Hit Evaluation of an α -Helical Peptide: Ala-scan, Truncation, and Sidechain-to-sidechain Macrocyclization of an RNA Polymerase Inhibitor

Author list: Ahmed Ashraf Moustafa Kamal^a, Monica Habib^{c, d}, Joerg Haupenthal^a, Rolf Wolfgang Hartmann^{a, b}, Martin Empting^{a, b*}

^{a)} Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Department of Drug Design and Optimization, Campus E8.1, 66123 Saarbrücken, Germany

^{b)} Department of Pharmacy, Saarland University, Campus E8.1, 66123 Saarbrücken, Germany.

^{c)} Department of Pharmaceutical Chemistry, Faculty of Pharmacy and Biotechnology, German University in Cairo, Cairo 11835, Egypt.

^{d)} Institute of Molecular Virology, Ulm University Medical Center, Ulm, Germany

To whom correspondence should be addressed: Martin Empting, Helmholtz Institute for Pharmaceutical Research Saarland, Department of Drug Design and Optimization, Campus E8.1, 66123 Saarbrücken, Germany, E-Mail: martin.empting@helmholtz-hzi.de

Tel.: +49 681 98806-2031. Fax: +49 681 98806-2009

Running Title: Engineering an α -Helical Peptide

Abstract

RNA polymerase (RNAP) remains a relatively underexplored target with only rifampicin and fidaxomicin in clinical use. Hence, the concurrent rise in bacterial resistance rate urges the search for novel RNAP inhibitors with novel mode of action. In this work, we investigated the impact of several systematic modifications including sidechain-to-sidechain macrocyclization in α -helical content and biological activity of a previously identified inhibitory sigma factor fragment. Ala-scan results, peptide truncation from both the *N*- and *C*- terminus, and modifications inspired by other RNAP inhibitors revealed novel structure activity relationships but did not yield a superior sequence. Additionally, four insertion points for non-natural amino acids bearing side chains required for macrocyclization were explored. Linear precursors showed improved stabilization of α -helical content compared to the original sequence as demonstrated by CD spectroscopy. However, this increase in α -helicity did not translate into improved biological activity. Instead, complete abolishment of RNAP inhibitory activity occurred. We hypothesize three possible reasons for such discrepancy and offer basis for further optimization efforts for this peptidic RNAP inhibitor.

Keywords

alpha helix; anti-infective; click chemistry; macrocyclic peptides; RNA polymerase

Introduction

Over the last few years, there has been an increased interest in investigating the underexplored chemical space of peptides beyond Lipinski's rule of five as a source of novel drugs (Doak et al. 2014). The presence of more than 140 peptide-based therapeutics in clinical trials in 2014 emphasizes this notion (Fosgerau and Hoffmann 2015). The modular assembly of peptides provides high structural diversity combined with the potential to display large interaction surface areas (Craik et al. 2013; Fosgerau and Hoffmann 2015). This enables to generate peptides with high affinity as well as specificity to their targets. Furthermore, metabolic degradation products of peptides are amino acids providing low risk of toxicity (Craik et al. 2013; Fosgerau and Hoffmann 2015). Additionally, peptides are considered superior in pursuing targets perceived as "undruggable" such as protein-protein interactions (PPIs) compared to conventional small organic molecules (SOMs) (Arkin et al. 2014; Cardote and Ciulli 2016; Pelay-Gimeno et al. 2015). Thus, peptides hit a "sweet spot" showing specificity and potency of larger protein-based biologics yet, being smaller in size. However, peptides still possess some undesirable attributes such as lack of oral bioavailability, susceptibility to proteolytic degradation, and rapid clearance (Otvos and Wade 2014). Additionally, short peptide sequences stripped from native proteins can unfold, resulting in loss of biological activity. One reason for the loss of the ordered geometry can be the inability of the peptide to retain its secondary structure i.e. α -helix without the stabilizing effect of the protein environment (Cardote and Ciulli 2016; Driggers et al. 2008; Goldflam and Ullman 2015; Vlieghe et al. 2010).

Strategies to overcome such physicochemical and pharmacokinetic deficits have been extensively researched through for example truncation, incorporation of unnatural or D-configured amino acids, and macrocyclization (Driggers et al. 2008; Goldflam and Ullman

2015; Vlieghe et al. 2010). In this study, we focus on sidechain-to-sidechain macrocyclization as an approach to enhance conformational stability, biological activity, and proteolytic stability of α -helical peptides.

If installed in a suitable manner, macrocyclization will constrain a peptide structure locking it in the bioactive conformation. Such decrease in the rotational degrees of freedom should lead to the reduction of the entropy “penalty” present upon binding (Marsault and Peterson 2011; Martí-Centelles et al. 2015). Additionally, a non-natural cross-linkage has been shown to protect the backbone amide moieties improving resistance to degradation *via* proteases (Lau et al. 2015; Li et al. 2013).

Over the years, several approaches for sidechain-to-sidechain macrocyclization have been devised such as disulfide bridges and lactone formations, Ruthenium-catalyzed ring-closing olefin metathesis (RCM), molybdenum-catalyzed ring-closing alkyne metathesis (RCAM) and Copper (I) as well as Ruthenium (II)-catalyzed alkyne-azide cycloaddition (CuAAC and RuAAC, respectively) (Avrutina et al. 2009; Cardote and Ciulli 2016; Empting et al. 2011; Pelay-Gimeno et al. 2015; Tischler et al. 2012; White and Yudin 2011). The latter technique has been used in this study as the strategy of choice as elaborated orthogonal protection schemes are not required when employing the prominent copper-click reaction. Additionally, a triazole moiety as a cross-linkage instead of long hydrocarbon chains would yield peptides with better physicochemical properties (Li et al. 2013; Pasini 2013).

The recent surge in antibiotic resistance accompanied with the lack of discovery of new antibiotics is quite troubling. Thus, the search for “outside of the box” antibiotics with novel mechanisms of action is very intriguing. In a previous study from our lab, we discovered through a peptide-scanning approach a novel nonadecapeptide inhibiting *E. coli* RNA polymerase (Hüsecken et al. 2013). The peptide sequence referred to as P07 was derived

from sigma factor 70 (σ^{70} , region 2.2) and showed activity in the micromolar range (Figure 1). An abortive transcription assay and ELISA demonstrated that P07 obstructs transcription initiation and inhibits binding of σ^{70} to the core enzyme. Additionally, molecular dynamics simulations and mutagenesis studies showed that P07 has the ability to address a new hot spot for the protein–protein interaction inhibition namely the lid-rudder-system (LRS). This part of the enzyme is a conformationally flexible region located within the β' subunit of RNAP performing important functions in the transcription process (Kuznedelov et al. 2002; Touloukhonov and Landick 2006). In the present study, we performed Alanine scan and truncation to identify the core sequence for activity and modifiable residues in the peptide sequence. Afterwards, we designed and synthesized macrocyclic derivatives through CuAAC-mediated sidechain-to-sidechain macrocyclization with the aim to improve potency and conformational stability of the parent peptide.

Results and Discussion

Alanine-Scan and Structural Insights

An Ala-scan was undertaken not only to reveal amino acids with high contribution to binding but to additionally give insight into the peptide's binding profile. It is expected that substitution of key amino acids, which are important for target interaction, by Ala would yield peptides with lower biological activity. Whereas, amino acids exposed to the surface when substituted would retain activity unless they have an effect on the whole dynamics or conformation of the peptide. Thus, this would provide us with data needed to identify sequence positions suitable for introduction of unnatural amino acids required for macrocyclization and further optimization.

The Ala-modified peptides did not yield any sequence of improved activity over the original sequence (Table 1). However, several positions – Thr1, Asn2, residues 5-10, residues 12-14, Gly17, and Met19 – have been exchanged without a big loss of activity (<2-folds), which could be used for sequence modification or as insertion points for unnatural amino acids. Peptide 10 has been previously found to have an $IC_{50} > 50 \mu M$ (Hüsecken et al. 2013). Key activity loss was observed (>5 folds) upon swapping of Arg3, Ile11, Ile16, and Leu18 (highlighted in red, Table 1). In case of Ile11, Ile16, and Leu18, this hints towards the importance of the hydrophobic sidechain for protein binding. While for Arg3 the loss in activity can be attributed to either loss of the hydrophilic and/or basic properties of the side chain. It cannot be excluded that the sidechain also plays a role in stabilizing the whole peptide conformation via intramolecular hydrogen bonding to neighboring residues. Lastly, four peptides showed a moderate loss (2-5 folds). The full biological results from the Ala-scan are compiled in Table 1. When mapping these results onto the X-ray structure of $\sigma 70$ in complex with RNAP (Figure 2), it becomes obvious that residue Ile16 presumably forms one critical hydrophobic interaction with the target protein. According to this model however, Ile11 and Leu18, which are also critical for activity, appear not to be in direct contact with the protein. It is likely that the system undergoes conformational adjustments to allow for hydrophobic contacts between these important sidechains and RNAP. Arg3 has been proposed earlier by molecular dynamics simulations to be important for disrupting the lid-rudder-system (Hüsecken et al. 2013), which is not resembled in the shown X-ray coordinates.

Aggregation Test

All peptides synthesized so far possessed micromolar activities. Peptides were tested for aggregation to exclude this unspecific phenomenon as the reason for either activity gain or loss. We chose five peptides for this experiment. The testing set included compounds **1**, **2**, **8** and **14**, which showed inhibitory activity in the single-digit micromolar range, while **17** was picked as an example for an inactive sequence. The measurements were taken at two-time points, after 10 min of incubation (similar to the biological test) and 24 h. In both cases, no aggregation was observed *via* light scattering measurements at OD₆₀₀ until the maximum soluble concentration of 100 μM. Values are summarized in Table S1 (Supporting Information).

Shortening of the Peptide Sequence and Modifications Inspired by Other RNAP inhibitors

One by one truncation of the peptide from both the *N*- and *C*- terminus was attempted in order to achieve the minimum active sequence with better physiochemical properties. Deletion of Thr1 at the *N*-terminus led to peptide **21** with similar activity to the parent peptide. *C*-terminal truncation of Met19 led to 3-folds loss of activity as seen in peptide **22** (Table 2). This could be attributed to either the importance of the hydrophobic interaction mediated by the methionine side chain or the terminal –COOH group of this residue. Replacement of Met19 with the unnatural amino acid norleucine (**23**) possessing a C4-alkyl chain or with alanine (**20**) led to peptides with IC₅₀s of 7.1 μM and 7.7 μM, respectively. In other words, increasing or decreasing the hydrophobicity of the sidechain did not cause significant changes in affinity. Considering that truncation of this residue did strongly impair binding, one can conclude that then positioning of the terminal –COOH group is essential, although is difficult to explain from the binding mode shown in Figure 2. As different decorations seemed to be accepted at this position, we explored one further modification.

Literature reports showed that a class of small organic molecules possesses an indole moiety that binds to the same region of RNAP as the corresponding methionine residue does in σ^{70} (Ma et al. 2013). Hence, we were inspired to synthesize peptide **24** with a tryptophan instead of Met19 combined with deletion of Thr1. **24** displayed an IC_{50} of 4.7 μ M similar to peptide **1**.

As no true activity gains were achieved upon deletion of the hydrophilic amino acid Thr1, or replacement of Met19, the original sequence was used as a template for further modifications and future syntheses.

Macrocyclic Peptide Syntheses

Solid-phase peptide syntheses (SPPS) of the desired peptide sequences were carried out incorporating the azide and alkyne bearing amino acids, which were used afterward for on-resin CuAAC (Scheme 1).

In order to provide proof for the formation of the desired triazole moiety, the consumption of the azide functionality was monitored using IR spectroscopy as described before (Empting et al. 2011). The azide group gives a distinctive band around 2100 cm^{-1} , which is well separated from other IR signals of peptides. Upon triazole formation, loss of this signal will provide an easy method for monitoring progression of the CuAAC reaction. However, there is the possibility of dimer formation (intermolecular reaction) instead of the desired monomer (intramolecular reaction). A successful CuAAC-mediated macrocyclization proceeds with no change in molecular weight compared to the linear precursor, while a possible dimer formation would lead to a doubled mass. Hence, cyclic monomers and triazole-linked dimers can be distinguished through examining the isotope pattern of the MS signals using high-resolution mass spectrometry. In the case of a dimeric product, the

distances between individual isotopic species would be half of the monomeric variants with the same m/z ratio (refer to Figure S1 in the Supporting Information for details). This methodology was used to prove successful installation of the macrocycle in peptides **29-32** (Table 3).

Macrocyclization Attempts and Their Effects on α -Helicity and Biological Activities

According to our hypothesis, stapling of α -helical peptides was expected to increase α -helical content, thus, enhancing peptide affinity and potency. The secondary structure (e.g. α -helix) of a short linear peptide sequence is usually not stable and multiple populations of the folded and unfolded peptide coexist in the solution. Hence, the peptide possesses a considerable degree of conformational freedom, which is lost upon binding to the target in a specific conformation. Hence, the system pays a tremendous entropy penalty, which can be reduced via rigidification through e.g. macrocyclization. This is a common strategy in medicinal chemistry and has been demonstrated for peptide optimization as well (Marsault and Peterson 2011; Martí-Centelles et al. 2015). However, in order to insert the macrocycle at the appropriate side of the α -helical peptide structural data is required. Attempts at resolving an X-ray structure of peptide **1** bound to the holoenzyme were not fruitful until now. So, our design was based on the results from the Ala scan and the X-ray structure of the σ factor in complex with RNAP (PDB ID: 4YG2) (Figure 2) (Murakami 2013). Accordingly, four positions were chosen. Two of these were based on the co-crystal structure: (i) to (i+4) side-chain-to-side-chain macrocyclizations were proposed between Ile11-Asn15 and Phe7-Ile11. The other two proposals were based on the Ala-scan results where amino acids showing no huge loss were assumed to be solvent exposed and not involved in binding. Accordingly, Gly14-Gly17 and Leu10-Gly14 were chosen as points for macrocycle insertion.

The staple between Gly14-Gly17 represents an (i) to (i+3) instead of the previously used (i) to (i+4) pattern. Linear precursor peptides harboring non-natural amino acids suitable for CuAAC macrocyclization at the suggested insertion points were synthesized and tested for biological activity. The four linear precursor peptides **25**, **26**, **27** and **28** showed inhibitory activity with IC₅₀s of 9.0, 43, 5.1, and 4.2 μ M, respectively (Table 3). Peptide **26** showed the biggest loss in activity of 10-folds, while the other variants basically retained potency. This fits with the Ala-scan results. When replacing Ile11 with a longer carbon chain bearing an azide moiety, resulting peptide **25** retained the activity. A greater activity loss was observed when replacing Ile11 with a shorter less flexible alkyne moiety leading to less active derivative **26**. Peptides **27** and **28** possessed activities similar to peptide **1** indicating that indeed the replaced amino acids might not be essential for binding. Most of the linear precursor peptides already displayed an increase in α -helicity determined through CD spectroscopy measurement (Table 3 and Supporting Information). Again, the only exception was observed for **26**, which showed lower content than **1**. Notably, peptide **27** showed almost a 2-fold increase in α -helical content (56.7% vs. 24.8%) while retaining activity. Finally, upon macrocyclization, almost complete abolishment of activity was observed regardless of the increase in the α -helical content compared to **1** and position of the macrocycle. Activity loss of **29** and **30** can be explained by the constraintment of the alkyl chains in position 11 that was important for interaction as seen from the Ala-scan results. However, this could not explain the activity loss in case of peptides **31** and **32**. This complete abolishment of activity upon stapling using CuAAC-mediated side-chain-to-side-chain macrocyclization was rather unexpected. However, it could be reasoned by one of the following explanations:

1. The (perfectly ordered) α -helical conformation might not be the active conformation during binding as seen by the presence of an atypical helix turn with extended vertical distances between consecutive turns of the helix bigger than 5.4 Å (PDB ID: 4YG2, see Supporting Information) (Murakami 2013). Thus, a more conformationally flexible peptide might be required for induced-fit binding. This conformational flexibility might be required for such a complex multi-subunit machinery, which undergoes pronounced structural changes for DNA unwinding and RNA synthesis. However, this hypothesis is difficult to validate experimentally and would be best addressed by an X-ray crystallographic analysis of the peptide-protein complex.
2. It could not be excluded that the actual binding mode is different from the one we proposed from the holoenzyme co-crystal structure and Ala-scan data. In this scenario, the active conformation could still be α -helical as expected but the triazolyl linkage might result in a steric clash between the peptide and the enzyme interface.
3. In the literature, a study by Czabotar *et al.* described similar findings with stapled peptides having an unexpected reduced affinity for their targets (Okamoto et al. 2013). In this case, the loss in affinity was attributed to disruption of a network of intramolecular interactions stabilizing the native peptide by the stapling motif leading to perturbation of the peptide structure.

Conclusions

In conclusion, we explored the possibility of entropic optimization of previously described peptidic inhibitor P07 (peptide **1**), which interferes with binding of RNAP core enzyme to σ^{70} . While sidechain-to-sidechain macrocyclization of peptides proved beneficial for α -helical content and proteolytic stability, in the present case, the reduced flexibility does not

translate into improvement of biological activity. These results might change by employing other macrocyclization motifs and lengthier linkers to strike a compromise between flexibility probably required for induced-fit binding and conformational stability. The Ala-scan results provide a basis for future rational optimization with 13 residues amenable for medicinal-chemistry-driven modifications (positions 1, 2, 5-10, 12-14, 17, and 19). The successful insertion of unnatural amino acids in the peptide sequence can be further exploited for the introduction of new interaction abilities not possible by natural amino acids and combinatorial tuning of the peptide e.g. by decorating it with additional moieties *via* copper-catalyzed click chemistry (Fittler et al. 2013). Among others CuAAC as well as olefin-metathesis-mediated sidechain-to-sidechain macrocyclization has been shown to be a valuable tool for peptide optimization (Martí-Centelles et al. 2015) and promising cyclic peptides for targeting PPI of RNAP core enzyme to sigma factors have been reported (El-Mowafi et al. 2015). Thus, targeting protein interaction surfaces is still to be considered an untapped source for novel drugs, yet more efforts have to be put into better understanding of the exact mechanism and interaction modes between RNAP as a multi-domain complex and the peptidic inhibitors described herein.

Materials and Methods

General

Solvents

All solvents were synthesis, HPLC, or peptide synthesis grade. DMF, MTBE, DIEA, TES, TIS, Anisole, TFA, Acetonitrile, DCM, Methanol and Piperidine were obtained from Acros Organics (Heidelberg, Germany), Sigma-Aldrich (Germany), Carl Roth (Karlsruhe, Germany) or VWR (Darmstadt, Germany).

Chemicals and Reagents

All reagents were synthesis or peptide synthesis grade. HATU, HBTU, and TBTU were obtained from IRIS biotech (Marktredwitz, Germany) or Novabiochem (Darmstadt, Germany).

Preloaded resins for SPPS were obtained from Rapp Polymere (Tübingen, Germany).

All Fmoc-protected natural amino acids were obtained from Novabiochem (Darmstadt, Germany) while unnatural Fmoc-protected amino acids were obtained from IRIS biotech (Marktredwitz, Germany).

Instrumentation and Analysis

Microwave synthesis system

Microwave irradiation for SPPS was achieved by using Discover[®] microwave synthesis system from CEM (Kamp-Lintfort, Germany) and an additional external unit was used for solvent removal by suction. The synthesizer was operated using 25 mL polypropylene vessels equipped with polypropylene filters (Roland Vetter, Laborbedarf OHG).

HPLC/MS

Crude and isolated peptides were analyzed using a Thermo Fisher Accela HPLC equipped with a degasser, a quaternary pump, and an autosampler, coupled to an Accela photodiode array detector (PDA) followed by a TSQ Quantum Access Max mass spectrometer with heated electrospray ionization source (HESI-II)(Thermo Fisher). For gradient elution, an Accucore RP-MS column (150 × 2.1 mm, 2.6 μm) was used with a mobile phase consisting of HPLC grade water containing 1% TFA (v/v) and acetonitrile containing 1% TFA (v/v) and a

flow rate of 800 $\mu\text{L}/\text{min}$ with a gradient of 10% Acetonitrile to 80% in 20 minutes. The injection volumes were 10 and 25 μL . The PDA was operated at 220 and 280 nm wavelength. The MS conditions used were electrospray ionization (ESI) in positive mode. Data acquisition and analysis were achieved using Xcalibur 2.2 software.

LC-HRMS

LC-HRMS data was performed on one of the following machines:

A) A Dionex Ultimate 3000 RSLC system using a Waters BEHC18, 100 x 2.1 mm, 1.7 μm dp column. All samples were constituted in acetonitrile and water (1:1). Separation of samples was achieved by a linear gradient with (A) H_2O + 0.1 % formic acid A to (B) ACN + 0.1 % formic acid at a flow rate of 600 $\mu\text{L}/\text{min}$ and 45 $^{\circ}\text{C}$. The gradient was initiated after a 0.5 min isocratic step at 5 % B, followed by an increase to 95 % B in 18 min to end up with a 2 min step at 95 % B before re-equilibration with initial conditions. UV spectra were recorded by a DAD in the range from 200 to 600 nm. The LC flow was split to 75 $\mu\text{L}/\text{min}$ before entering the maXishr-ToF mass spectrometer Bruker Daltonics (Bremen, Germany) using the standard ESI source.

B) A Dionex Ultimate 3000 RSLC system using a ThermoFisher Scientific Accucore phenyl-hexyl, 100x2.1 mm, 2.6 μm dp column. All samples were constituted in acetonitrile and water (1:1). Separation of samples was achieved by a linear gradient with (A) H_2O + 0.1 % formic acid to (B) Acetonitrile + 0.1 % formic acid at a flow rate of 500 $\mu\text{L}/\text{min}$ and 40 $^{\circ}\text{C}$. The gradient was initiated after 0.5 min isocratic step at 5 % B, followed by an increase to 100 % B in 23 min to end up with a 2 min step at 100 % B before re-equilibration with initial conditions. All data acquisition was done in positive ion mode and full-scan mode ranging from 500-1500 m/z at a resolution of 70,000 using a Thermo Fisher (Dreieich, Germany) Q

Exactive Focus system equipped with heated electrospray ionization (HESI)-II source and Xcalibur (Version 4.0.27.19) software. Mass calibration was done prior to analysis according to the manufacturer's recommendations using external mass calibration.

Semi-preparative HPLC

Purification of crude peptides was performed using one of the following preparative HPLCs:

A) An Agilent 1200 HPLC system with an automated fraction collector (Agilent Technologies, Inc. Headquarters, Santa Clara, CA, USA) in "time-based" mode. ChemStation® software was used for control, report and data analysis. The sample was manually injected. A Nucleocel 100-7 C18 HD column (250 / 16 mm) was used as stationary phase (Macherey-Nagel GmbH, Dueren, Germany). The solvent system consisted of HPLC grade water containing 1‰ TFA (v/v) (A) and acetonitrile containing 1‰ TFA (v/v) (B). HPLC-Method: Flow rate 5 mL/min with a gradient run from 10% B to 60% in 40 minutes.

B) A Waters HPLC system equipped with a Phenomenex C18 Luna column (5 mm pore size, 100 Å particle size, 250, 21.2 mm) in an acetonitrile/water gradient under acidic conditions. The peptide mass was verified by liquid chromatography mass spectroscopy (LCMS; Waters).

Lyophilization

For the removal of RP-HPLC eluents or aqueous solvents from peptidic samples one of the following instrumentations was used:

A) An Alpha-2- 4-LD plus freeze dryer with condensation temperature of -85 °C (Martin Christ Gefriertrocknungsanlagen GmbH, Osterode, Germany) equipped with a high vacuum pump Ilmvac, type 109012 (Ilmvac GmbH, Ilmenau, Germany).

B) A Labconco freeze dryer.

Centrifugation

Sedimentation and washing of precipitated peptides was performed using a Thermo Fisher Scientific Heraeus Multifuge 15-R centrifuge (Schwerte, Germany).

FT-IR Spectroscopy

Fourier-transform infrared (FT-IR) spectra were recorded on a PerkinElmer (Rodgau, Germany) Spectrum 100 FT-IR spectrometer.

Circular Dichroism Spectroscopy

All single Far-UV-CD spectra were obtained at 24 °C using a Jasco-1500 spectropolarimeter (Gross-Umstadt, Germany). All spectra were recorded using a 1-mm pathlength rectangular quartz cell, 1-nm bandwidth, 100 nm/min scan speed, 5 accumulations, within a wavelength range of 185 to 260 nm. Peptides were dissolved to a final concentration of 50 µM in deionized water (milliQ, Millipore) containing 30% 2,2,2-trifluoroethanol (TFE). After blank subtraction and smoothing, the percentage of α -helicity of the peptides was computed *via* the implemented CONTIN algorithm.

Biological Testing

The RNAP inhibition assay was performed as previously described. In brief, inhibitory effects of our peptides on RNAP transcription and respective IC₅₀ values were measured through a transcription inhibition assay. *E. coli* RNAP holoenzyme was incubated with 3H-UTP, T7A1 promoter PCR fragment as a DNA template to which tested compounds were added afterward. Then transcription reaction is stopped and the residual unincorporated 3H-UTP is

quantified by liquid scintillation counting. All peptides were dissolved in a stock solution of 1:1 acetonitrile and MilliQ water before testing. Three different concentrations of a compound/peptide were chosen for the determination of an IC_{50} value (two samples for each concentration). The calculation of the IC_{50} value was performed by plotting the percent inhibition versus the concentration of inhibitor. Every compound was measured in at least two independent experiments (replicates) with number of measurements at least being four.

Data Analysis

Data analysis was performed using GraphPad Prism 6. The different effect of the compounds was plotted as dose–response curves such that each individual point represents the average of a single experiment. Lines of best fit and associated IC_{50} values as well as respective 95% confidence intervals (CI) were calculated for dose–response curves using the nonlinear fit variable slope (four parameters) function with a bottom constraint = 0 and top constraint = 1. Determined 95% CIs indicate that values which lay outside this interval can be rejected as plausible values for the given compound at $p=0.05$ level. In cases when no precise IC_{50} values and no confidence intervals are provided, determination was not possible due to low solubility and/or low potency.

Aggregation Test

A master mix of *E. coli* RNAP holoenzyme is incubated with T7A1 promoter PCR fragment then mixed with serial dilutions of test compounds into a transparent F-bottom 384-well microtiter plate (Greiner Bio-One, Germany, cat. no. 781901) in duplicates. Absorbance was measured on a PheraStar FS (BMG Labtech, Germany) plate reader at 600 nm after 10 min at

37 °C and 24 h at room temperature of incubation. The data was analyzed using BMG Labtech MARS software.

Synthesis

Peptides **1-24** were supplied by our *in-house* peptide platform (HZI Braunschweig). Peptides **25, 26, 29** and **30** were synthesized according to synthesis protocol 1. Peptides **27, 28, 31** and **32** were synthesized according to synthesis protocol 2. Macrocyclization was performed using the protocol described below.

General Protocol for Microwave-assisted Fmoc- SPPS

Synthesis Protocol 1

The synthesis was performed on a TentaGel® S AC-Met-Fmoc resin (scale 0.24/4 mmol) using coupling protocol 1, except in the case of unnatural amino acids, coupling protocol 2 was applied to assure more efficient couplings.

Resin Swelling. Prior to synthesis, the dry resin was allowed to swell in DCM for 30 min followed by washing (3x with DCM and 3x with DMF).

Deprotection. Removal of the *N*-terminal protecting group was performed by the addition of a solution containing 20% (v/v) piperidine in DMF followed by MW-irradiation (40 °C, 30 W) for 5 min. After washing and re-addition of 20% piperidine in DMF irradiation was applied for 15 min. Afterward, the reaction mixture was removed by filtration and the resin was subsequently washed with DMF (7x).

Coupling protocol 1. First, 4 eq of the Fmoc protected amino acid (scaled to the resin loading) and 3.9 eq of the coupling reagent (HBTU) were dissolved in a minimal amount of DMF. Then, 8 eq DIEA were added for activation. After 15 min of pre-activation, the resulted

mixture was added to the *N*-terminal de-protected peptide-resin and irradiated with MW (20 min, 40 °C, 30 W). Excess reagents were removed by filtration and the peptide resin was washed thoroughly. These steps were repeated 3 times.

Coupling protocol 2. First, 2 eq of the unnatural Fmoc protected amino acid (scaled to the resin loading) and 1.9 eq of a coupling reagent (HATU) in a minimal amount of DMF is dissolved. Then, 4 eq DIEA were added for activation. After 15 min of pre-activation, the resulted mixture was added to the *N*-terminal de-protected peptide-resin and irradiated with MW (20 min, 40 °C, 30 W). Excess reagents were removed by filtration and the peptide resin was washed thoroughly. These steps were repeated 3 times.

Cleavage and workup of the assembled peptide. After completion of synthesis, the resins were washed with DMF (7x), DCM (3x), and diethyl ether (3x) then stored in a desiccator and dried under vacuum. To cleave the assembled peptide off the resin and achieve full deprotection of amino acid side chain protecting groups, the peptide resin was treated with a mixture of TFA/H₂O/TIS 95:3:2 v/v for 3 h. The cleaved peptide was precipitated in cold MTBE, centrifuged and washed with MTBE 3 times. At the end, the crude peptide was stored in a desiccator and dried under vacuum. Crude peptides were purified via semi-preparative HPLC (see above).

Synthesis Protocol 2

The synthesis was performed using microwave-equipped automated peptide synthesizer (Liberty 1; CEM) on a 0.1 mmol scale using standard Fmoc-based solid phase peptide synthesis techniques. A resin preloaded with methionine was used and provided in the reactor.

Resin Swelling. Prior to synthesis, the dry resin was washed with DMF.

Deprotection. Removal of the *N*-terminal protecting group was performed *via* 20% (v/v) piperidine in DMF and initialized by MW-irradiation. The resin was subsequently washed with DMF.

Coupling protocol. Amino acids were added in 2 eq to the reactor, HBTU in 5 eq was dosed into the amino acid solution. After that 2 eq of DIEA was added to the resin. The coupling reaction was irradiated with microwaves for a few minutes then the resin was washed with DMF.

Cleavage and workup of the assembled peptide. After completion of synthesis, the peptides were cleaved in 95% (v/v) trifluoroacetic acid (TFA), 2.5% (v/v) triisopropyl silane (TIS), and 2.5% (v/v) H₂O for one hour. The cleaved peptide was precipitated in cold diethyl ether, centrifuged and washed with diethyl ether. The peptide precipitate was then allowed to dry under vacuum to remove residual ether. Crude peptides were purified via semi-preparative HPLC (see above).

On-resin Side-chain Macrocyclization using CuAAC.

On-resin CuAAC was performed using CuBr (1.0 eq), sodium ascorbate (1.0 eq) dissolved in water (20 mg/mL), 2,6-lutidine (10 eq), and DIPEA (10 eq) in 1 mL of degassed DMSO for 18 h at RT to give resin-bound intermediate. Afterward, the solvent was removed *via* lyophilization. Cleavage and workup were performed as described above.

Molecule Modeling

Protein data were collected from the Protein Data Bank (PDB) and prepared using the MOE 2014 LigX function and Amber99 force field. Distance measurements were done using MOE 2014.

Supplementary Materials

Peptides HPLC chromatograms, mass spectra, IR spectra, CD spectra, distance modellings, compound characterization and purities are summarized in supporting information.

Acknowledgments

We are grateful for Werner Tegge for managing our in-house peptide platform operations. We are thankful for Richard Ebright for his attempts in obtaining X-ray co-crystals. We thank Jeanine Jung and Tabea Schramm for technical assistance and diligence in performing the biological assays. We are appreciative for Michael Hoffmann assistance in obtaining HRMS spectra.

References

- Arkin MR, Tang Y, Wells JA (2014). Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.* **21**, 1102–1114.
- Avrutina O, Empting M, Fabritz S, Daneschdar M, Frauendorf H, Diederichsen U, Kolmar H (2009). Application of copper(I) catalyzed azide-alkyne 3+2 cycloaddition to the synthesis of template-assembled multivalent peptide conjugates. *Org. Biomol. Chem.* **7**, 4177–4185.
- Cardote TAF, Ciulli A (2016). Cyclic and Macrocyclic Peptides as Chemical Tools To Recognise Protein Surfaces and Probe Protein-Protein Interactions. *ChemMedChem* **11**, 787–794.
- Craik DJ, Fairlie DP, Liras S, Price D (2013). The future of peptide-based drugs. *Chem. Biol. Drug. Des.* **81**, 136–147.
- Doak BC, Over B, Giordanetto F, Kihlberg J (2014). Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chem. Biol.* **21**, 1115–1142.
- Driggers EM, Hale SP, Lee J, Terrett NK (2008). The exploration of macrocycles for drug discovery—an underexploited structural class. *Nat Rev Drug Discov* **7**, 608–624.
- El-Mowafi SA, Sineva E, Alumasa JN, Nicoloff H, Tomsho JW, Ades SE, Keiler KC (2015). Identification of inhibitors of a bacterial sigma factor using a new high-throughput screening assay. *Antimicrob. Agents Chemother.* **59**, 193–205.
- Empting M, Avrutina O, Meusinger R, Fabritz S, Reinwarth M, Biesalski M, Voigt S, Buntkowsky G, Kolmar H (2011). “Triazole bridge”: disulfide-bond replacement by ruthenium-catalyzed formation of 1,5-disubstituted 1,2,3-triazoles. *Angew. Chem. Int. Ed. Engl.* **50**, 5207–5211.

- Fittler H, Avrutina O, Glotzbach B, Empting M, Kolmar H (2013). Combinatorial tuning of peptidic drug candidates: high-affinity matriptase inhibitors through incremental structure-guided optimization. *Org. Biomol. Chem.* **11**, 1848–1857.
- Fosgerau K, Hoffmann T (2015). Peptide therapeutics: current status and future directions. *Drug Disc. Today* **20**, 122–128.
- Goldflam M, Ullman CG (2015). Recent Advances Toward the Discovery of Drug-Like Peptides De novo. *Front. Chem.* **3**, 69.
- Hüsecken K, Negri M, Fruth M, Boettcher S, Hartmann RW, Haupenthal J (2013). Peptide-based investigation of the Escherichia coli RNA polymerase $\sigma(70)$:core interface as target site. *ACS Chem. Biol.* **8**, 758–766.
- Kuznedelov K, Korzheva N, Mustaev A, Severinov K (2002). Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *EMBO J* **21**, 1369–1378.
- Lau YH, Andrade P de, Wu Y, Spring DR (2015). Peptide stapling techniques based on different macrocyclisation chemistries. *Chem. Soc. Rev.* **44**, 91–102.
- Li H, Aneja R, Chaiken I (2013). Click chemistry in peptide-based drug design. *Molecules* **18**, 9797–9817.
- Ma C, Yang X, Kandemir H, Mielczarek M, Johnston EB, Griffith R, Kumar N, Lewis PJ (2013). Inhibitors of bacterial transcription initiation complex formation. *ACS Chem. Biol.* **8**, 1972–1980.
- Marsault E, Peterson ML (2011). Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem.* **54**, 1961–2004.
- Martí-Centelles V, Pandey MD, Burguete MI, Luis SV (2015). Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization. *Chem. Rev.* **115**, 8736–8834.
- Murakami KS (2013). X-ray crystal structure of Escherichia coli RNA polymerase $\sigma 70$ holoenzyme. *J Biol Chem.*, 9126–9134.
- Okamoto T, Zobel K, Fedorova A, Quan C, Yang H, Fairbrother WJ, Huang DCS, Smith BJ, Deshayes K, Czabotar PE (2013). Stabilizing the pro-apoptotic BimBH3 helix (BimSAHB) does not necessarily enhance affinity or biological activity. *ACS Chem. Biol.* **8**, 297–302.
- Otvos L, Wade JD (2014). Current challenges in peptide-based drug discovery. *Front. Chem.* **2**, 62.
- Pasini D (2013). The click reaction as an efficient tool for the construction of macrocyclic structures. *Molecules* **18**, 9512–9530.
- Pelay-Gimeno M, Glas A, Koch O, Grossmann TN (2015). Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes. *Angew. Chem. Int. Ed. Engl.* **54**, 8896–8927.
- Tischler M, Nasu D, Empting M, Schmelz S, Heinz DW, Rottmann P, Kolmar H, Buntkowsky G, Tietze D, Avrutina O (2012). Braces for the peptide backbone: insights into structure-activity relationships of protease inhibitor mimics with locked amide conformations. *Angew. Chem. Int. Ed. Engl.* **51**, 3708–3712.
- Touloukhonov I, Landick R (2006). The role of the lid element in transcription by E. coli RNA polymerase. *J Mol Biol* **361**, 644–658.
- Vlieghe P, Lisowski V, Martinez J, Khrestchatsky M (2010). Synthetic therapeutic peptides: science and market. *Drug Disc. Today* **15**, 40–56.
- White CJ, Yudin AK (2011). Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **3**, 509–524.

Table 1. Alanine scan results for RNAP inhibition.

Peptide name	Sequence ^a	IC ₅₀ (μM) or %inhibition ^b	95% CI ^c
1	<i>H</i> -TNRGLQFLDLIQEGNIGLM- <i>OH</i>	5.2	4.6 - 5.8
2	<i>H</i> - A NRGLQFLDLIQEGNIGLM- <i>OH</i>	6.2	4.9 - 7.8
3	<i>H</i> -T A RGLQFLDLIQEGNIGLM- <i>OH</i>	16	14 - 20
4	<i>H</i> -TN A GLQFLDLIQEGNIGLM- <i>OH</i>	28	22 - 36
5	<i>H</i> -TNR A LQFLDLIQEGNIGLM- <i>OH</i>	12	10 - 14
6	<i>H</i> -TNRG A QFLDLIQEGNIGLM- <i>OH</i>	6.5	5.4 - 7.8
7*	<i>H</i> -TNRGL A FLDLIQEGNIGLM- <i>OH</i>	5.2	4.7 - 5.7
8	<i>H</i> -TNRGLQ A LDLIQEGNIGLM- <i>OH</i>	8.2	5.7 - 12
9	<i>H</i> -TNRGLQF A DLIQEGNIGLM- <i>OH</i>	8.9	7.3 - 11
10*	<i>H</i> -TNRGLQFL A LIQEGNIGLM- <i>OH</i>	7.1	6.6 - 7.6
11	<i>H</i> -TNRGLQFLD A IQEGNIGLM- <i>OH</i>	11	9.8 - 12
12	<i>H</i> -TNRGLQFLDL A QEGNIGLM- <i>OH</i>	>50	--
13*	<i>H</i> -TNRGLQFLDL A EGNIGLM- <i>OH</i>	9.6	8.3 - 11
14*	<i>H</i> -TNRGLQFLDLIQ A GNIGLM- <i>OH</i>	9.1	6.0 - 14
15	<i>H</i> -TNRGLQFLDLIQE A NIGLM- <i>OH</i>	5.4	4.6 - 6.4
16	<i>H</i> -TNRGLQFLDLIQEG A IGLM- <i>OH</i>	13	10 - 16
17	<i>H</i> -TNRGLQFLDLIQEGN A GLM- <i>OH</i>	>50	-
18	<i>H</i> -TNRGLQFLDLIQEGNI A LM- <i>OH</i>	5.0	4.6 - 5.5
19	<i>H</i> -TNRGLQFLDLIQEGNIG A M- <i>OH</i>	56	38 - 81
20	<i>H</i> -TNRGLQFLDLIQEGNIGL A - <i>OH</i>	7.7	7.0 - 8.5

^a) Substituted residues are indicated in bold red. ^b) % Inhibition mean values and SD are given.

^c) 95% CI means a 0.95 probability of containing the population mean. Highlighted in red are results with more than 5-folds loss and in orange between 2-5 folds loss compared to parent peptide. *Peptides were previously reported in (Hüsecken et al. 2013).

Table 2. Biological activity of terminal amino acid substitutions and truncation.

Peptide name	Sequence ^a	IC ₅₀ (μM)	95% CI ^b
1	<i>H-TNRGLQFLDLIQEGNIGLM-OH</i>	5.2	4.6 - 5.8
21	<i>H-NRGLQFLDLIQEGNIGLM-OH</i>	6.7	5.6 - 8.1
22	<i>H-TNRGLQFLDLIQEGNIGL---OH</i>	14	11 - 18
23	<i>H-TNRGLQFLDLIQEGNIGLNor-OH</i>	7.1	6.2 - 8.2
24	<i>H-NRGLQFLDLIQEGNIGLW-OH</i>	4.7	3.8 - 5.9

^a) Substituted residues are indicated in bold red and deleted residues are indicated by bold red dashes ^b) 95% CI means a 0.95 probability of containing the population mean. ^c) Nor stands for un-natural amino acid norleucine.

Table 3. Peptide sequences and results of macrocyclized derivatives (**29-32**) and their linear precursors (**25-28**).

Peptide	Sequence ^a	IC ₅₀ or % inhibition (μM) ^b	95% CI ^c	% α helicity ^d
1	<i>H</i> -TNRGLQFLDLIQEGNIGLM- <i>OH</i>	5.2	4.6 - 5.8	24.8
25		9.0	7.9 - 10	34.6
26		43	30 - 63	6.9
27		5.1	4.4 - 6.0	56.7
28		4.5	3.7 - 5.5	36.2
29		14 ± 0.9% @ 50 μM		33.7
30		9.8 ± 1.4% @ 50 μM		31.8
31*		50 ± 1.5% @ 25 μM		33

^{a)} Substituted residues are indicated in bold red ^{b)} % inhibition mean values and SD are given

^{c)} 95% CI means a 0.95 probability of containing the population mean. ^{d)} Percentage of α-helicity was computed *via* CONTIN algorithm. *Results shown are from a 3:2 mixture of **31** and **27**.

Figure 1. Co-crystal of RNA polymerase and the σ 70 factor (PDB ID: 4YG2) with P07 highlighted in yellow.

Figure 2. Close-up of peptide **1** (P07) sequence bound to RNA polymerase (In red, yellow, and blue are amino acids with 5-, 5-2, <2 folds of activity loss, respectively) as found in the coordinates of σ 70 protein (PDB ID 4YG2). Highlighted with black dotted lines are the amino acid positions used for macrocyclization.

Scheme 1. Description of the general procedure for on-resin side-chain macrocyclization using CuAAC. Reagents and conditions: CuBr (1.0 eq), sodium ascorbate (1.0 eq) dissolved in water (20 mg/mL), 2,6-lutidine (10 eq), and DIPEA (10 eq) in 1 mL of degassed DMSO for 18 h at RT.