.2.5 Design of Cul3-stapled peptides

With the aim to test the effect of the enhancement of the helical content of Cul3 analogues on binding affinity with KCTD11, two further peptides were designed: a cyclic peptide containing a hydrocarbon bridge in the center of native sequence of Cul3-wt peptide (Cul3-stapled) and its linear analogue (Cul3-stapline). The bridge was inserted substituting Leu57 and Ala61 located at the center of putative  $\alpha$ -helix of Cul3 with two 2(S)-(4'- pentenyl) alanine (S5) residues. The metathesis reaction by Grubb's catalyst allowed the formation of this hydrocarbon bond between the two C4' of side chains of the  $\alpha,\alpha$ -disubstituted amino acids.

Cul3-stapled and Cul3-stapline peptides were synthesized in acetylated and biotinylated forms to analyze both conformational character and ability to bind KCTD11 by ELISA assay, respectively.

Table 10

Peptide	Sequence
Cul3-stapline	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-S5-Tyr-Arg-Asn-S5-Tyr-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>
Cul3-stapled	Ac-Asn-Ser-Gly-Leu-Ser-Phe-Glu-Glu-S5-Tyr-Arg-Asn-S5-Tyr-Thr-Met-Val-Leu-His-Lys-NH <sub>2</sub>

**Red Line** S5 residues for Ring Closing Metathesis

**Blue Line** Tyr key residues for interaction between POZ/BTB and Cul3

### 3.2.6 Synthesis of Cul3 stapled

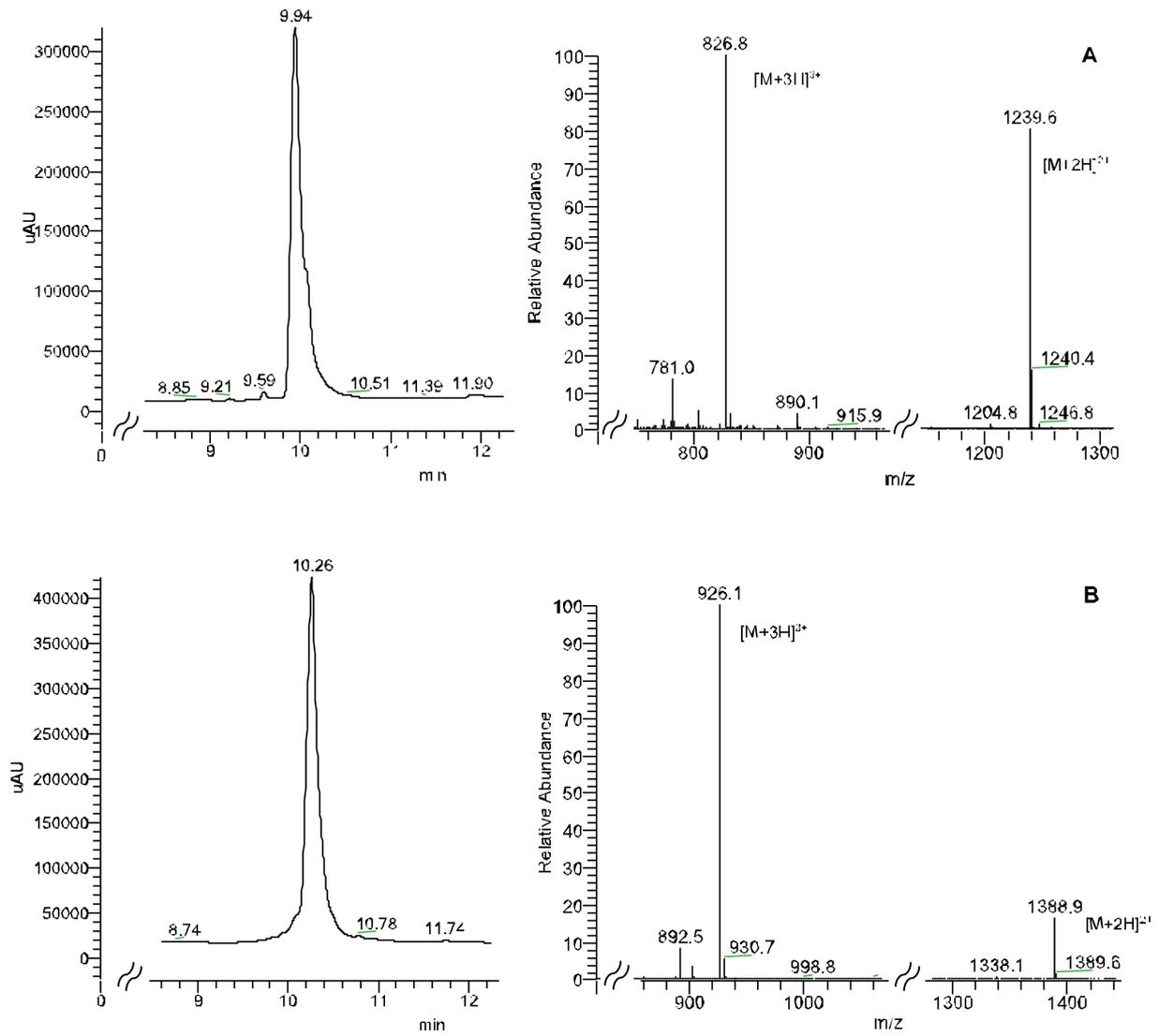
The Cul3-stapled, Cul3-stapline peptides and their biotinylated derivatives were synthesized by Fmoc solid phase strategy.

The metathesis reaction was performed on solid phase. The reaction was followed by LC/MS on specimens of peptide obtained after cleavage from the resin.

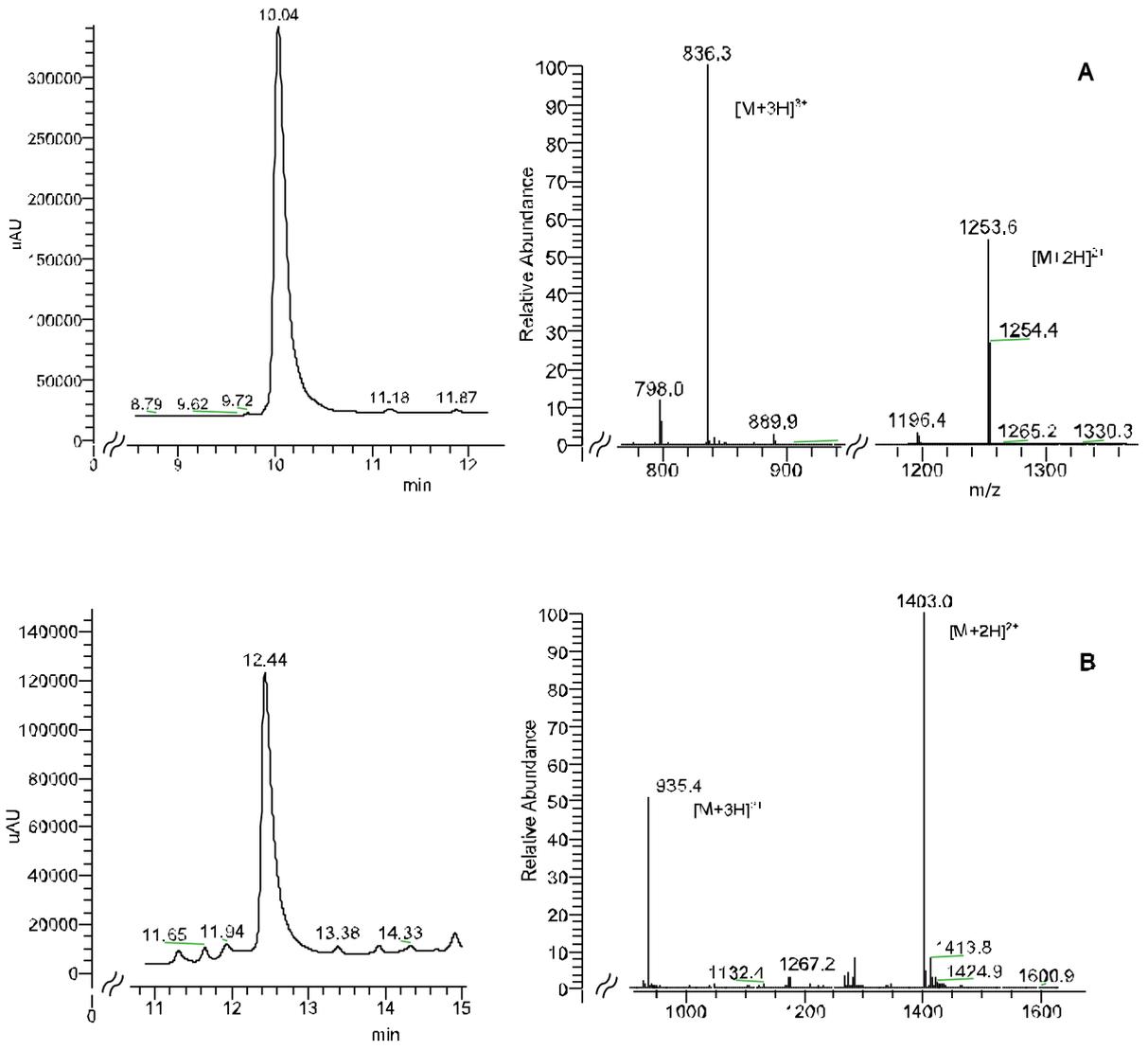
The identities and the purities (99%) of Cul3-stapled and Cul3-stapline peptides were assessed by LC-MS.

Table 11

Peptide	Yield (%)	Theoretical Mass (Da)	Molecular ion peak
AcCul3-stapled	16,7	2478,2	2477,2
BiotinylCul3-stapled	15,9	2777,2	2775,8
AcCul3-stapline	16,3	2506,3	2505,2
BiotinylCul3-stapline	15,7	2805,3	2804,4



**Figure 28. LC-MS (A) of AcCul3-stapled and (B) of BiotinCul3-stapled**



**Figure 29. LC-MS (A) of AcCul3-stapline of (B) BiotinCul3-stapline**

### 3.2.7 CD Studies

The CD spectra of Cul3-stapled and Cul3-stapline (Figure 30) are almost identical. They show two deep minima at ~207-208 nm and 222 nm with similar intensity, a cross over at 201 nm, and a maximum at a wavelength lower than 192 nm. The general features of the spectra of both peptides are indicative of a good helical content.

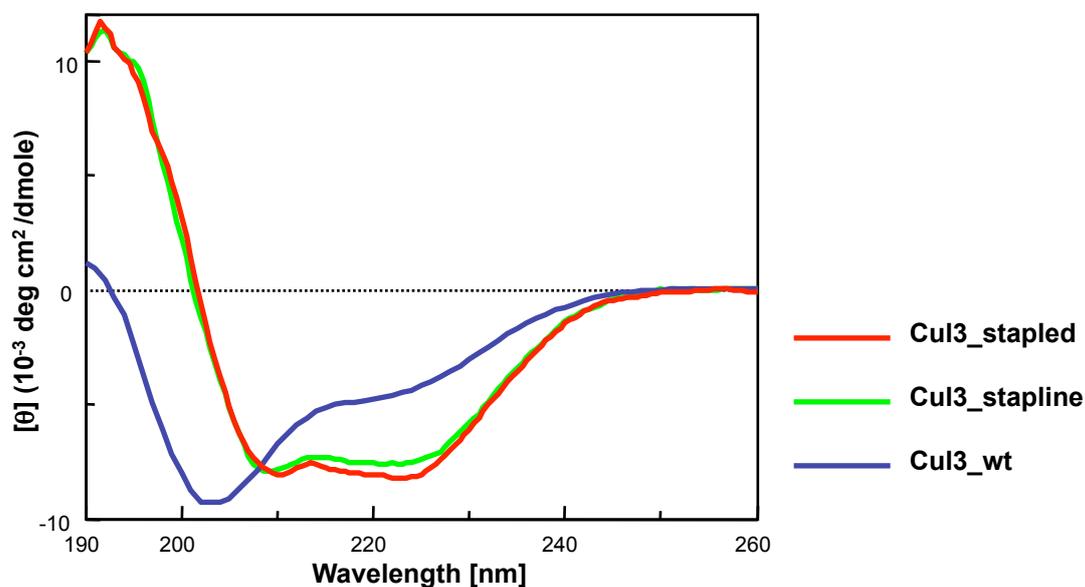


Figure 30. CD spectra of Cul3-stapled (red line), Cul3-stapline (green line), Cul3-wt (blue line)

The data obtained (Table 12) highlighted that the ratios between the CD signal at 222 nm and the signal of the other minimum at lower wavelength are 0.98 and 0.95 for Cul3-stapled and Cul3-stapline peptides, respectively.

Noticeably, these values indicate a higher  $\alpha$ -helix content of the peptides compared to Cul3-wt peptide.

Table 12

Peptide	$\theta_{222}$	$\theta_{\min}$	$\theta_{222}/\theta_{\min}$	$\theta_{190}$	cross
Cul3-stapled	-8065	-8234	0,979	11675	201
Cul3-stapline	-7570	-7924	0,955	11314	201

### 3.2.8 ELISA Assay

The assays were performed with biotinylated Cul3-stapled and Cul3-stapline and KCTD11-Trx fused protein. Trx protein was used as negative control. The trend of binding curve of ELISA assays (Figure 31) displayed that Cul3-stapled and Cul3- stapline bind to KCTD11 with similar affinity.

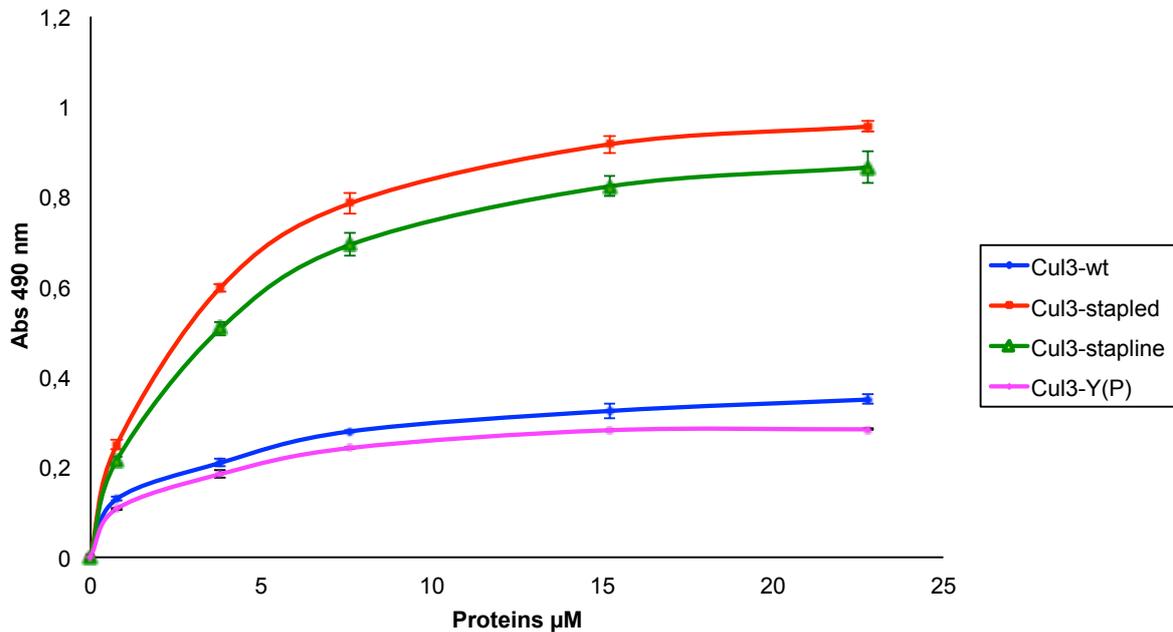


Figure 31. Binding curves obtained from ELISA assays on the proteins KCTD11 using the peptides BiotinCul3-wt (blue line), BiotinCul3-stapline (yellow line), BiotinCul3-stapled (pink line) and BiotinCul3-wt (cyan line)

#### 4. CONCLUSION AND FUTURE PERSPECTIVES

The present PhD thesis pertained to the development of new molecules able to interfere with relevant cellular activities.

The starting point was the analysis of molecular features at the basis of some PPI involving helical motifs. In particular, the research was focused on the design, the synthesis and the biological characterization of peptide molecules mimicking the helical surfaces of selected PPI. The design strategy consisted of keeping fixed the three dimensional conformation of the residues relevant for the interaction and introducing molecular tools to stabilize the helical secondary structure.

Malfunctions in transcriptional regulation are associated with many human diseases and there is considerable interest in biotechnological field in developing artificial transcription factors that must minimally contain a DBD and a TAD (Figure 32).

Designing artificial TADs in an efficient and predictable manner has proven difficult [60]. Efforts to design artificial TADs are hampered by the fact that TADs interact with multiple target proteins as part of their normal function and that there are only a limited number of high-resolution structures of TADs in complex with their target proteins.

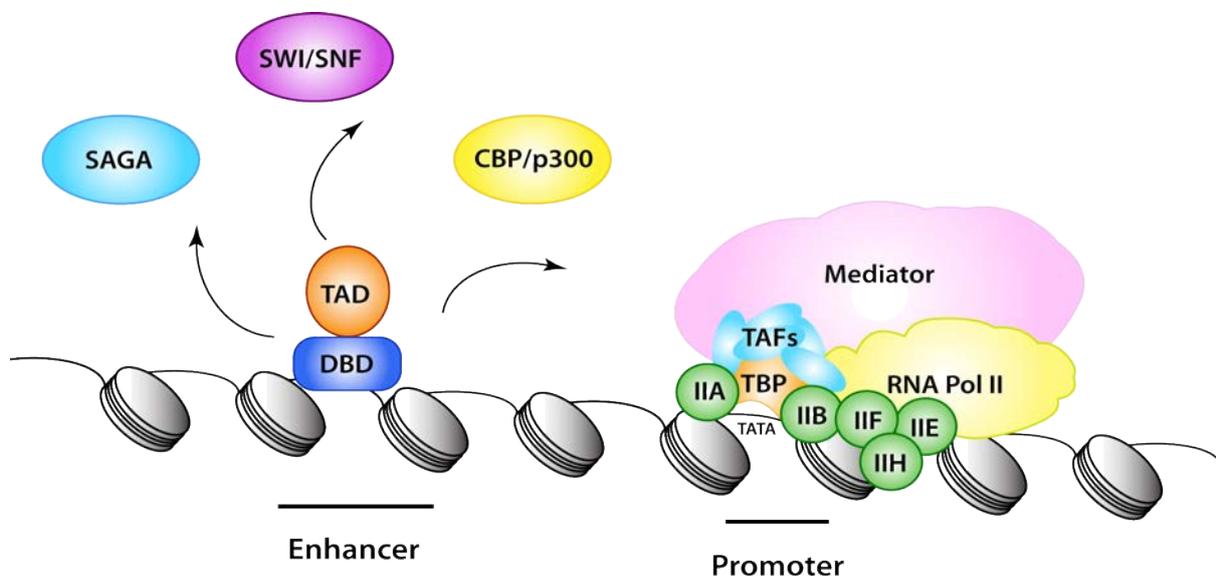


Figure 32. Activators in transcription

The p53 protein has a fundamental role in cellular survival due to its transcriptional activator function. Recent structural studies gave insights into interaction between p53 TAD2 and tfb1 subunit of TFIIF and contributed to the understanding of the way p53 protein works in transcriptional activation. On the basis of this information, peptide molecules mimicking the helical fragment of p53 TAD2, able to bind tfb1 and to function as TADs, were designed and synthesized. Our hypothesis was that if we could enhance the helical propensity of p53 TAD2, this would yield a more potent artificial TAD.

All the synthesized peptides showed an increased helical content relative to native p53-13 peptide as highlighted by CD experiments.

ITC experiments showed that E-Cap peptide has a higher binding affinity than other peptides and similar to p53<sub>25-65</sub> protein. In particular, peptide E-Cap forms a 9-residue  $\alpha$ -helix when in complex with tfb1 and binds along the same interface, as evident by NMR spectra. The binding interface is a shallow groove surrounded by positively-charged residues that help position the negatively-charged TAD so that its hydrophobic residues can participate in a series of van der Waals contacts with hydrophobic pockets dispersed along the groove.

Moreover, the increased *in vivo* potency of E-Cap relative to the native p53-13 peptide is directly linked to its enhanced helical stability since both the N-Capping motif and the two leucines *i*, *i*+3 positions are required for maximal activity.

The introduction of the hydrophobic interaction on the opposite face to the tfb1 interacting interface represents a potentially valuable strategy in designing artificial TADs with increased *in vivo* activity. Previous experimental studies with model peptides suggested that hydrophobic amino acids with an *i*, *i*+4 spacing pattern had a slightly higher helix stabilizing effect than those with an *i*, *i*+3 spacing pattern, but that either pattern can significantly increase the helical propensity of peptides relative to alanine.

In the case of the E-Cap peptide, mutation studies provide evidence that Leu6 can participate in *i*, *i*+3 interaction with Leu9 and *i*, *i*+4 interaction with Trp10 and thus both types of interactions are functioning. The fact that the NMR studies of E-Cap in complex demonstrates the presence of both the *i*, *i*+3 and the *i*, *i*+4 interactions further confirms the important role of these interactions for the *in vitro* binding and *in vivo* activity.

In conclusion, a new peptide with high content of helical conformation was designed using only proteinogenic amino acids. This result will be used to develop a new interesting biotechnological system such as "programmable ATF", linking obtained TAD to new synthetic DBD.

BTB/POZ domain-containing proteins are identified in diverse cellular locations and participate in different processes ranging from transcriptional repression/activation to cytoskeletal organization. In particular, KCTD are a class of protein containing BTB/POZ domain with different biological function as KCTD11 that inhibits tumor growth in medulloblastoma.

The understating of the structural features required for the interaction of the Cul3 with KCTDs is an intriguing challenge and is the basis for the future design of molecules with agonist/antagonist function for therapeutic applications.

To achieve this goal the attention was focused on KCTD11 that mediates interaction with Cul3 by its POZ/BTB domain at N-terminus. Starting from the proposed model of KCTD11-Cul3 interaction, peptide molecules mimicking the helical fragment of Cul3 49-68, were designed.

To evaluate the role of some aromatic residues of Cul3 in the interaction with KCTD11 two mutant peptides were developed.

Subsequently, in order to test if the enhance of the helical propensity match with a gain in binding affinity, opportunely constrained sequences were designed and synthesized. Cul3 stapled peptide, with a hydrocarbon bridge in central position of sequence, showed indeed a higher helical structure than the native Cul3-wt peptide. Preliminary ELISA assays indicate that both Cul3-wt and Cul3 stapled peptides were able to interact with KCTD11.

Future studies by ITC and NMR will allow the evaluation of the  $K_d$  and the molecular features governing the interaction with KCTD11.

## 5. BIBLIOGRAPHY

1. De Las Rivas, J. and C. Fontanillo, *Protein-protein interactions essentials: key concepts to building and analyzing interactome networks*. PLoS Comput Biol, 2010. **6**(6): p. e1000807.
2. Perkins, J.R., et al., *Transient protein-protein interactions: structural, functional, and network properties*. Structure, 2010. **18**(10): p. 1233-43.
3. Edwards, T.A. and A.J. Wilson, *Helix-mediated protein--protein interactions as targets for intervention using foldamers*. Amino Acids, 2011. **41**(3): p. 743-54.
4. Nooren, I.M. and J.M. Thornton, *Diversity of protein-protein interactions*. EMBO J, 2003. **22**(14): p. 3486-92.
5. Jones, S. and J.M. Thornton, *Principles of protein-protein interactions*. Proc Natl Acad Sci U S A, 1996. **93**(1): p. 13-20.
6. Moreira, I.S., P.A. Fernandes, and M.J. Ramos, *Hot spots--a review of the protein-protein interface determinant amino-acid residues*. Proteins, 2007. **68**(4): p. 803-12.
7. Thorn, K.S. and A.A. Bogan, *ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions*. Bioinformatics, 2001. **17**(3): p. 284-5.
8. Yin, H. and A.D. Hamilton, *Strategies for targeting protein-protein interactions with synthetic agents*. Angew Chem Int Ed Engl, 2005. **44**(27): p. 4130-63.
9. Keskin, O., B. Ma, and R. Nussinov, *Hot regions in protein--protein interactions: the organization and contribution of structurally conserved hot spot residues*. J Mol Biol, 2005. **345**(5): p. 1281-94.
10. Doig, A.J., *Stability and design of alpha-helical peptides*. Prog Mol Biol Transl Sci, 2008. **83**: p. 1-52.
11. Andrews, M.J.I., *Forming Stable Helical Peptides Using Natural and Artificial Amino Acids*. Tetrahedron, 1999. **55**: p. 32.
12. Bullock, B.N., A.L. Jochim, and P.S. Arora, *Assessing helical protein interfaces for inhibitor design*. J Am Chem Soc, 2011. **133**(36): p. 14220-3.
13. Keskin, O., et al., *Principles of protein-protein interactions: what are the preferred ways for proteins to interact?* Chem Rev, 2008. **108**(4): p. 1225-44.
14. Jochim, A.L. and P.S. Arora, *Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors*. ACS Chem Biol, 2010. **5**(10): p. 919-23.
15. Eichler, J., *Peptides as protein binding site mimetics*. Curr Opin Chem Biol, 2008. **12**(6): p. 707-13.
16. Garner, J. and M.M. Harding, *Design and synthesis of alpha-helical peptides and mimetics*. Org Biomol Chem, 2007. **5**(22): p. 3577-85.
17. Henchey, L.K., A.L. Jochim, and P.S. Arora, *Contemporary strategies for the stabilization of peptides in the alpha-helical conformation*. Curr Opin Chem Biol, 2008. **12**(6): p. 692-7.
18. Venkatraman, J., S.C. Shankaramma, and P. Balaram, *Design of folded peptides*. Chem Rev, 2001. **101**(10): p. 3131-52.
19. Phelan, J.C., *A General Method for Constraining Short Peptides to an alpha-Helical Conformation*. J Am Chem Soc, 1997. **119**(3): p. 455-460.
20. Leduc, A.M., et al., *Helix-stabilized cyclic peptides as selective inhibitors of steroid receptor-coactivator interactions*. Proc Natl Acad Sci U S A, 2003. **100**(20): p. 11273-8.

21. Blackwell, H.E., Grubbs, R. H., *Highly Efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis*. *Angew Chem Int Ed Engl*, 1998. **37**(23): p. 3281-3284.
22. Schafmeister, C.E., J. Po, and G.L. Verdine, *An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides*. *Journal of the American Chemical Society*, 2000. **122**(24): p. 5891-5892.
23. Walensky, L.D., et al., *Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix*. *Science*, 2004. **305**(5689): p. 1466-1470.
24. Jacobsen, O., et al., *Stapling of a 3(10)-Helix with Click Chemistry*. *Journal of Organic Chemistry*, 2011. **76**(5): p. 1228-1238.
25. Ingale, S. and P.E. Dawson, *On Resin Side-Chain Cyclization of Complex Peptides Using CuAAC*. *Organic Letters*, 2011. **13**(11): p. 2822-2825.
26. Horne, W.S. and S.H. Gellman, *Foldamers with Heterogeneous Backbones*. *Accounts of Chemical Research*, 2008. **41**(10): p. 1399-1408.
27. Wu, Y.D., et al., *Theoretical Analysis of Secondary Structures of beta-Peptides*. *Accounts of Chemical Research*, 2008. **41**(10): p. 1418-1427.
28. Becerril, J. and A.D. Hamilton, *Helix mimetics as inhibitors of the interaction of the estrogen receptor with coactivator peptides*. *Angew Chem Int Ed Engl*, 2007. **46**(24): p. 4471-3.
29. Cummings, C.G. and A.D. Hamilton, *Disrupting protein-protein interactions with non-peptidic, small molecule alpha-helix mimetics*. *Current Opinion in Chemical Biology*, 2010. **14**(3): p. 341-346.
30. Fisher, D.E., *The p53 tumor suppressor: critical regulator of life & death in cancer*. *Apoptosis*, 2001. **6**(1-2): p. 7-15.
31. Morrison, R.S. and Y. Kinoshita, *The role of p53 in neuronal cell death*. *Cell Death and Differentiation*, 2000. **7**(10): p. 868-879.
32. Levine, A.J., *p53, the cellular gatekeeper for growth and division*. *Cell*, 1997. **88**(3): p. 323-331.
33. Romer, L., et al., *p53 - A natural cancer killer: Structural insights and therapeutic concepts*. *Angewandte Chemie-International Edition*, 2006. **45**(39): p. 6440-6460.
34. Oren, M., *Decision making by p53: life, death and cancer*. *Cell Death and Differentiation*, 2003. **10**(4): p. 431-42.
35. Di Lello, P., et al., *Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIID and the activation domain of p53*. *Mol Cell*, 2006. **22**(6): p. 731-40.
36. Leveillard, T., et al., *Functional interactions between p53 and the TFIID complex are affected by tumour-associated mutations*. *EMBO J*, 1996. **15**(7): p. 1615-24.
37. Di Lello, P., et al., *p53 and TFIIDalpha share a common binding site on the Tfb1/p62 subunit of TFIID*. *Proc Natl Acad Sci U S A*, 2008. **105**(1): p. 106-11.
38. Langlois, C., et al., *NMR structure of the complex between the Tfb1 subunit of TFIID and the activation domain of VP16: structural similarities between VP16 and p53*. *J Am Chem Soc*, 2008. **130**(32): p. 10596-604.
39. Zollman, S., et al., *The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila*. *Proc Natl Acad Sci U S A*, 1994. **91**(22): p. 10717-21.
40. Bardwell, V.J. and R. Treisman, *The POZ domain: a conserved protein-protein interaction motif*. *Genes Dev*, 1994. **8**(14): p. 1664-77.

41. Melnick, A., et al., *In-depth mutational analysis of the promyelocytic leukemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions*. Mol Cell Biol, 2000. **20**(17): p. 6550-67.
42. Kang, M., et al., *Small potassium ion channel proteins encoded by chlorella viruses*. Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5318-24.
43. Aravind, L. and E.V. Koonin, *Fold prediction and evolutionary analysis of the POZ domain: structural and evolutionary relationship with the potassium channel tetramerization domain*. J Mol Biol, 1999. **285**(4): p. 1353-61.
44. Ahmad, K.F., C.K. Engel, and G.G. Prive, *Crystal structure of the BTB domain from PLZF*. Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12123-8.
45. Kreuzsch, A., et al., *Crystal structure of the tetramerization domain of the Shaker potassium channel*. Nature, 1998. **392**(6679): p. 945-948.
46. Dementieva, I.S., et al., *Pentameric assembly of potassium channel tetramerization domain-containing protein 5*. J Mol Biol, 2009. **387**(1): p. 175-91.
47. Van Bogaert, P., et al., *Mutation of a potassium channel-related gene in progressive myoclonic epilepsy*. Ann Neurol, 2007. **61**(6): p. 579-86.
48. Jesch, S.A., et al., *Mitotic phosphorylation of Golgi reassembly stacking protein 55 by mitogen-activated protein kinase ERK2*. Mol Biol Cell, 2001. **12**(6): p. 1811-7.
49. Feinstein, T.N. and A.D. Linstedt, *Mitogen-activated protein kinase kinase 1-dependent Golgi unlinking occurs in G2 phase and promotes the G2/M cell cycle transition*. Mol Biol Cell, 2007. **18**(2): p. 594-604.
50. Canettieri, G., et al., *Histone deacetylase and Cullin3-REN(KCTD11) ubiquitin ligase interplay regulates Hedgehog signalling through Gli acetylation*. Nat Cell Biol, 2010. **12**(2): p. 132-42.
51. Bayon, Y., et al., *KCTD5, a putative substrate adaptor for cullin3 ubiquitin ligases*. FEBS J, 2008. **275**(15): p. 3900-10.
52. Azizieh, R., et al., *Progressive myoclonic epilepsy-associated gene KCTD7 is a regulator of potassium conductance in neurons*. Mol Neurobiol, 2011. **44**(1): p. 111-21.
53. Correale, S., et al., *Molecular organization of the cullin E3 ligase adaptor KCTD11*. Biochimie, 2011. **93**(4): p. 715-24.
54. Zheng, N., et al., *Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex*. Nature, 2002. **416**(6882): p. 703-9.
55. Aurora, R. and G.D. Rose, *Helix capping*. Protein Sci, 1998. **7**(1): p. 21-38.
56. Vuilleumier, S. and M. Mutter, *Synthetic peptide and template-assembled synthetic protein models of the hen egg white lysozyme 87-97 helix: importance of a protein-like framework for conformational stability in a short peptide sequence*. Biopolymers, 1993. **33**(3): p. 389-400.
57. Diana, D., et al., *Structural determinants of the unusual helix stability of a de novo engineered vascular endothelial growth factor (VEGF) mimicking peptide*. Chemistry, 2008. **14**(14): p. 4164-6.
58. Rush, J., et al., *Immunoaffinity profiling of tyrosine phosphorylation in cancer cells*. Nat Biotechnol, 2005. **23**(1): p. 94-101.
59. Pirone, L., et al., *Design, synthesis and characterization of a peptide able to bind proteins of the KCTD family: implications for KCTD-cullin 3 recognition*. J Pept Sci, 2011. **17**(5): p. 373-6.

60. Lee, L.W. and A.K. Mapp, *Transcriptional switches: chemical approaches to gene regulation*. J Biol Chem, 2010. **285**(15): p. 11033-8.

## SUPPLEMENTARY INFORMATION

### PUBLICATIONS

Del Gatto A, De Simone M, de Paola I, Saviano M, Zaccaro L. (2011)  
Investigation of the Best Conditions to Obtain c(RGDfK) Peptide on Solid Phase  
INTERNATIONAL JOURNAL OF PEPTIDE RESEARCH AND THERAPEUTICS  
17:39-45 doi: 10.1007/s10989-010-9238-4

Pirone L, Correale S, de Paola I., Zaccaro L, De Simone G, Vitagliano L, Pedone E,  
Di Gaetano S (2011)  
Design, synthesis and characterization of a peptide able to bind proteins of the KCTD  
family: implications for KCTD-cullin 3 recognition  
JOURNAL OF PEPTIDE SCIENCE, vol. 17; p. 373-376 doi:10.1002/psc.1366

Langlois, C; Del Gatto, A; Arseneault, G; Lafrance-Vanasse, J; De Simone, M;  
Morse, T; de Paola, I; Lussier-Price, M; Legault, P; Pedone, C; Zaccaro, L;  
Omichinski, J.  
"Structure-Based Design of a Potent Artificial Transactivation Domain Based on p53"  
Journal of the American Chemical Society  
Submitted Manuscript (Minority Revision) ID: ja-2011-08999e

### POSTER COMMUNICATION

Del Gatto A., Avvakumova S., de Paola I., De Simone M., Leone M., Saviano M.,  
Pedone C., Scari G., Porta F., Zaccaro L.  
"A novel dual function peptide for gold nanoparticles stabilization and integrin  
targeting"  
XXIV SCI congress, Lecce, Italy, September, 2011.

Pirone L., Correale S., Di Gaetano S., de Paola I., Zaccaro L., De Simone G.,  
Vitagliano L., Pedone C. and Pedone E.  
A Cul3-derived peptide binding different members of the KCTD family  
12th Naples Workshop on Bioactive Peptides, Napoli, Italy, June, 2010

Pirone L, Correale S, de Paola I, Zaccaro L, De Simone G, Vitagliano L, Pedone E,  
Pedone C, Di Gaetano S  
KCTD7: a new substrate-specific adaptor for Cul3 involved in Progressive Myoclonic  
Epilepsy  
Italian Chemical Society Division of Systems Biology, San Vito di Cadore, Italy  
(2010)

### ATTENDED LABORATORIES

Dipartimento di Scienze Biologiche - IBB/CNR Università degli studi di Napoli  
"Federico II" - Via Mezzocannone, 16 Napoli

11/2010-12/2010

Sanford-Burnham Medical Research Institute, lab Prof. Pellecchia M.

North Torrey Pines Road, La Jolla, Ca (USA)

"Virtual screening for design of new selective molecules for PDK1"

## Investigation of the Best Conditions to Obtain c(RGDfK) Peptide on Solid Phase

Annarita Del Gatto · Mariarosaria De Simone ·  
Ivan de Paola · Michele Saviano · Laura Zaccaro

Accepted: 30 December 2010 / Published online: 9 January 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** The cyclic pentapeptide c(RGDfK) is a high affinity ligand of  $\alpha$ V $\beta$ 3 integrin. It was an analog of Cilengitide (EMD 121974) developed to be employed as tracer for cancer diagnosis and therapy by functionalisation of its Lys side-chain. Solution-phase and solid-phase synthetic approaches were previously reported. In the attempt to improve solid-phase synthesis of the cyclopeptide circumventing cyclodimerisation reactions, a systematic study of the synthetic conditions was performed, evaluating and optimising parameters directly involved in the ring closure step. The three-dimensional orthogonal solid-phase strategy developed in this study yields the desired c(RGDfK) peptide with no cyclodimerisation by-products. The protocols described allow the modification of the peptide directly on the solid support in order to obtain novel derivatives for biomedical applications.

**Keywords** Optimisation · Synthesis · Reaction · Pharmaceuticals

### Abbreviations

All	Allyl
tBu	Tert-butyl
DCM	Dichloromethane

A. Del Gatto · L. Zaccaro (✉)  
Istituto di Biostrutture e Bioimmagini—CNR,  
Via Mezzocannone 16, 80134 Naples, Italy  
e-mail: lzaccaro@unina.it

M. De Simone · I. de Paola  
Dipartimento delle Scienze Biologiche, Università degli Studi di  
Napoli “Federico II”, Via Mezzocannone 16, 80134 Naples,  
Italy

M. Saviano  
Istituto di Cristallografia—CNR, Via Giovanni Amendola 122/O,  
70126 Bari, Italy

DIPEA	<i>N,N</i> -diisopropylethylamine
DMF	<i>N,N</i> -dimethylformamide
EDT	Ethanedithiol
f	<i>D</i> -phenylalanine
LC/ESI/MS	Liquid chromatography electrospray ionisation mass spectrometry
Fmoc	9H-fluoren-9- ylmethyloxycarbonyl
HBTU	2-(1H-benzotriazole-1-yl)- 1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	1-Hydroxybenzotriazole
ivDde	1-(4,4-Dimethyl-2,6-dioxo- cyclohexylidene)-3-methyl- butyl
MALDI-TOF spectrometry	Matrix-assisted laser desorption ionization time- of-flight mass spectrometry
MeIm	1-Methylimidazole
MSNT	(1-(Mesitylene-2-sulfonyl)- 3-nitro-1H-1,2,4-triazole)
Mtt	Methyltrityl
Pbf	2,2,4,6,7-Pentamethyl-2, 3-dihydrobenzofuran- 5-sulfonyl
PEG	Polyethylene glycol
PyBop	Benzotriazole-1-yl-oxy-tris- pyrrolidino-phosphonium hexafluorophosphate
RP-HPLC	Reverse phase-high performance liquid chromatography
TFA	Trifluoroacetic acid
TIS	Triisopropylsilane

## Introduction

RGD-peptides are known to be used for selective targeting and delivery of drugs to tumour cells. This is due to the ability of RGD motif to interact with integrin membrane receptors, such as  $\alpha$ V $\beta$ 3 that is closely associated with tumour progression, playing an important role in tumour-induced angiogenesis and metastasis (Brooks et al. 1994a, b; Felding-Habermann et al. 2001; Reinmuth et al. 2003; Zhao et al. 2007; Eliceiri and Chersesh 1999).

Based on the RGD sequence present in the  $\alpha$ V $\beta$ 3 natural ligands, a series of small RGD linear and cyclic peptides was designed to antagonise the function of this integrin (Haubner et al. 1997; Zaccaro et al. 2007, 2009) and, therefore, to promote tumour regression by apoptosis of angiogenic blood vessels (Brooks et al. 1994a, b) and to prevent metastases (Clezzardin 1998). Kessler and coworkers, for example, developed a new cyclic peptide c(RGDfNMeV) (Cilengitide, EMD 121974) (Dechantsreiter et al. 1999) which entered clinical phase III studies as an angiogenesis inhibitor for glioblastoma therapy (Eskens et al. 2003; Haubner et al. 2005; Friess et al. 2006; Beekman et al. 2006; Nabors et al. 2007; Garanger et al. 2007; MacDonald et al. 2008). Furthermore a large variety of labeled analogs was developed in order to employ them as tracers for cancer diagnosis and therapy (Zaccaro et al. 2009). To this end, RGD cyclic peptides carrying a Lys residue were developed in order to use the Lys side chain to introduce several tracing molecules. In particular, c(RGDfK) was synthesised on solid-phase by Kessler and coworkers (Haubner et al. 1996) by employing an *o*-chlorotriptyl chloride resin and a Fmoc strategy. In this method the cleavage of the fully protected linear peptide from the support and cyclisation under high dilution conditions were described in order to avoid cyclodimerisation and cyclooligomerisation side reactions. With the aim to circumvent solubility problems and intermediate purification, later on, other groups reported the on-resin cyclisation and lysine functionalisation (Boturyn and Dumy 2001; McCusker et al. 2002).

It has been described that the cyclisation reaction represents the key point in the synthesis of constrained head-to-tail cyclopeptides (Rovero 2000; Lambert et al. 2001) because of the high tendency of the corresponding linear precursors to oligomerise (Bodansky 1984). Although solid-phase cyclisation has represented a step forward respect with the solution phase counterpart in circumventing cyclodimerisation and/or cyclooligomerisation side reactions, it does not always completely solve the problem.

On the other hand, having a cyclic peptide as single product of reaction is a desirable goal in order to obtain more complex systems starting from it directly on solid phase. Many examples of c(RGDfK) conjugated with

chelator for radionuclides (Zaccaro et al. 2007, 2009) or functionalised with fluorescent probes (Goldshaid et al. 2010) are reported in literature.

In this paper a systematic study on the synthesis of c(RGDfK) peptide, by on-resin head-to-tail cyclisation, is described by using the Fmoc/tBu/OAll three-dimensional orthogonal solid-phase protection scheme. A comparative study was undertaken with the aim to evaluate the best combination of solid support and amino acid protecting groups that could play an important role in the ring-closure reaction, avoiding the possible cyclodimerisation side reaction.

## Materials and Methods

### Materials

Polypropylene reaction vessels and sintered polyethylene frits were supplied by Alltech Italia. Wang resin, DPhe amino acid, MeIm, TFA and scavengers were purchased from Fluka; NovaSyn TGA and 2-chlorotriptyl chloride resins, coupling reagents and all amino acids were from Novabiochem. DIPEA was purchased from Romil; piperidine from Biosolve; PhSiH<sub>3</sub> and Pd(PPh<sub>3</sub>)<sub>4</sub> from Sigma-Aldrich.

### Peptide Synthesis and Cyclisation Reaction

Assembly of fully protected peptides was carried out manually in a polypropylene reaction vessel fitted with a sintered polyethylene frit using the Fmoc/tBu solid-phase strategy (0.1 mmol). The syntheses were carried out on Wang, NovaSyn TGA and 2-chlorotriptyl chloride resins (loading 1, 0.26 and 1.3 mmol/g, respectively), using all standard amino acids except for D-Phe. Resins were swollen in DCM for 30 min prior to synthesis. The first amino acid was bound to the supports carrying hydroxymethyl groups by treatment with Fmoc-Asp(OAll)-OH (5 equiv)/MSNT (5 equiv)/MeIm (3.75 equiv) in DCM for 3 h. In the case of trityl resin, the loading was performed by using 1.2 equiv of Fmoc-Asp(OAll)-OH respect with the resin and 4 equiv of DIPEA relative to the amino acid in DCM for 4 h. In all cases the yield of the loading after the anchorage of first amino acid onto the resin was 100%. Fmoc protecting group was removed by treatment with 30% piperidine in DMF (3 × 10 min). Coupling reactions of the following four amino acids (Fmoc-Gly-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Lys(Mtt)-OH or Fmoc-Lys(ivDde)-OH, Fmoc-DPhe-OH) were performed by using 10 equiv of Fmoc protected derivative activated in situ with HBTU (9.8 equiv)/HOBT (9.8 equiv)/DIPEA (20 equiv) in DMF for 30 min. The coupling efficiency was assessed by Kaiser

test. Each step was followed by resin washings ( $3 \times 5$  min). Before the final Fmoc deprotection, selective  $\alpha$ -carboxyl deprotection of the Asp residue from the allyl group was carried out by treatment of the peptidyl resin with  $\text{PhSiH}_3$  (24 equiv)/ $\text{Pd}(\text{PPh}_3)_4$  (0.25 equiv) in DCM ( $3 \times 30$  min). The cyclisation between  $\alpha\text{NH}$  of D-Phe and  $\alpha\text{CO}$  of Asp was performed with PyBop (1.5 equiv)/HOBt (1.5 equiv)/DIPEA (2 equiv) in DMF for 5 h at millimolar pseudo-dilution.

Finally the peptides were cleaved off the resin and deprotected using a mixture of TFA/ $\text{H}_2\text{O}$ /EDT/TIS (94:2.5:2.5:1 v/v/v/v) for 3 h. The resins were then filtered, and the white solid peptides were obtained by precipitation from cold anhydrous diethyl ether. In the case of the peptides containing Lys residue protected with ivDde, this group was removed by using a solution of 2% hydrazine in DMF ( $10 \times 2$  min) before cleavage reaction.

### Peptide Purification and Characterisation

The crude products were purified by preparative RP-HPLC on the Shimadzu LC-8A system, equipped with an UV-Vis detector SPD-10A using a Phenomenex Jupiter Proteo column ( $21.2 \times 250$  mm;  $4 \mu\text{m}$ ;  $90 \text{ \AA}$ ) and a linear gradient of  $\text{H}_2\text{O}$  (0.1% TFA)/ $\text{CH}_3\text{CN}$  (0.1% TFA) from 5 to 70% of  $\text{CH}_3\text{CN}$  (0.1% TFA) in 30 min at a flow rate of 20 ml/min. The purified peptides were analysed and characterised by an ESI-LC-MS instrument (ThermoFinnigan) equipped with a diode array detector combined with an electrospray ion source and a quadrupole mass analyzer, using a Phenomenex Jupiter Proteo column ( $4.60 \times 150$  mm;  $4 \mu\text{m}$ ;  $90 \text{ \AA}$ ) and the linear gradient of the preparative scale at a flow rate of 0.8 ml/min. The peptide characterization was also performed by MALDI-TOF spectrometry on a MALDI-TOF Voyager-DE (Perseptive Biosystem) spectrometer.

## Result and Discussion

A systematic study on the synthesis of c(RGDfK), by evaluating and optimising parameters directly involved in the cyclisation step to circumvent cyclodimerisation reactions, is reported. To this end, six syntheses of RGD-containing cyclopeptides were performed by using the three-dimensional protection scheme Fmoc/tBu/OAll to allow the orthogonal deprotection of the C- and N-termini for the subsequent on-resin head-to-tail cyclisation. Three different resins were used (Wang (I), NovaSyn TGA (II) and 2-chlorotriyl chloride (III)), for each of them two syntheses were performed by using Fmoc-Lys(Mtt)-OH and Fmoc-Lys(ivDde)-OH residues, respectively.

All the peptide syntheses were started by anchoring the side-chain of an Asp residue to the chosen resin and by

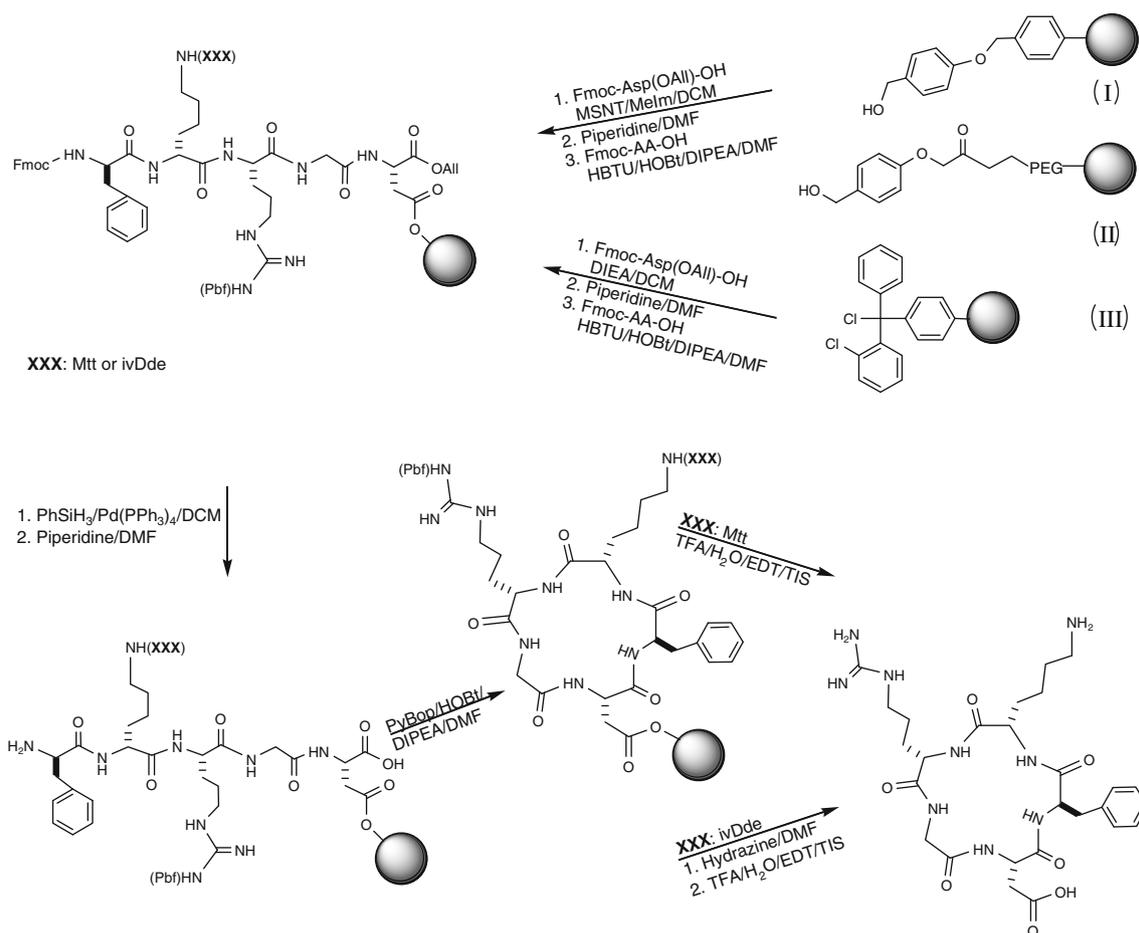
coupling the following amino acids by HBTU/HOBt/DIPEA procedure according to the Fmoc solid-phase chemistry (Fig. 1).

Once completed the synthesis, selective  $\alpha$ -carboxyl deprotection of the Asp residue from the allyl group was carried out with  $\text{PhSiH}_3$ / $\text{Pd}(\text{PPh}_3)_4$  in DCM, and the Fmoc group of the N-terminal position of the peptide was removed. Head-to-tail cyclisation between the  $\alpha\text{NH}$  of D-Phe and  $\alpha\text{CO}$  of Asp was performed with PyBop/HOBt/DIPEA in DMF. Final deprotection and cleavage from the resin were achieved with TFA and scavengers. Purification step was realized by RP-HPLC and purity and identity of the peptides were confirmed by ESI-LC-MS and MALDI-TOF mass spectrometry.

The first synthesis was performed on Wang resin employing Fmoc-Lys(Mtt)-OH residue. The crude peptide obtained gave two chromatographic peaks at RP-HPLC coupled with ESI-MS (Fig. 2). The largest peak ( $t_R = 11.08$  min) corresponded to the desired monomeric form of c(RGDfK), as shown by the mass ion  $[\text{M} + \text{H}]^+$  at  $604.31 \text{ m/z}$  (calculated value  $603.68 \text{ m/z}$ ), by the related diprotonated form  $[\text{M} + 2\text{H}]^{2+}$  at  $303.31 \text{ m/z}$ , and by a species at  $1,208.10 \text{ m/z}$  that could be assigned to an adduct formed by two c(RGDfK) molecules ( $\text{M}_2$ ) (Fig. 2b). The second peak at 11.97 min was about 35% of the first, and can be assigned to the cyclodimerisation product of c(RGDfK) ( $2\text{M}$ ), as the spectrum indicated the presence of three mass values potentially corresponding to  $[2\text{M} + \text{H}]^+ = 1,207.91 \text{ m/z}$  (calculated value  $1,207.36 \text{ m/z}$ ), and to the  $[2\text{M} + 2\text{H}]^{2+}$  and  $[2\text{M} + 3\text{H}]^{3+}$  forms at  $402.42$  and  $604.07 \text{ m/z}$ , respectively (Fig. 2c). In order to confirm the hypothesis of the obtainment of dimeric form of c(RGDfK), MALDI-TOF mass spectra of the purified fractions were performed (Fig. 3). The spectrum relative to the chromatographic peak at 11.08 min gave the monoprotinated ( $[\text{M} + \text{H}]^+ = 603.474$ ) and the sodium ( $[\text{M} + \text{Na}]^+ = 625.476$ ) adduct ions of the cyclomeric form (Fig. 3a); while the one at 11.97 min gave the monoprotinated ion at  $1,206.84 \text{ m/z}$ , and one and two sodium adduct ions of the cyclodimeric species at  $1,228.66$  and  $1,250.48 \text{ m/z}$ , respectively (Fig. 3b).

The same result was obtained on Wang resin when an ivDde-protected Lys residue was used in place of Fmoc-Lys(Mtt)-OH, and even changing pseudo-dilution conditions and coupling agents for cyclisation step.

On the basis of this result, c(RGDfK) synthesis using a different kind of resin was carried on. The resin chosen was NovaSyn TGA that presents a benzyl alcohol function as Wang resin with more a PEG spacer. The chromatographic profile of the crude peptide obtained from this resin, used together with Mtt-protected Lys, again indicated the presence of two peaks, as reported above, with the second about 40% of the first. On the contrary, the combination of



**Fig. 1** Synthetic route of c(RGDfK)

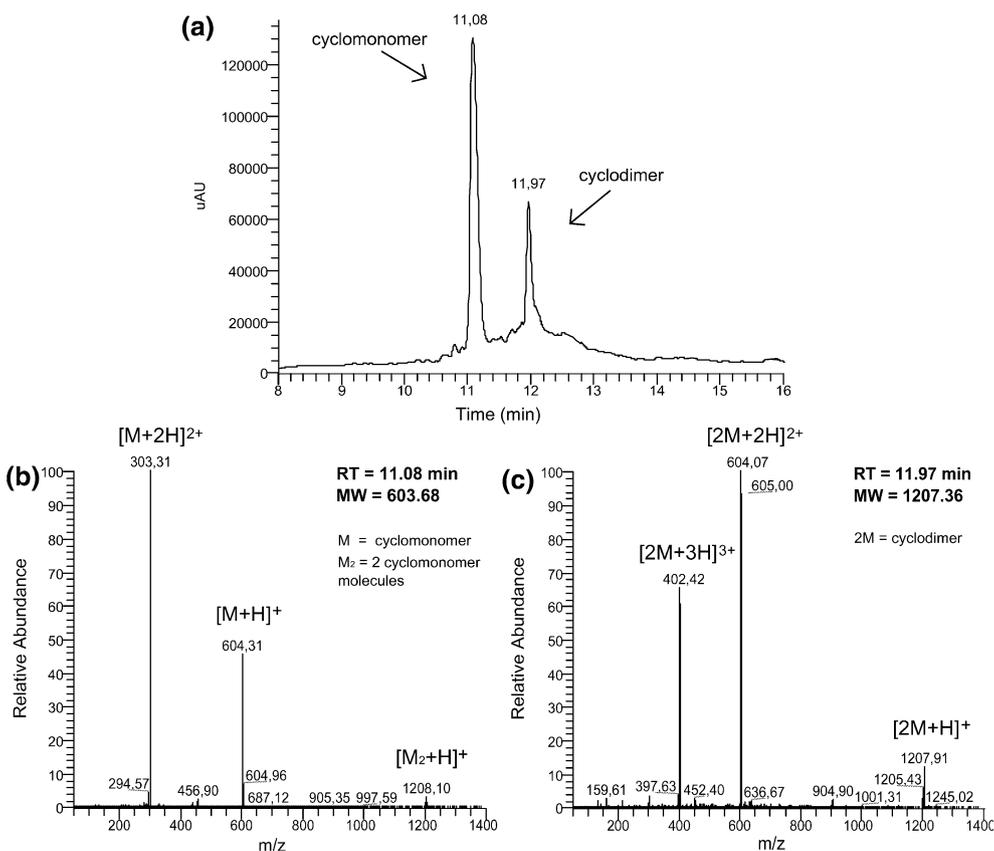
NovaSyn TGA with ivDde-protected Lys gave only the desired monomeric c(RGDfK) form as shown by mono-protonated  $[M + H]^+ = 604.51 \text{ m/z}$  and diprotonated  $[M + 2H]^{2+} = 303.60 \text{ m/z}$  molecular ions (Fig. 4). The yield of c(RGDfK) was 66% based on estimated loading of peptide-resin.

In the attempt to provide a feasible explanation for the obtained results by using different combinations of resin-Lys(protecting group), the two syntheses were repeated on a third kind of resin. The 2-chlorotrityl chloride resin was chosen with the aim to investigate the consequences of employing a resin strongly hindered and chemically different respect with those previously used. In this case, the RP-HPLC analysis of the cyclised product obtained evidenced the presence of a single chromatographic peak of c(RGDfK), both using Lys(Mtt) and Lys(ivDde) residues, just as for NovaSyn TGA resin combined with Lys(ivDde). Nevertheless it has to be stressed that this kind of solid support is not compatible with the use of Lys(Mtt) as the mild acid deprotection conditions of Mtt group lead to peptide cleavage from the resin.

Putting together all these data a possible explanation for the different behaviour observed can benefit of the following considerations. In the process of bending to give head-to-tail cyclisation, the peptide sequence bumps into a steric hindrance more or less prominent depending on the type of resin, Lys protecting group and combination of the two factors. In more detail, in the case of NovaSyn TGA resin together with Lys(Mtt) residue, because of the bulky methyltrityl group, the peptide sequence experiences an obstacle in the intracyclisation step that leads to a certain percentage of inter-cyclic- respect with the intracyclic-product. On the other side, being ivDde a relatively small protecting group compared with Mtt, the steric hindrance is definitely less prominent, therefore the intracyclisation is the favored reaction.

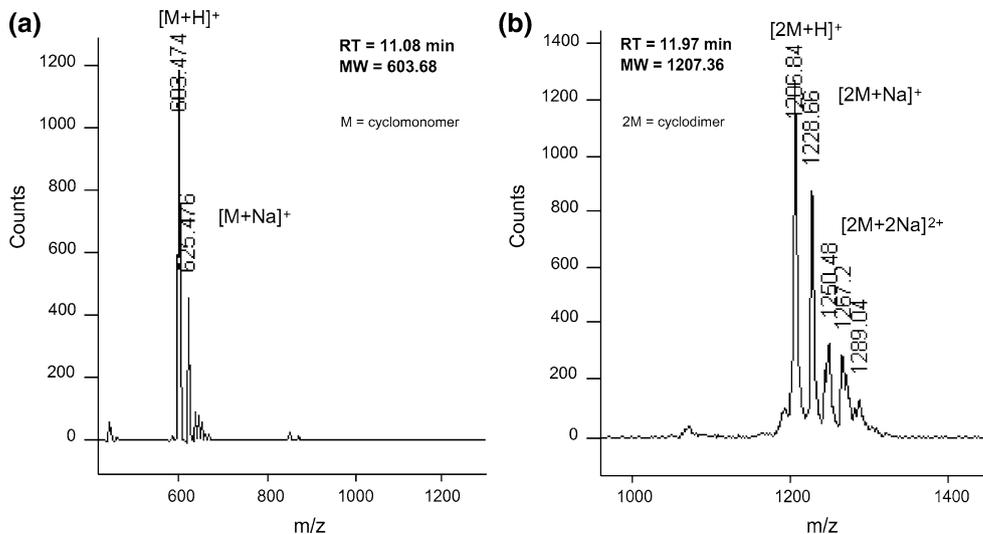
Interestingly when 2-chlorotrityl chloride is used, the bulky triphenyl group of the resin totally prevents inter-cyclisation side reaction both with Mtt and ivDde groups, giving only the desired cyclomonomeric-product.

Finally about Wang resin, even if this presents chemical structure similar to NovaSyn TGA that would have



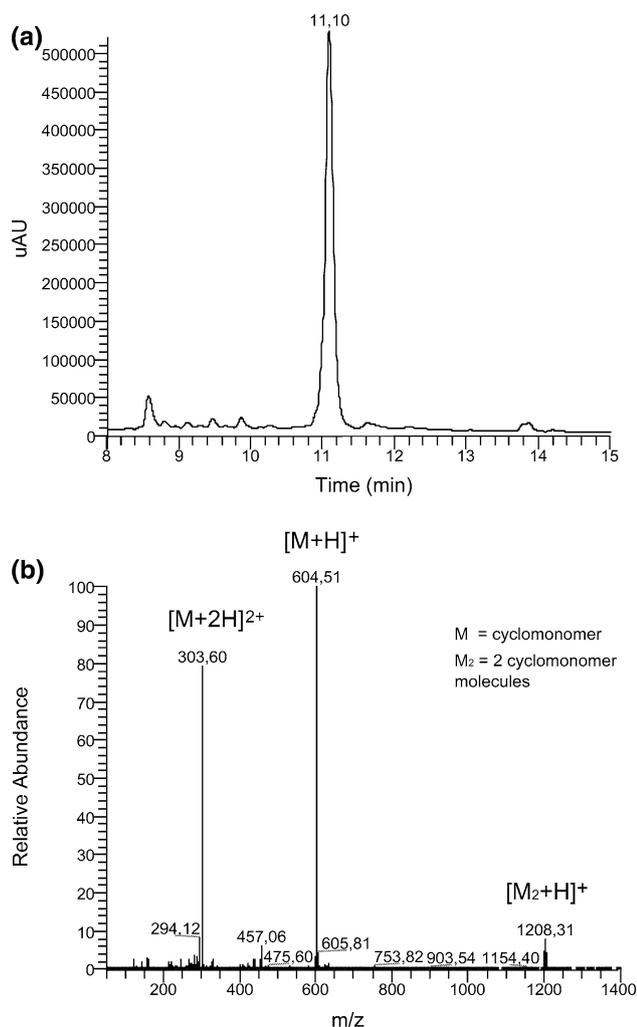
**Fig. 2** a RP-HPLC chromatographic profile of crude c(RGDfK) peptide obtained from the synthesis on Wang resin using Mtt as Lys-protecting group. b and c ESI-MS profiles corresponding to peaks at  $t_R$ : 11.08 and 11.97 min, respectively

**Fig. 3** MALDI-TOF spectra of purified fractions of c(RGDfK) peptide corresponding to peaks at a  $t_R$ : 11.08 min and b 11.97 min, respectively



foreseen an analogous behaviour, the intercyclc side product is obtained with both Lys-protecting groups. This could be probably ascribed to the higher loading of Wang respect with NovaSyn TGA resin that gets peptide chains closer in the space making the inter-reaction easier to occur.

In conclusion, head-to-tail cyclisation of RGDfK peptide sequence performed using the combination of NovaSyn TGA resin with ivDde-protected Lys or 2-chlorotriyl chloride resin both with ivDde or Mtt-protected Lys gives the best results in term to prevent cyclodimerisation reaction.



**Fig. 4** **a** RP-HPLC and **b** ESI-MS profiles of crude c(RGDfK) peptide obtained from the synthesis on NovaSyn TGA using ivDde as Lys-protecting group

### Concluding Remarks

c(RGDfK) is an  $\alpha_V\beta_3$  integrin ligand very promising for biotechnological purposes. It is indeed a good tumor marker and it can be conjugated with a large variety of chelators for radionuclides or functionalised with fluorescent probes for both diagnosis and therapy, through the lysine side-chain without significant modification of its binding capacity.

Therefore, the optimisation of condition to cyclise c(RGDfK) peptide on solid phase allows to obtain a single compound useful to reduce the following on resin reaction steps and to increase the final product yield.

In the present paper the conditions to optimize the cyclisation reaction of c(RGDfK) peptide on solid-phase, were investigated. The synthesis of c(RGDfK) was performed using combinations of different resins and Lys-protecting groups in order to evaluate the role of these

parameters and the influence of their steric hindrance on the ring closure reaction.

On the basis of the obtained results, it can be concluded that the use of NovaSyn TGA or 2-chlorotrityl chloride resins with ivDde-protected Lys residue gives the best combination of solid support and amino acid protecting group. This synthetic procedure provides the required product, usable to carry on further reactions onto the resin, without any side cyclodimeric adduct.

**Acknowledgments** This work was supported by DFM Center of Regione Campania and MIUR (PRIN2008F5A3AF). We thank Dr. G. Perretta and Mr. L. De Luca for technical assistance.

### References

- Beekman KW, Colevas AD, Cooney K, Dipaola R, Dunn RL, Gross M, Keller ET, Pienta KJ, Ryan CJ, Smith D, Hussain M (2006) Phase II evaluations of cilengitide in asymptomatic patients with androgen-independent prostate cancer: scientific rationale and study design. *Clin Genitourin Cancer* 4:299–302
- Bodansky M (1984) Principle of peptide synthesis. Springer Verlag, Berlin
- Boturn D, Dumy P (2001) A convenient access to  $\alpha_V\beta_3:\alpha_V\beta_5$  integrin ligand conjugates: regioselective solid-phase functionalisation of an RGD based peptide. *Tetrahedron Lett* 42:2787–2790
- Brooks PC, Montgomery AM, Rosenfeld M, Reisfeld RA, Hu T, Klier G, Cheresch DA (1994a) Integrin  $\alpha_V\beta_3$  antagonists promote tumor regression by inducing apoptosis of angiogenic blood vessels. *Cell* 79:1157–1164
- Brooks PC, Clark RA, Cheresch DA (1994b) Requirement of vascular integrin  $\alpha_V\beta_3$  for angiogenesis. *Science* 264:569–571
- Clezzardin P (1998) Recent insights into the role of integrins in cancer metastasis. *Cell Mol Life Sci* 54:541–548
- Dechantsreiter MA, Planker E, Matha B, Lohof E, Holzemann G, Jonczyk A, Goodman SL, Kessler H (1999) N-methylated cyclic RGD peptides as highly active and selective  $\alpha_V\beta_3$  integrin antagonists. *J Med Chem* 42:3033–3040
- Eliceiri BP, Cheresch DA (1999) The role of  $\alpha_V$  integrins during angiogenesis: insights into potential mechanisms of action and clinical development. *J Clin Invest* 103:1227–1230
- Eskens FA, Dumez H, Hoekstra R, Perschl A, Brindley C, Bottcher S, Wynendaele W, Dreves J, Verweij J, van Oosterom AT (2003) Phase I and pharmacokinetic study of continuous twice weekly intravenous administration of cilengitide (EMD 121974), a novel inhibitor of the integrins  $\alpha_V\beta_3$  and  $\alpha_V\beta_5$  in patients with advanced solid tumours. *Eur J Cancer* 39:917–926
- Felding-Habermann B, O'Toole TE, Smith JW, Fransvea E, Ruggeri ZM, Ginsberg MH, Hughes PE, Pampori N, Shattil SJ, Saven A, Mueller BM (2001) Integrin activation controls metastasis in human breast cancer. *Proc Natl Acad Sci USA* 98:1853–1858
- Friess H, Langrehr JM, Oettle H, Raedle J, Niedergethmann M, Dittrich C, Hossfeld DK, Stoger H, Neyns B, Herzog P, Piedbois P, Dobrowolski F, Scheithauer W, Hawkins R, Katz F, Balcke P, Vermorken J, van Belle S, Davidson N, Esteve AA, Castellano D, Kleeff J, Tempia-Caliera AA, Kovar A, Nippgen JA (2006) Randomized multi-center phase II trial of the angiogenesis inhibitor cilengitide (EMD 121974) and gemcitabine compared with gemcitabine alone in advanced unresectable pancreatic cancer. *BMC Cancer* 6:285
- Garanger E, Boturn D, Dumy P (2007) Tumor targeting with RGD peptide ligands-design of new molecular conjugates for imaging

- and therapy of cancers. *Anticancer Agents Med Chem* 7:552–558
- Goldshaid L, Rubinstein E, Brandis A, Segal D, Leshem N, Brenner O, Kalchenko V, Eren D, Yechezkel T, Salitra Y, Salomon Y, Scherz A (2010) Novel design principles enable specific targeting of imaging and therapeutic agents to necrotic domains in breast tumors. *Breast Cancer Res* 12:1–18
- Haubner R, Gratiás R, Diefenbach B, Goodman SL, Jonczyk A, Kessler H (1996) Structural and functional aspects of RGD-containing cyclic pentapeptides as highly potent and selective integrin  $\alpha$ V $\beta$ 3 antagonists. *J Am Chem Soc* 118:7461–7472
- Haubner R, Finsinger D, Kessler H (1997) Stereoisomeric peptide libraries and peptidomimetics for designing selective inhibitors of the  $\alpha$ V $\beta$ 3 integrin for a new cancer therapy. *Angew Chem Int Ed Engl* 36:1374–1389
- Haubner R, Weber WA, Beer AJ, Vabulienė E, Reim D, Sarbia M, Becker KF, Goebel M, Hein R, Wester HJ, Kessler H, Schwaiger M (2005) Noninvasive visualization of the activated  $\alpha$ V $\beta$ 3 integrin in cancer patients by positron emission tomography and [ $^{18}$ F]galacto-RGD. *PLoS Med* 2:e70
- Lambert JN, Mitchell JP, Roberts KD (2001) The synthesis of cyclic peptides. *J Chem Soc Perkin Trans* 1:471–484
- MacDonald TJ, Stewart CF, Kocak M, Goldman S, Ellenbogen RG, Phillips P, Lafond D, Poussaint TY, Kieran MW, Boyett JM, Kun LE (2008) Phase I clinical trial of cilengitide in children with refractory brain tumors: Pediatric Brain Tumor Consortium Study PBTC-012. *J Clin Oncol* 26:919–924
- McCusker CF, Kocienski PJ, Boyle FT, Schatzlein AG (2002) Solid-phase synthesis of c(RGDfK) derivatives: on-resin cyclisation and lysine functionalisation. *Bioorg Med Chem Lett* 12:547–549
- Nabors LB, Mikkelsen T, Rosenfeld SS, Hochberg F, Akella NS, Fisher JD, Cloud GA, Zhang Y, Carson K, Wittmer SM, Colevas AD, Grossman SA (2007) Phase I and correlative biology study of cilengitide in patients with recurrent malignant glioma. *J Clin Oncol* 25:1651–1657
- Reinmuth N, Liu W, Ahmad SA, Fan F, Stoeltzing O, Parikh AA, Bucana CD, Gallick GE, Nickols MA, Westlin WF, Ellis LM (2003)  $\alpha$ V $\beta$ 3 integrin antagonist S247 decreases colon cancer metastasis and angiogenesis and improves survival in mice. *Cancer Res* 63:2079–2087
- Rovero P (2000) In: Kates SA, Albericio F (eds) *Solid-phase synthesis*. Dekker, New York
- Zaccaro L, Del Gatto A, Pedone C, Saviano M (2007) Integrin receptor family: promising target for therapeutic and diagnostic applications. *Curr Top Biochem Res* 9:45–56
- Zaccaro L, Del Gatto A, Pedone C, Saviano M (2009) Peptides for tumour therapy and diagnosis: current status and future directions. *Curr Med Chem* 16:780–795
- Zhao Y, Bachelier R, Treilleux I, Pujuguet P, Peyruchaud O, Baron R, Clement-Lacroix P, Clezardin P (2007) Tumor  $\alpha$ V $\beta$ 3 integrin is a therapeutic target for breast cancer bone metastases. *Cancer Res* 67:5821–5830

# Design, synthesis and characterization of a peptide able to bind proteins of the KCTD family: implications for KCTD – cullin 3 recognition<sup>‡</sup>

Luciano Pirone,<sup>a,b§</sup> Stefania Correale,<sup>a,b§</sup> Ivan de Paola,<sup>b</sup> Laura Zaccaro,<sup>a</sup> Giuseppina De Simone,<sup>a</sup> Luigi Vitagliano,<sup>a</sup> Emilia Pedone<sup>a\*</sup> and Sonia Di Gaetano<sup>a\*</sup>

**Pox virus Zinc/Bric-à-brac, Tramtrack and Broad (POZ/BTB) is a widespread domain detected in proteins involved in a variety of biological processes. Human genome analyses have unveiled the presence of POZ/BTB domain in a class of proteins (KCTD) whose role as important players in crucial biological processes is emerging. The development of new molecular entities able to interact with these proteins and to modulate their activity is a field of relevant interest. By using molecular modeling and literature mutagenesis analyses, we here designed and characterized a peptide that is able to interact with submicromolar affinities with two different members (KCTD11 and KCTD5) of this family. This finding suggests that the tetrameric KCTD11 and the pentameric KCTD5 are endowed with a similar cavity at the subunit–subunit interface deputed to the Cul3 binding, despite their different oligomeric states. Copyright © 2011 European Peptide Society and John Wiley & Sons, Ltd.**

**Keywords:** protein–protein recognition; structure–function relationships; ubiquitination; POZ/BTB domains

## Introduction

The Pox virus Zinc/Bric-à-brac, Tramtrack and Broad (POZ/BTB) domain is a module that has been found in a large number of functionally unrelated proteins [1–3]. The domain has been originally identified as a conserved motif present in the *Drosophila melanogaster* BTB complex transcription regulators, and in many POZ finger proteins [1–3]. Proteins containing the POZ/BTB domain are involved in diversified biological processes, which include transcription repression [4], cytoskeleton regulation [5], mediators of oligomerization and gating of ion channels [6] and protein ubiquitination/degradation [7–9]. It is commonly assumed that this domain plays a major role in protein–protein interactions. Structural characterizations of POZ/BTB domains isolated from different proteins have revealed the occurrence of variations on a common theme [1]. In all structures of POZ/BTB domains, so far reported it is possible to identify a common motif characterized by the presence of a three-stranded  $\beta$ -sheet and five  $\alpha$ -helices [1]. Recent human genome investigations have highlighted the presence of the POZ/BTB domain in a new class of human proteins. The POZ/BTB domain of these proteins share a significant sequence identity (~30%) with the tetramerization domain of the K<sup>+</sup> voltage-gated channel T1 (T1-Kv) proteins (<http://btb.uhnres.utoronto.ca/index.html>). On this basis, this class of proteins has been denoted as KCTD (K<sup>+</sup> channel tetramerization domain containing proteins). Although KCTD proteins share a common POZ/BTB domain in their N-terminal regions, they present diversified C-terminal domains [10]. Although the biological role of these proteins is yet to be determined, recent investigations suggest that they are involved in

important biological processes. It has been shown that a missense mutation of KCTD7 is linked to neurodegeneration and progressive myoclonic epilepsy [11]. Moreover, KCTD15 variants have been associated with adult obesity risk [12]. Very recently, it has been shown that KCTD12, previously identified as an important factor involved in the maturation of ear neurons, along with KCTD8 and KCTD16, plays a role in the pharmacology and kinetics of the gamma-aminobutyric acid (GABA(B)) receptor response [13,14]. The two members of the family that have been characterized in details are KCTD5 and KCTD11. KCTD5, the only member of the

\* Correspondence to: Emilia Pedone, Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, 80134, Napoli, Italy. E-mail: empedone@unina.it

Sonia Di Gaetano, Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, 80134, Napoli, Italy. E-mail: digaetan@unina.it

a Institute of Biostructures and Bioimaging, CNR, 80134 Napoli, Italy

b Department of Biological Sciences, "Federico II" University, 80134 Napoli, Italy

‡ Special issue devoted to contributions presented at the 12th Naples Workshop on Bioactive Peptides and 2nd Italy-Korea Symposium on Antimicrobial Peptides, 4–7 June 2010, Naples, Italy.

§ These authors equally contributed.

**Abbreviations used:** KCTD, potassium channel tetramerization domain containing proteins; PDB, Protein data bank; Cul3, cullin 3; KCTD11 BTB and KCTD5 BTB, POZ/BTB domains of KCTD11 and KCTD5, respectively; TrxA, *Escherichia coli* thioredoxin; Cul3wt-pep, peptide corresponding to the 49–68 region of Cul3; Cul3KK-pep, variant of Cul3wt-pep in which two Tyre residues were replaced with Lys residues.

family whose three-dimensional structure has been determined experimentally, interacts with human Golgi reassembly stacking protein 55 (GRASP55) [10,15]. It has also been shown that KCTD5 also interacts with cullin 3 (Cul3) and is potentially involved in the E3 ubiquitin ligase complex [15].

KCTD11, a tumor suppressor, is an important antagonist of the Hedgehog pathway which is frequently deleted in human medulloblastoma [16–19]. Very recent studies have shown that it is widely downregulated in other human cancers [20]. KCTD11 is involved in the deacetylation and activation of Gli1 and Gli2, two key transcription factors of Hh signaling. Indeed, Gli1 is deacetylated by histone deacetylase 1 (HDAC1), which is ubiquitinated for degradation by the E3 ubiquitin ligase, formed by Cul3 and KCTD11 [21].

Taking into account the role played by KCTD in important biological processes, the development of new molecular entities able to interact with these proteins and to modulate their activity is a field of relevant interest. We here report the design and the characterization of a Cul3-based peptide that is able to interact either with the pentameric KCTD5 or with the tetrameric KCTD11.

## Experimental

### Materials

All amino acids, Rink-amide 4-methylbenzhydrylamine (MBHA) resin, HBTU, HOBt and HATU were purchased from Novabiochem. DIPEA, TFA, piperidine, *N*-(+)-biotinyl-6-aminocaproic acid, acetic anhydride and scavengers were supplied from Fluka. The solvents used in the synthesis, purification and characterization of the peptides were from Romil.

### Peptide Synthesis

Cul3wt\_pep and Cul3KK\_pep derived from Cul3 sequence (see below) were obtained by Fmoc solid-phase strategy (0.1 mmol). To mimic the charge status of the fragment within the parent protein, the *N*-terminus and the *C*-terminus of the peptides were acetylated and amidated, respectively. The syntheses were carried out with Rink-amide MBHA resin (substitution 0.51 mmol g<sup>-1</sup>), using all standard amino acids. Coupling reactions were performed by using 10 equiv of Fmoc protected amino acids activated *in situ* with HBTU (9.8 equiv)/HOBt (9.8 equiv)/DIPEA (20 equiv) in DMF for 1 h.

Fmoc protecting group was removed by treatment with 30% piperidine in DMF two times for 10 min. Before the cleavage from the resin, both peptides were acetylated or biotinylated at *N*-terminus to obtain the corresponding derivatives. The acetylation reaction was carried out two times for 10 min using a solution of acetic anhydride (0.5 M)/DIPEA (0.015 M)/HOBt (0.125 M) in DMF (4.7 : 4 : 91.3 v/v/v). Biotinylated peptides were obtained using a solution of *N*-(+)-biotinyl-6-aminocaproic acid (2 equiv)/HATU (1.9 equiv)/DIPEA (4 equiv) in DMF overnight.

All peptides were cleaved off the resin by treatment with a mixture of TFA/H<sub>2</sub>O/Ethanedithiol (EDT)/Triisopropylsilane (TIS) (94 : 2.5 : 2.5 : 1 v/v/v/v) for 2 h at room temperature. The resins were filtered and the crude peptides were precipitated with diethyl ether, dissolved in a H<sub>2</sub>O/CH<sub>3</sub>CN (1 : 1) solution and lyophilized.

The peptides were purified by preparative RP-HPLC on the Shimadzu system equipped with a UV-Vis detector SPD10A using a Phenomenex Jupiter Proteo C<sub>12</sub> column (21.2 × 250 mm; 4 μm; 90 Å) and a linear gradient of H<sub>2</sub>O (0.1% TFA)/CH<sub>3</sub>CN (0.1% TFA)

from 20 to 70% of CH<sub>3</sub>CN (0.1% TFA) in 20 min at flow rate of 5 ml/min. The collected fractions containing the peptides were lyophilized. The identity and purity of the peptides were assessed by an ESI-LC-MS instrument (ThermoFinnigan, NY, USA) equipped with a diode array detector combined with an electrospray ion source and ion trap mass analyzer using a Phenomenex Jupiter Proteo C<sub>12</sub> column (150 × 2 mm; 4 μm; 90 Å) and a linear gradient of H<sub>2</sub>O (0.1% TFA)/CH<sub>3</sub>CN (0.1% TFA) from 20 to 80% of CH<sub>3</sub>CN (0.1% TFA) in 20 min at flow rate of 200 μl/min.

## CD Spectroscopy Studies

Far-UV CD spectra were registered in the 190–260 nm range on a Jasco J-810 spectropolarimeter at 25 °C using a quartz cell with a 0.1-cm path length. The peptides were dissolved in 10 mM phosphate buffer (pH = 7.1). The final concentrations of Cul3wt\_pep and Cul3KK\_pep were 5.56 × 10<sup>-5</sup> and 5.05 × 10<sup>-5</sup> M, respectively.

## Expression and Purification of POZ/BTB domains of KCTD11 and KCTD5

Expression and purification of the KCTD11 BTB were performed as described elsewhere [22]. Briefly, the His tag recombinant domain was expressed as a recombinant fusion protein with thioredoxin A (TrxA). A homogenous product was obtained upon purification.

The recombinant plasmid containing the POZ/BTB sequence of KCTD5 was a gift of Prof. Goldstein (Department of Pediatrics and Institute of Molecular Pediatric Sciences, University of Chicago). The His-tag recombinant protein was expressed and purified as already reported [10,22].

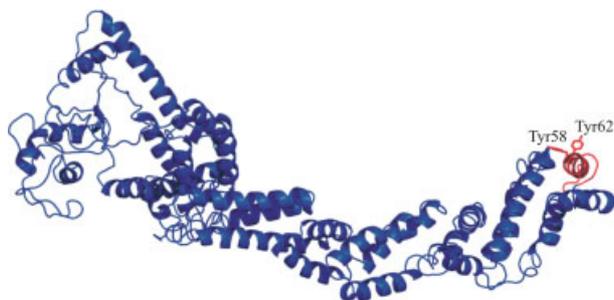
### ELISA Assays

For ELISA assays, 5 μg/ml streptavidin in phosphate/citrate buffer pH 5.0 was incubated overnight at 37 °C for coating. First, binding was executed with 0.8 μM biotinylated Cul3 peptide in Phosphate Buffered Saline (PBS) 1X for 1 h at room temperature. Second, binding was performed with different concentrations of His-TrxA/KCTD11 BTB or His/KCTD5 BTB (0.8, 1.5, 3.8, 7.6 and 15.6 μM) in PBS 1X. His-TrxA was used as negative control in the same concentrations. As blocking solution 1% BSA, 0.05% Tween-20 in PBS 1X was used for 1 h at room temperature. To reveal the occurred interaction mouse anti-His monoclonal antibody was incubated in 1 : 1000 dilution at room temperature for 2 h; then, horseradish peroxidase-conjugated anti-mouse antibody (Pierce) was diluted 1 : 10 000 in PBS 1X and incubated at room temperature for 1 h. The colorimetric reaction has been carried out with SIGMAFAST OPD reagent (Sigma Aldrich), according to the manufacturer's instructions. Finally, a Model 680 Microplate Reader (Bio-Rad, Hercules, CA-USA) has been used for readings at 490 nm; data were processed by a Microplate Manager 5.2 program. The reported data are mean values of triplicate experiments.

## Results and Discussion

### Design of a Cul3-Based Peptide as a Potential KCTD11 Binder

Previous modeling and mutagenesis studies have provided some clues on the regions that mediate the recognition between Cul3 and KCTD11 BTB [21,22]. The complex is stabilized by a cluster



**Figure 1.** Location of the 49–68 region (red), corresponding to the sequence of the Cul3wt\_pep, in the three-dimensional structure of Cul3. Side chains of the residues Tyr58 and Tyr62 directly involved in KCTD11 are also shown.

**Table 1.** Analytical data of synthesized peptides obtained by ESI-LC-MS

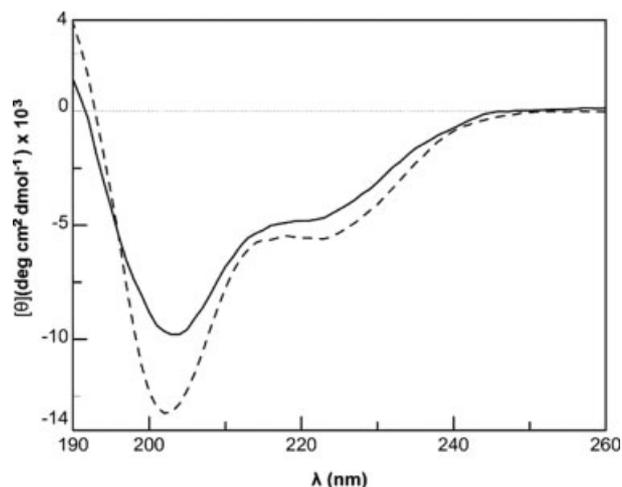
Peptide	$T_r$ (min)	Theoretical mass (Da)	Found mass (ESI-MS) (Da)
Ac-Cul3_wt	15.24	2412.2	2413.1
Biotin-Cul3_wt	15.40	2709.3	2711.8
Ac-Cul3_KK	11.51	2342.2	2343.6
Biotin-Cul3_KK	13.06	2639.4	2640.9

of aromatic residues which include Tyr58, Ty62 and Tyr125 of the cullin and Phe102 and Tyr103 of KCTD11 BTB. It has been shown that mutations of these residues inhibit the formation of the complex. An important role in Cul3-KCTD11 interaction is played by the helix 55–65 of Cul3 that is tightly bound at the interface of KCTD11 tetramer. Although other Cul3 residues are involved in KCTD11 recognition, we focused our attention on the role played by this region. A Cul3-based peptide as a potential KCTD11 binder was therefore designed by considering the peptide NSGLSFEELYRNAYTMVLHK, corresponding to the 49–68 region in Cul3 sequence (designed as Cul3wt\_pep hereafter). The location of this fragment within Cul3 three-dimensional structure is shown in Figure 1.

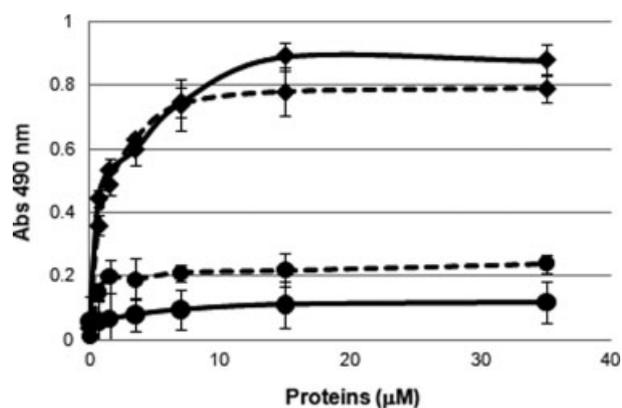
### Synthesis, Purification, and Characterization of the Cul3-Based Peptides

To verify the effectiveness of the designed peptide, a biotinylated form of Cul3wt\_pep was synthesized on Fmoc solid-phase and then purified by RP-HPLC. To check the role of aromatic side chains in Cul3 – KCTD11 recognition, a Cul3wt\_pep variant (Cul3KK\_pep) in which Tyr58 and Tyr62 were substituted with Lys residues, was also prepared. Cul3KK\_pep was synthesized and purified following the same procedures used for Cul3wt\_pep. The identities of the obtained compounds were assessed by ESI-LC-MS (Table 1).

The CD spectra of both peptides, dissolved in a phosphate buffer (pH = 7.1), show a deep minimum at ~202–203 nm, a second minimum at ~222 nm, a cross over at ~192–193, and a maximum at a wavelength lower than 190 nm (Figure 2). The general features of the spectra are indicative for both peptides of the presence of a low, but significant, helical secondary structure content. The ratio between the CD signal at 222 nm and the signal of the other minimum at lower wavelength, which is often used as a mean to confirm the presence of helical content [23], is 0.48 and



**Figure 2.** CD spectra of Cul3wt\_pep (—) and Cul3KK\_pep (---) registered in phosphate buffer at pH = 7.1.  $\theta$  is expressed as mean residue ellipticity  $[\theta]$ .



**Figure 3.** Binding curves obtained from ELISA assays on the proteins KCTD11 BTB (continue line) and KCTD5 BTB (dashed line) using the peptides Biotin Cul3wt\_pep (diamond) and Biotin Cul3KK\_pep (circle).

0.42 for Cul3wt\_pep and Cul3KK\_pep, respectively. These values, although much smaller than those observed for samples with elevated helical contents (close to the unity), are indicative of the presence of helical conformations. To quantify the helical content, we performed spectral deconvolution analysis using CDNN and CDPRO programs. The estimation of the helical content provided by these analyses was ~13% for both peptides.

### Cul3 Peptide Interacts with POZ/BTB Domains with Different Oligomeric Organization

The ability of biotinylated Cul3wt\_pep to bind the recombinant KCTD11 BTB domain, expressed as a fused protein with TrxA, was investigated by ELISA assay, following the procedure illustrated in the Methods section. The binding of the peptide to TrxA was monitored as negative control. As shown in Figure 3, the trend of the binding curve demonstrates that Cul3wt\_pep specifically binds KCTD11 BTB with an apparent affinity constant of 0.8  $\mu\text{M}$ . Similar results were obtained either coating KCTD11 BTB or the biotinylated peptide through streptavidin. The inability of the peptide variant Cul3KK\_pep (Figure 3) to bind KCTD11 BTB

corroborates the indications provided by the modeling on the important role of the aromatic residues in the recognition process.

It has been reported that, in addition to KCTD11, another member of the KCTD, KCTD5, is able to bind Cul3. Intriguingly, the BTB domain of KCTD11 and KCTD5 exhibit different oligomeric states, being KCTD11 and KCTD5 tetrameric and pentameric, respectively. In this framework, we checked the binding of Cul3wt\_peg to KCTD5 BTB. The binding curve (Figure 3) indicates that the peptide is also able to bind to KCTD5 BTB with an apparent affinity constant that is similar to that observed for the association with KCTD11 BTB.

## Conclusion

We here show that a rather small peptide derived from a Cul3 fragment is able to interact with BTB domains of KCTD proteins despite their different oligomeric states. As KCTD are involved in crucial biological processes, the development of peptides able to interact with these proteins has important potential applications. Present findings also suggest strategies for future peptide optimization. Indeed, we here show that Cul3wt\_peg only partially retains the helical state that the fragment assumes in the parent protein. It is likely that an increase in the helical propensity of Cul3wt\_peg variants may be associated with an increased affinity for KCTD proteins. Several literature successful examples on other systems conducted using this strategy provide an indirect support to this approach [24,25].

## Acknowledgements

The authors thank the Italian MIUR for financial support (PRIN 2007 and FIRB contract RBRN07BMCT).

## References

- 1 Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG. Sequence and structural analysis of BTB domain proteins. *Genome Biol.* 2005; **6**: R82.
- 2 Zollman S, Godt D, Prive GG, Couderc JL, Laski FA. The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 1994; **91**: 10717–10721.
- 3 Bardwell VJ, Treisman R. The POZ domain: a conserved protein-protein interaction motif. *Genes Dev.* 1994; **8**: 1664–1677.
- 4 Melnick A, Ahmad KF, Arai S, Polinger A, Ball H, Borden KL, Carlile GW, Prive GG, Licht JD. In-depth mutational analysis of the promyelocytic leukemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions. *Mol. Cell Biol.* 2000; **20**: 6550–6567.
- 5 Kang MI, Kobayashi A, Wakabayashi N, Kim SG, Yamamoto M. Scaffolding of keap1 to the actin cytoskeleton controls the function of nrf2 as key regulator of cytoprotective phase 2 genes. *Proc. Natl. Acad. Sci. U.S.A.* 2004; **101**: 2046–2051.
- 6 Minor DL, Lin YF, Mobley BC, Avelar A, Jan YN, Jan LY, Berger JM. The polar t1 interface is linked to conformational changes that open the voltage-gated potassium channel. *Cell* 2000; **102**: 657–670.
- 7 Pintard L, Willis JH, Willems A, Johnson JL, Srayko M, Kurz T, Glaser S, Mains PE, Tyers M, Bowerman B, Peter M. The BTB protein mel-26 is a substrate-specific adaptor of the cul-3 ubiquitin-ligase. *Nature* 2003; **425**: 311–316.
- 8 Furukawa M, He YJ, Borchers C, Xiong Y. Targeting of protein ubiquitination by BTB-cullin 3-roc1 ubiquitin ligases. *Nat. Cell Biol.* 2003; **5**: 1001–1007.
- 9 Geyer R, Wee S, Anderson S, Yates J, Wolf DA. BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. *Mol. Cell* 2003; **12**: 783–790.
- 10 Dementieva IS, Tereshko V, McCrossan ZA, Solomaha E, Araki D, Xu C, Grigorieff N, Goldstein SA. Pentameric assembly of potassium channel tetramerization domain-containing protein 5. *J. Mol. Biol.* 2009; **387**: 175–191.
- 11 Van Bogaert P, Azizieh R, Desir J, Aeby A, De Meirleir L, Laes JF, Christiaens F, Abramowicz MJ. Mutation of a potassium channel-related gene in progressive myoclonic epilepsy. *Ann. Neurol.* 2007; **61**: 579–586.
- 12 Elks CE, Loos RJ, Sharp SJ, Langenberg C, Ring SM, Timpson NJ, Ness AR, Davey Smith G, Dunger DB, Wareham NJ, Ong KK. Genetic markers of adult obesity risk are associated with greater early infancy weight gain and growth. *PLoS Med.* 7: e1000284.
- 13 Schwenk J, Metz M, Zolles G, Turecek R, Fritzius T, Bildl W, Tarusawa E, Kulik A, Unger A, Ivankova K, Seddik R, Tiao JY, Rajalu M, Trojanova J, Rohde V, Gassmann M, Schulte U, Fakler B, Bettler B. Native GABA(B) receptors are heteromultimers with a family of auxiliary subunits. *Nature* 2010; **465**: 231–235.
- 14 Bartoi T, Rigbolt KT, Du D, Kohr G, Blagoev B, Kornau HC. GABAB receptor constituents revealed by tandem affinity purification from transgenic mice. *J. Biol. Chem.* 2010; **285**: 20625–20633.
- 15 Bayon Y, Trinidad AG, de la Puerta ML, Del Carmen Rodriguez M, Bogetz J, Rojas A, De Pereda JM, Rahmouni S, Williams S, Matsuzawa S, Reed JC, Crespo MS, Mustelin T, Alonso A. Kctd5, a putative substrate adaptor for cullin3 ubiquitin ligases. *FEBS J.* 2008; **275**: 3900–3910.
- 16 Gallo R, Zazzeroni F, Alesse E, Mincione C, Borello U, Buanne P, D'Eugenio R, Mackay AR, Argenti B, Gradini R, Russo MA, Maroder M, Cossu G, Frati L, Screpanti I, Gulino A. Ren: a novel, developmentally regulated gene that promotes neural cell differentiation. *J. Cell Biol.* 2002; **158**: 731–740.
- 17 De Smaele E, Di Marcotullio L, Ferretti E, Screpanti I, Alesse E, Gulino A. Chromosome 17p deletion in human medulloblastoma: a missing checkpoint in the hedgehog pathway. *Cell Cycle* 2004; **3**: 1263–1266.
- 18 Di Marcotullio L, Ferretti E, De Smaele E, Argenti B, Mincione C, Zazzeroni F, Gallo R, Masuelli L, Napolitano M, Maroder M, Modesti A, Giangaspero F, Screpanti I, Alesse E, Gulino A. Ren(kctd11) is a suppressor of hedgehog signaling and is deleted in human medulloblastoma. *Proc. Natl. Acad. Sci. U.S.A.* 2004; **101**: 10833–10838.
- 19 Ferretti E, De Smaele E, Di Marcotullio L, Screpanti I, Gulino A. Hedgehog checkpoints in medulloblastoma: The chromosome 17p deletion paradigm. *Trends Mol. Med.* 2005; **11**: 537–545.
- 20 Mancarelli MM, Zazzeroni F, Ciccocioppo L, Capece D, Po A, Murgo S, Di Camillo R, Rinaldi C, Ferretti E, Gulino A, Alesse E. The tumor suppressor gene kctd11ren is regulated by sp1 and methylation and its expression is reduced in tumors. *Mol. Cancer* 2010; **9**: 172.
- 21 Canetti G, Di Marcotullio L, Greco A, Coni S, Antonucci L, Infante P, Pietrosanti L, De Smaele E, Ferretti E, Miele E, Pelloni M, De Simone G, Pedone EM, Gallinari P, Giorgi A, Steinkuhler C, Vitagliano L, Pedone C, Schinin ME, Screpanti I, Gulino A. Histone deacetylase and cullin3-ren(kctd11) ubiquitin ligase interplay regulates hedgehog signalling through gli acetylation. *Nat. Cell Biol.* 2010; **12**: 132–142.
- 22 Correale S, Pirone L, Di Marcotullio L, De Smaele E, Greco A, Mazzà D, Moretti M, Alterio V, Vitagliano L, Di Gaetano S, Gulino A, Pedone EM. Molecular organization of the cullin e3 ligase adaptor kctd11. *Biochimie* 2011; **93**: 715–724.
- 23 Vuilleumier S, Mutter M. Synthetic peptide and template-assembled synthetic protein models of the hen egg white lysozyme 87–97 helix: importance of a protein-like framework for conformational stability in a short peptide sequence. *Biopolymers* 1993; **33**: 389–400.
- 24 D'Andrea LD, Iaccarino G, Fattorusso R, Sorriento D, Carannante C, Capasso D, Trimarco B, Pedone C. Targeting angiogenesis: structural characterization and biological properties of a de novo engineered VEGF mimicking peptide. *Proc. Natl. Acad. Sci. U.S.A.* 2005; **102**: 14215–14220.
- 25 Walensky LD, Kung AL, Escher I, Malia TJ, Barbuto S, Wright RD, Wagner G, Verdine GL, Korsmeyer SJ. Activation of apoptosis in vivo by a hydrocarbon-stapled bh3 helix. *Science* 2004; **305**: 1466–1470.

**Structure-Based Design of a Potent Artificial Transactivation Domain Based on p53**

Journal:	<i>Journal of the American Chemical Society</i>
Manuscript ID:	ja-2011-08999e
Manuscript Type:	Article
Date Submitted by the Author:	24-Sep-2011
Complete List of Authors:	Langlois, Chantal; Universite de Montreal, Biochimie Del Gatto, Annarita; Istituto di Biostrutture e Bioimmagini-CNR Arseneault, Genevieve; Universite de Montreal, Biochimie Lafrance-Vanasse, Julien; Universite de Montreal, Biochimie De Simone, Mariarosaria; Università degli Studi di Napoli "Federico II", Scienze Biologiche Morse, Thomas; Universite de Montreal, Biochimie de Paola, Ivan; Università degli Studi di Napoli "Federico II", Scienze Biologiche Lussier-Price, Mathieu; Universite de Montreal, Biochimie Legault, Pascale; Universite de Montreal, Biochimie Pedone, Carlo; Università degli Studi di Napoli "Federico II", Scienze Biologiche Zaccaro, Laura; Istituto di Biostrutture e Bioimmagini-CNR Omichinski, James; Universite de Montreal, Biochimie

SCHOLARONE™  
Manuscripts

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

# Structure-Based Design of a Potent Artificial Transactivation Domain Based on p53

*Chantal Langlois<sup>1</sup>, Annarita Del Gatto<sup>2</sup>, Geneviève Arseneault<sup>1</sup>, Julien Lafrance-Vanasse<sup>1</sup>,  
Marianosaria De Simone<sup>2</sup>, Thomas Morse<sup>1</sup>, Ivan de Paola<sup>2</sup>, Mathieu Lussier-Price<sup>1</sup>, Pascale  
Legault<sup>1</sup>, Carlo Pedone<sup>2</sup>, Laura Zaccaro<sup>2\*</sup> and James G. Omichinski<sup>1\*</sup>*

<sup>1</sup>Département de Biochimie, Université de Montréal C.P. 6128 Succursale Centre-Ville,  
Montréal, QC H3C 3J7 Canada

<sup>2</sup>Institute of Biostructures and Bioimaging, CNR and Department of Biological Sciences,  
University of Naples “Federico II”, via Mezzocannone 16, 80134, Napoli, Italy

**\* Correspondence should be addressed to J.G.O. or L. Z.:**

Email: [jg.omichinski@umontreal.ca](mailto:jg.omichinski@umontreal.ca) or [lzaccaro@unina.it](mailto:lzaccaro@unina.it)

**Running Title:** Design of a potent artificial transactivation domain.

**Abstract**

Malfunctions in transcriptional regulation are associated with a number of critical human diseases. As a result, there is considerable interest in designing artificial transcription activators (ATAs) that specifically control genes linked to human diseases. Like native transcriptional activator proteins, an ATA must minimally contain a DNA-binding domain (DBD) and a transactivation domain (TAD), and although there are several reliable methods for designing artificial DBDs, designing artificial TADs has proven difficult. In this manuscript, we present a structure-based strategy for designing short peptides containing natural amino acids that function as artificial TADs. Using a segment of the TAD of p53 as the scaffolding, modifications are introduced to increase the helical propensity of the peptides. The most active artificial TAD, termed E-Cap-(LL), is a 13-mer peptide that contains four key residues from p53, an N-capping motif and a dileucine hydrophobic bridge. In vitro analysis demonstrates that E-Cap(LL) interacts with several known p53 target proteins, while in vivo studies in a yeast model system show that it is a 20-fold more potent transcriptional activator than the native p53-13 peptide. These results demonstrate that structure-based design represents a promising approach for developing artificial TADs that can be combined with artificial DBDs to create potent and specific ATAs.

## Introduction

Activation of transcription is regulated by a complex network of macromolecular interactions that leads to enhanced rates of gene expression, and one of the key components of this network are transcriptional activators<sup>1</sup>. Transcriptional activators are proteins minimally composed of a DNA-binding domain (DBD) and a transcription activation domain (TAD), and these two domains can exist either within the same protein or be assembled through protein-protein interactions<sup>1-3</sup>. The DBD functions to direct the activator to specific sites on DNA, whereas the TAD participates in several protein-protein interactions with multiple components of the transcriptional machinery, including nucleosome-remodeling complexes, histone acetyltransferases (HATs) and general transcription factors (TFIIB, TBP, TFIID)<sup>4-12</sup>.

Malfunctions in transcriptional regulation are associated with many human diseases and there is considerable interest in developing artificial transcription factors (ATFs) that can either activate or repress a specific gene<sup>13-18</sup>. Given the recent success using RNAi to repress expression of specific genes<sup>19-21</sup>, most current efforts are now focused in developing artificial transcriptional activators (ATAs). Like their natural counterparts, ATAs must minimally contain a DBD and a TAD. The design of artificial DBDs has benefited from the wealth of structural information available for DNA and protein:DNA complexes. The most successful examples of artificial DBDs include pyrole-imidazole polyamides (PIP)<sup>22-24</sup>, peptide nucleic acids (PNA)<sup>25-28</sup> and engineered zinc-finger proteins<sup>29-31</sup>. In contrast to artificial DBDs, designing artificial TADs in an efficient and predictable manner has proven difficult<sup>13</sup>. Efforts to design artificial TADs are hampered by the fact that TADs interact with multiple target proteins as part of their normal function<sup>4-12</sup> and that there are only a limited number of high-resolution structures of TADs in complex with their target proteins. Structural studies of TADs are limited by the fact that most

1  
2  
3 native TADs are intrinsically unstructured domains that must transition from a disordered to an  
4  
5 ordered state to bind their targets<sup>32-38</sup>. This intrinsic flexibility has made it more difficult to  
6  
7 crystallize TADs in complexes and this has severely limited the available structural information.  
8  
9

10  
11 The most practical solution in designing ATAs has been to attach the sequence from a  
12  
13 native TAD to the artificial DBD of choice<sup>26,39-42</sup>. However, native TADs vary tremendously in  
14  
15 size and complexity (ranging from 14-300+ amino acids) and often generate variable responses  
16  
17 when incorporated into ATAs. Attempts to design shorter artificial TADs have generally started  
18  
19 either by concatenating short sections from native TADs<sup>14,41,43-46</sup> or by screening peptide/peptoid  
20  
21 libraries<sup>44,45,47-50</sup>. Unfortunately, these methods are limited by the fact that it is difficult to  
22  
23 improve their design in a systematic manner in the absence of structural information. In addition,  
24  
25 attempts to synthetically prepare small molecules that function as artificial TADs have proven to  
26  
27 be very challenging<sup>44,51-53</sup>. This is also due to the shortage of structural information and the fact  
28  
29 that TADs often bind over a large surface area when in complex with their partners making it  
30  
31 difficult to define a scaffold from which to base the design.  
32  
33  
34  
35  
36  
37

38 The Herpes Viral protein 16 (VP16) and the human tumour suppressor protein p53 are  
39  
40 two of the most potent transcriptional activator proteins known, and their TADs share several  
41  
42 common features<sup>54-56</sup>. This includes being very acidic, containing two subdomains  
43  
44 (p53TAD1/p53TAD2 and VP16N/VP16C), and interacting with many of the same target proteins  
45  
46<sup>54,55,57</sup>. We have previously determined the structures of the second subdomains from the TADs  
47  
48 of p53 (p53TAD2) and VP16 (VP16C) in complex with the Tfb1 subunit of the general  
49  
50 transcription factor IIIH (TFIIH)<sup>32,33</sup>. These structures demonstrated that p53TAD2 and VP16C  
51  
52 both transition from an unstructured state to form a 9-residue amphipathic  $\alpha$ -helix in complex  
53  
54 with the Pleckstrin Homology domain of Tfb1 (Tfb1PH). Comparison of the two structures  
55  
56  
57  
58  
59  
60

1  
2  
3 showed that three hydrophobic residues and one acidic residue located at positions four, five,  
4 seven and eight in the helices of p53 (residues Ile50, Glu51, Trp53 and Phe54) and VP16  
5 (residues Phe475, Glu476 Met478, Phe479) are crucial to formation of the interface with  
6 Tfb1PH. Although the three hydrophobic residues are not strictly conserved between p53 and  
7 VP16, they form similar interactions with Tfb1PH, and this suggests that either of these two  
8 structures could serve as a general template for the structure-based design of artificial TADs.  
9  
10  
11  
12  
13  
14  
15  
16

17  
18 In this paper, we report the design of an artificial TAD [E-Cap(LL)] based on the structure  
19 TAD of p53 in complex with Tfb1PH. E-Cap (LL) is a 13-residue peptide composed of natural  
20 amino acids, which preserves the four key residues of p53 that form the interface with Tfb1PH.  
21 In addition, E-Cap (LL) contains an N-terminal capping motif (N-Cap) and two leucines spaced  
22 in an i, i+3 manner to increase its helical propensity. In vitro studies indicate that E-Cap (LL)  
23 functions like p53TAD2 in a number of binding and competition assays. In addition, E-Cap (LL)  
24 is an extremely potent in vivo transcriptional activator in yeast. E-Cap-(LL) is the first artificial  
25 TAD designed based on a known structure of a TAD bound to its target and its potent in vivo  
26 activity indicates that structure-based design represents a promising approach for developing  
27 artificial TADs to be used in ATAs.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Experimental Procedures

### Chemical peptide synthesis and purification.

The p53-13, NC15, NC17, A-Cap (LL), E-Cap (LL) and W-Cap (LL) peptide were synthesized on solid phase and purified by HPLC (For further details see **Supplementary Experimental Procedures, Supporting Information**). The identity and purity of the peptides were verified by LC/MS spectrometry.

### Cloning of recombinant proteins for purification.

The bacterial expressed p53 peptide analogs [E-Cap-(LL), p53-13, p53-13-(LL) and mutants] were constructed by inserting the BamHI-EcoRI-digested DNA (IDT) into pGEX-2T plasmid. Tfb1PH was cloned as previously described<sup>32</sup>. Mutants of Tfb1PH, E-Cap-(LL) and p53-13-(LL) were prepared using the QuickChange II site-directed mutagenesis kit (Stratagene).

### Protein expression and purification.

The p53 peptide analogs and Tfb1PH, were expressed as GST-fusion proteins in *E. coli* host strain TOPP2 and bound to GSH-resin (GE Healthcare) as previously described<sup>32</sup>. The resin bound protein was incubated overnight with thrombin (Calbiochem). After cleavage, the supernatant was purified by FPLC over a Q-Sepharose (p53 analogs) or a SP-Sepharose (Tfb1PH) High Performance column. Uniformly (>98%) <sup>15</sup>N-labeled and <sup>15</sup>N/<sup>13</sup>C-labeled proteins were prepared in minimal media containing <sup>15</sup>NH<sub>4</sub>Cl with or without <sup>13</sup>C<sub>6</sub>-glucose as the sole nitrogen and carbon source. The CBP KIX domain (provided by Alanna Schepartz, Yale University, New Haven, CT) was expressed as His-Tag fusion protein in *E. Coli* host strain BL21

1  
2  
3 (DE3) and purified to homogeneity (See **Supplementary Experimental Procedures,**  
4  
5 **Supporting Information** for details).  
6  
7

### 8 9 10 **Circular Dichroism Studies.**

11  
12  
13 Circular dichroism (CD) studies were performed with synthetic peptides on a Jasco J-810  
14 spectropolarimeter at 25°C in 10 mM sodium phosphate (pH=7.1). All peptide concentrations  
15 were determined by  $A_{280}$ . The results are reported as mean residue molar ellipticity  $[\theta]$ . The  
16 intensities of  $[\theta]$  at 215, 207, 190nm, the cross over and the  $\theta_{222}/\theta_{207}$  ratio are reported for all  
17 peptides (**Supplementary Table 1, Supporting Information**).  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27

### 28 **Isothermal titration calorimetry (ITC) studies.**

29  
30  
31 The ITC experiments were performed as previously described <sup>32</sup>, in 20 mM TRIS (pH=7.5) for  
32 Tfb1PH, or in 20 mM sodium phosphate (pH=6.5) for the CBP KIX domain. The protein  
33 concentrations were determined from  $A_{280}$ . All titrations were done at least in duplicates and were  
34 fit to a single binding site mechanism with 1:1 stoichiometry.  
35  
36  
37  
38  
39  
40  
41  
42

### 43 **Media, plasmids and strains.**

44  
45  
46 All yeast strains were grown in synthetic complete media (SC; 0,67% yeast nitrogen base w/o  
47 amino acids, 2% glucose and amino acids drop-out mix) lacking uracil and histidine. The EGY48  
48 (Mat alpha leu2-3 his3-11,15 trp1-1 ura3-1 6lexAops-LEU2) strain was transformed with the  
49 LexA operator-Lac-Z fusion plasmid pSH18-34 combined with either pEG202NLS (pEG202  
50 derivative with SV40 nuclear localization sequence between LexA and polylinker) as a negative  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 control, pSH17-4 (GAL4-activation domain cloned into pEG202 backbone) as a positive control  
4  
5 or pEG202NLS with LexA fused to the activation domains to be tested <sup>58</sup>.  
6  
7

### 8 9 10 **β-Galactosidase activation assay.**

11  
12 Liquid β-galactosidase assays were performed as previously described <sup>58</sup>. Results are presented as  
13  
14 the mean of the percentages of the β-galactosidase units of the tested peptides on the β-  
15  
16 galactosidase units of the GAL4 positive control ± standard error of the mean (SEM). Western  
17  
18 blot analysis was performed with an anti-LexA antibody to verify expression of all LexA-fused  
19  
20 peptides.  
21  
22  
23  
24  
25  
26

### 27 28 **NMR samples.**

29  
30 For the NMR chemical shift mapping studies with labelled Tfb1PH, the samples consisted of 0.5  
31  
32 mM <sup>15</sup>N-Tfb1PH in 20 mM sodium phosphate (pH=6.5), 1 mM EDTA and 90% H<sub>2</sub>O/10% D<sub>2</sub>O;  
33  
34 unlabeled p53 analogs [NC17, NC15, E-Cap-(LL), W-Cap-(LL) or A-Cap-(LL)] were added to a  
35  
36 final ratio of 1:2. For the chemical shift mapping studies with labeled E-Cap-(LL) peptide, the  
37  
38 sample consisted of 0.5 mM of <sup>15</sup>N-E-Cap (LL) in 20 mM sodium phosphate (pH=6.5), 1 mM  
39  
40 EDTA and 90% H<sub>2</sub>O/10% D<sub>2</sub>O to which unlabeled Tfb1PH was added to a final ratio of 1:2. For  
41  
42 the competition experiment, an HSQC was first collected with a sample containing 0.8 mM of  
43  
44 <sup>15</sup>N-p53<sub>40-73</sub> (p53TAD2) in 20 mM sodium phosphate (pH=6.5), 1.0 mM EDTA and 90%  
45  
46 H<sub>2</sub>O/10% D<sub>2</sub>O. Then 1 mM of unlabeled Tfb1PH was added and a second HSQC was collected.  
47  
48  
49  
50 Finally, 0.8 mM of unlabeled E-Cap-(LL) peptide was added and a third HSQC spectrum was  
51  
52 recorded. The structural studies of the E-Cap-(LL) peptide in complex with Tfb1PH were  
53  
54 performed on two samples. The first contained 0.5 mM of <sup>15</sup>N-E-Cap-(LL) and 0.5 mM unlabeled  
55  
56  
57  
58  
59  
60

1  
2  
3 Tfb1PH in 20 mM sodium phosphate (pH=6.5), 1 mM EDTA and 90% H<sub>2</sub>O/10% D<sub>2</sub>O. The  
4  
5 second sample contained 0.5 mM <sup>15</sup>N/<sup>13</sup>C-Tfb1PH and 0.5 mM unlabeled E-Cap-(LL) in 20 mM  
6  
7 sodium phosphate (pH=6.5), 1 mM EDTA. For studies in D<sub>2</sub>O, the sample was dissolved in  
8  
9 99.996% D<sub>2</sub>O.  
10  
11  
12  
13

### 14 15 **NMR spectroscopy experiments.**

16  
17  
18 The NMR experiments were carried out at 295 K on Varian Unity Inova 500, 600 and 800 MHz  
19  
20 spectrometers. For the chemical shift mapping studies, 2D 1H-15N HSQC were recorded.  
21  
22

23 Intramolecular nuclear Overhauser effects (NOEs) for E-Cap (LL) were obtained from 3D 15N-  
24  
25 edited NOESY-HSQC ( $\tau_m=140$  ms)<sup>59</sup> and 2D <sup>13</sup>C/<sup>15</sup>N-{F1/F2}-filtered <sup>1</sup>H-<sup>1</sup>H NOESY ( $\tau_m=40$   
26  
27 and 100 ms)<sup>60</sup>. The NMR data were processed with NMRPipe/NMRDraw<sup>61</sup> and analyzed with  
28  
29 CcpNmr<sup>62</sup>.  
30  
31  
32  
33  
34

### 35 **Structures calculations.**

36  
37  
38 The NOE-derived distance restraints were divided into four classes defined as strong (1.8-2.8 Å),  
39  
40 medium (1.8-3.4 Å), weak (1.8-5.0 Å) and very weak (3.3-6.0 Å). Backbone dihedral angles were  
41  
42 derived with the program TALOS<sup>63</sup>. The structures of E-Cap-(LL) was calculated using the  
43  
44 program CNS<sup>64</sup>, with a combination of torsion angle and Cartesian dynamics<sup>65</sup> and starting from  
45  
46 an extended structure with standard geometry. The quality of structures was assessed using  
47  
48 PROCHEK-NMR<sup>66</sup>. The figures were generated with the program PyMol (<http://www.pymol.org>).  
49  
50  
51

### 52 **Results**

53  
54  
55  
56  
57  
58  
59  
60

### p53TAD analogs with an N-Cap and a C-Cap motif.

Based on the NMR structure of the p53TAD2/Tfb1PH complex<sup>32</sup>, we set out to design short peptide analogs that mimic both the in vitro binding properties and the in vivo activity of p53TAD2. We hypothesize that peptides analogs with increased helical propensity which, retain four key residues (IExWF) in the  $\alpha$ -helix of p53TAD2 (residues 47 to 55 in p53) will mimic p53TAD2 and be stronger in vivo activators. In the first approach, capping motifs are introduced both at the N- and C-terminus (N-Cap and C-Cap) of the four key residues in two peptides (NC17 and NC15) (**Figure 1A**). NC17 consists of 17 residues with an N-Cap and a C-Cap Schellman motif which were chosen based on statistical preference<sup>67,68</sup>. In NC17, the native GLN at position 51 is replaced by a Glu to favor stabilization of an  $i, i+4$  salt bridge with the Lys in the C-Cap. NC15 is a 15-residue peptide, which contains the same N-Cap, but its C-Cap Schellman motif includes Trp53 and Phe54 from the native p53 sequence. NC15 and NC17 were prepared by solid phase peptide synthesis, and their helical character and affinity for Tfb1PH were analyzed using circular dichroism (CD) and isothermal calorimetry (ITC), respectively. The CD spectra demonstrate a significant increase in helical character for both NC15 and NC17 ( $\theta_{222}/\theta_{207}$  ratio of 0.59 and 0.73, respectively) in comparison to p53-13 ( $\theta_{222}/\theta_{207}$  ratio of 0.23), a control peptide corresponding to residues 45 to 57 of the TAD of p53 (**Supplementary Table S1, Supporting Information**). However, ITC experiments indicate that neither NC15 nor NC17 are able to bind Tfb1PH under the conditions tested (**Figure 2A-B**). The absence of binding by ITC is further supported by the lack of significant changes in chemical shifts during NMR titration experiments with Tfb1PH (data not shown). In comparison, we measure an apparent dissociation constant ( $K_d$ ) of  $1.6 \pm 0.3 \mu\text{M}$  for the interaction of p53-13 with Tfb1PH under identical

1  
2  
3 conditions (**Figure 2B**). Thus, although the addition of an N-Cap and a C-Cap increases the  
4  
5 helical propensity, both NC15 and NC17 possess significantly lower affinity for Tfb1PH.  
6  
7

### 8 9 10 **p53TAD analogs with an N-Cap and a leucine bridge.**

11  
12  
13 In the second approach, we combined an N-cap with a hydrophobic bridge involving side chains  
14  
15 from two leucine residues [(N-Cap-(LL) peptides] (**Figure 1B**). This approach is similar to the  
16  
17 one used to develop the VEGF mimetic peptide QK (**Figure 1B**)<sup>67,69</sup>. Three different N-Caps  
18  
19 were tested and the resulting peptides are referred to as the A-Cap-(LL), E-Cap-(LL) and W-Cap-  
20  
21 (LL) (LTAE, LTEE and LTWE N-Capping residues respectively). The E-Cap-(LL) motif was  
22  
23 selected based on statistical probability<sup>68</sup>, the W-Cap-(LL) was chosen based on homology with  
24  
25 the QK-peptide<sup>67,69</sup> and the A-Cap-(LL) was chosen based on homology with the N-Cap from  
26  
27 the helix formed by VP16C<sup>33</sup>. The leucines were inserted with an i, i+3 spacing at positions 6  
28  
29 and 9 so that the hydrophobic interaction would form in the center of the peptide, but on the  
30  
31 opposite face of the  $\alpha$ -helix relative to the Tfb1PH binding interface (**Figure 1B**). The three  
32  
33 peptides are all 13-residue long, and CD spectra recorded on the N-Cap (LL) peptides indicate  
34  
35 that they all possess approximately the same helical character [ $\theta_{222}/\theta_{207}$  ratios of 0.52, 0.54 and  
36  
37 0.52 for A-Cap-(LL), E-Cap-(LL) and W-Cap-(LL) respectively], which is significantly higher  
38  
39 than the p53-13 peptide (**Supplementary Table S1, Supporting Information**). Next, we  
40  
41 measured the  $K_d$  values of the three peptides for Tfb1PH by ITC. The A-Cap-(LL) and W-Cap-  
42  
43 (LL) have apparent  $K_d$  values of  $1.3 \pm 0.1 \mu\text{M}$  and  $1.9 \pm 0.3 \mu\text{M}$  respectively, which is similar to  
44  
45 p53-13 ( $K_d$  of  $1.6 \pm 0.3 \mu\text{M}$ ), whereas the E-Cap-(LL) peptide has an apparent  $K_d$  of  $0.24 \pm 0.03$   
46  
47  $\mu\text{M}$  (**Figure 2B**). Thus, the introduction of a N-Cap in combination with a dileucine bridge  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 improves the helical character of all three peptides and the E-Cap-(LL) peptide binds Tfb1PH  
4  
5 with the highest affinity.  
6  
7

### 10 **N-Cap-(LL) analogs and p53TAD2 share a common binding site on Tfb1PH.**

11  
12  
13 To determine the binding site on Tfb1PH for the N-Cap-(LL) peptides, we performed NMR  
14  
15 titration and displacement experiments. Addition of unlabeled E-Cap-(LL) to <sup>15</sup>N-labeled Tfb1PH  
16  
17 produces changes in both the <sup>1</sup>H and <sup>15</sup>N chemical shifts for several signals in the <sup>1</sup>H-<sup>15</sup>N-HSQC  
18  
19 spectrum of Tfb1PH (**Supplementary Figure 1A-B, Supporting Information**). Like p53TAD2  
20  
21 <sup>32</sup>, the residues that exhibited the most significant changes are located within the strands β5, β6  
22  
23 and β7 and in the loop between β5 and β6 when mapped on the structure of Tfb1PH (**Figure 3A-**  
24  
25 **B**). Similar changes in chemical shifts are also observed in titrations of Tfb1PH with both A-Cap-  
26  
27 (LL) (**Supplementary Figure 1C, Supporting Information**) and W-Cap-(LL) (**Supplementary**  
28  
29 **Figure 1D, Supporting Information**). In addition, NMR displacement experiments demonstrate  
30  
31 that p53TAD2 and E-Cap-(LL) (**Figure 3C-D**) compete for a common binding site on Tfb1PH.  
32  
33  
34  
35  
36

37 In a previous study <sup>32</sup>, we identified five mutants of Tfb1PH [Tfb1PH (Q49A), Tfb1PH  
38  
39 (K57E), Tfb1PH (M59A), Tfb1PH (R61E), and Tfb1PH (M88A)] that significantly perturb  
40  
41 binding to p53TAD2. The five point mutations are located on the surface of Tfb1PH within the  
42  
43 β5, β6 and β7 strands (**Figure 4A**) and do not alter the structure of Tfb1PH. By ITC, we are  
44  
45 unable to detect binding of E-Cap-(LL) to the R61A mutant ( $K_d \geq 100 \mu\text{M}$ ) using this assay  
46  
47 (**Figure 4B**). In addition, the Q49A and M88A mutants decrease the binding of Tfb1PH to E-  
48  
49 Cap-(LL) by over 20-fold ( $K_d = 2.4 \pm 0.2 \mu\text{M}$  and  $1.9 \pm 0.2 \mu\text{M}$  respectively), whereas the K57E  
50  
51 and the M59A mutants decrease binding by approximately 10-fold ( $K_d = 1.0 \pm 0.1 \mu\text{M}$  and  $0.9 \pm$   
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 0.09  $\mu\text{M}$  respectively). These results support that E-Cap-(LL) is forming very similar  
4  
5 interactions with Tfb1PH as seen with p53TAD2.  
6  
7  
8  
9

### 10 **E-Cap-(LL) binds to CBP/p300.**

11  
12  
13 CREB-binding protein (CBP) and p300 (CBP/p300) are two highly homologous HATs that have  
14  
15 been shown to play an important role in regulating a number of transcriptional activators  
16  
17 including p53<sup>70,71</sup>. Four domains of CBP/p300 (TAZ1/CH1, TAZ2/CH3, KIX and IBiD) have  
18  
19 been shown to interact with the TAD of p53 and acetylation of p53 by CBP/p300 is essential for  
20  
21 p53-dependent activation<sup>72,73</sup>. ITC studies demonstrate that E-Cap-(LL) is able to interact with  
22  
23 both the KIX and IBiD domains of CBP/p300 ( $K_d=3.8 \pm 0.8 \mu\text{M}$  and  $0.84 \pm 0.12 \mu\text{M}$   
24  
25 respectively; **Figure 4C**). These results indicate that like the TAD of p53, E-Cap-(LL) is able to  
26  
27 interact with domains of CBP/p300 .  
28  
29  
30  
31  
32  
33  
34

### 35 **E-Cap-(LL) is a potent activator in vivo.**

36  
37  
38 To verify that E-Cap-(LL) activates transcription *in vivo*, it was fused to the DNA-binding  
39  
40 domain (DBD) of LexA, and its activation potential was measured in yeast cells. The activity of  
41  
42 E-Cap-(LL) for the *lacZ* reporter gene is measured relative to a positive control (Gal4 TAD-  
43  
44 LexA-DBD) whose activity is established as 100 %. In this system, the E-Cap-(LL)-LexA-DBD  
45  
46 fusion protein activates transcription at  $161 \pm 9 \%$  of the positive control. In comparison, the  
47  
48 native p53-13 peptide fused to the LexA-DBD activates transcription at only  $8 \pm 3 \%$ . Thus, E-  
49  
50 Cap-(LL) is ~20 times more potent than p53-13 as a transcriptional activator in this *in vivo*  
51  
52 system (**Figure 5**). To compare the *in vivo* activity of the E-Cap-(LL) relative to other known  
53  
54 artificial TADs, we compared its activity relative to two model artificial TADs, the AH<sup>47</sup> and  
55  
56  
57  
58  
59  
60

1  
2  
3 VP2<sup>45</sup> peptides. The AH-LexA-DBD and VP2-LexA-DBD activates transcription at  
4  
5 approximately 5 % of the positive control (**Figure 5**) which is similar to p53-13, but over 30-fold  
6  
7 less than E-Cap-(LL). These results strongly support the idea that stabilizing the helical character  
8  
9 of short analogs of p53TAD2 can lead to a significant enhancement of their *in vivo*  
10  
11 transcriptional activity.  
12  
13  
14

### 15 16 17 **The role of leucines for the *in vivo* activity of the p53TAD analogs.**

18  
19  
20 To investigate the role that the two leucine residues play in the ability of E-Cap-(LL) to activate  
21  
22 transcription *in vivo*, we mutated the leucine residue at position 9 to the native glutamine residue  
23  
24 found in p53-13 to generate E-Cap-(LQ). This change lowers the *in vivo* activity to  $67 \pm 8$  % of  
25  
26 the positive control, and thus E-Cap-(LQ) is ~50% less active than E-Cap-(LL) ( $161 \pm 9$  %;  
27  
28 **Figure 5**). Next, we inserted two leucine residues at equivalent positions of p53-13 to generate  
29  
30 p53-13-(LL). The *in vivo* transcriptional activation of p53-13-(LL) is 10-fold higher than p53-13  
31  
32 ( $95 \pm 6$  % versus  $8 \pm 3$  %), but 50% less active than E-Cap-(LL) (**Figure 5**). These results support  
33  
34 the idea that both the N-Cap motif, and the dileucine-bridge contribute significantly to the *in vivo*  
35  
36 transcriptional activity of E-Cap-(LL).  
37  
38  
39  
40  
41  
42  
43  
44

### 45 **Peptides with leucines in the i, i+4 spacing are less active *in vivo***

46  
47  
48 Experimental and theoretical studies indicate that two hydrophobic amino acids separated by  
49  
50 either three (i, i+3 spacing pattern) or four residues (i, i+4 spacing pattern) enhance the helical  
51  
52 propensity of peptides through side-chain interactions<sup>74-78</sup>. Given that the insertion of the two  
53  
54 leucines with an i, i, i+3 spacing pattern enhances the *in vivo* activation of E-Cap-(LL) and p53-  
55  
56 13-(LL), we tested the role of leucines with an i, i+4 spacing pattern on *in vivo* activity. Leucine  
57  
58  
59  
60

1  
2  
3 residues were introduced at positions 4 and 8 in the p53-13 peptide [p53-13-(LL4)] and at  
4  
5 positions 5 and 9 in the E-Cap-(LL) peptide [E-Cap-(LL4)] (**Figure 5**). These positions are again  
6  
7 chosen in an attempt to place the leucine bridge in the center of the peptide, but on the backside  
8  
9 of the helix relative to the Tfb1PH binding interface. In the yeast activation assay, LexA-p53-13-  
10  
11 (LL4) displays  $51 \pm 8$  % of the activity of GAL4-LexA (**Figure 5**). This corresponds to a 6-7-  
12  
13 fold increase in activity compared to the native p53-13 ( $8 \pm 3\%$ ), but only about half of the  
14  
15 activity that we observe when the leucines are in the i, i+3 spacing pattern in p53-13-(LL) ( $95 \pm$   
16  
17  $6$  %). Similarly, E-Cap-(LL4) displays  $60 \pm 2$  % of the activity of GAL4-LexA, but this  
18  
19 corresponds to roughly 40% of the activity compared to E-Cap-(LL) ( $161 \pm 9$  %). The increased  
20  
21 in vivo activity observed when two leucines were inserted in either the i, i+3 or i, i+4 spacing  
22  
23 pattern of p53-13 is consistent with in vitro studies with model peptides showing that both  
24  
25 spacing patterns are able to increase their helical propensity/stability<sup>74-78</sup>.  
26  
27  
28  
29  
30  
31  
32  
33

### **E-Cap-(LL) forms a helix in complex with Tfb1PH.**

34  
35  
36  
37 We have previously shown that p53TAD2 transitions from an unstructured state to form a nine-  
38  
39 residue  $\alpha$ -helix upon binding to Tfb1PH<sup>32</sup>. In this work, we attempted to increase the helical  
40  
41 propensity of the same region of p53TAD2 by adding an N-Cap and a dileucine bridge.  
42  
43 Although CD studies indicate that E-Cap-(LL) possesses a higher helical content than p53-13  
44  
45 (**Supplementary Table 1, Supporting Information**), NMR experiments with E-Cap-(LL) did  
46  
47 not show the presence of NOE signals characteristic of an  $\alpha$ -helical conformation in the free form  
48  
49 (**Supplementary Figure 2, Supporting Information**). Additional NMR studies of E-Cap-(LL)  
50  
51 in complex with Tfb1PH show that like p53TAD2, it transitions to form a nine-residue  $\alpha$ -helix  
52  
53 from Glu3 to Phe11 (**Figure 6A-B and Supplementary Figure 2, Supporting Information**).  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 The structure of the E-Cap-(LL) peptide in complex with Tfb1PH is calculated from 105 NOE-  
4 derived distance restraints and 22 dihedral angle restraints. An analysis of the twenty lowest-  
5 energy structures indicates that they have no NOE violation greater than 0.2 Å, no backbone  
6 dihedral angle violation greater than 2° and low pairwise rmsd values (**Table 1**). The structure of  
7 E-Cap-(LL) in complex with Tfb1PH confirms that the side-chains of Leu6 and Leu9 are in close  
8 proximity to each other, and on the opposite side of the helix relative to the binding interface with  
9 Tfb1PH (**Figure 6C-D**). In addition, the side-chain of Leu6 is in position to further stabilize the  
10 helix through contacts with the aromatic ring of Trp10 in the i+4 position (**Figure 6E-F**).  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

25 **Both i, i+3 and i, i+4 side-chain interactions contribute to E-Cap-(LL) activity.**

26  
27  
28 The NMR studies clearly indicate that Leu6 in the E-Cap-(LL) peptide is in position to enhance  
29 the stability of the helix through both i, i+3 interactions (L6-L9) and i, i+4 interactions (L6-W10).  
30 In order to verify this observation, two additional mutants were tested. In p53-13-(LQ), the  
31 second leucine of p53-13-(LL) is replaced with the native glutamine residue, whereas in the E-  
32 Cap-(DL) the first leucine of E-Cap-(LL) is replaced with the native aspartic acid residue of p53.  
33 In the yeast activation assay, LexA-p53-13-(LQ) displays 45 ± 8 % of the activity of the control  
34 (**Figure 5**). This corresponds to a 5-fold increase in activity compared to the p53-13 peptide (8 ±  
35 3%), but less than half of the activity we observe when the leucines are in the i, i+3 spacing  
36 pattern in p53-13-(LL) (95 ± 6 %). Likewise, E-Cap-(DL) displays 35 ± 4 % of the activity of  
37 the GAL4-LexA positive control, and this corresponds to roughly one-half of the activity of the  
38 E-Cap-(LQ) (67 ± 8 %). Consistent with the NMR structure of the E-Cap-(LL) peptide bound to  
39 Tfb1PH, these results indicate that both i, i+3 interactions (L6-L9) and i, i+4 interactions (L6-  
40 W10) are contributing to the in vivo activity of the p53-13-(LL) and E-Cap-(LL) analogs.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Discussion

ATAs have enormous potential for use as either therapeutic agents for treating human diseases or as biological probes for investigating the correlation between aberrant transcription and human diseases<sup>13,14</sup>. In designing an ATA, the ultimate goal is to construct both the DBD and the TAD component as efficiently as possible, while maintaining both specificity and activity<sup>13</sup>. There are now several methods for designing artificial DBDs that depend extensively on the availability of high-resolution structural information of protein:DNA complexes<sup>22,23,25-27,29</sup>. In contrast, there is far less structural information available for TADs in complex with their target proteins, and this is mostly due to the fact that TADs are generally intrinsically disordered in their free state. Unfortunately, this flexibility and adaptability adds to the complexity of designing artificial TADs. Thus, the key step for preparing a minimal artificial TAD is to identify structures of TADs in complex with target proteins that could serve as models for the design of artificial TADs.

In this work, we used the structure of the TAD of p53 bound to Tfb1PH as a template for designing ATAs<sup>32</sup>. Our hypothesis was that if we could enhance the helical propensity of the Tfb1PH interacting region from the TAD of p53 this would yield a more potent artificial TAD. Like p53, our designed peptide E-Cap-(LL) forms a 9-residue  $\alpha$ -helix when in complex with Tfb1PH and binds along the same interface. The binding interface is a shallow groove surrounded by positively-charged residues that help position the negatively-charged TAD so that its hydrophobic residues can participate in a series of van der Waals contacts with hydrophobic pockets dispersed along the groove. Superposition of the hydrophobic residues from p53TAD2 and E-Cap-(LL) peptide when bound to Tfb1PH demonstrates that the residues from the IExWF motif [Ile6-Phe11 in E-Cap-(LL)] are located in virtually identical positions (**Figure 6D**), and E-

1  
2  
3 Cap-(LL) appears to make the same contacts with Tfb1PH as p53TAD2. The structure is further  
4 supported by our ITC results in which mutations of the Tfb1PH residues (Gln49, Lys57, Met59,  
5 Arg61 and Met88) contributing to the interface with p53TAD2<sup>32</sup> also disrupt binding of the E-  
6  
7  
8  
9  
10  
11 Cap-(LL) peptide (**Figure 4A-B**).

12  
13  
14 The increased in vivo potency of E-Cap-(LL) relative to the native p53-13 peptide appears  
15 to be directly linked to its enhanced helical stability since both the N-Cap motif and the two  
16 leucines are required for maximal activity. The introduction of the hydrophobic interaction on the  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
The introduction of the hydrophobic interaction on the face opposite to the Tfb1PH interacting interface represents a potentially valuable strategy in designing artificial TADs with increased in vivo activity. Previous experimental studies with model peptides suggested that hydrophobic amino acids with an *i, i+4* spacing pattern had a slightly higher helix stabilizing effect than those with an *i, i+3* spacing pattern, but that either pattern can significantly increase the helical propensity of peptides relative to alanine<sup>74-77</sup>. Likewise, Monte Carlo simulations predict that the *i, i+3* and *i, i+4* spacing of leucines can enhance helical stability, but differed from the experimental results by suggesting that *i, i+3* spacing should be more effective<sup>78</sup>. In the case of the E-Cap-(LL) peptide, our mutation studies provide evidence that Leu6 can participate in *i, i+3* interaction with Leu9 and *i, i+4* interaction with Trp10 and thus both types of interactions are functioning. The fact that the NMR studies of E-Cap-(LL) in complex demonstrates the presence of both the *i, i+3* and the *i, i+4* interactions further confirms the important role of these interactions for the in vitro binding and in vivo activity.

52  
53  
54  
55  
56  
57  
58  
59  
60  
The importance of the spacing patterns for bridging interactions has also been observed with other helical stabilizing procedures such as stapled peptides and  $\beta$ -peptides<sup>79-81</sup>. It is clear that the location plays a huge factor when introducing helix-stabilizing modifications, and that

1  
2  
3 structural characterization can aid tremendously in selecting the location. In the case of E-Cap-  
4 (LL), it must still undergo a transition from partially unstructured to the helical conformation that  
5 binds Tfb1PH. However, the activation energy required for this transition is lowered by the  
6 presence of the N-Cap and the leucines, and this ultimately leads to a significantly higher level of  
7 activity in vivo. The key question that still remains is whether we can design a peptide locked in a  
8 helical conformation that would further enhance the activity in vivo beyond what we observe  
9 with E-Cap-(LL). To function in vivo TADs must interact with multiple targets using a  
10 “flycasting” like method<sup>82</sup>, and it may be crucial that, like E-Cap-(LL), they retain a minimal  
11 amount of flexibility to bind optimally to their different targets. Future studies are required to test  
12 this possibility using more constrained analogs of E-Cap-(LL) and structural studies will be  
13 crucial to developing and optimizing such constrained analogs.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29

## 30 **Acknowledgements**

31  
32  
33 We would like to thank Dr. G. Perretta and Mr. L. De Luca for technical assistance, and Drs.  
34 Paola Di Lello and Luca D’Andrea for helpful comments. We thank Alanna Schepartz for the  
35 CBP-Kix clone used in this work. This work was supported by a grant from the Canadian  
36 Cancer Society (J.G.O.). C.L. is a recipient of a Postdoctoral Fellowship from the FRSQ. J.L-V.  
37 is a Vanier Canada Graduate Scholar from the CIHR. T. M. is the recipient of a graduate  
38 fellowship from the CIHR. P.L. is a recipient of a Canadian Research Chair in Structural Biology  
39 and Engineering of RNA.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

## 53 **SI Available**

54  
55 Supplementary procedures describe details of peptide synthesis and protein purification.  
56  
57  
58  
59  
60

1  
2  
3 Supplementary Table 1 shows the results of circular dichroism studies. Supplementary Figure 1  
4  
5 shows HSQC spectra from NMR chemical shift perturbations studies of E-Cap-(LL) and A-Cap-  
6  
7 (LL) with  $^{15}\text{N}$ -labeled Tfb1PH. Supplementary Figure 2 shows HSQC spectra from NMR  
8  
9 chemical shift perturbations studies of Tfb1PH with  $^{15}\text{N}$ -labeled E-Cap-(LL). This information is  
10  
11 available free of charge via the Internet at <http://pubs.acs.org>.  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- 1
- 2
- 3
- 4
- 5
- 6
- 7 (1) Ptashne, M.; Gann, A. A. F. *Nature* **1997**, *386*, 569.
- 8 (2) Giniger, E.; Ptashne, M. *Nature* **1987**, *330*, 670.
- 9 (3) Sadowski, I.; Ma, J.; Triezenberg, S.; Ptashne, M. *Nature* **1988**, *335*, 563.
- 10 (4) Goodrich, J. A.; Hoey, T.; Thut, C. J.; Admon, A.; Tjian, R. *Cell* **1993**, *75*, 519.
- 11 (5) Brown, C. E.; Howe, L.; Sousa, K.; Alley, S. C.; Carrozza, M. J.; Tan, S.; Workman, J. L.  
12 *Science* **2001**, *292*, 2333.
- 13 (6) Prochasson, P.; Neely, K. E.; Hassan, A. H.; Li, B.; Workman, J. L. *Mol. Cell* **2003**, *12*, 983.
- 14 (7) Yang, F.; DeBeaumont, R.; Zhou, S.; Naar, A. M. *Proc. Natl. Acad. Sci. USA* **2004**, *101*,  
15 2339.
- 16 (8) Blau, J.; Xiao, H.; McCracken, S.; O'Hare, P.; Greenblatt, J.; Bentley, D. *Mol. Cell. Biol.*  
17 **1996**, *16*, 2044.
- 18 (9) Tansey, W. P.; Ruppert, S.; Tjian, R.; Herr, W. *Genes & Dev.* **1994**, *8*, 2756.
- 19 (10) Black, J. C.; Choi, J. E.; Lombardo, S. R.; Carey, M. *Mol. Cell* **2006**, *23*, 809.
- 20 (11) Gutierrez, J. L.; Chandy, M.; Carrozza, M. J.; Workman, J. L. *EMBO J.* **2007**, *26*, 730.
- 21 (12) Brown, S. A.; Weirich, C. S.; Newton, E. M.; Kingston, R. E. *EMBO J.* **1998**, *17*, 3146.
- 22 (13) Lee, L. W.; Mapp, A. K. *J. Biol. Chem.* **2010**, *285*, 11033.
- 23 (14) Graslund, T.; Li, X.; Magnenat, L.; Popkov, M.; Barbas, C. F., 3rd *J. Biol. Chem.* **2005**, *280*,  
24 3707.
- 25 (15) Rodriguez-Martinez, J. A.; Peterson-Kaufman, K. J.; Ansari, A. Z. *Biochim. Biophys. Acta-*  
26 *Gene Reg. Mech.* **2010**, *1799*, 768.
- 27 (16) Sera, T. *Adv Drug Deliv. Rev.* **2009**, *61*, 513.
- 28 (17) Visser, A. E.; Verschure, P. J.; Gommans, W. M.; Haisma, H. J.; Rots, M. G. *Adv. Genet.*  
29 **2006**, *56*, 131.
- 30 (18) Verschure, P. J.; Visser, A. E.; Rots, M. G. *Adv. Genet.* **2006**, *56*, 163.
- 31 (19) Hannon, G. J. *Nature* **2002**, *418*, 244.
- 32 (20) Rana, T. M. *Nature Rev. Mol. Cell Biol.* **2007**, *8*, 23.
- 33 (21) Kim, D. H.; Rossi, J. J. *Nat. Rev. Gen.* **2007**, *8*, 173.
- 34 (22) Dervan, P. B.; Doss, R. M.; Marques, M. A. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*,  
35 373.
- 36 (23) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363.
- 37 (24) Muzikar, K. A.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16598.
- 38 (25) Nielsen, P. E. *Quart. Rev. Biophys.* **2005**, *38*, 345.
- 39 (26) Chen, J.; Peterson, K. R.; Iancu-Rubin, C.; Bieker, J. J. *P Proc. Natl. Acad. Sci. USA* **2010**,  
40 *107*, 16846.
- 41 (27) Liu, B.; Han, Y.; Corey, D. R.; Kodadek, T. *J. Am. Chem. Soc.* **2002**, *124*, 1838.
- 42 (28) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
- 43 (29) Beerli, R. R.; Barbas, C. F., 3rd *Nat. Biotechnol.* **2002**, *20*, 135.
- 44 (30) Beerli, R. R.; Dreier, B.; Barbas, C. F., 3rd *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1495.
- 45 (31) Beerli, R. R.; Schopfer, U.; Dreier, B.; Barbas, C. F., 3rd *J. Biol. Chem.* **2000**, *275*, 32617.
- 46 (32) Di Lello, P.; Jenkins, L. M. M.; Jones, T. N.; Nguyen, B. D.; Hara, T.; Yamaguchi, H.;  
47 Dikeakos, J. D.; Appella, E.; Legault, P.; Omichinski, J. G. *Mol. Cell* **2006**, *22*, 731.
- 48 (33) Langlois, C.; Mas, C.; Di Lello, P.; Jenkins, L. M. M.; Legault, P.; Omichinski, J. G. *J. Am.*  
49 *Chem. Soc.* **2008**, *130*, 10596.
- 50 (34) Uesugi, M.; Nyanguile, O.; Lu, H.; Levine, A. J.; Verdine, G. L. *Science* **1997**, *277*, 1310.
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

- 1  
2  
3 (35) Radhakrishnan, I.; PerezAlvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright,  
4 P. E. *Cell* **1997**, *91*, 741.  
5 (36) Dames, S. A.; Martinez-Yamout, M.; Guzman, R. N. D.; Dyson, H. J.; Wright, P. E. *Proc.*  
6 *Natl. Acad. Sci. USA* **2002**, *99*, 5271.  
7 (37) Wojciak, J. M.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *EMBO J.* **2009**, *28*, 948.  
8 (38) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N.  
9 P. *Science* **1996**, *274*, 948.  
10 (39) Mattei, E.; Corbi, N.; Di Certo, M. G.; Strimpakos, G.; Severini, C.; Onori, A.; Desantis, A.;  
11 Libri, V.; Buontempo, S.; Floridi, A.; Fanciulli, M.; Baban, D.; Davies, K. E.; Passananti, C. *PLoS One* **2007**, *2*,  
12 e774.  
13 (40) Lu, Y.; Tian, C.; Danialou, G.; Gilbert, R.; Petrof, B. J.; Karpati, G.; Nalbantoglu, J. *J. Biol.*  
14 *Chem.* **2008**, *283*, 34720.  
15 (41) Stanojevic, D.; Young, R. A. *Biochemistry* **2002**, *41*, 7209.  
16 (42) Di Certo, M. G.; Corbi, N.; Strimpakos, G.; Onori, A.; Luvisetto, S.; Severini, C.;  
17 Guglielmotti, A.; Batassa, E. M.; Pisani, C.; Floridi, A.; Benassi, B.; Fanciulli, M.; Magrelli, A.; Mattei, E.;  
18 Passananti, C. *Human Mol. Genet.* **2010**, *19*, 752.  
19 (43) Wilber, A.; Tschulena, U.; Hargrove, P. W.; Kim, Y. S.; Persons, D. A.; Barbas, C. F.;  
20 Nienhuis, A. W. *Blood* **2010**, *115*, 3033.  
21 (44) Rowe, S. P.; Casey, R. J.; Brennan, B. B.; Buhrlage, S. J.; Mapp, A. K. *J. Am. Chem. Soc.*  
22 **2007**, *129*, 10654.  
23 (45) Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M. *Chem. Biol.* **2001**, *8*,  
24 583.  
25 (46) Beltran, A. S.; Sun, X.; Lizardi, P. M.; Blancafort, P. *Mol. Cancer Ther.* **2008**, *7*, 1080.  
26 (47) Ma, J.; Ptashne, M. *Cell* **1987**, *51*, 113.  
27 (48) Xiao, X.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew Chem. Int. Ed. Engl.* **2007**, *46*,  
28 2865.  
29 (49) Liu, B.; Alluri, P. G.; Yu, P.; Kodadek, T. *J. Am. Chem. Soc.* **2005**, *127*, 8254.  
30 (50) Xiao, X.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *J. Comb. Chem.* **2007**, *9*, 592.  
31 (51) Jung, D. J.; Shimogawa, H.; Kwon, Y.; Mao, Q.; Sato, S.; Kamisuki, S.; Kigoshi, H.; Uesugi,  
32 M. J. *Am. Chem. Soc.* **2009**, *131*, 4774.  
33 (52) Jung, D. J.; Choi, Y. M.; Uesugi, M. *Drug Discovery Today* **2006**, *11*, 452.  
34 (53) Minter, A. R.; Brennan, B. B.; Mapp, A. K. *J. Am. Chem. Soc.* **2004**, *126*, 10504.  
35 (54) Candau, R.; Scolnick, D. M.; Darpino, P.; Ying, C. Y.; Halazonetis, T. D.; Berger, S. L.  
36 *Oncogene* **1997**, *12*, 807.  
37 (55) Triezenberg, S. J.; Kingsbury, R. C.; McKnight, S. L. *Genes & Dev.* **1988**, *2*, 718.  
38 (56) Unger, T.; Nau, M. M.; Segal, S.; Minna, J. D. *EMBO J.* **1992**, *11*, 1383.  
39 (57) Sullivan, S. M.; Horn, P. J.; Olson, V. A.; Koop, A. H.; Niu, W.; Ebright, R. H.; Triezenberg, S.  
40 *J. Nucleic Acid Res.* **1998**, *26*, 4487.  
41 (58) Di Lello, P.; Jenkins, L. M. M.; Mas, C.; Langlois, C.; Malitskaya, E.; Fradet-Turcotte, A.;  
42 Archambault, J.; Legault, P.; Omichinski, J. G. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 106.  
43 (59) Marion, D.; Kay, L. E.; Sparks, S. W.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*,  
44 1515.  
45 (60) Ikura, M.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 2433.  
46 (61) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**,  
47 *6*, 277.  
48 (62) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, P.; Ulrich, E. L.;  
49 Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins* **2005**, *59*, 687.  
50 (63) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (64) Brunger, A. T.; Adams, P. D.; Clore, G. M.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.;  
4 Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta. Cryst.*  
5 **1998**, *D54*, 905.  
6 (65) Choy, W.-Y.; Tollinger, M.; Mueller, G. A.; Kay, L. E. *J. Biomol. NMR* **2001**, *21*, 31.  
7 (66) Laskowski, R. A.; Antoon, J.; Rullmann, C.; Macarthur, M. W.; Kaptein, R.; Thornton, J. M.  
8 *J. Biomol. NMR* **1996**, *8*, 477.  
9 (67) D'Andrea, L. D.; Iaccarino, G.; Fattorusso, R.; Sorriento, D.; Carannante, C.; Capasso, D.;  
10 Trimarco, B.; Pedone, C. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14215.  
11 (68) Aurora, R.; Rose, G. D. *Protein Sci.* **1998**, *7*, 21.  
12 (69) Diana, D.; Ziaco, B.; Colombo, G.; Scarabelli, G.; Romanelli, A.; Fedone, C.; Fattorusso, R.;  
13 D'Andrea, L. D. *Chemistry-a Eur. J.* **2008**, *14*, 4164.  
14 (70) Gu, W.; Roeder, R. G. *Cell* **1997**, *90*, 595.  
15 (71) Avantaggiati, M. L.; Ogryzko, V.; Gardner, K.; Giordano, A.; Levine, A. S.; Kelly, K. *Cell*  
16 **1997**, *89*, 1175.  
17 (72) Ferreón, J. C.; Lee, C. W.; Arai, M.; Martínez-Yamout, M. A.; Dyson, H. J.; Wright, P. E.  
18 *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6591.  
19 (73) Teufel, D. P.; Freund, S. M.; Bycroft, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2007**, *104*,  
20 7009.  
21 (74) Padmanabhan, S.; Baldwin, R. L. *Protein Sci.* **1994**, *3*, 1992.  
22 (75) Padmanabhan, S.; Baldwin, R. L. *J. Mol. Biol.* **1994**, *241*, 706.  
23 (76) Luo, P.; Baldwin, R. L. *Biophys. Chem.* **2002**, *96*, 103.  
24 (77) Munoz, V.; Serrano, L. *Nature Struct. Biol.* **1994**, *1*, 399.  
25 (78) Creamer, T. P.; Rose, G. D. *Protein Sci.* **1995**, *4*, 1305.  
26 (79) Kutchukian, P. S.; Yang, J. S.; Verdine, G. L.; Shakhnovich, E. I. *J. Am. Chem. Soc.* **2009**,  
27 *131*, 4622.  
28 (80) Bautista, A. D.; Appelbaum, J. S.; Craig, C. J.; Michel, J.; Schepartz, A. *J. Am. Chem. Soc.*  
29 **2010**, *132*, 2904.  
30 (81) Kim, Y. W.; Kutchukian, P. S.; Verdine, G. L. *Org. Lett.* **2010**, *12*, 3046.  
31 (82) Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8868.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Tables

**Table 1. Structural Statistics for E-Cap-(LL) in Complex with Tfb1PH<sup>a</sup>**

Restrains used for the structure calculations	
Total number of NOE distances restraints	105
Short-range (intraresidue)	49
Medium-range ( $ i-j  \leq 4$ )	56
Long-range	0
Number of dihedral angle restraints ( $\varphi, \psi$ )	22
Structural statistics	
Rms deviations from idealized geometry	
Bonds (Å)	0.0027±0.00007
Angles (deg)	0.4025±0.0038
Impropers (deg)	0.231745±0.0111
Rms deviations from distance restraints (Å)	0.0240±0.0005
Rms deviations from dihedral restraints (deg)	1.1482±0.015
Ramachandran statistics (%) <sup>b</sup>	
Residues in most favored regions	69.2
Residues in additional allowed regions	30.8
Residues in generously allowed regions	0
Residues in disallowed regions	0
Coordinate precision <sup>c</sup>	
Atomic pair wise rmsd (Å)	
Tfb1PH/E-Cap-(LL) complex	
Backbone atoms (C', C <sup>α</sup> , N)	0.50±0.17
All heavy atoms	1.72±0.32

<sup>a</sup> The 20 conformers with the lowest energies were selected for statistical analysis.

<sup>b</sup> Based on PROCHECK-NMR analysis.

## Figure Legends

**Figure 1. First and second designs of the p53TAD2 mimetics.** (A) Sequence alignments of p53-13 with the NC15 and the NC17 peptides. The four key residues (IExWF) of the helical binding interface of p53TAD2/Tfb1PH complex are highlighted in grey. The residues included in N- and C-capping motifs are boxed. (B) Sequence alignments of p53-13 peptide with the VEGF helical analog (QK), A-Cap-(LL), E-Cap-(LL) and W-Cap-(LL). As above, the key residues (IExWF) of the helical binding interface of p53TAD2/Tfb1PH complex are highlighted in grey and the N- and C-capping motifs are boxed. The two leucine residues forming the bridge with an  $i, i+3$  spacing are shown connected by a bridge.

**Figure 2. Dissociation constants ( $K_d$ ) for the interaction between Tfb1PH and the p53TAD2 analogs.** (A) Representative ITC thermogram obtained by successive additions of E-Cap-(LL) into a solution of Tfb1PH. (B) Dissociation constants ( $K_d$ ) for the binding of p53TAD2 peptide analogs to Tfb1PH as determined by ITC. All injections fit the single binding site mechanism with 1:1 stoichiometry.

**Figure 3. E-Cap-(LL) peptide share a common binding site on Tfb1PH with p53TAD2.** (A-B) Ribbon model of the NMR structure of free Tfb1PH<sup>32</sup>. Residues that undergo significant chemical shift changes in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Tfb1PH upon formation of the Tfb1PH/p53TAD2 complex are mapped in yellow (A), or upon formation of the Tfb1PH/ E-Cap-(LL) complex are mapped in red (B). (C) Overlay of a selected region from the two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for a 0.8 mM sample of  $^{15}\text{N}$ -labeled p53TAD2 in the free form (black) and in the presence of 1.0 mM unlabeled Tfb1PH (red). (D) Overlay of a selected region from the

1  
2  
3 two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for a 0.8 mM sample of  $^{15}\text{N}$ -labeled p53TAD2 in the free  
4 form (black), in the presence of 1.0 mM unlabeled Tfb1PH (red), and after addition of 0.8 mM  
5 unlabeled E-Cap-(LL) peptide (blue). Signals of  $^{15}\text{N}$ -labeled p53TAD2 that undergo significant  
6 changes in  $^1\text{H}$ - and  $^{15}\text{N}$ -chemical shifts upon formation of the complex with Tfb1PH, (C) and that  
7 return towards their original position following the addition of E-Cap-(LL) peptide (D) are  
8 indicated by arrows. See **Supplementary Figure 1** for spectra of titration of Tfb1PH with E-Cap-  
9 (LL), A-Cap-(LL) and W-Cap-(LL).  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21

22 **Figure 4. Comparison of the dissociation constants ( $K_d$ ) of E-Cap-(LL) binding to p53 target**

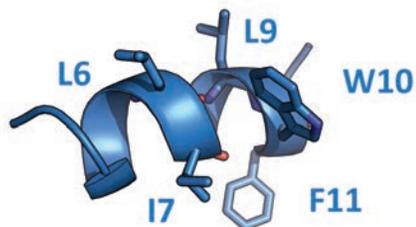
23 **proteins.** (A) Ribbon model of the NMR structure of Tfb1PH. The five residues located on the  
24 surface of Tfb1PH within the  $\beta 5$ ,  $\beta 6$  and  $\beta 7$  strands that are mutated are highlighted. (B)  
25 Comparison of the dissociation constants ( $K_d$ ) for the interaction of the E-Cap-(LL) analog with  
26 Tfb1PH and the five mutants as determined by ITC measurements. (C) Comparison of the  
27 dissociation constants ( $K_d$ ) for the interaction of the E-Cap-(LL) analog with Tfb1PH and the five  
28 mutants as determined by ITC measurements.  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40

41 **Figure 5. E-Cap-(LL) functions as a potent activation domains in yeast.** LexA-peptide fusion

42 proteins were co-transformed into yeast with the reporter LexA operator-Lac-Z fusion plasmid  
43 pSH18-34. Results are presented as the mean of the percentages of the  $\beta$ -galactosidase units of  
44 the tested fusion proteins on the  $\beta$ -galactosidase units of the LexA-GAL4TAD positive control.  
45 Error bars represent standard error about the mean of a minimum of three independent  
46 experiments  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 **Figure 6. NMR structure of E-Cap-(LL) peptide in complex with Tfb1PH.** (A) Overlay of the  
4 10 lowest-energy structures of the E-Cap-(LL) peptide in complex with Tfb1PH. The structures  
5 were superimposed using the backbone atoms C', C<sup>α</sup>, and N. (B) Ribbon model of the 10 lowest  
6 energy conformers of the E-Cap-(LL) peptide. (C) Overlay of the 10 lowest energy conformers of  
7 the E-Cap-(LL) peptide showing the relative position of the three hydrophobic residues I7, W10  
8 and F11. (D) Overlay of residues 45-57 of p53TAD2 (in orange) and the average structure of E-  
9 Cap-(LL) peptide (in blue). The three key hydrophobic residues I50, W53 and F54 of p53TAD2  
10 are located in similar orientations as I7, W10 and F11 of the E-Cap-(LL) peptide. (E) Ribbon  
11 model of the 10 lowest energy conformers of the E-Cap-(LL) peptide from the complex with  
12 Tfb1PH. The side chains of the leucines (L6 and L9) that form the bridge are highlighted. (F)  
13 Ribbon model of the average structure of the E-Cap-(LL) peptide from the complex with  
14 Tfb1PH. The side chains of the three hydrophobic residues (I7, W10 and F11) and the leucines  
15 (L6 and L9) are highlighted to show that they are on opposite faces of the helix. See  
16 **Supplementary Figure 2** for NMR spectra of E-Cap-(LL) in the absence and presence of  
17 Tfb1PH.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## TOC Graphic

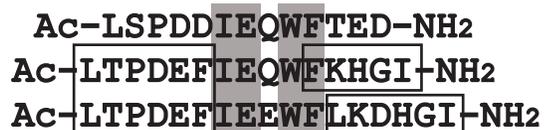


**A**

p53-13

NC15

NC17

**B**

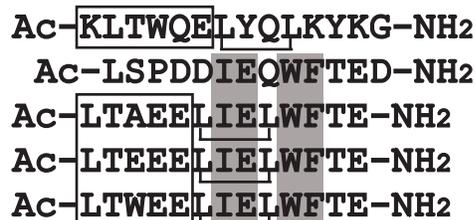
QK

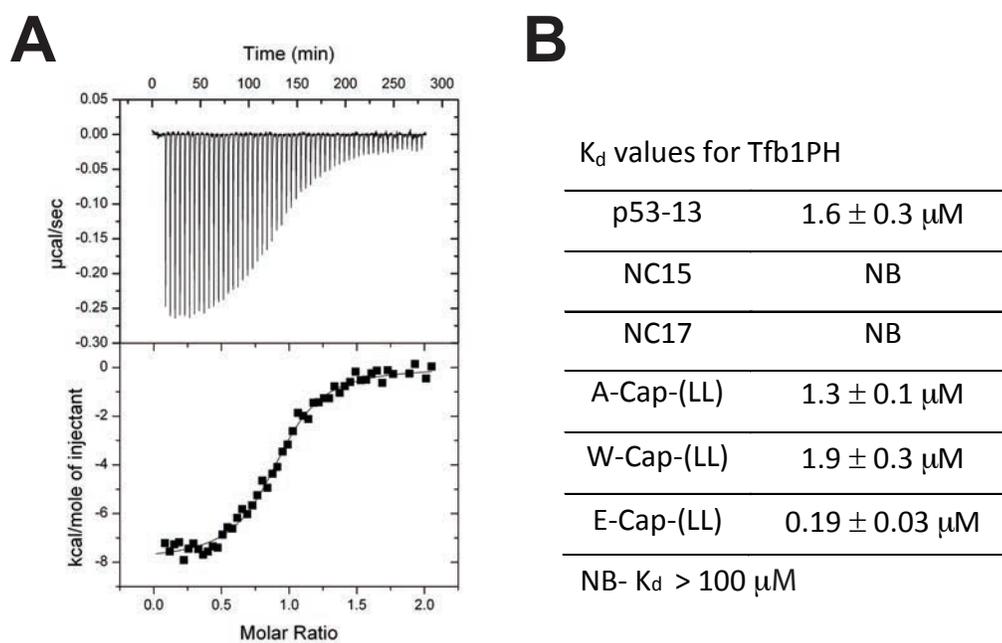
p53-13

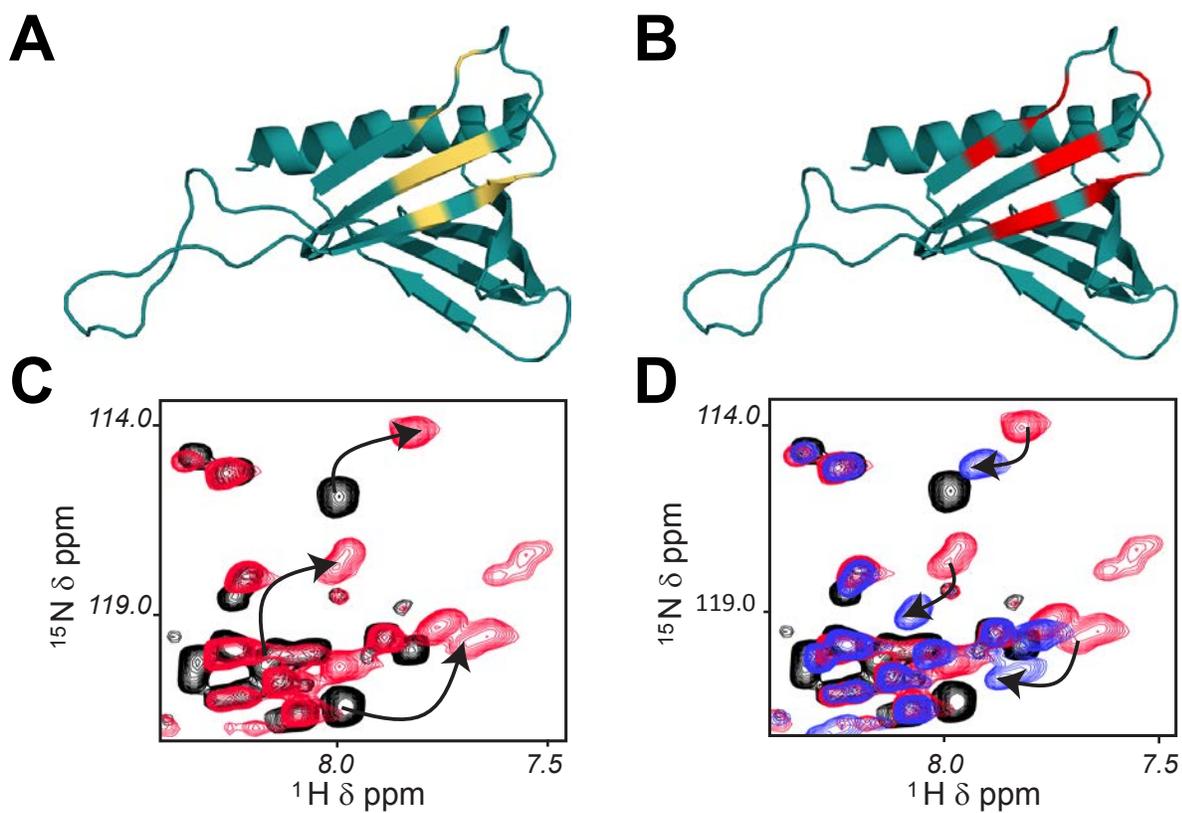
A-Cap- (LL)

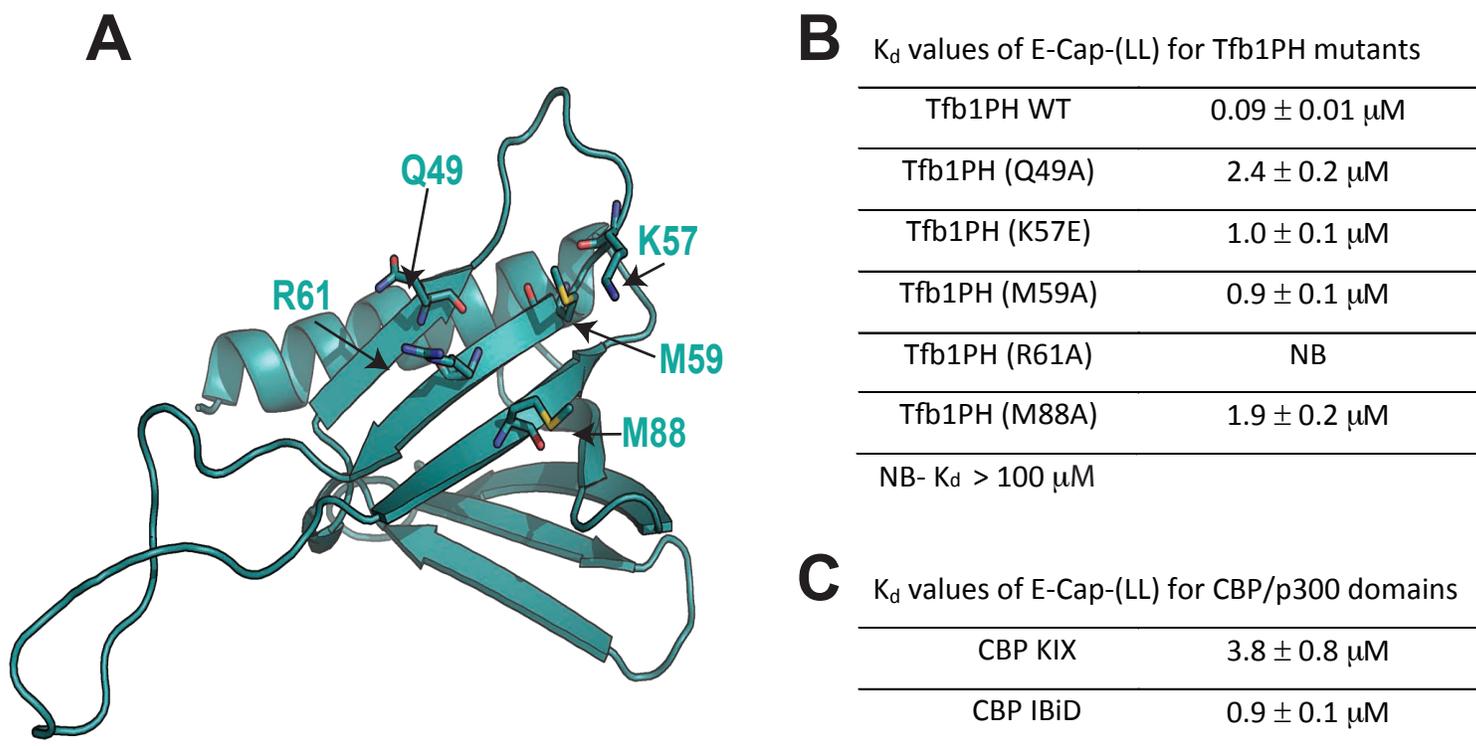
E-Cap- (LL)

W-Cap- (LL)

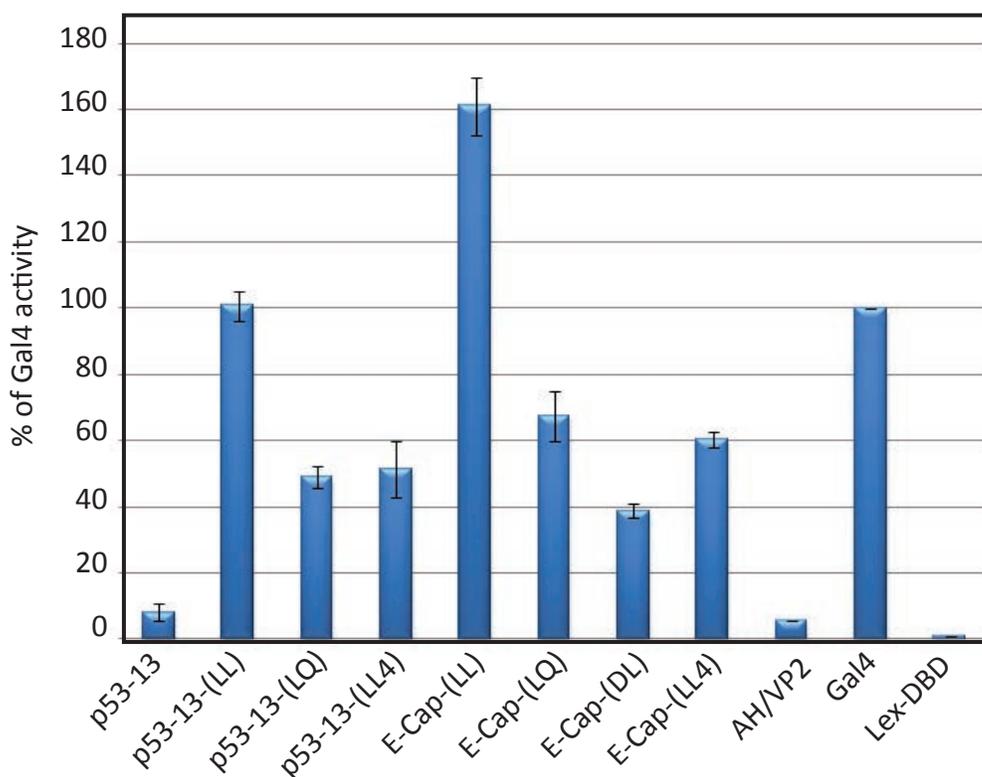
**Figure 1**

**Figure 2**

**Figure 3**

**Figure 4**



**Figure 5**

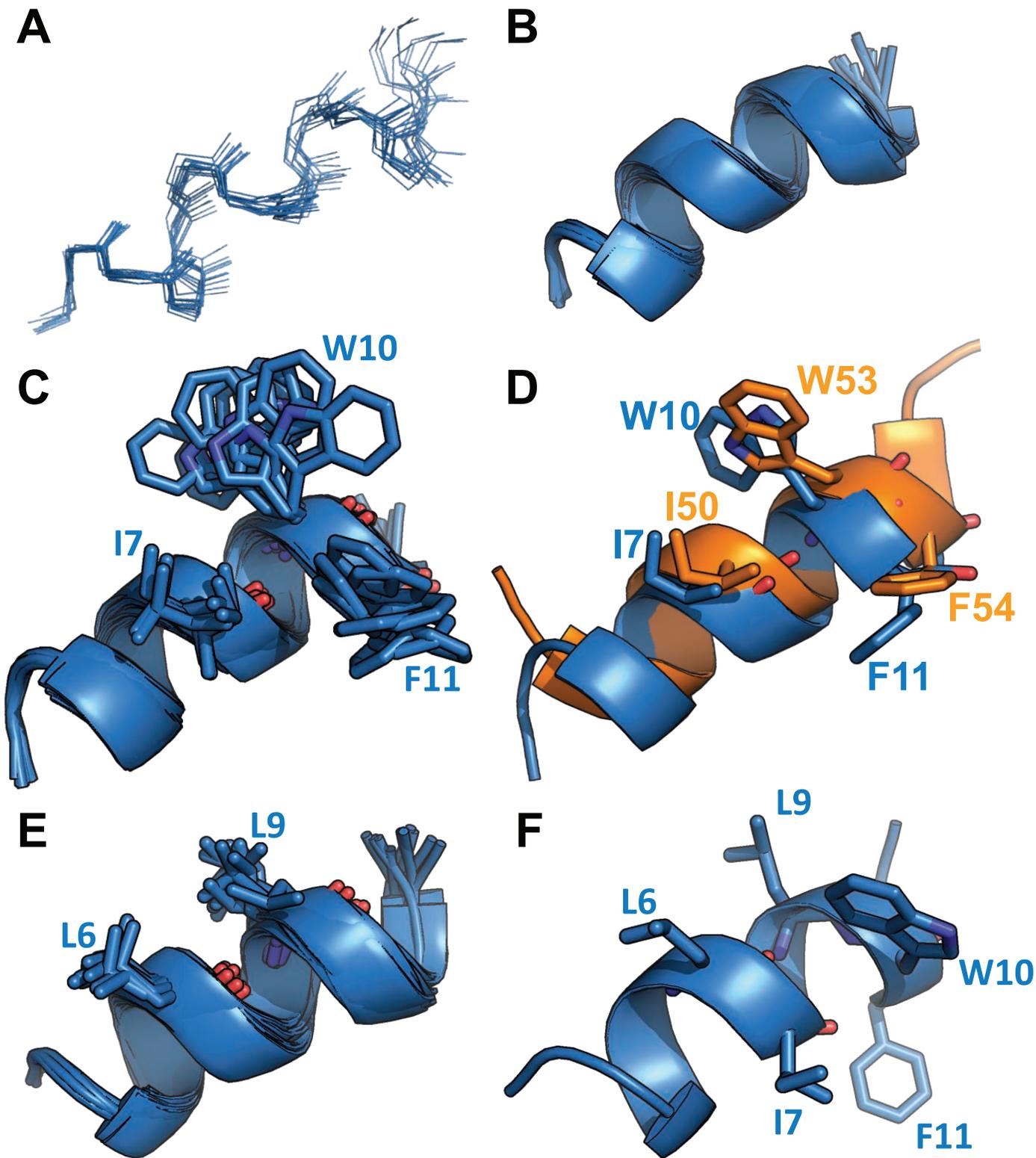


Figure 6