Rational design of cyclic peptide inhibitors of U2AF homology motif (UHM) domains to modulate pre-mRNA splicing

Pravin Kumar Ankush Jagtap†,‡,#, Divita Garg†,‡,#, Tobias G. Kapp‡,§, Cindy L Will$, Oliver Demmer‡,§, Reinhard Lührmann$, Horst Kessler‡,§ and Michael Sattler†,‡,\*

†Institute of Structural Biology, Helmholtz Zentrum München

‡Center for Integrated Protein Science Munich (CIPSM), Department Chemie, Technische Universität München, Lichtenbergstrasse 4, 85747 Garching, Germany.

§Institute for Advanced Study (IAS), Technische Universität München, Lichtenbergstrasse 4, 85747 Garching, Germany.

$Max Planck Institute for Biophysical Chemistry, Department of Cellular Biochemistry, Am Fassberg 11, 37077 Göttingen, Germany

KEYWORDS. Splicing regulation • UHM domains • cyclic peptides • NMR • protein-protein interactions.

**ABSTRACT**. U2AF homology motifs (UHMs) are atypical RNA Recognition Motif (RRM) domains that mediate critical protein-protein interactions during the regulation of alternative pre-mRNA splicing and other processes. The recognition of UHM domains by UHM Ligand Motif (ULM) peptide sequences plays important roles during early steps of spliceosome assembly. Splicing factor 45 kDa (SPF45) is an alternative splicing factor implicated in breast and lung cancer and splicing regulation of apoptosis-linked pre-mRNAs by SPF45 was shown to depend on interactions of its UHM domain with ULM motifs in constitutive splicing factors. We have developed cyclic peptide inhibitors that target UHM domains. By screening a focused library of linear and cyclic peptides and performing structure-activity relationship (SAR) analysis, we designed cyclic peptides with 4-fold improved binding affinity for the SPF45 UHM domain compared to native ULM ligands and 270-fold selectivity to discriminate UHM domains from alternative and constitutive splicing factors. These inhibitors are useful tools to modulate and dissect mechanisms of alternative splicing regulation.

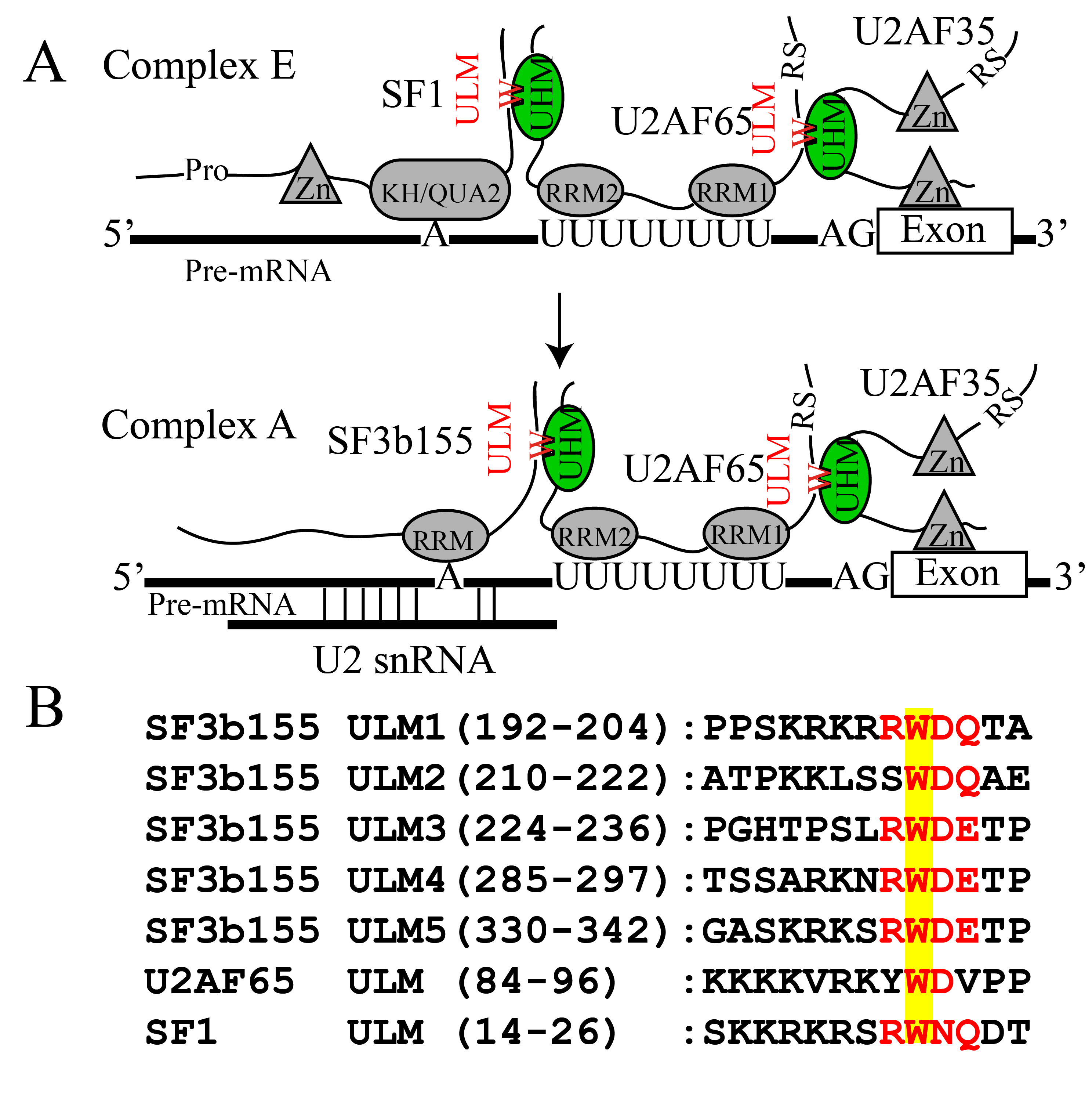
**Introduction**

Alternative splicing (AS) is an essential cellular process and greatly expands the coding capacity of eukaryotic genomes by generating multiple protein isoforms from a single primary transcript[1](#_ENREF_1), [2](#_ENREF_2). The regulation of AS involves the recognition of *cis* regulatory elements, i.e. short RNA sequence motifs, by *trans*-acting factors, i.e. RNA binding proteins[3-5](#_ENREF_3). Aberrant splicing has been implicated in human diseases[6](#_ENREF_6), [7](#_ENREF_7), and the role of AS in cancer progression suggests that targeting splicing regulation may be effective for cancer therapy[8](#_ENREF_8), [9](#_ENREF_9).

Splicing factor SPF45 was identified as a component of the spliceosome[10](#_ENREF_10) and has been shown to activate a cryptic 3’ splice site in -thalassemia[11](#_ENREF_11). It harbors a C-terminal U2AF homology motif (UHM) domain that mediates protein-protein interactions. The SPF45 UHM domain is required to regulate alternative splicing of the pre-mRNA encoding the apoptosis promoting factor FAS[12](#_ENREF_12). Given that the switch in expression of pro- and anti-apoptotic isoforms of FAS is tightly regulated[13](#_ENREF_13), an imbalance in the FAS isoforms by overexpression of SPF45 could provide a means for tumor cells to escape apoptosis. This links overexpression of SPF45 to breast, lung, colon and ovarian tumors[14](#_ENREF_14) and to the reported multi-drug resistance to anticancer drugs[14](#_ENREF_14), [15](#_ENREF_15).

UHM domains are atypical RNA recognition domains containing a signature Arg-Xaa-Phe (RXF) amino acid motif (**Figure S1**). They were identified as protein recognition motifs in both subunits of the U2AF heterodimer[16](#_ENREF_16), 17 and have since then been found in many other proteins with diverse biological functions[12](#_ENREF_12), [18-22](#_ENREF_18). UHM domains recognize UHM ligand motif (ULM) peptide sequences that harbor a conserved tryptophan residue flanked by a stretch of basic and negatively-charged residues[12](#_ENREF_12), [18](#_ENREF_18), [21](#_ENREF_21). UHM-ULM interactions play a crucial role in early spliceosome assembly (**Figure 1A,B**) and a given UHM domain can bind to various ULM ligands. For example, SPF45 UHM binds to SF3b155, SF1 and U2AF65 ULMs[12](#_ENREF_12) and U2AF65 UHM binds to all five ULMs from SF3b155 with low micro-molar affinity[23](#_ENREF_23), [24](#_ENREF_24).

Here, we have developed a cyclic peptide UHM inhibitor that targets the SPF45 UHM with high selectivity as a molecular probe to delineate the contributions of UHM domains in splicing regulation. The cyclic peptide inhibitor is optimized to target the SPF45 UHM domain, and exhibits 4-fold higher affinity and 270-fold selectivity compared to the constitutive splicing factor U2AF65. We show that the inhibitor is an efficient tool for studying molecular mechanisms of splicing regulation during early stages of spliceosome assembly.



**Figure 1**. (A) Schematic overview of UHM/ULM interactions during spliceosome assembly. UHM domains and ULM peptide motifs are shown in green and red colors respectively. (B) Alignment of ULM sequences.

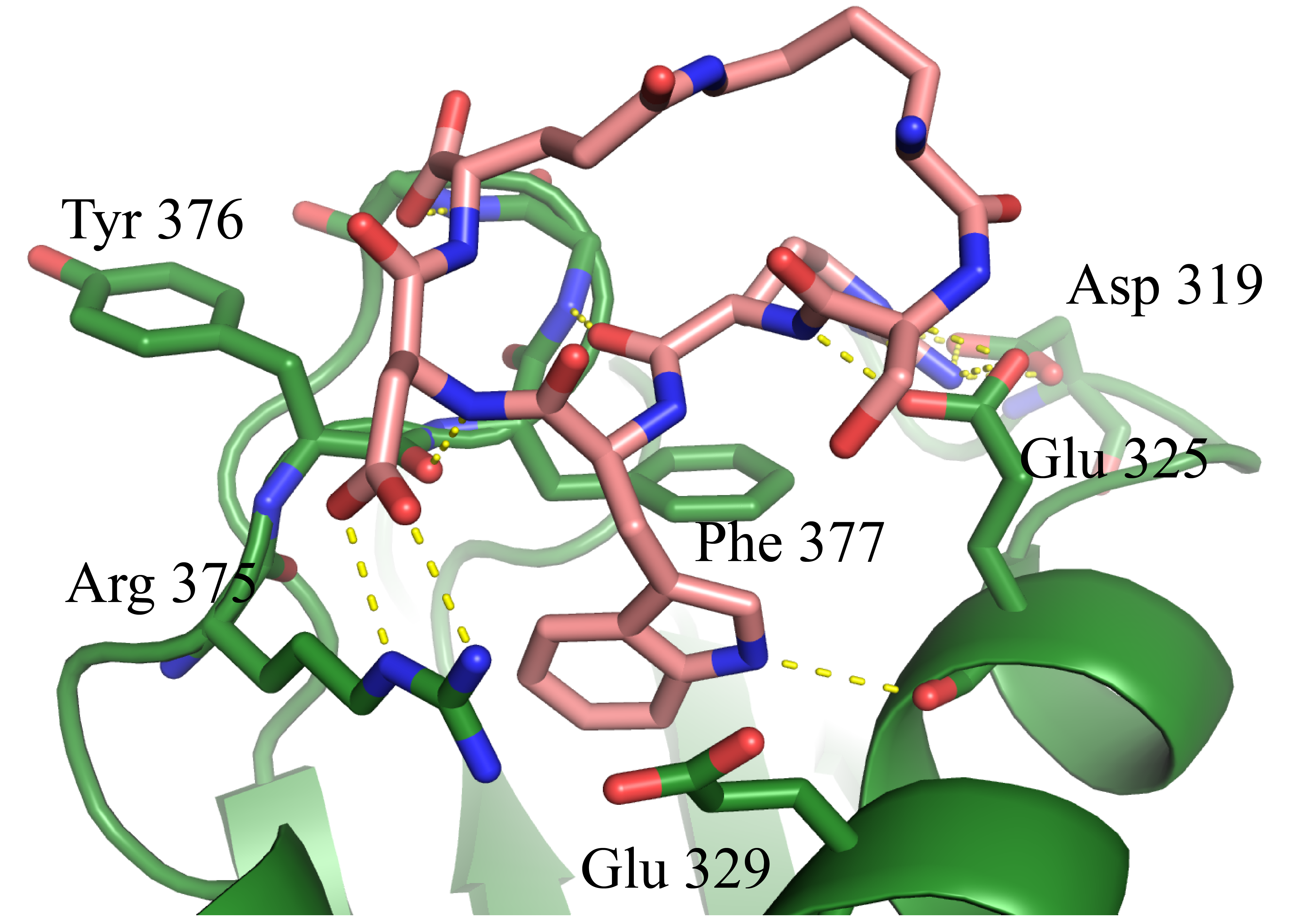
**Results and Discussion**

**Design of minimal ULM cyclic peptide.** Analysis of the crystal structure of SPF45 UHM-SF3b155 ULM5 complex[12](#_ENREF_12) (PDB accession: 2PEH) reveals that the ULM peptide adopts a -turn with a hydrogen bond between the backbone amide of Glu340 and carbonyl of Arg337 (**Figure S2**). Ser336 stabilizes this -turn by forming an intra-peptide hydrogen bond between its side chain hydroxyl and the backbone amide of Trp338 though it does not contribute directly to the interaction with the UHM domain. To understand therequirement of a pre-defined tertiary structure of the ULM peptide, we replaced Ser336 by alanine and analyzed the thermodynamic contributions to binding by isothermal titration calorimetry (ITC).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 1**. Binding affinity (determined by isothermal titration calorimetry) of different cyclic peptides synthesized to explore different linker lengths for cyclization. | | | | |
| Peptide | Sequence[a] | *K*D (M) | H[b] | -TS[b] |
| **0** | KSRWDE | 15.7±0.98 | -10.4±36 | 3.8±.17 |
| **1** | [sc,bb(KSRWDE)] | 122.3±3.9 | -7.8±.23 | 2.5±.21 |
| **2** | [sc,sc(OrnSRWDE)] | 2.0±.3 | -16.5±.55 | 8.7±.55 |
| **3** | [sc,sc(KSRWDE)] | 1.4±.27 | -18.3±1.3 | 10.3±1.4 |
| *a*Errors represent standard deviation of fitting errors calculated by error propagation. *b*kcal/mol | | | | |

Removal of serine shows an 8-fold decrease in binding affinity owing to a simultaneous loss of binding enthalpy and entropy compared to the native ULM sequence (RKSRWDETP vs. RKARWDETP; **Table S1**). In order to further understand the role of Ser336, we compared the binding affinity of two minimal ULM peptides, SRWDET and RWDET (**Table S1**). Inclusion of serine leads to a significant gain of favorable entropy, however at the cost of binding enthalpy. This reflects enthalpy-entropy compensation effects and highlights the need to maintain sufficient peptide flexibility in order to reduce enthalpy losses by restricting the conformation of the peptide. As Ser336 supports the -turn conformation, we hypothesized that the ULM peptide -turn can be further stabilized via side chain cyclization of Lys335 and Glu340 which are not in direct contact with the UHM domain. Different types of cyclizations were used: forming an amide bond between the backbone carboxyl group of Glu340 (peptide **1**) or the side chain carboxyl group of Glu340 (peptide **3**) to the side chain of Lys335 or the side chain of Glu340 to ornithine instead of Lys335 (peptide **2**). The rationale of these cyclizations was to explore different linker lengths, possible cyclization options and to maintain sufficient flexibility in the cyclized peptide (**Table 1**). Although decreasing the linker length of the cyclized part in peptide **1** significantly reduced the unfavorable binding entropy contribution compared to the non-cyclized peptide **0**, it also reduced the binding enthalpy. Nevertheless, peptide **2** and peptide **3** exhibit higher binding enthalpy compared to peptide **1**. This highlights the importance of peptide flexibility to enable optimal contacts with the UHM domain. As peptide **3** has the highest binding affinity with a dissociation constant *K*D = 1.4 M, it was chosen for further optimization.

**Crystal structure of SPF45 UHM - peptide 3 complex.** To investigate the structural basis of the peptide **3** interaction, we determined the crystal structure of the SPF45 UHM domain bound to peptide **3** (**Table S2**). The structure shows clear electron density for the cyclic peptide in the ULM binding pocket of the UHM (**Figure S3**). Overall, there are only minor structural differences in the UHM domain when bound to the linear versus the cyclic peptide. The backbone amides and carbonyls of the cyclic peptide **3** form an extensive network of hydrogen bonds with the backbone amides of the UHM domain. The Trp338 side-chain inserts into the hydrophobic tryptophan binding pocket and forms parallel and T-stacking interactions with the side chains of Arg375 and Phe377 in the UHM domain, respectively. Thus, critical interactions with the UHM domain are preserved between the linear and cyclic peptide bound forms (**Figure 2**).



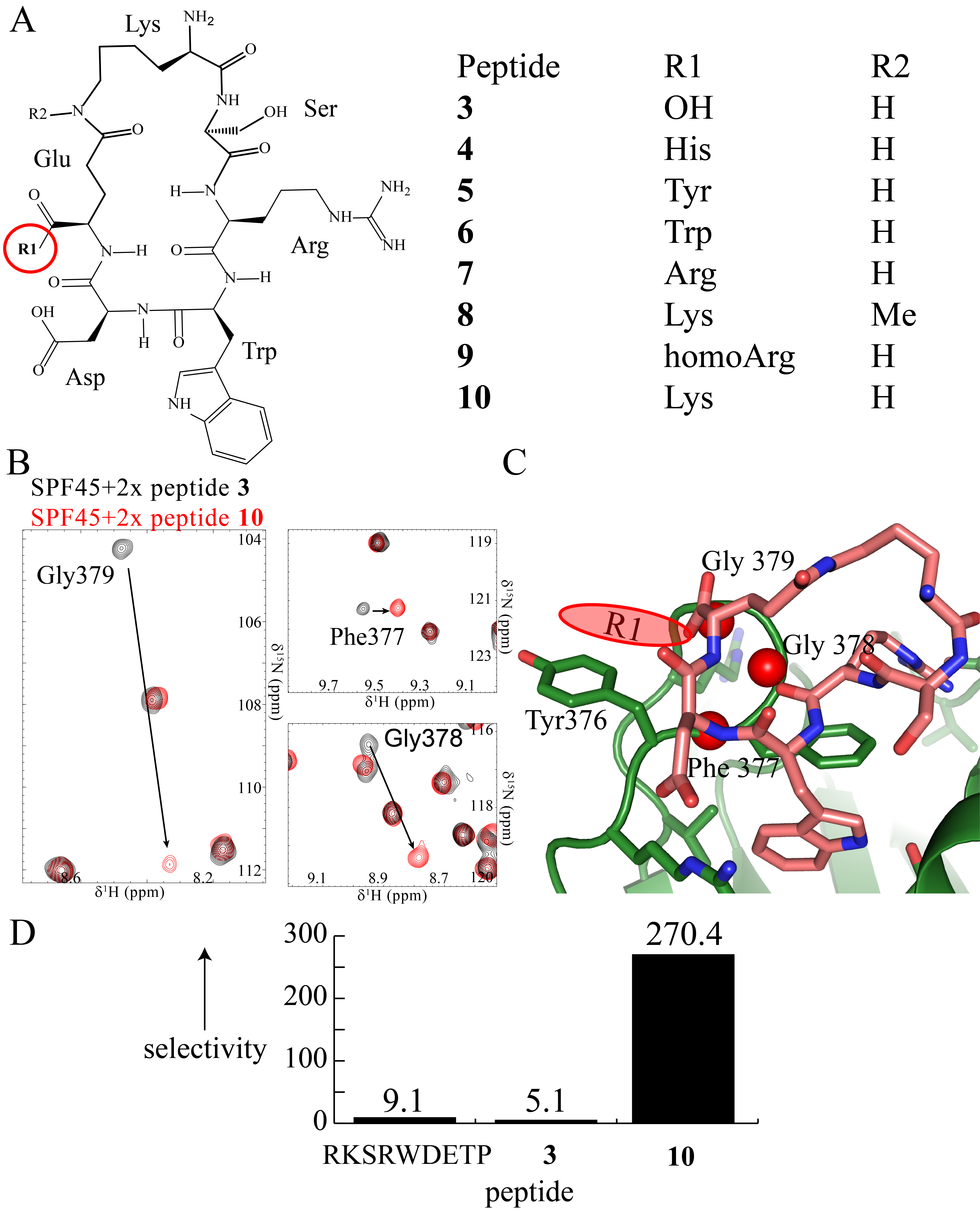
**Figure 2.** Co-crystal structure of SPF45 UHM in complex with peptide **3** inhibitor. The peptide and the UHM domain are shown in salmon and green, respectively. Yellow dashed lines indicate hydrogen bonds.

The conformation of the cyclic peptide shows some differences compared to the linear peptide when bound to the SPF45 UHM domain (backbone coordinate r.m.s.d. 0.97 Å; **Figure S4**). As expected, most of the structural differences are limited to the Lys335 and Glu340 side-chains, which mediate the cyclization. Interestingly, when bound to the SPF45 UHM domain (PDB accession: 2PEH), the side-chain of Lys335 does not show visible electron density, while the terminal atoms of Lys335 and Glu340 are clearly visible with the cyclic peptide. This indicates a significant rigidification of the peptide induced by the cyclization.

**Structure based optimization of peptide 3.** For further optimization of the affinity of peptide **3**, a focused library of seven peptides with different amino acid additions at carboxyl termini of Glu340 were synthesized (position R1 in **Figure 3A**) to enable either - or cation- interactions with Tyr376 in the UHM domain (**Figure 3C**). Addition of aromatic amino acids at position R1 exhibits a strong effect on favorable binding entropy (**Table 2**). This presumably is linked to the hydrophobic effect[25](#_ENREF_25) that arises from aromatic stacking of Tyr376 and the aromatic side chain in the cyclic peptide. However, as observed with other peptide ligands, the affinity increase obtained from favorable binding entropy is often counteracted by a decrease in binding enthalpy. Adding lysine at position R1 (peptide **10**) maintains the binding enthalpy and yields a ~4 fold increase in binding affinity (*K*D = 180 nM) for the cyclic peptide compared to the native linear peptide (**Table 2**).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2**. Binding affinity (determined by isothermal titration calorimetry) of cyclic peptides synthesized during optimization. | | | | |
| Peptide | Sequence[a] | *K*D (M) | H[b] | -TS[b] |
| **4** | [sc,sc(KSRWDE)]-H | 0.96±.08 | -10.0±.19 | 1.8±.19 |
| **5** | [sc,sc(KSRWDE)]-Y | 1.85±.21 | -9.2±.31 | 1.4±.32 |
| **6** | [sc,sc(KSRWDE)]-W | 0.80±.14 | -11.0±.54 | 2.6±.55 |
| **7** | [sc,sc(KSRWDE)]-R | 0.51±.12 | -13.7±.76 | 5.2±.77 |
| **8** | [sc,sc((NƐMe)KSRWDE)]-K | 1.26±.11 | -9.2±.20 | 1.1±.21 |
| **9** | [sc,sc(KSRWDE)]-homoR | 0.75±.05 | -20.2±.19 | 11.9±.19 |
| **10** | [sc,sc(KSRWDE)]-K | 0.18±.02 | -16.4±.18 | 7.2±.19 |
| *a*Errors represent standard deviation of the fitting errors calculated by error propagation *b*kcal/mol | | | | |

To test whether the lysine at position R1 in peptide **10** indeed contacts Tyr376, we monitored chemical shift perturbations (CSPs) in 1H,15N NMR correlation experiments upon addition of peptide **3** and peptide **10** to 15N-labeled SPF45 UHM domain. Overlay of the two spectra at saturating concentrations of the cyclic peptides shows that many signals exhibit similar chemical shifts in the bound state. However, significant chemical shift differences are found for the backbone amides of residues Phe377, Gly378 and Gly379 which flank Tyr376 (**Figure 3B,C**). These differential CSPs may reflect ring current effects induced by a conformational rearrangement of the Tyr376 sidechain caused by stacking of attached lysine at position R1 in the peptide **10**, and changes in hydrogen bond strengths of the amide groups of these residues.

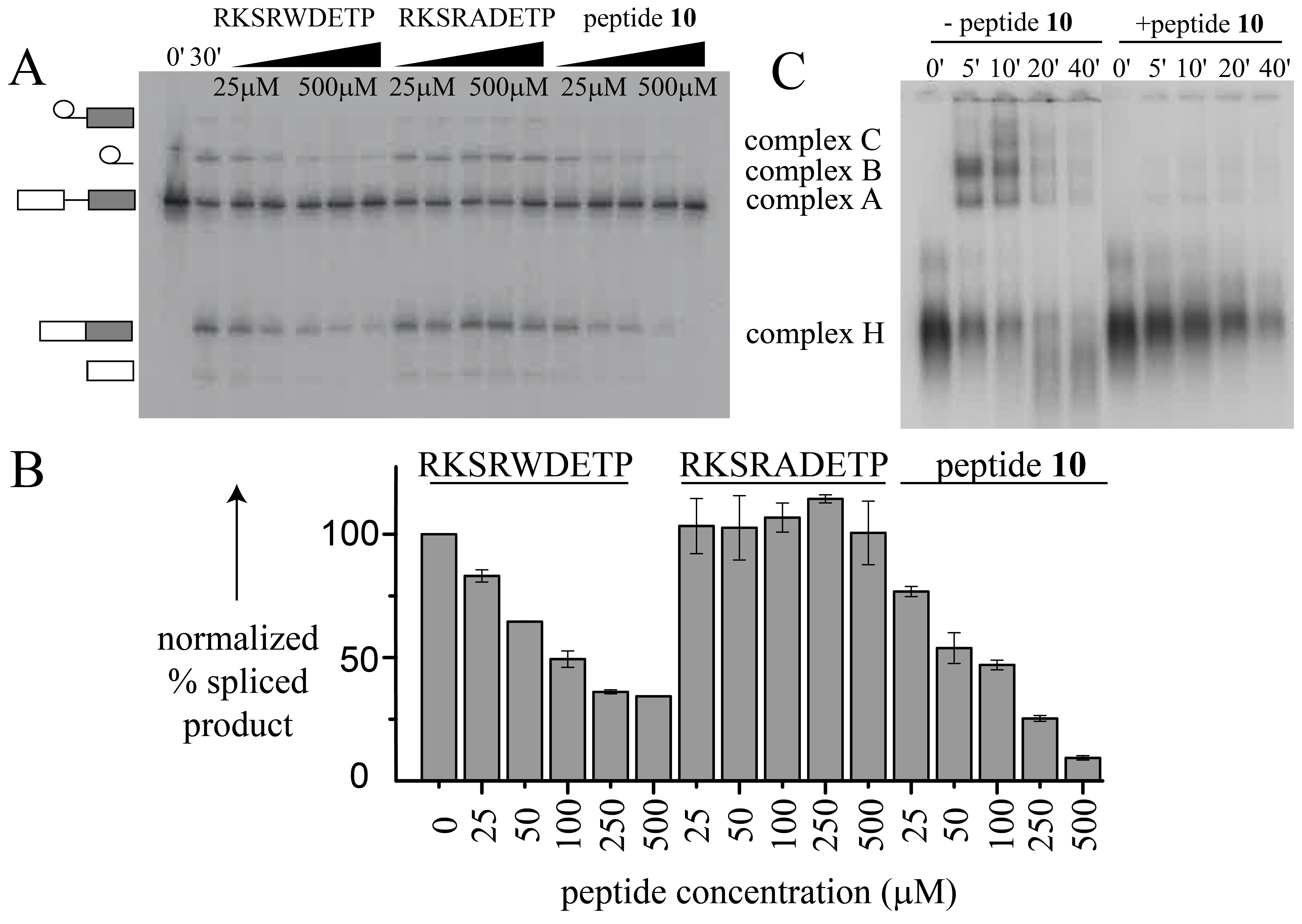


**Figure 3**. (A) Schematic overview of the synthesized focused library of the cyclic peptides. The different amino acids and functional groups added to the peptide **3** at R1 and R2 positions, respectively, are listed. (B) 1H, 15N HSQC spectra superposition of SPF45 UHM domain bound to peptide **3** (black) and peptide **10** (red) is shown. In the SPF45 UHM domain, residues Gly379, Gly378 and Phe377 exhibit large chemical shift changes induced by peptide **10** binding compared to peptide **3**. These residues are in close proximity to Tyr376. (C) Backbone amides of Gly379, Gly378 and Phe377 residues are shown with red spheres on the SPF45 UHM- peptide **3** complex structure. A red oval shows the attachment position of amino acid R1 in peptide **3** along with the direction in which it is expected to extend. (D) Selectivity profile of the designed cyclic peptides for SPF45 and U2AF65 UHM domains (Selectivity=*K*DU2AF65/*K*DSPF45; **Table S3)**.

**Selectivity of peptide 10.** The linear ULM peptide exhibits some selectivity (~9-fold) towards the U2AF65 UHM (**Figure 3D**). The cyclization of the peptide (peptide **3**) alone does not lead to an increase in selectivity, as the cyclic peptide does not establish novel interactions with the UHM domain but mainly stabilizes the -turn conformation. However, the lysine side chain variant of the cyclic peptide (peptide **10**) affords 270-fold selectivity against the U2AF65 UHM domain. Analysis of a multiple sequence alignment of UHM domains (**Figure S1**) shows that besides the variable X position (Tyr376 in the SPF45 UHM) in the RXF motif, the residues following this motif (Gly378, Gly379) are variable amongst different UHMs, which might be exploited to discriminate UHMs and increase selectivity of inhibitors. Notably, amide signals of these two residues exhibit largest chemical shift differences upon titration with the cyclic peptide **10**, suggesting that peptide **10** indeed exploits interactions with this region. The Lys addition in peptide **10** targets this variable region and increases affinity of the cyclic peptides, thereby providing enhanced selectivity and affinity against the SPF45 UHM. This strategy may be used for the design of more potent, selective cyclic peptides against specific UHM domains.

***In-vitro* splicing assays.** We next tested the functional activity of peptide **10** using *in vitro* splicing assays. Assembly of the early spliceosomal complex E, which precedes formation of the spliceosomal A complex, depends on UHM-ULM interactions involving U2AF35, U2AF65, SF1 and SF3b155[23](#_ENREF_23), [24](#_ENREF_24). The SF3b155-ULM5 has been previously shown to bind to both the SPF45-UHM[12](#_ENREF_12) and the U2AF65-UHM[23](#_ENREF_23), [24](#_ENREF_24) with *K*D = 0.7 M and 6.7 M, respectively (Supplementary **Table S1** and **S3**). On the other hand, the U2AF35-ULM binds to U2AF65-UHM with nanomolar affinity (*K*D = 135 nM)[12](#_ENREF_12), [16](#_ENREF_16). This suggests that the cyclic peptide inhibitor may interfere with UHM-ULM interactions involving the SF3b155 ULM and thereby modulate splicing. To test this hypothesis we monitored splicing and splicing complex formation *in vitro* with an IgM pre-mRNA as splicing substrate[26](#_ENREF_26) in the presence of peptide **10**. The native ULM5 of SF3b155 and its W338A variant served as positive and negative controls respectively (**Figure 4A,B**). Splicing was nearly abolished at the highest concentration of peptide **10** tested. Similar results were observed with the MINX pre-mRNA, which possesses a stronger 3’ splice site compared to IgM pre-mRNA (**Figure S5**). However, the extent of splicing inhibition for IgM and MINX pre-mRNA differed significantly, i.e. 9% spliced product formed relative to the no peptide control vs 29%, respectively, at the highest tested concentrations of peptide **10**. This is consistent with a differential requirement of UHM-ULM interactions for splicing of substrates with weak versus strong 3’ splice sites.

To confirm that peptide **10** indeed modulates UHM-ULM interactions during spliceosome assembly and rule out that the observed splicing inhibition is due to other factors, splicing complex assembly was monitored on an agarose gel (**Figure 4C**). In the absence of the peptide inhibitor, splicing complexes A, B and C are formed during the course of the reaction. In the presence of peptide **10**, assembly is stalled already at the A complex stage, where SF1 is replaced by SF3b155[27](#_ENREF_27), [28](#_ENREF_28), indicating that the cyclic peptide indeed inhibits UHM-ULM interactions, consistent with previous results[26](#_ENREF_26).



**Figure 4.** (A) *In vitro* splicing inhibition assays with peptide **10**. SF3bULM (RKSRWDETP) and its mutant (RKSRADETP) were used as positive and negative controls, respectively. (B) Quantification of the splicing assay showing % spliced product (normalized to the control without added peptide) formed vs. the peptide concentration. Error bars represent standard deviation of the % spliced product measured in two independent experiments. (C) Spliceosome complex formation analyzed by agarose gel electrophoresis. Peptide **10** inhibits complex A formation during spliceosome assembly, where U2AF UHM-ULM interactions play a crucial role.

**Conclusion**

Aberrant splicing regulation is the cause of a growing number of pathological conditions including cancer and neuro-degenerative diseases. Early spliceosome assembly offers an attractive target for inhibiting and modulating splicing as most of the splicing regulation takes place during this stage. Moreover, stalling of early spliceosome assembly provides novel opportunities for the structural analysis of these early complexes, which have been difficult to capture so far. Here, we have developed inhibitors of protein-protein interactions between spliceosomal UHM domains and ULM peptides, which play a crucial role during early spliceosome assembly.

We have designed a cyclic peptide that selectively binds the UHM domain of the alternative splicing factor SPF45 and discriminates against the UHM domain of the essential splicing factor U2AF65[29](#_ENREF_29). By targeting the less conserved region in the UHM domains flanking the RXF motif, our cyclic peptide inhibitor not only shows better selectivity but also exhibits increased affinity for the SPF45 UHM domain compared to the linear peptides. This approach could also be used to design cyclic ULM peptide inhibitors with variable selectivity by targeting less conserved residues on the UHM binding surface. Such selective UHM inhibitors can thus be used to delineate the functions of individual UHM domains without interfering with the roles of other domains in multi-domain UHM proteins.

Our *in vitro* splicing data confirm that the novel inhibitors are able to modulate UHM-ULM interactions and stall the spliceo­some assembly during or prior to complex A formation. Surprisingly, the affinity of peptide **10** as well as the linear peptide is 500-1000 fold lower in cell extracts compared to *in vitro* binding studies. This is likely due to non-specific interaction of the peptide with other cellular proteins. Further experiments will aim to optimize the activity of these peptides *in vitro* and *in vivo*. Nevertheless we show that peptide **10** inhibits splicing at an early stage prior to the formation of the spliceosomal A complex. This presumably involves an inhibition of the U2AF65 UHM domain, which plays a critical role at this stage. Thus, although peptide **10** has higher affinity and selectivity for the SPF45 UHM domain *in vitro*, it is nevertheless able to inhibit early spliceosome assembly. On the other hand, the functional data may also indicate a potential role for SPF45 in early spliceosome assembly. This is an interesting possible function of SPF45 that is suggested by our data and will be the starting point for future studies.

Our results represent a first proof of principle showing that by targeting UHM-ULM interactions, spliceosome assembly can be stalled at initial stages, where most of the splicing regulation takes place. A number of splicing inhibitors have been reported in recent years[8](#_ENREF_8), [30](#_ENREF_30). However, the molecular targets and mode of action of these inhibitors are poorly understood. Notably, most of these inhibitors are natural products, which renders their chemical modification and optimization difficult. Also, most of these inhibitors interfere with the splicing reaction at later stages. In this respect, our UHM inhibitor is distinct in two aspects: 1) the structural and rational design approach is based on detailed knowledge of the target, and the mode of action involving UHM inhibition, which is confirmed by our splicing assays, 2) the inhibitor opens novel ways to modulate splicing and interfere with spliceosome assembly at early stages, where alternative splicing is typically regulated by alternative splicing factors. Importantly, the fact that our UHM inhibitor is able to stall spliceosome assembly at an early stage opens the possibility for biochemical and structural studies of very early splicing complexes.

**Experimental Section**

**Protein expression and purification**

SPF45 was expressed and purified as mentioned previously[12](#_ENREF_12). Briefly, protein was expressed in BL21 (DE3) cells as a His-Ztag fusion protein with overnight induction at 20 ˚C in either LB or minimal media supplemented with 0.5 g/L 15NH4Cl. Protein expression was induced with 0.5 mM IPTG. Protein was purified by affinity chromatography using Ni-NTA resin, followed by overnight TEV cleavage of the His-Ztag, ion exchange and size exclusion chromatography. Purified protein was then concentrated to 10 mg/mL and stored in aliquots at -80 ˚C until further use.

**Crystallization of the SPF45 UHM-peptide 3 complex**

Crystals were obtained by mixing 2 L (10 mg/mL in 20 mM Tris pH 7, 150 mM NaCl and 1 mM DTT) protein (SPF45 UHM domain and peptide **3** mixed in 1:1.5 molar ratio) and 2 L reservoir solution (50 mM MES pH 6, 70% MPD) by hanging drop method. Thin plate crystals were obtained in 5-7 days which were flash frozen in liquid nitrogen for data collection. Datasets were collected at the PXIII beam line at Swiss light source and were integrated and scaled with the XDS package[31](#_ENREF_31). The structure was solved via molecular replacement using the native structure of SPF45 as a search model (PDB id: 2PE8) using Phaser[32](#_ENREF_32" \o "McCoy, 2007 #47) and refined in Refmac[33](#_ENREF_33" \o "Murshudov, 1997 #48) from the CCP4 suite[34](#_ENREF_34).

**Isothermal titration calorimetry experiments**

Isothermal titration calorimetry experiments were performed with an ITC200 Microcal system. SPF45 UHM domain protein concentration ranging from 5-20 M was used in the cell and was titrated with 10-fold concentrated peptides in the syringe with 1.5 L injection volume. The data was fitted to a one site-binding model with the Origin software provided with the instrument.

**NMR titrations**

1H,15N HSQC NMR spectra were acquired at 298K using an AVIII600 Bruker NMR spectrometer equipped with a cryogenic probe. 50 M of 15N-labelled SPF45 UHM domain in 50 mM Kpi, 150 mM NaCl , 1 mM DTT , pH 6.8 and 10% D2O was titrated with 2x cyclic peptide. All spectra were processed using NMRPipe/Draw[35](#_ENREF_35) and analyzed using CCPN analysis[36](#_ENREF_36) software.

***In-vitro* splicing assays**

Splicing reactions contained 45% HeLa nuclear extract prepared essentially as described previously[37](#_ENREF_37), 60 mM KCl, 3 mM MgCl2, 2 mM ATP, 20 mM creatine phosphate, and 10 nM uniformly 32P-labeled, m7G-capped MINX[38](#_ENREF_38) or IgM[26](#_ENREF_26) pre-mRNA, and were incubated at 30°C for the indicated times. Reactions were supplemented with RKSRWDETP, RKSRADETP or peptide **10** to a final concentration of 25-500 M and incubated for 20 min on ice prior to the addition of the pre-mRNA and subsequent incubation at 30°C. RNA was recovered at the indicated time points and separated on a 14 % denaturing polyacrylamide gel. Unspliced pre-mRNA, and splicing intermediates and products were detected using a Typhoon PhosphoImager (GE Healthcare). Spliceosomal complex formation was analyzed by agarose gel electrophoresis in the presence of 0.65 μg/μL heparin[39](#_ENREF_39) and bands were visualized by autoradiography or with a Typhoon PhosphoImager. The pre-mRNA, and splicing intermediates and products (spliced mRNA and excised lariat intron) were quantified using ImageQuantTL (GE Healthcare). The percent spliced product was calculated by dividing the amount of mRNA and spliced lariat intron by the amount of pre-mRNA/lariat-intermediate/splicing products and multiplying by 100. Percent spliced product values were then normalized to the control (i.e. no added peptide) spliced product value, which was set to 100%.

**Synthesis of linear peptides**

The peptides RWDET, SRWDET, RKARWDETP and KSRWDE were synthesized on solid support (trityl chloride resin, Intavis, Germany) using standard Fmoc-procedure. Couplings were performed using 2 eq. amino acid, 2 eq. of HATU, 2 eq. HOAt and 5 eq. DIPEA. The peptide was cleaved from the resin using a mixture containing 95 % TFA, 2.5 % H2O and 2.5 % TIPS for 1 h. After evaporation of the solvent, the peptide was taken up in a mixture of ACN/H2O and purified using semipreparative HPLC. The peptide was obtained in good purity (>95 %) as white powder.

**Synthesis of peptide 1**

The orthogonally protected, linear peptide Boc-Lys(Fmoc)-Ser(*t*Bu)-Arg(Pbf)-Trp(Boc)-Asp(*Ot*Bu)-Glu(*O*tBu)-OH was synthesized on solid support (trityl chloride resin, Intavis, Germany) using standard Fmoc-procedure. Couplings were performed using 2 eq. of amino acid, 2 eq. of HATU, 2 eq. HOAt and 5 eq. DIPEA. After the final Fmoc-deprotection on solid support, the peptide was cleaved from the resin using a solution of 20% hexafluoroisopropanol in DCM. After evaporation of the solvent, peptide cyclization was achieved by dissolution of the peptide in DMF (app. 10-2 M), followed by subsequent addition of solid NaHCO3  (5 eq.) and DPPA (3 eq) [40](#_ENREF_40). Compared to procedures using standard peptide coupling reagents also for cyclization, this technique allows a reaction under mild conditions corresponding to the dilution principle in the cyclization as the low solubility of the base allows cyclization only at the interface between solvent and basic solid. This avoids cyclodimerization reactions and leads to relative low racemization rate despite long reaction times of approximately 12 h. After evaporation of the solvent, the peptide was subsequently treated with a solution of 85% TFA, 10% DCM, 2.5% TIPS and 2.5% H2O to deprotect the acid labile side chain protecting groups. Afterwards, the peptide was purified by semi-preparative HPLC and lyophilized, yielding a colourless powder. The peptide was obtained in good purity (>95 %).

**Synthesis of peptide 2 and peptide 3**

The orthogonally protected, linear peptides Boc-Lys(Fmoc)-Ser(*t*Bu)-Arg(Pbf)-Trp(Boc)-Asp(*Ot*Bu)-Glu(*O*Allyl) and Boc-Orn(Fmoc)-Ser(*t*Bu)-Arg(Pbf)-Trp(Boc)-Asp(*Ot*Bu)-Glu(*O*Allyl) were synthesized on solid support, respectively, (trityl chloride resin, Intavis, Germany) using standard Fmoc-procedure. Couplings were performed using 2 eq. of amino acid, 2 eq. of HATU, 2 eq. HOAt and 5 eq. DIPEA. After final Fmoc and Allyl-deprotection on solid support, the side chains of the peptide’s lysine (or ornithine, respectively) and glutamic acid were cyclized on the resin by incubating the orthogonally deprotected peptide for 1 h with a solution of 2 eq. HATU, 2 eq. HOAt and 5 eq. DIPEA in DMF. After completion of the reaction (monitoring by LC-MS), the peptide was cleaved from the resin using 20% HFIP-solution in DCM. After evaporation of the solvent, the peptide was treated with a solution of 85% TFA, 10% DCM, 2.5% TIPS and 2.5% H2O to deprotect the acid labile side chain protecting groups. Afterwards, the peptide was purified by semi-preparative HPLC and lyophilized, yielding a colourless powder. The peptide was obtained in good purity (>95 %).

**Synthesis of the precursor [*sc,sc*(Boc-KS(tBu)R(Pbf)W(Boc)D(OtBu)E)]-OH**

The orthogonally protected, linear peptide Boc-Lys(Fmoc)-Ser(*t*Bu)-Arg(Pbf)-Trp(Boc)-Asp(*Ot*Bu)-Glu(*O*Allyl) was synthesized on solid support (trityl chloride resin, Intavis, Germany) using standard Fmoc-procedure. Couplings were performed using 2 eq. of amino acid, 2 eq. of HATU, 2 eq. HOAt and 5 eq. DIPEA. After final Fmoc and Allyl-deprotection on solid support, the side chains of the peptide’s lysine and glutamic acid were cyclized on the resin by incubating the orthogonally deprotected peptide for 1 h with a solution of 2 eq. HATU, 2 eq. HOAt and 5 eq. DIPEA in DMF. After completion of the reaction (monitoring by LC-MS), the peptide was cleaved from the resin using 20% HFIP-solution in DCM. After evaporation of the solvent, the product was dissolved in H2O/ACN and lyophilized overnight. The peptide [*sc,sc*(Boc-KS(tBu)R(Pbf)W(Boc)D(OtBu)E)]-OH was obtained in good purity (>90 %) as white powder.

**Synthesis of peptide 4-10**

The corresponding amino acid (e.g. His for peptide **4**) was loaded onto trityl chloride resin (Intavis, Germany) and Fmoc-deprotected. Subsequently, the resin was incubated with a solution of the compound [*sc,sc*(Boc-KS(tBu)R(Pbf)W(Boc)D(OtBu)E)]-OH (1.5 eq.), HATU (1.5 eq.), HOAt (1.5 eq.) and 3 eq. DIPEA in DMF until all resin bound starting material was consumed (at least 3h). The reaction was monitored using LC-MS. After completion of the reaction, the coupled peptide was cleaved from the resin (20% HFIP/DCM) and subsequently treated with a solution of 85% TFA, 10% DCM, 2.5% TIPS and 2.5% H2O to deprotect the acid labile side chain protecting groups. Afterwards, the peptide was purified by semi-preparative HPLC and lyophilized, yielding a colourless powder. The peptides were obtained in good purity (>95 %).

ASSOCIATED CONTENT

**Supporting Information**.

This material is available free of charge via the Internet at http://pubs.acs.org. Tables for dissociation constants, X-ray crystallography statistics, analytical data for synthesized peptides and supplementary figures. Coordinates and structure factors for the SPF45 UHM domain in complex with cyclic peptide **3** have been deposited in the PDB with accession number 5LSO.

AUTHOR INFORMATION

Corresponding Author

\*Phone: +49-89-289-13418; E-mail: sattler@helmholtz-muenchen.de

Author Contributions

D.G, O.D, M.S and H.K designed the early cyclic peptides, O.D synthesized the peptides and D.G performed the ITC experiments. P.K.A.J, T.G.K, M.S and H.K designed the cyclic peptides in the later part of the project, T.G.K synthesized the peptides and P.K.A.J performed the ITC experiments. P.K.A.J crystalized and determined the structure of UHM-peptide **3** complex and performed NMR experiments. C.L.W performed the *in vitro* splicing assays. P.K.A.J and M.S wrote major parts of the manuscript. All authors have given approval to the final version of the manuscript.

#These authors contributed equally.

**ACKNOWLEDGMENT**

We thank Gabi Heyne for help with in vitro splicing assays. P.K.A.J acknowledges Boehringer Ingelheim Fonds for a doctoral fellowship. This work was supported by the DFG (grants SFB1035, GRK1721 to M.S and SFB806, project A1 to R.L).

ABBREVIATIONS

UHM, U2AF homology motif; ULM, U2AF ligand motif; *K*D, equilibrium dissociation constant; H, enthalpy; S, entropy.

**REFERENCES**

1. Nilsen, T. W.; Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **2010,** *463*, 457-463.

2. Wahl, M. C.; Will, C. L.; Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **2009,** *136*, 701-718.

3. Xiong, H. Y.; Alipanahi, B.; Lee, L. J.; Bretschneider, H.; Merico, D.; Yuen, R. K.; Hua, Y.; Gueroussov, S.; Najafabadi, H. S.; Hughes, T. R.; Morris, Q.; Barash, Y.; Krainer, A. R.; Jojic, N.; Scherer, S. W.; Blencowe, B. J.; Frey, B. J. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* **2015,** *347*, 1254806.

4. Ray, D.; Kazan, H.; Cook, K. B.; Weirauch, M. T.; Najafabadi, H. S.; Li, X.; Gueroussov, S.; Albu, M.; Zheng, H.; Yang, A.; Na, H.; Irimia, M.; Matzat, L. H.; Dale, R. K.; Smith, S. A.; Yarosh, C. A.; Kelly, S. M.; Nabet, B.; Mecenas, D.; Li, W.; Laishram, R. S.; Qiao, M.; Lipshitz, H. D.; Piano, F.; Corbett, A. H.; Carstens, R. P.; Frey, B. J.; Anderson, R. A.; Lynch, K. W.; Penalva, L. O.; Lei, E. P.; Fraser, A. G.; Blencowe, B. J.; Morris, Q. D.; Hughes, T. R. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **2013,** *499*, 172-177.

5. Hennig, J.; Gebauer, F.; Sattler, M. Breaking the protein-RNA recognition code. *Cell Cycle* **2014,** *13*, 3619-3620.

6. Cooper, T. A.; Wan, L.; Dreyfuss, G. RNA and disease. *Cell* **2009,** *136*, 777-793.

7. David, C. J.; Manley, J. L. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.* **2010,** *24*, 2343-2364.

8. Bonnal, S.; Vigevani, L.; Valcarcel, J. The spliceosome as a target of novel antitumour drugs. *Nat. Rev. Drug Discov.* **2012,** *11*, 847-859.

9. Daguenet, E.; Dujardin, G.; Valcarcel, J. The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. *EMBO Rep.* **2015,** *16*, 1640-1655.

10. Neubauer, G.; King, A.; Rappsilber, J.; Calvio, C.; Watson, M.; Ajuh, P.; Sleeman, J.; Lamond, A.; Mann, M. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat. Genet.* **1998,** *20*, 46-50.

11. Lallena, M. J.; Chalmers, K. J.; Llamazares, S.; Lamond, A. I.; Valcarcel, J. Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. *Cell* **2002,** *109*, 285-296.

12. Corsini, L.; Bonnal, S.; Basquin, J.; Hothorn, M.; Scheffzek, K.; Valcarcel, J.; Sattler, M. U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. *Nat. Struct. Mol. Biol.* **2007,** *14*, 620-629.

13. Izquierdo, J. M.; Majos, N.; Bonnal, S.; Martinez, C.; Castelo, R.; Guigo, R.; Bilbao, D.; Valcarcel, J. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol. Cell* **2005,** *19*, 475-484.

14. Sampath, J.; Long, P. R.; Shepard, R. L.; Xia, X.; Devanarayan, V.; Sandusky, G. E.; Perry, W. L., 3rd; Dantzig, A. H.; Williamson, M.; Rolfe, M.; Moore, R. E. Human SPF45, a splicing factor, has limited expression in normal tissues, is overexpressed in many tumors, and can confer a multidrug-resistant phenotype to cells. *Am. J. Pathol.* **2003,** *163*, 1781-1790.

15. Perry, W. L., 3rd; Shepard, R. L.; Sampath, J.; Yaden, B.; Chin, W. W.; Iversen, P. W.; Jin, S.; Lesoon, A.; O'Brien, K. A.; Peek, V. L.; Rolfe, M.; Shyjan, A.; Tighe, M.; Williamson, M.; Krishnan, V.; Moore, R. E.; Dantzig, A. H. Human splicing factor SPF45 (RBM17) confers broad multidrug resistance to anticancer drugs when overexpressed--a phenotype partially reversed by selective estrogen receptor modulators. *Cancer Res.* **2005,** *65*, 6593-6600.

16. Kielkopf, C. L.; Rodionova, N. A.; Green, M. R.; Burley, S. K. A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell* **2001,** *106*, 595-605.

17. Selenko, P.; Gregorovic, G.; Sprangers, R.; Stier, G.; Rhani, Z.; Kramer, A.; Sattler, M. Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. *Mol. Cell* **2003,** *11*, 965-976.

18. Kielkopf, C. L.; Lucke, S.; Green, M. R. U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.* **2004,** *18*, 1513-1526.

19. Manceau, V.; Kielkopf, C. L.; Sobel, A.; Maucuer, A. Different requirements of the kinase and UHM domains of KIS for its nuclear localization and binding to splicing factors. *J. Mol. Biol.* **2008,** *381*, 748-762.

20. Maucuer, A.; Ozon, S.; Manceau, V.; Gavet, O.; Lawler, S.; Curmi, P.; Sobel, A. KIS is a protein kinase with an RNA recognition motif. *J. Biol. Chem.* **1997,** *272*, 23151-23156.

21. Corsini, L.; Hothorn, M.; Stier, G.; Rybin, V.; Scheffzek, K.; Gibson, T. J.; Sattler, M. Dimerization and protein binding specificity of the U2AF homology motif of the splicing factor Puf60. *J. Biol. Chem.* **2009,** *284*, 630-639.

22. Loerch, S.; Maucuer, A.; Manceau, V.; Green, M. R.; Kielkopf, C. L. Cancer-relevant splicing factor CAPERalpha engages the essential splicing factor SF3b155 in a specific ternary complex. *J. Biol. Chem.* **2014,** *289*, 17325-17337.

23. Spadaccini, R.; Reidt, U.; Dybkov, O.; Will, C.; Frank, R.; Stier, G.; Corsini, L.; Wahl, M. C.; Luhrmann, R.; Sattler, M. Biochemical and NMR analyses of an SF3b155-p14-U2AF-RNA interaction network involved in branch point definition during pre-mRNA splicing. *RNA* **2006,** *12*, 410-425.

24. Thickman, K. R.; Swenson, M. C.; Kabogo, J. M.; Gryczynski, Z.; Kielkopf, C. L. Multiple U2AF65 binding sites within SF3b155: thermodynamic and spectroscopic characterization of protein-protein interactions among pre-mRNA splicing factors. *J. Mol. Biol.* **2006,** *356*, 664-683.

25. Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **2005,** *437*, 640-647.

26. Guth, S.; Martinez, C.; Gaur, R. K.; Valcarcel, J. Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). *Mol. Cell. Biol.* **1999,** *19*, 8263-8271.

27. Gozani, O.; Potashkin, J.; Reed, R. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell. Biol.* **1998,** *18*, 4752-4760.

28. Das, R.; Zhou, Z.; Reed, R. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol. Cell* **2000,** *5*, 779-787.

29. Banerjee, H.; Rahn, A.; Gawande, B.; Guth, S.; Valcarcel, J.; Singh, R. The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF 65) is essential in vivo but dispensable for activity in vitro. *RNA* **2004,** *10*, 240-253.

30. Zaharieva, E.; Chipman, J. K.; Soller, M. Alternative splicing interference by xenobiotics. *Toxicology* **2012,** *296*, 1-12.

31. Kabsch, W. Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66*, 125-132.

32. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007,** *40*, 658-674.

33. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997,** *53*, 240-255.

34. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011,** *67*, 235-242.

35. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995,** *6*, 277-293.

36. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **2005,** *59*, 687-696.

37. Dignam, J. D.; Lebovitz, R. M.; Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **1983,** *11*, 1475-1489.

38. Zillmann, M.; Zapp, M. L.; Berget, S. M. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **1988,** *8*, 814-821.

39. Das, R.; Reed, R. Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels. *RNA* **1999,** *5*, 1504-1508.

40. Heckmann, D.; Kessler, H. Design and chemical synthesis of integrin ligands. *Methods Enzymol.* **2007,** *426*, 463-503.

**Table of Contents graphic**

