EPFL

Thèse n° 9431

Generating macrocyclic inhibitors of protein-protein interactions

Présentée le 1er avril 2022

Faculté des sciences de base
Laboratoire de protéines et peptides thérapeutiques
Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

Gontran SANGOUARD

Acceptée sur proposition du jury

Prof. K. Severin, président du jury
Prof. C. Heinis, directeur de thèse
Prof. T. Kodadek, rapporteur
Prof. P. Arora, rapporteur
Dr R. Hovius, rapporteur

École polytechnique fédérale de Lausanne

2022

"Science is not only a disciple of reason but also one of romance and passion."

Stephen Hawking

Acknowledgments

First, I would like to thank Prof. Christian Heinis, who gave me the opportunity to work on exciting projects in a dynamic and fascinating environment. Your advice and scientific guidance allowed me to learn, grow, and become a better scientist overall.

I would also like to thank the members of my thesis committee Prof. Kay Severin, Prof. Thomas Kodadek, Prof. Paramjit Arora, and Dr. Ruud Hovius. In addition, I am grateful to the members of my candidacy exam committee who were Prof Christian Heinis, Prof. Gerardo Turcatti, and Dr. Ruud Hovius.

A big thank you to Béatrice Bliesener-Tong for all your help during these years. On a smaller note, also, thank you for the delicious cakes, tiramisu, and cookies often brought in the lab.

To all members of the lab, both past (Ale, Carl, Ganesh M., Ganesh S., Hitesh, Jonas, Jun, Maurice, Patrick, Sangram, Vanessa, Xudong, Yuteng), and current (Alexander, Bo, Cristina, Ed, Manuel, Mischa, Sevan, Xinjian, Zsolt), thank you very much for the fantastic ambiance in our group, the Laboratory of Therapeutic Proteins and Peptides (LPPT). It was always a pleasure to work with such friendly colleagues. Thank you also for always being there if I needed some advice or help. An extra thank you to Mischa for the home-made devices, the desserts, and the numerous conversations we had about so many topics of common interest (space, movies, snowboarding, etc.). I am also grateful to Jonas, who initially was my semester project supervisor during my master thesis and stayed in touch after and till now for numerous events. It was really nice to have you in the lab at the beginning of my Ph.D. Thank you as well to Sangram, who supervised me during the start of my thesis.

I am thankful to everyone who helped contribute to my project. At LPPT, Mischa, Sangram, Cristina, Ed, Maurice, Ale, and Yuteng. To my students, Edouard and Simon, thank you for your hard work and dedication. It was very nice working with you. From the BSF, thank you to Jonathan and Julien for your contribution on the experiments and all your advice about HTS. You were both very nice, and it was always a pleasure to come to your facility. Thank you to Soraya and Laurence for your help in expressing IL-23R in mammalian cells. Thank you also to Prof. David Hacker for your advice regarding protein expression and purification.

To all the staff from the Magasin in BCH (Antonio, Florent, Kosta, Maurizio), thank you for your help in handling and receiving all the compounds. A special thank you to Maurizio for your warm welcoming in the shop as well as the numerous conversations we had.

Thank you to all my family and friends for the support you have given me and all the great moments we spent together.

Last but not least, a huge thank you to my amazing girlfriend, Yara. You have always encouraged me, believed in me and were there for me. Your love and support were really important during all these years. I am very lucky to have you in my life.

Abstract

Protein-protein interactions (PPIs) play important roles in many diseases and their modulation is an attractive strategy for developing new therapeutics. However, the relatively large and often flat binding surfaces of PPIs makes the development of inhibitors based on classical small molecule drugs rather challenging. Macrocycles, a class of ring-shaped chemical structures, are able to bind to challenging targets such as PPIs, while having the potential to be cell-permeable or even orally available. The potential of targeting intracellular PPIs by macrocyclic compounds has triggered a great deal of interest in the pharmaceutical industry. A challenge with developing macrocycle-based drugs to new targets, however, is the identification of ligands, the reason for this mainly being the lack of large libraries of macrocyclic compounds that could be screened.

The goal of my Ph.D. thesis project was to develop macrocyclic inhibitors for several protein-protein interactions. Towards this end, I contributed to establishing novel strategies for synthesizing and screening large macrocyclic compound libraries. Additionally, I established reported or developed new bioassays suited for the high-throughput screening of compound libraries, and I applied them for screening macrocycle libraries generated with the new methods.

In my first project, I applied acoustic droplet ejection technology (ADE) for the synthesis and screening of large macrocyclic compound libraries at a picomole scale. Towards this aim, I used a synthetic strategy based on the combinatorial diversification of *m* N_{α}-bromoacetamide linear peptides with *n* primary amines followed by cyclization with *o* bis-electrophile linkers yielding $m \times n \times o$ macrocycles. I used ADE to perform all liquid transfers, which proved efficient and yielded macrocycles at picomole scale. I performed the screening by directly adding assay reagents to the synthesis plates containing crude macrocycle products. As a proof-of-concept, I assembled a 2,700-member target-focused macrocycle library and screened it against the MDM2:p53 protein-protein interaction, and was able to identify macrocyclic inhibitors with micromolar to sub-micromolar affinity. Most importantly, this work validated the feasibility to synthesize combinatorial macrocycle libraries using ADE at the picomole scale, which was subsequently applied to several other approaches later developed in our lab.

In my second project, I aimed at finding macrocyclic inhibitors of the PPI IL-23:IL-23R, which is an important target of several autoimmune disorders. To that end, I developed a high-throughput screening assay for identifying inhibitors of the interaction. I initially developed a screening assay based on fluorescence polarization, which turned out to be prone to artifacts.

In a second attempt, I established an assay based on TR-FRET that proved reliable for the identification of IL-23:IL-23R inhibitors, and in particular also for the screening of crude products of macrocyclization reactions. I subsequently synthesized a pilot-scale library by combinatorially cyclizing 384 linear peptides with 8 cyclization linkers to generate 3,072 target-focused macrocyclic compounds. Screening the library using the TR-FRET-based assay identified single-digit micromolar inhibitors of the IL-23:IL-23R PPI, that however, were based on linear side-products of the macrocyclization reactions rather than macrocyclic structures. While no potent macrocyclic inhibitors could be identified from the first library, the results validated the robustness of the assay and its applicability for high-throughput screening of IL-23R antagonists.

In my third project, I aimed at finding macrocyclic inhibitors of KRAS(G12D):effectors PPIs, known to be key drivers in numerous cancers. To that aim, I synthesized a target-focused library of 92,160 macrocyclic compounds using a strategy that was developed just before in our lab and is based on the combinatorial diversification of a large number of cyclic peptide scaffolds via a lateral amino group, using a large number of carboxylic acids. The library assembly was performed at a picomole scale using ADE and directly screened using an assay based on fluorescence polarization. While the screen did not identify any binders for KRAS(G12D), the large macrocycle library was later screened by other members of the lab and yielded ligands to a PPI target.

Overall, my work contributed to developing and evaluating new methods for the synthesis and screening of large macrocycle libraries at a picomole scale. Additionally, my work led to the establishment of an assay suited for the high-throughput screening of compounds against IL-23:IL-23R PPI. The future will tell if the methods developed yield drug candidates against therapeutically important PPI targets.

Résumé

Les interactions protéine-protéines sont impliquées dans de nombreuses maladies, et leurs modulations est une stratégie attractive pour le développement de nouveaux médicaments. Cependant, à cause de leurs relativement larges surfaces d'interactions, le développement d'inhibiteurs basés sur des petites molécules traditionnelles a été compliqué. Les macrocycles, une classe de composé ayant une structure chimique cyclique, sont capable d'inhiber des interactions entre des protéines, tout en pouvant traverser la membrane cellulaire et être délivrés oralement. Cette combinaison de propriétés fait que cette classe de molécule attire beaucoup d'intérêt dans l'industrie pharmaceutique. Cependant, leur développement a été difficile à cause d'un manque important de librairies de composés macrocycliques qui pourraient être testées.

Le but de ma thèse était de développer des inhibiteurs macrocycliques pour plusieurs interactions protéine-protéines. Pour cela, j'ai contribué au développement d'une nouvelle stratégie synthétique qui permet d'accéder rapidement à des grandes librairies de composés macrocycliques. De plus, j'ai développé plusieurs tests, nouveaux ou basés sur des rapports publiés, permettant de simuler les interactions entre les protéines d'intérêts et de tester à très grands débits les librairies de composés macrocycliques synthétisées avec la nouvelle méthode.

Dans mon premier projet, j'ai appliqué l'utilisation de la technologie d'éjection acoustique de gouttelettes pour la synthèse et le test d'une large librairie de composés macrocycliques à l'échelle du picomole. Pour cela, j'ai utilisé une stratégie combinatoire de synthèse qui implique la diversification de *m* peptides linéaires contenant un N_{α}-bromoacétamide, avec *n* amines primaires suivit de la cyclisation avec *o* réactifs, pour donner $m \times n \times o$ macrocycles. J'ai utilisé l'éjection acoustique de gouttelettes pour tous les transferts de liquide, ce qui a prouvé être une méthode efficace et a permis la génération de macrocycles à l'échelle du picomole. J'ai ensuite ajouté les composés requis au test directement dans la plaque de synthèse contenant le mélange réactionnel avec les macrocycles non-purifiés. Pour valider le concept, j'ai assemblé une librairie de 2,700 macrocycles axés sur la protéine cible et je les ai testés contre l'interaction MDM2:p53. Cela a permis de découvrir des inhibiteurs avec une activité de l'ordre du bas-micromolaire et en dessous. Mon travail a donc permis de valider la possibilité de synthétiser de façon combinatoire des librairies de macrocycle en utilisant l'éjection acoustique de gouttelette à l'échelle du picomole, qui a plus tard été appliquée à d'autres stratégies de synthèse qui ont été développées dans notre laboratoire.

Dans mon deuxième projet, mon but était de trouver des inhibiteurs macrocycliques contre l'interaction IL-23:IL-23R, qui est impliquée dans de nombreuses maladies autoimmunes. Pour cela, j'ai développé un test compatible avec les tests à très haut débit de composés en parallèle dans le but d'identifier des molécules inhibitrices de l'interaction d'intérêt. J'ai initialement développé un test basé sur la polarisation de la fluorescence, qui a prouvé être sensible à des effets indésirables des composés testés. Dans un second essai, j'ai développé un test basé sur le TR-FRET qui a prouvé être robuste pour le test de molécules et l'identification d'inhibiteurs de l'interaction IL-23:IL-23R, et en particulier aussi du test du mélange réactionnel contenant les macrocycles. J'ai alors synthétisé et testé une librairie en cyclisant de façon combinatoire 384 peptides linaires avec 8 réactifs pour donner 3,072 composés macrocycliques axés sur la protéine cible. Le test de la librairie utilisant le test basé sur le TR-FRET a permis l'identification d'inhibiteurs avec une activité de l'ordre du bas-micromolaire, qui cependant, étaient basés sur des produits secondaires linéaires de la réaction de cyclisation et non liés aux composés macrocycliques. Bien qu'aucun inhibiteur macrocyclique ayant une forte activité n'a été découvert, les résultats ont validés la fiabilité du test développé et son application pour le test à très haut débit de nombreux composés contre l'intéraction IL-23:IL-23R.

Dans mon troisième projet, mon but était de de trouver un inhibiteur macrocyclique contre l'interaction entre KRAS(G12D) et des protéines permettant son activation, connue pour être un facteur clé dans de nombreux cancers. Pour ce faire, j'ai synthétisé une librairie de 92,160 macrocycles axés sur la protéine cible utilisant une approche développée dans notre laboratoire et basée sur la diversification combinatoire d'un grand nombre de peptides cycliques via un groupe amine latéral, en utilisant un grand nombre d'acides carboxyliques. Cette librairie a été assemblée à l'échelle du picomole en utilisant l'éjection acoustique de gouttelettes, et directement testée contre un test basé sur la polarisation de la fluorescence. Bien qu'aucun inhibiteur de KRAS(G12D) n'a été identifié, cette librairie de macrocycle a plus tard été utilisée par d'autres collègues et a permis de découvrir des inhibiteurs d'autres interactions protéine-protéine.

Dans l'ensemble, mon travail a aidé au développement et à l'évaluation de nouvelles méthodes de synthèse de large librairies de macrocycle à l'échelle du picomole. De plus, mon travail a permis la génération d'un test fiable pour tester de nombreux composés contre IL-23:IL-23R. Le futur nous dira si le format moléculaire synthétisé peut générer des candidats de médicaments contre des interactions protéine-protéines d'intérêts.

Table of contents

Abbreviations

5(6)-FAM	5(6)-carboxyfluorescein
AC_2O	Acetic anhydride
ACN	Acetonitrile
AcOH	Acetic acid
ADE	Acoustic droplet ejection
ALPHA	Amplified luminescent proximity homogeneous assay
Amp	Ampicillin
Arg	L-arginine
Asp	L-aspartic acid
AU	Arbitrary unit
Bad	BCL-2 associated agonist of cell death
Bax	BCL-2-associated X
BCL-2	B-cell lymphoma 2
BDT	1,4-butanedithiol
Boc	Ter-butoxycarbonyl
BrAc	2-bromoacetamide
bRO5	Beyond rule of five space
BSA	Bovine serum albumin
BSF	Biomolecular sreening facility
CCR5	C-C Motif chemokine receptor 5
CD40	Cluster of differentiation 40
CD40-L	CD40 ligand
Cdc37	Hsp90 co-chaperone Cdc37
CDKs	Cyclin dependent kinases
CFP	Cyan fluorescent protein
CHO	Chinese hamster ovary
cLog P	Calculated partition coefficient
c-Myc or MYC	Cellular myelocytomatosis oncogene protein
CREB	C-AMP response element-binding protein
Cy5	Cyanine 5
cys	Cysteamine
Cys*	Cysteamine
DABCO	1,4-diazabicyclo[2.2.2]octane
DCM	Dichloromethane
DIC	N,N’-diisopropylcarbodiimide
DMF	N,N-dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTS	DNA-template synthesis
DTT	Dithiothreitol
E6AP	E6-associated protein
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid

EGF	Epidermal growth factor
EPFL	Swiss federal institute of technology, Lausanne
ESI-MS	Electrospray ionization mass spectroscopy
Et_2O	Diethyl ether
Eu	Europium
Fc-domain	Fragment crystallizable region
FGF	Fibroblast growth factor
FIT	Flexible *in vitro* translation system
Fmoc	9-fluorenylmethoxycarbonyl
FP	Fluorescence polarization
FRET	Föster (or fluorescence) resonance energy transfer
FTL3	Fms-like tyrosine kinase-3
GDP	Guanosine diphosphate
Glu	L-glutamine
Gly	L-glycine
Gp120	Envelope glycoprotein GP120
GPCR	G protein-coupled receptors
GST	Glutathion S-transferase
GTP	Guanosine triphosphate
HATU	Hexafluorophosphate azabenzotriazole tetramethyl uronium
H-bond	Hydrogen bond
HBTU	Hexafluorophosphate benzotriazole tetramethyl uranium
HCV	Hepatitis C virus3
HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His_6	6 × L-histidines
HPLC	High-pressure liquid chromatography
HPV-E1	Human Papillomavirus E1 protein
HPV-E2	Human Papillomavirus E2 protein
HRAS	Harvey-RAS
Hsp40	Heat shock protein 40 kDa
Hsp70	Heat shock protein 70 kDa
Hsp90	Heat shock protein 90 kDa
HTS	High-throughput screening
IBD	Inflammatory bowel disease
IC_{50}	Half maximal inhibitory concentration
ICAM-1	Intercellular adhesion molecule 1
IL-12	Interleukin-12
IL-12Rβ1	IL-12 receptor subunit β1
IL-17A	Interleukin-17A
IL-17AR	IL-17A receptor
IL-2	Interleukin-2
IL-23	Interleukin-23
IL-23R	IL-23 receptor
IL-2R	IL-2 receptor
Ile	L-isoleucine

IPTG	Isopropyl β-D-1-thiogalactopyranoside
JAK2	Janus kinase-2
Kan	Kanamycin
K_D	Dissociation constant
KEAP1	Kelch-like ECH-associated protein 1
K_i	Inhibition constant
KRAS	Kristen rat sarcoma isoform
KRAS	Kristen-RAS
LB	Luria Broth
LC-MS	Liquid chromatography-mass spectrometry
LDLR	Low density lipoprotein receptor
LDV	Low dead volume
Leu	L-leucine
LFA-1	Lymphocyte function-associated antigene 1
LOCI	Luminescent oxygen channeling immunoassay
Lys(Dde)-OH	N-α-Fmoc-N-ε-1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl-L-lysine
Max	Myc-associated factor X
MDM2	Mouse double minute 2
Mea	Mercaptoethylamine (cysteamine)
MeOH	Methanol
MMP	Matrix metalloproteinase
MPA	3-mercaptopropionic acid
mRNA	Messenger RNA
MW	Molecular weight
MWCO	Molecular weight cut-off
NMM	N-methyl morpholine
NMP	N-methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
NRAS	Neuroblastoma-RAS
NRF2	Nuclear factor (erythroid-derived 2)-like 2
OBOC	One-bead-one-compound
OD_{600}	Optical density at 600 nm
p53	Tumor protein 53
PCR	Polymerase chain reaction
PCSK9	Proprotein convertase subtilisin/kexin type 9
PD-1	Programmed cell death protein 1
PDAC	Pancreatic ductal adenocarcinomas
PD-L1	Programmed death-ligand 1
PED-MS	Partial Edman degradation-mass spectrometry
PMSF	Phenylmethylsulfonyl fluoride
PP	Polypropylene
PPIs	Protein-protein interactions
PS	Polystyrene
PSA	Polar surface area
RAS	Rat sarcoma

RNA	Ribonucleic acid
RO5	Lipinski's rule of five
RP-HPLC	Reverse phase HPLC
rpm	Round per minute
RT	Room temperature
SDs	Standard deviations
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	L-serine
SOS1	Son of sevenless homolog 1
SPPS	Solid phase peptide synthesis
SPR	Surface plasmon resonance
TAE	Tris-acetate-EDTA
TBTA	Tris(benzyltriazolylmethyl)amine
tBu	Ter-butyl
TCEP	Tris(2-carboxyethyl)phosphine
Tcf4	Transcription factor 4
TEA	Triethylamine
TEV	Tobacco etch virus
TFA	Trifluoro acetic acid
Th17	IL-17 secreting CD4+ T helper cells
TIC	Total ion count
TIS	Triisopropylsilane
TNFα	Tumor necrosis factor-α
TNFα-R	TNFα receptor
TR-FRET	Time-resolved FRET
TRIS	Tris(hydroxymethyl)aminomethan
tRNA	Transfer RNA
Trp	L-tryptophan
$Trp^{6\text{-}Cl}$	6-chloro-L-tryptophan
Tyr	L-tyrosine
UHPLC	Ultra high-pressure liquid chromatography
UV	Ultraviolet
XIAP	X-linked inhibitor of apoptosis protein
YFP	Yellow fluorescent protein

1. Introduction

1.1 Protein-protein interactions

1.1.1 Protein-protein interactions: an important source of drug targets

The human proteome was estimated to contain around 20,000 canonical proteins, among which 3,000 were associated with diseases.[1] Out of these, only 600 to 1,500 proteins have been considered amenable for manipulation with traditional orally available small molecule drugs (typically < 500 Da, properties described in section 1.1.3), which target well-defined binding sites such as enzymes or G protein-coupled receptors (GPCRs).[1] As only a subset of proteins seems reachable by traditional drug format, therapeutic capabilities could appear restricted. However, another target class being protein-protein interactions (PPIs), offers numerous potential target sources for drug discovery. PPIs are involved in countless vital cellular functions such as cellular growth/death, DNA replication, transcriptional activation, translation, and transmembrane signal transduction. Due to their key biological role, dysregulation of PPIs was associated with numerous diseases such as cancer,[2–7] infectious diseases,[8,9] neurological disorders,[10,11] heart failure,[12,13] and inflammation[14–16] (see examples in Table 1). Dysregulation of PPIs is caused by the loss of an essential interaction or the formation/stabilization of a protein complex at an inappropriate time and location.[17] PPIs were approximated to be 650,000 in humans, among which 15 to 40% were estimated to be druggable (see sections 1.1.2 and 1.1.3), yielding between 97,500 to 260,000 potential new targets.[18,19] Therefore, PPIs offer the most important source of targets for drug discovery. Targeting PPIs using molecules that can modulate their interactions and activities could lead to numerous novel treatments for a broad spectrum of diseases.

Table 1. Examples of PPIs involved in diseases. Table 1 continues next page.

PPI	Related disease	Reference
MDM2/p53	Cancer	[3]
β-Catenin/CREB binding protein	Cancer	[4]
Bcl-2/Bax	Cancer	[5]
CD40/CD40-L	Cancer	[6]
PD-1/PD-L1	Cancer	[7]
Gp120/CCR5	HIV	[8]
PCSK9/LDLR	Cardiovascular disease	[12]

Table 1 continued

PPI	Related disease	Reference
IL-2/IL-2R	Auto-immune disease	[14]
TNFα-/TNFα-R	Auto-immune disease	[15]
KEAP1/NRF2	Cancer	[20]
c-Myc/Max	Cancer	[21]
HPV-E2/HPV-E1	Viral infection leading to cancer	[22]
IL-17A/IL-17AR	Auto-immune disease	[16]
Amyloid-β aggregation	Alzheimer's disease	[11]
LFA-1/ICAM-1	Dry eye	[23]
Bromodomain/Histone	Cancer	[24]

1.1.2 Classifying protein-protein interactions and their druggability

PPIs are defined by the interaction between one (or several) proteins with one (or several) others. They can be homomeric (same proteins involved) or heteromeric (different proteins involved).[25] Depending on their properties, the interaction between two proteins can be divided into several classes, among which some are more easily druggable with traditional small molecules.

PPIs can be classified into several classes based on their structural features. The first class consists of the interaction between two globular proteins. Depending on the protein pair, the binding event can be characterized by no conformational change due to pre-formed surfaces or important conformational changes due to significant structural reorganization (Figure 1a & 1b).[26] The surface interaction between two globular proteins is usually large, flat, and relatively featureless.[27] Due to the lack of clearly defined pockets, this class of PPIs is particularly challenging to access with small molecules, which usually bind narrow and deep pockets. The second class of PPIs is characterized by the binding between a globular protein with a single peptide chain of another protein (Figure 1c & 1d).[26] Upon binding, a single peptidic part of a protein can adopt one or more secondary structural elements, such as α-helices or β-sheets, to interact with the surface of the globular protein. These secondary structures could already be pre-formed before binding or might need a folding to reach the correct conformation. This type of interaction is characterized by smaller interaction surfaces compared to the interaction between two globular proteins.[27] Defined surface pockets to anchor the peptidic part of one of

the proteins are present on the epitope of the globular protein, making this class of PPIs more reachable to interfering compounds that can bind in the same pockets and disrupt the interaction.[26] Due to a relatively tight binding domain and defined pockets, this PPI class is the most accessible for therapeutic intervention. It was estimated to represent around 15 to 40% of all PPIs.[19] The third class of PPIs is defined by the interaction between two peptidic regions of two proteins (Figure 1e).[26] One or both peptidic regions can be intrinsically disordered when uncomplexed and need folding into the required secondary structure upon binding. This aspect makes this class of PPIs particularly difficult to drug due to the potential lack of predefined binding sites.

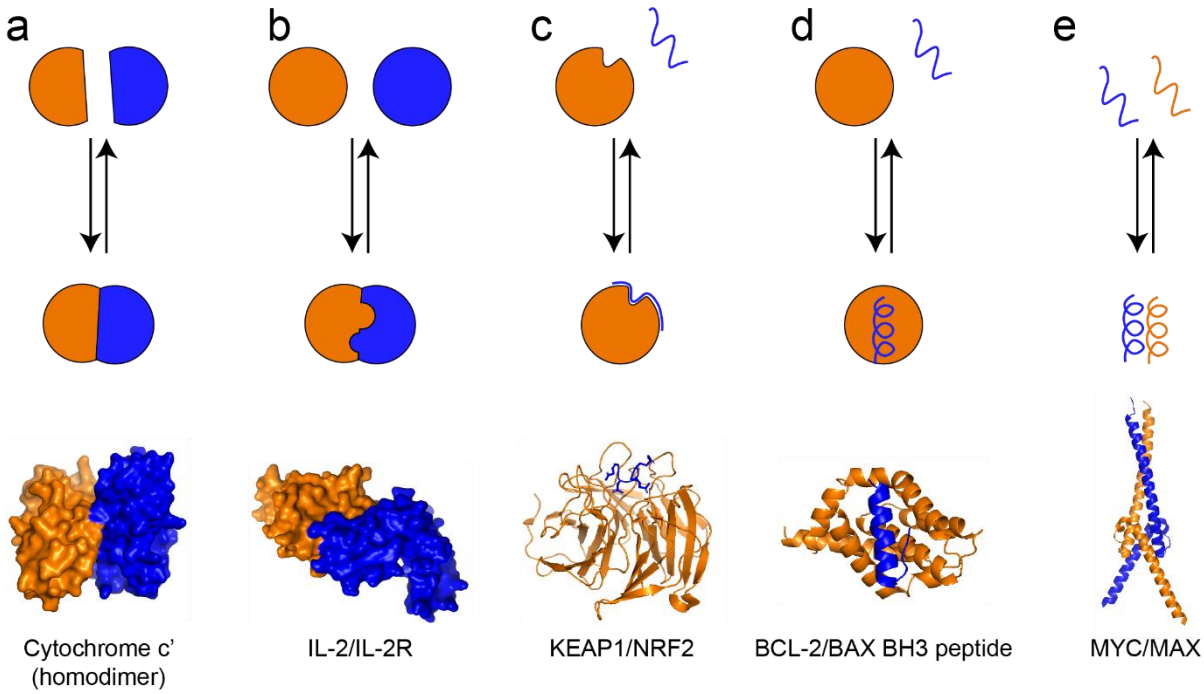


Figure 1. PPIs structural classification. The upper part of the figure shows a simplification of the different PPI classes. The lower part shows the crystal structure of a corresponding example. (a) Interaction between two globular proteins presenting pre-formed surfaces (PBD entry 2ccy). (b) Interaction between two globular proteins with surface rearrangement (PDB entry 1z92). (c) Interaction between a globular protein and a peptidic part of a protein (PDB entry 2dyh). (d) Interaction between a globular protein upon rearrangement of a peptidic part of a protein (PDB entry 2xa0). (e) Interaction between two peptides (PDB entry 1nkp). This figure was reproduced from the reference paper [26]. Permission conveyed through Copyright Clearance Center, RightsLink®. (license ID: 5255800924553).

Another way to categorize PPIs is based on whether they are narrow/wide and tight/loose. Based on their binding affinity (K_D) and their buried surface area, PPIs can be categorized into four categories (Figure 2).[28] "Tight and Narrow" and "Loose and Narrow" PPIs are classified by buried surface area lower than 2,500 Å^2 and K_D in the low nanomolar range (below 200 nM) or high nanomolar range (above 200 nM), respectively. Alternatively, "Tight and Wide" and "Loose and Wide" PPIs are characterized with large buried surface area (above 2,500 Å^2)

and K_D in the low nanomolar range (below 200 nM) or high nanomolar range (above 200 nM), respectively. Little correlation between the buried surface area and the binding affinity of PPIs was established. The main interaction driving the affinity of PPIs was often linked to certain key residues called "hot spots".[29] These often involve large amino acids such as Tyrosine, Arginine, and Tryptophan that bind in small pockets across the surface and contribute to a major part of the binding energy. These hot spots can be spatially separated but interact cooperatively for interface stability. Considering the importance of hot spots, it is understandable that "Tight and Narrow" PPIs have proven to be the most amenable to inhibition as around 80% of reported therapeutic molecules target this class.[30] This can be explained by the presence of relatively deep pockets in one of the proteins engaged by only a few amino acids with spatially close "hot spots" amino acids of the other protein. This property makes it easier to design molecules that can tightly bind in the same binding pocket by mimicking the key interactions of the "hot spots" amino acids. Alternatively, finding potent small molecule inhibitors against loose or wide PPIs proved challenging.[28] For loose PPI, this could be explained by their low affinity, making it hard to develop molecules with high affinity even if they mimic potential hot spots interaction. For wide PPI, the challenge in generating inhibitors could be explained by spatially distant hot spots, making it hard for molecules of small size to mimic the large PPIs binding interaction. If the PPI is loose and wide, both limitations are combined, making it the most challenging target class among all the PPI.

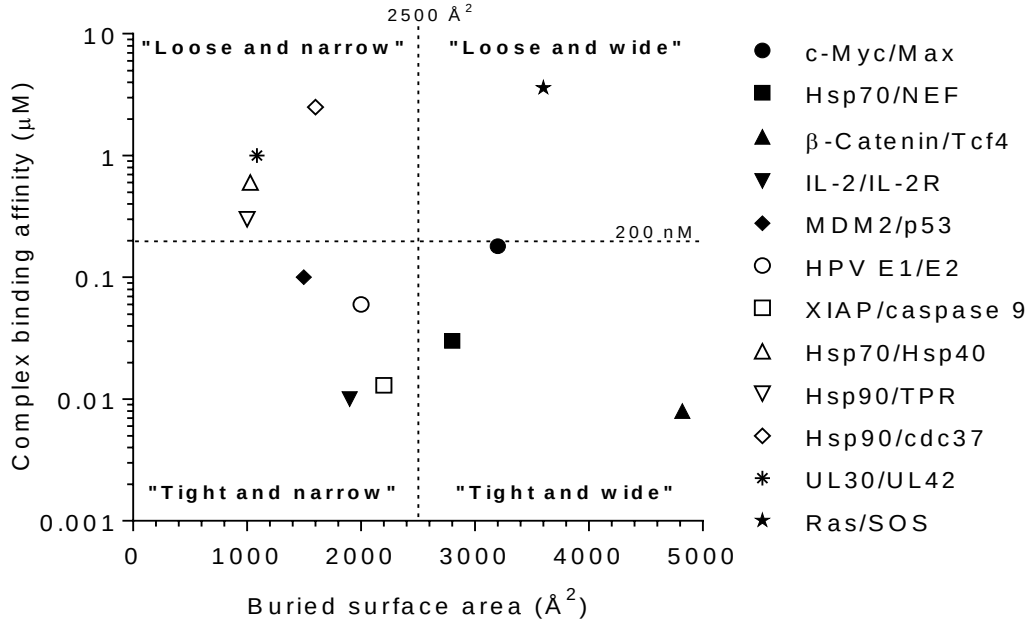


Figure 2. Categorization of PPIs. Arbitrary categories of PPIs based on their corresponding binding affinities and buried surface areas. The values were obtained from the reference paper [28].

1.1.3 Modalities to inhibit protein-protein interactions

In most cases, inhibition of PPIs is achieved by orthosteric modulation. The inhibitor binds in the binding site of one of the two proteins, resulting in the dissociation of the other protein. PPIs inhibitors are classified into three main categories being small molecules, peptides, and antibodies described below. Due to their properties, these different classes of inhibitors are not suitable for all types of PPIs, with some inhibitors capable of targeting more difficult-to-drug PPIs than others. It is worth emphasizing that among all the existing modalities to inhibit PPIs, the most important challenge remains to find inhibitors that can interfere with intracellular PPIs. In this case, if capable of interfering with the PPIs, the inhibitor needs to be cell-permeable, which is not the case for most of the inhibitors discussed. Additionally, one other significant challenge in targetting PPIs is generating orally available compounds, which is often the preferred delivery strategy but has been difficult to achieve mainly due to the size of the PPIs inhibitors.

The first class of PPIs inhibitors is the small molecules. They usually fall into the chemical space defined by Lipinski's rule, making them an attractive drug format as they can be orally delivered. This rule, also named "rule of five" (RO5), predicts poor passive permeation/oral bioavailability of compounds presenting more than 5-H-bond donors, 10 H-bond acceptors, a molecular weight greater than 500 Da, and a calculated partition coefficient (cLog P) greater than 5.[31] This molecular format was traditionally used for targets with small deep pockets such as enzyme active sites.[32] Due to their limited size, small molecules were often considered unsuitable for disrupting PPIs. However, they proved capable of inhibiting some PPIs, especially those presenting tight and narrow interfaces.[30,33] In this type of PPI, the small molecule can still bind in the defined hot spot area, which is key to the interaction. In addition, due to their usual high membrane permeability, small molecules can target intracellular PPIs. Due to their attractive properties, small molecules appear to be the most common format chosen to target PPIs.[30] Successful PPIs inhibition by small molecules highlights this drug format potential (see examples in Table 2).

The second category of PPIs inhibitors is peptides. Peptides are derived from the protein region involved in the binding to its protein partner.[34] They are characterized by molecular weights traditionally higher than 1,000 Da. Due to their size, they can interact with relatively large surface areas, making some large PPIs accessible (buried surface area > 2,500 $Å^2$).[35] This molecular format is especially suitable for modulation of PPIs characterized by single peptidic chain binding to the other protein. Peptides are usually characterized by a strong binding affinity and selectivity towards the desired proteins. However, rapid metabolic instability, poor membrane permeability, and oral availability are often important limitations in

their development as a therapeutic.[30] To improve their properties, several options can be considered, such as introducing non-natural amino acids, a conformational restricted secondary structure such as with stapled peptides, or even short peptide mimetics.[30,36,37] However, their usual low oral bioavailability and poor membrane permeability limit their use. While traditionally limited to extracellular targets, PPI modulation using peptide as drug format has been successful (see examples in Table 2). Additionally, due to their capacity to mimic the secondary structure required in many PPIs, peptides can also be used as a tool to study PPIs modulation followed by their conversion into more drug-like molecules.[38]

The third type of PPI inhibitor is monoclonal antibodies. They are characterized by a large molecular weight (on average 150 kDa) and their capacity to bind broad domain, making numerous PPIs accessible.[17] They are characterized by a high target specificity and high affinity. While they are traditionally unable to target intracellular PPIs due to their size and are not orally available, they have been successfully used to target extracellular targets. (see examples in Table 2).

Table 2. Examples of PPI inhibitors.

	PPI	Related disease	Name	Developer	Reference
Small molecules	MDM2/p53	Acute myeloid leukemia	Idasanutlin	Roche	[39]
	Bcl-2/Bax	Chronic lymphocytic leukemia	ABT-199	Novartis	[40]
	XIAP/caspase-9	Solid tumors, lymphoma	ASTX-660	Astex	[41]
	PD-1/PD-L1	Prostatic neoplasms	CA-170	Astellas	[42]
Peptides	MDM2/p53	Advanced solid tumors, lymphomas	ALRN-6924	Aileron	[43]
	IL-23/IL23-R	Inflammatory bowel disease	PTG-200	Protagonist Therapeutics	[44]
Antibodies	CD40/CD40-L	Multiple myeloma	Lucatumumab	Novartis	[6]
	PD-1/PD-L1	Metastatic breast cancer	Avelumab	EMD Serono Inc.	[45]

1.2 Macrocycles as protein-protein interaction inhibitors

1.2.1 Macrocycles properties

Macrocycles are cyclic compounds containing a ring size of 12 atoms or more.[46] In terms of molecular weights, they traditionally range between 500 to 2,000 daltons, making them larger than most classical small molecules while overlapping with some peptides. Due to their size, macrocycles can engage with their targets through multiple spatially distant or rather flat interaction points, making relatively large PPIs accessible (> 2,500 Å^2).[47] In addition, some macrocycles can be cell-permeable and orally available. From a study of around 100 macrocyclic drugs and clinical candidates, Giordanetto and Kihlberg showed that macrocycles just beyond the rule of five space (bRO5) can be membrane-permeable and orally bioavailable (Figure 3).[48] The upper limit was fixed to macrocycles with molecular weight up to 1,000 g/mol and polar surface area up to 250 Å^2. This combination of features makes macrocycles an attractive drug format for generating orally available PPIs inhibitors and targeting intracellular PPIs, which proved difficult to reach by other inhibitor classes (see section 1.1.3). Moreover, the cyclic nature of macrocycles confers them numerous attractive properties, which are discussed below.

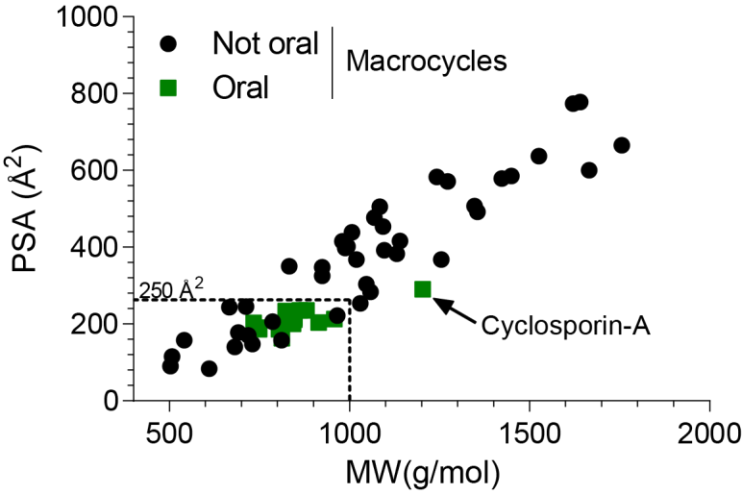


Figure 3. Orally available macrocycles. A representation of macrocyclic-based drugs considering their polar surface area (PSA) and molecular weight (MW). Despite cyclosporin-A, all orally deliverable macrocycles fit below 1,000 g/mol and 250 Å^2. The data were taken from the reference paper [48].

Macrocycles' cyclic shapes give them conformational rigidity. This structural preorganization can increase macrocycle affinity compared to their linear analog.[49] A rigid structure minimizes the energy required to reorganize a molecule in the desired configuration upon binding to a target. Having the molecule already locked in the correct format makes the interaction

energetically more favorable as it minimizes the entropy loss generated by the reorganization of a molecule upon binding.[50] For example, cyclization of a linear molecule could improve its activity against farnesyltransferase by 55,000-fold (Figure 4a).[51] Moreover, due to their rigid structures, macrocycles can be more selective than their linear homologs.[52] Their conformational rigidity minimizes their flexibility to accommodate other binding sites. Hence, cyclizing linear inhibitors can be used to specifically inhibit proteins that might present similar binding sites to other proteins, such as the catalytic sites in enzymes. For example, selective matrix metalloproteinase 8 (MMP-8) inhibitors were obtained by cyclizing an initially non-specific linear precursor (Figure 4b).[53] In addition, the cyclic nature of macrocycles can confer them high protease and metabolic stability compared to their linear analogs. For example, it was shown that cyclizing linear peptides could drastically improve their stability *in vivo*.[54]

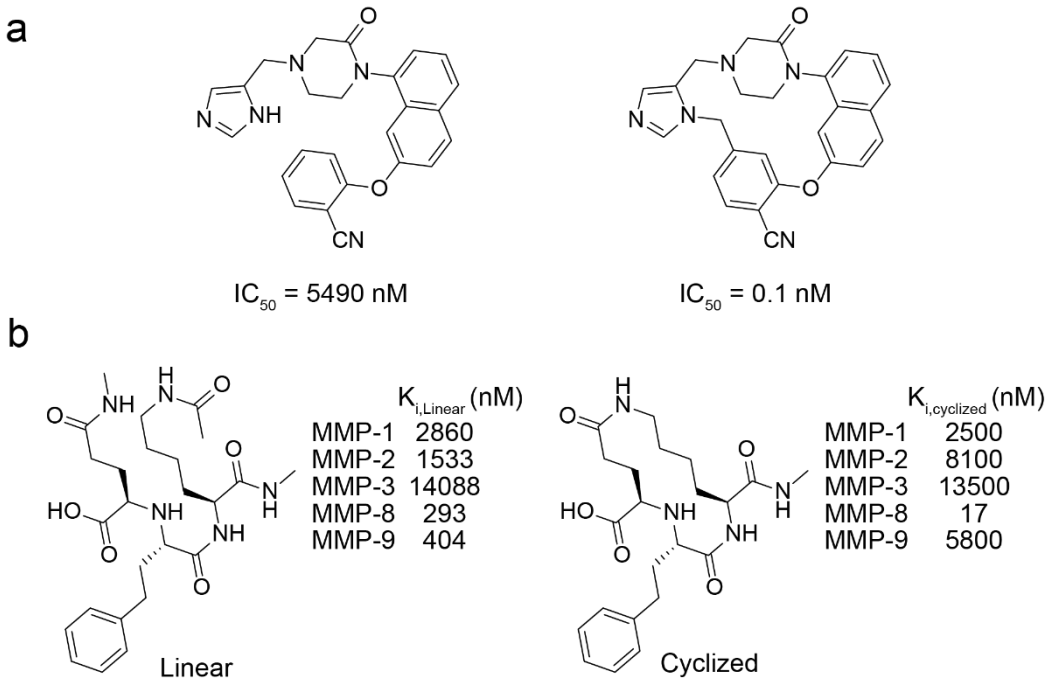


Figure 4. Macrocyclization impact on linear molecules. (a) Improving the potency of farnesyltransferase inhibitors through macrocyclization. Data were taken from the reference paper [51]. (b) Improving selectivity of MMP through cyclization. Data were taken from the reference paper [53].

One key property making macrocycles an attractive drug format is their possibility to be membrane-permeable and orally available. The unusually high membrane permeability of some macrocycles can be explained by the formation of intramolecular H-bonds in the macrocycle ring.[55] During passive diffusion, these interactions can shield some polar groups (H-bond donors and acceptors) from solvent exposure, reducing the three-dimensional exposed polar surface area. In addition, due to their rigidity, macrocycles have only a limited amount of rotatable bonds, which negatively impact passive diffusion.[56] This combination of

properties allows some macrocycles beyond the RO5 to present good membrane permeability.

Macrocycle overall properties are linked to their cyclic structure, defined by a complex system of intramolecular H-bond, hydrophobic and steric interactions, as well as ring strain. Therefore, any change in their backbone can lead to drastic conformational change.[57] A single point modification in the ring can lead to a three-dimensional transmission of structural adjustment, which can affect the macrocycle affinity, selectivity, stability, or permeability. Therefore, when optimization of a macrocycle is considered, numerous cyclic backbones have to be tested to find one that locks the molecule in the desired configuration.

1.2.2 Macrocycle therapeutics

As of 2014, 68 macrocycles are being used for therapeutic purposes, and more than 30 are in clinical development.[48] The vast majority of macrocyclic drugs were derived from natural sources such as bacteria, fungi, animals, or plants. Despite being less represented, synthetic macrocycles have also demonstrated several successes emphasizing their potential. Macrocycles found numerous therapeutical applications, such as antibiotics, immune system modulators, anticancer drugs, or antiviral agents discussed below. However, despite their significant therapeutic potential, macrocycles have been underexploited.[58] This can be explained by their limited discovery due to the lack of large libraries of macrocycles to screen. Additionally, it could be linked to their synthetic difficulties and challenging optimization. Hence, there is much room to expand macrocycles' potential, especially concerning their use to target intramolecular PPIs and to generate orally available PPIs inhibitors.

Several macrocycles have found some applications as antibiotics. Common examples include the two orally available erythromycin and rifampicin, and vancomycin, all isolated from bacterial sources. Erythromycin, isolated from the bacteria *Saccharapolyspora erythraea*, binds the inner surface of the bacteria ribosomal tunnel preventing the exit of nascent protein (Figure 5a).[59] Rifampicin, isolated from the bacteria *Amycolatopsis*, is an allosteric inhibitor of the deoxyribonucleic acid (DNA) dependent ribonucleic acid (RNA) and prevents the elongation of RNA.[60] Rifampicin was found to be a powerful molecule against tuberculosis.[61] Alternatively, vancomycin, isolated from the bacteria *Amycolatopsis orientalis* proved to be an efficient antibiotic against gram-positive bacteria (Figure 5b).[62] Vancomycin binds with the bacterial peptidoglycan precursors, therefore preventing new bacterial wall synthesis.[63]

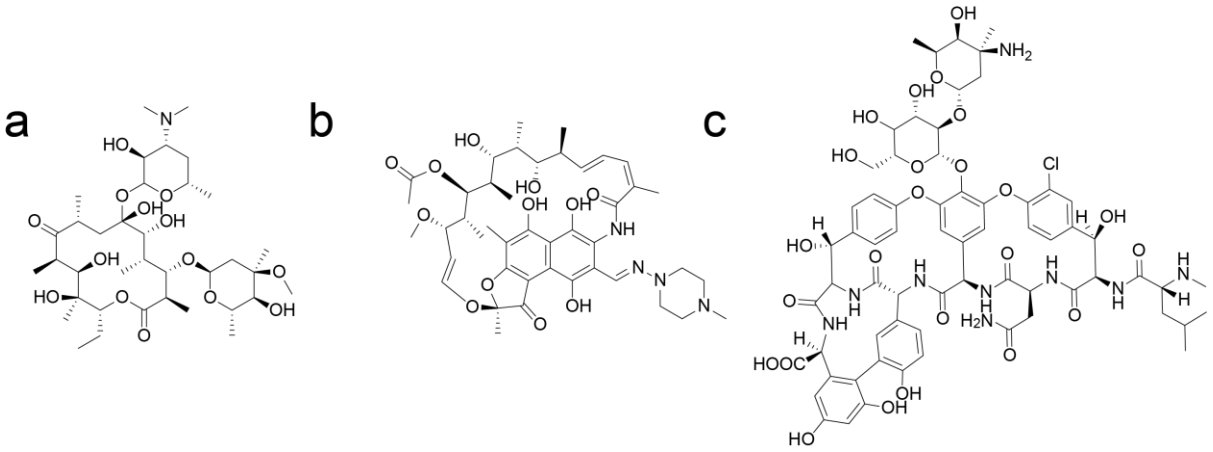


Figure 5. Examples of macrocycles used as antibiotics. (a) Erythromycin. (b) Rifampicin. (c) Vancomycin.

Several macrocycles have found applications as immune system modulators. Perhaps the most common macrocycles include the orally available cyclosporin-A and tacrolimus isolated from fungus and bacterial sources, respectively. Cyclosporin-A was isolated from the fungus *Tolypocladium inflatum* and proved a potent immunosuppressive agent. Its mechanism of action is based on the inhibition of calcineurin, which lowers T-cell activity (Figure 6a).[64] Cyclosporin-A proved helpful in preventing organ transplantation rejection.[65] Alternatively, tacrolimus, isolated from the bacteria *Streptomyces tsukubaensis*, also acts as an immunosuppressive drug. Its mode of action is similar to cyclosporine-A, in which calcineurin activity is inhibited, resulting in the inhibition of T-cell activity.[66]

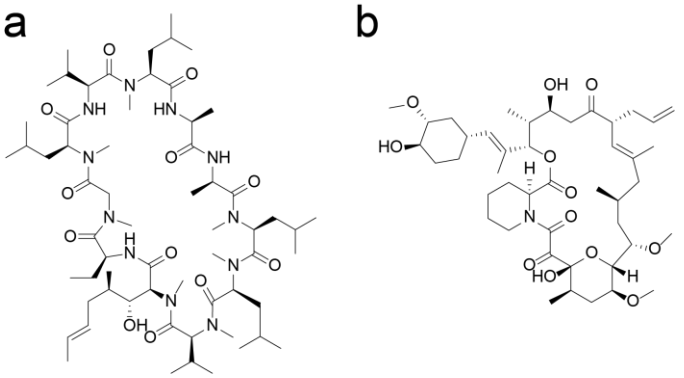


Figure 6. Examples of macrocycles used as immune system modulators. (a) Cyclosporin-A. (b) Tacrolimus.

In addition, macrocycles proved to be efficient for cancer treatment. From natural sources, examples include peloruside A, isolated from a marine sponge, and ixabepilone, a semi-synthetic derivative of epothiline B, isolated from bacteria. Synthetic sources include the two orally available macrocycles pacritinib and zotiraciclib, which are currently being studied in

clinical phase. Peluroside-A, isolated from the marine sponge *Mycale hentschieli* proved to have anticancerous activity (Figure 7a).[67] Its mechanism of action resides in the microtubule stabilization resulting in arrested tumor cell division and apoptosis. Similarly, ixabepilone, a semi-synthetic analog of epothilone B, a natural product identified from myxobacterium *Sorangium cellulosum*, also proved to have anticancerous activity by stabilizing microtubules (Figure 7b).[68] Alternatively, pacritinib, a pyrimidine-based synthetic macrocycle, showed activity against cancer by inhibition of Janus Kinase-2 (JAK2) and fms-like tyrosine kinase-3 (FLT3).[69] Zotiraciclib, another pyrimidine-based macrocycle, showed a similar effect by inhibiting JAK2, FLT3, and cyclin-dependent kinases (CDKs).[70]

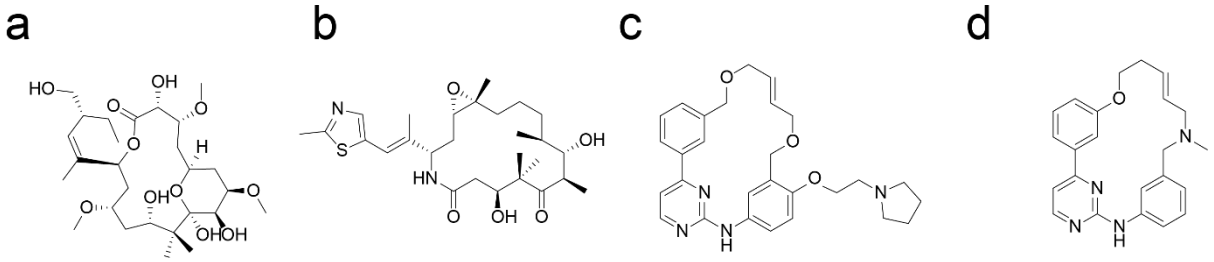


Figure 7. Examples of macrocycles used as anticancer drugs. (a) Peloruside A. (b) Ixabepilone. (c) Pacritinib. (d) Zotiraciclib.

Some macrocycles were also used as an antiviral drug. Typical examples include the two orally available synthetic macrocycles simeprevir and vaniprevir (Figure 8a & 8b). Both macrocycles showed antiviral activity against hepatitis C virus (HCV).[71,72] Their mode of action is based on the inhibition of NS3/4A, a protease essential for HCV viral replication.

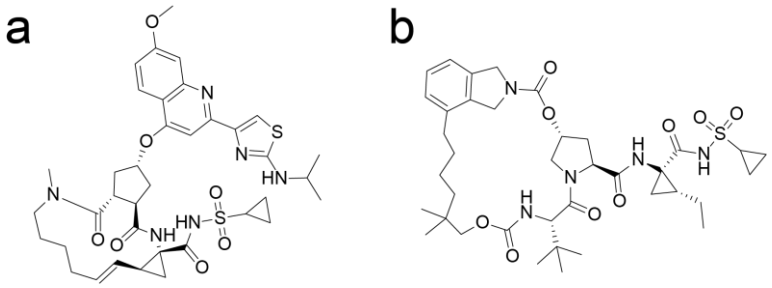


Figure 8. Examples of macrocycles used as antiviral drugs. (a) Simeprevir. (b) Vaniprevir.

1.3 Methods for generating macrocyclic inhibitors

1.3.1 Screening macrocyclic compounds

A few libraries of purified macrocycle libraries are available for purchase in the market (see Table 3). These macrocycles can be tested individually in microwell plates against PPIs using traditional biochemical assays (discussed in section 1.4). Some macrocycle libraries focus on macrocycles with the potential to be cell-permeable and orally available, therefore making them potentially applicable for intracellular PPIs. For example, ChemBridge, Asinex, Enamine, and Polyphor (with MacroFinder) synthesized macrocycles with a molecular weight below 1,000 Da.[73–76] Alternatively, some libraries have larger macrocycles sizes (1,000 to 2,000 Da), which will likely restrict their application to extracellular targets, such as the library PEMfinder from Polyphor.[76] Despite presenting interesting features, all these libraries are somewhat limited in size. This limitation could be explained by the difficulty of synthesizing and purifying tens of thousands of individual compounds, which can be time, labor, and resource-intensive.

Table 3. Commercially available macrocycle libraries.

Company	Library size	Library description	Reference
ChemBridge	13,000	< 800 Da, 11-27 atoms ring size, focus on potential permeable macrocycles	[73]
Asinex	10,091	< 600 Da focus on potential permeable macrocycles (many tested by PAMPA)	[74]
ChemDiv	2,335	14-22 atoms ring size, membered macrocyclic (cyclized by lactamization or by click chemistry)	[77]
Polyphor	50,000	MacroFinder (500 to 1,000 Da), PEMfinder (1,000 to 2,000 Da)	[76]
Enamine	592	< 500 Da, cyclized by click chemistry, macrolacton/lactamization, ring closure metathesis)	[75]

1.3.2 One-bead-one-compound (OBOC) libraries

The OBOC synthesis methodology allows for the construction of large libraries of macrocycles on resin, each bead containing many copies of the same compound. In this strategy, linear precursors, such as peptides or peptoids, are combinatorially synthesized on beads following the split-and-pool synthesis methodology, followed by their cyclization on resin (Figure 9a).[78] The split-and-pool methodology is a step-wise process based on the separation of a defined amount of resin in several reaction vessels, which undergoes couplings with different building blocks before pooling the resin back together. This principle performed multiple times allows the exponential growth of linear precursors, with each bead containing many copies of a single compound. The cyclization of the numerous linear precursors can then be achieved on beads following several strategies. For example, it can be performed by coupling the N-terminal amine with the side chain of a C-terminal glutamic acid or with the C-terminus of a glutamic acid immobilized through its side chain.[79,80]

The screening of the macrocycle library can be performed all at once by incubating the entire library with a target of interest. Hit identification can be obtained following a colorimetric approach, in which positive hits are identified by a striking change of color.[78] For example, the target can be fluorescently labeled with a dye or fused with a fluorescent protein.[80] The structure of the hit is traditionally performed by mass analysis.[81] As the determination of cyclic structures by mass can be challenging, each bead usually contains several copies of the uncyclized linear precursors as encoding tags. These linear species are then easier to analyze, such as using partial Edman degradation-mass spectrometry (PED-MS), often used for the sequential truncation and mass analysis of linear peptides.[82] These linear species can be obtained by using different protecting groups for the building block undergoing the cyclization, one of these protecting groups preventing the final cyclization reaction. For example, for a head-to-tail cyclization using a C-terminal glutamic acid, using 95% Alloc/5% tert-Butyl protected glutamic acid will yield 5% of the uncyclized compound.[83] Alternatively, technics were also developed to spacially segregate the beads into outer and inner layers, allowing a difference in reactivity to specifically introduce building blocks undergoing the cyclization in the outer layer while introducing another building block preventing the cyclization in the inner core.[84] As a result, the inner core of each bead would contain the corresponding linear encoding sequence, and the outer layer would contain the macrocycle.

OBOC methodology proved to be powerful in identifying macrocyclic inhibitors. For example, OBOC was used for the generation of nanomolar inhibitors of K-RAS(G12V)/effectors and TNFα/TNFα-R from libraries of 6×10^6 and 3×10^5 macrocycles, respectively (Figure 9b & 9c).[80,85]

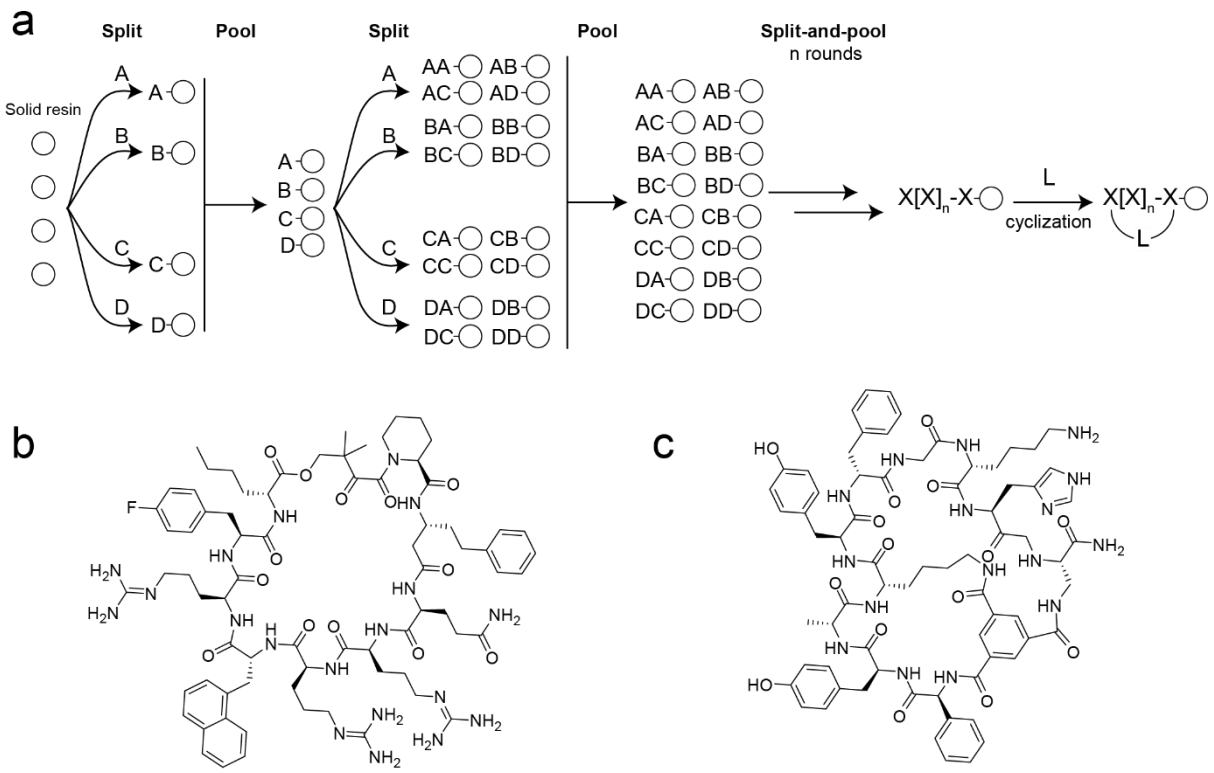


Figure 9. Generation of macrocyclic compounds using OBOC. (a) Split-and-pool synthesis principle of linear precursors followed by their cyclization. (b) Nanomolar macrocyclic inhibitor (K_D of 830 nM) against K-RAS(G12V)/effectors found using OBOC library. (c) Nanomolar macrocyclic inhibitor (K_D of 450 nM) against TNFα/TNFα-R found using OBOC library.

1.3.3 Phage display libraries

Phage display allows the generation of large libraries of peptides that can then be cyclized at a later stage. In this technique, exogenous peptide sequences are transfected into filamentous phages.[86] These peptides are then displayed on the phage surface as a fusion with coating proteins. Cyclization of the peptides can be performed by several techniques. For example, cyclic peptides can be obtained through the oxidation of two cysteines present in each peptide (Figure 10a).[87] Alternatively, cyclization of the peptides could also be reached by exploiting the unique nucleophilicity of the cysteine side chains. For example, the generation of bicyclic peptides was performed with peptides containing three cysteines, which reacted with one tris(bromomethyl)benzene linker (Figure 10b).[88] Similarly, libraries of bicyclic peptides were also generated with peptides containing four cysteines reacting with two bis-electrophile linkers.[89]

The library of cyclic peptides on phage is screened all at once against an immobilized target. After affinity selection, the peptide sequences that could bind to the desired target can be discovered through bacterial infection, amplifying the corresponding phages, and sequencing their DNA.[90] Therefore, the phage acts as a tag used for sequence identification of the peptide.

Despite being mainly limited to natural amino acids, phage display proved to be a powerful methodology to discover macrocyclic binders. For example, it was used for the generation of nanomolar binders against Fc/ protein A and against plasma kallikrein from libraries of 4×10^9 and 4.4×10^9 macrocycles, respectively (Figure 10c & 10d).[87,88]

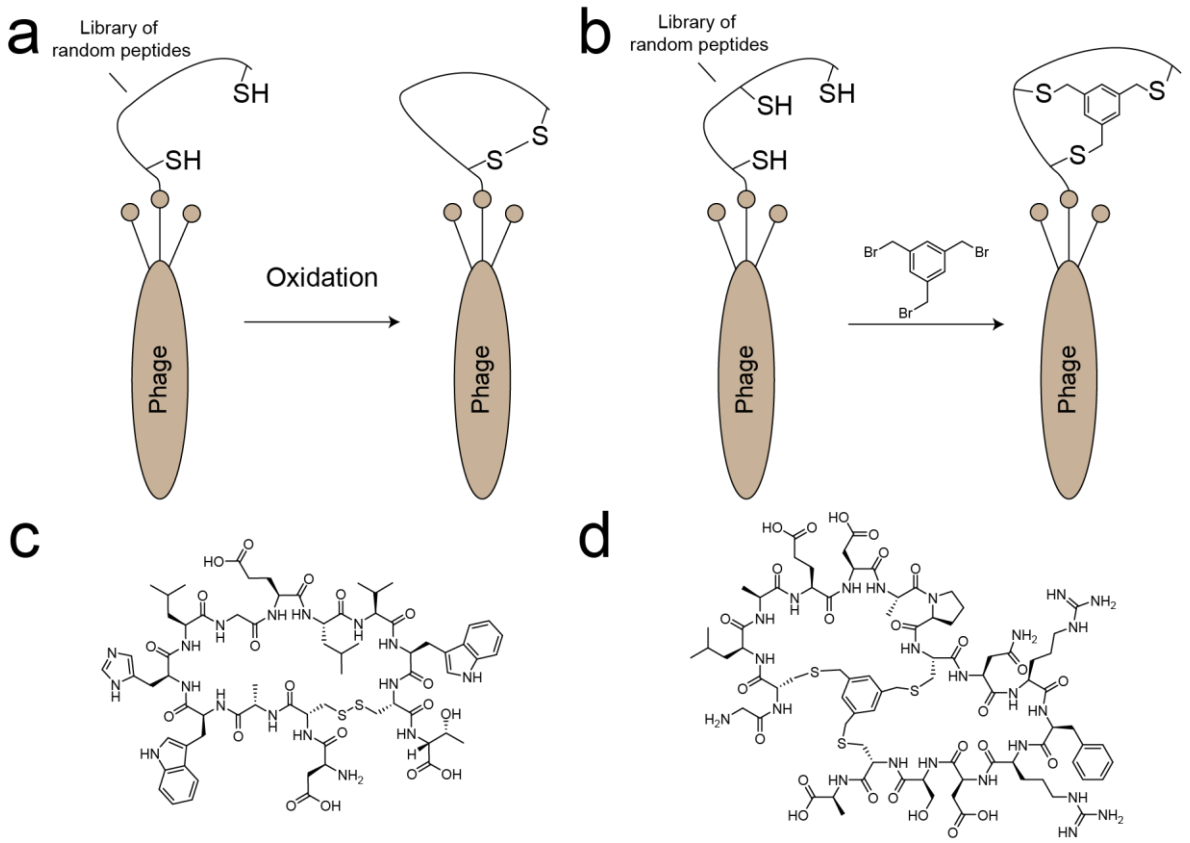


Figure 10. Generation of macrocyclic compounds using phage display. (a) A library of random peptides containing two cysteines expressed on the coat protein of phages. The cyclization is performed by oxidation of the two cysteines. (b) A library of random peptides containing three cysteines expressed coat protein of phages. Cyclization is performed through the reaction of the three cysteines with tris(bromomethyl)benzene. (c) Nanomolar inhibitor (K_D = 25 nM) against Fc/protein A inhibitor found using phage display. (d) Nanomolar inhibitor (IC_{50} = 1.5 nM) against Plasma Kallikrein found using phage display.

1.3.4 mRNA display libraries

mRNA (messenger RNA) display allows the generation of numerous linear peptides that can then be cyclized. This methodology involves the *in vitro* transcription and translation of a DNA library coding for the peptides.[91] The central feature of this method is that each peptide is covalently attached to the 3' end of its own mRNA through puromycin and a short DNA linkage. Because mRNA display is based on a cell-free system, unnatural amino acids can be introduced more easily than with phage display. For example, this can be performed using the

flexible *in vitro* translation system (FIT), which uses artificial ribozymes (flexizyme) to efficiently load tRNA with non-canonical amino acids, which are tolerated by the ribosomal translation.[92] Cyclization of the peptides can be achieved with several techniques. For example, cyclic peptides using two cysteines reacting with bis-electrophile linkers can be performed (Figure 11a).[93] Alternatively, peptides can be cyclized by the reaction between an amino acid bearing a 2-chloroacyl group at the N-terminus and one cysteine side chain at the C-terminus of the peptide (Figure 11b).[94] Spontaneous intramolecular cyclization generates the desired macrocycles.

The library of cyclic peptides linked to their mRNA can be screened all at once against an immobilized target. After affinity selection, the sequence of the peptide that could bind the target can be obtained by reverse transcription of the mRNA, amplification by PCR, and sequencing.[95]

mRNA display is a powerful methodology to generate macrocyclic compounds. For example, it was used to discover nanomolar binders against thrombin and E6AP/p53 from libraries of 2×10^{13} and 6×10^{12} macrocycles, respectively (Figure 9a & 9b).[93,94]

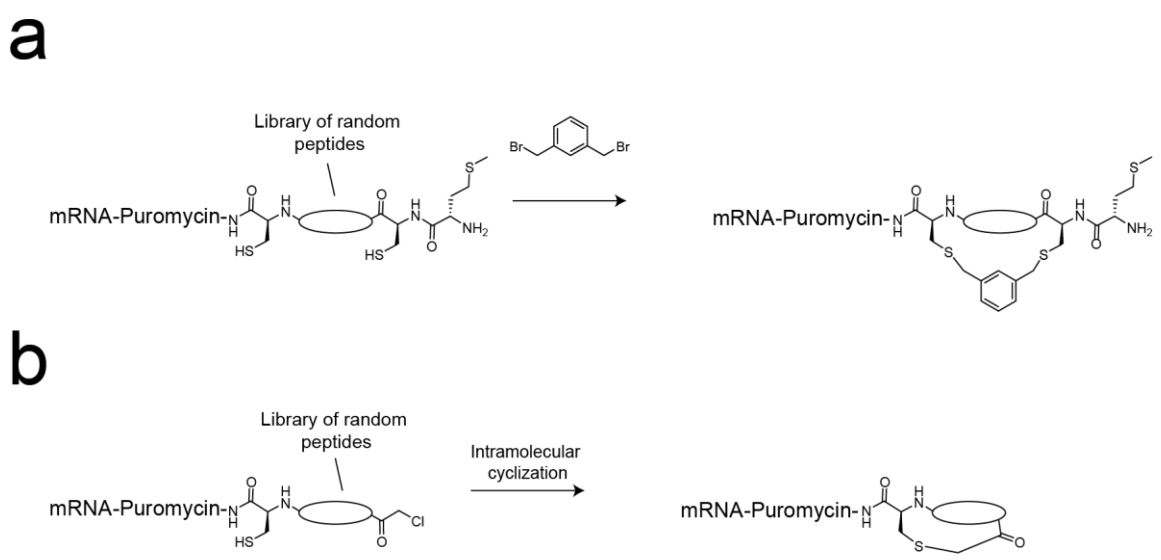


Figure 11. Generation of macrocyclic compounds using mRNA display. (a) A library of random peptides flanked with two cysteines for late-stage cyclization using bis-electrophile linker (dibromoxylene). (b) A library of random peptides with the incorporation of an N-terminal chloro-acetyl and a C-terminal cysteine used for intramolecular cyclization. Figure 11 continues next page.

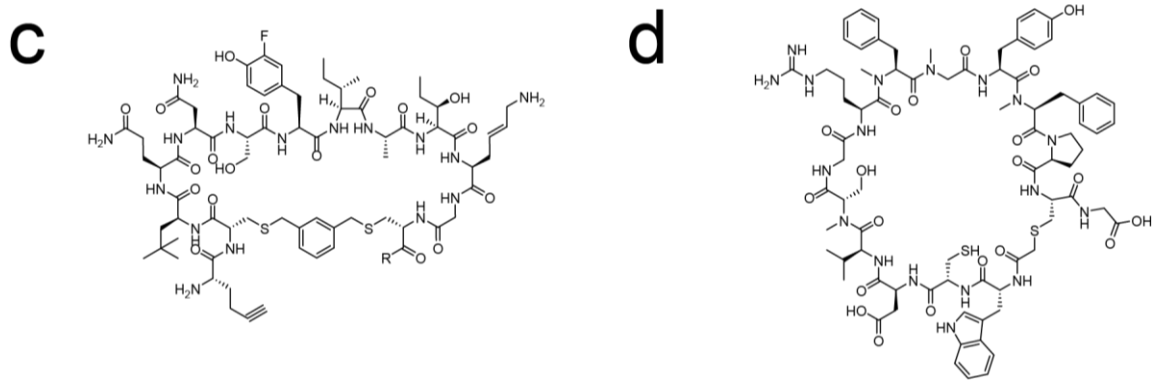


Figure 11 continued. (c) Nanomolar (K_D = 4.5 nM) thrombin macrocyclic binder obtained using mRNA display. R = linker. (d) E6AP/p53 subnanomolar (K_D = 0.6 nM) macrocyclic inhibitor found by mRNA display.

1.3.5 DNA-encoded libraries

DNA-encoded libraries are synthetic methodologies allowing the synthesis of numerous macrocycles. They can be separated into two main categories: DNA-templated synthesis (DTS) and DNA-recorded synthesis.[96]

DTS is based on the mediation of DNA sequences to bring together reactants from complex mixtures allowing their efficient reaction by close proximity (Figure 12a).[97] The methodology requires a DNA template comprising all coding sequences for each library member and corresponding oligonucleotide sequences attached by cleavable linkers to reactants. Library synthesis is then achieved in a step-wise fashion. Briefly, the DNA templates are mixed with the first set of complementary oligonucleotides bearing different reactants, which are then chemically attached to the template. The linker between the reactant and the short oligonucleotide is then cleaved, and this process is repeated several times until cyclization of the library is performed. As a result, a library of macrocycles linked to their corresponding DNA templates is obtained. For example, pentapeptide macrocycles were synthesized by DTS followed by a final cyclization via triazole formation under Huisgen conditions (i.e, alkyne azide click chemistry).[98] Alternatively, tetrapeptides were synthesized followed by Wittig reaction for cyclization.[99]

DTS macrocycle libraries can be screened all at once against an immobilized target. After affinity selection, the macrocycles structures can be revealed upon PCR-amplification followed by sequencing of the corresponding DNA. The sequences obtained then reveal which building blocks were used in the macrocycle assembly.

DTS methodology proved to be capable of generating macrocyclic binders. For example, nanomolar inhibitors of XIAP BIR3 domain and insulin-degrading enzyme were discovered from libraries of 160,000 and 256,000 macrocycles, respectively (Figure 12b & 12c).[98,99]

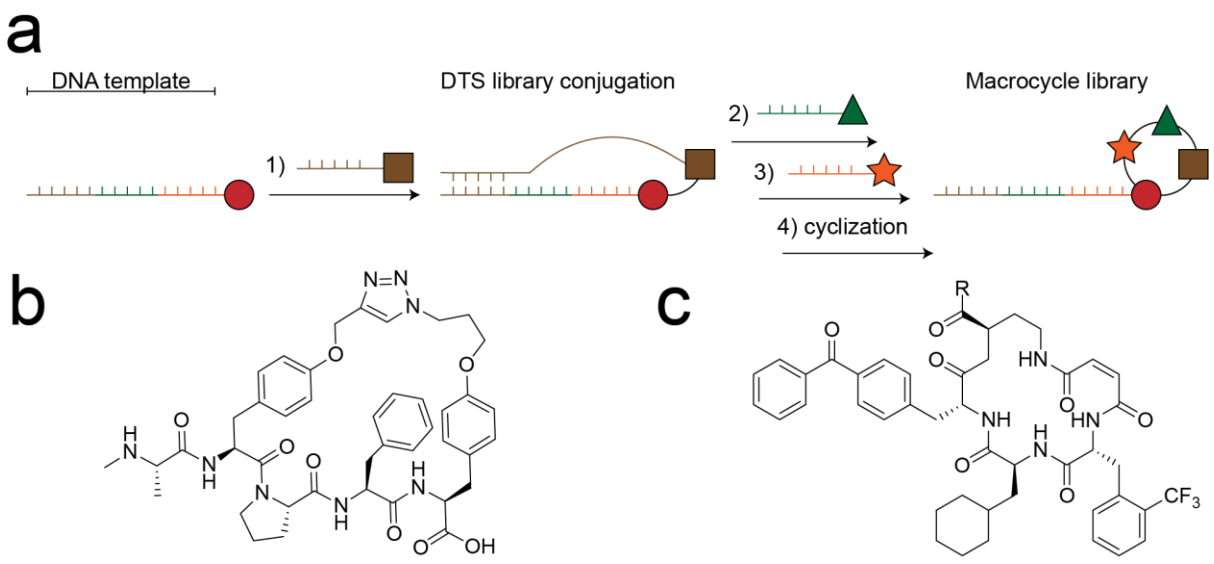


Figure 12. Generation of macrocyclic compounds using DNA-template synthesis. (a) DTS general principle. Building blocks react by close proximity through the binding with the DNA encoding template. (b) Nanomolar macrocyclic inhibitor (IC_{50} of 366 nM) found using DTS macrocycle library against XIAP BIR3 domain. (c) Nanomolar macrocyclic inhibitor (IC_{50} of 40 nM) found by found using DTS macrocycle library against the insulin-degrading enzyme. R = $NH(CH_2)_2O(CH_2)_2NH_2$.

The other type of DNA-encoded libraries can be called DNA-recorded libraries. Their assembly is similar to the OBOC library, as they follow the combinatorial synthesis by split-and-pool (see section 1.3.2). The difference resides in the addition of a coding DNA sequence used as a tag for each building block used for the macrocycle assembly. These libraries initially require a set of compounds attached to the chemically modified extremity of a universal oligonucleotide. The other building blocks are then added to the first set of compounds by split-and-pool, followed by their encoding by a unique corresponding DNA sequence to the remote end of the universal oligonucleotide (Figure 13a). The addition of building blocks and encoding DNA sequences is performed step-wise until the desired size is reached. The final step traditionally involves a cyclization yielding the desired macrocycles. This library assembly can be performed in solution or on beads.[100,101] For example, OBOC DNA-encoded library principle was used for the generation of a library of macrocycles through the construction of linear peptoid-inspired conformationally constrained oligomers cyclized in a final step by thioether formation.[102] Alternatively, this synthesis strategy was used in solution for the synthesis of a library of macrocycles containing three amino acids and one last building block used for cyclization via triazole formation under Huisgen conditions (i.e alkyne azide click chemistry).[103]

The screening of the macrocycle library can be performed all at once by incubating the entire library with a target of interest. After affinity selection, the macrocycles hits structures can be revealed upon amplification and sequencing of their encoding DNA. The sequences obtained then reveal which building blocks were used in the macrocycle assembly.

DNA encoded library proved to be capable of generating macrocyclic binders. For example, nanomolar binder of streptavidin and micromolar inhibitors of human serum albumin were discovered from libraries of 580,000 and 1,254,838, respectively (Figure 13b & 13c).[102,103]

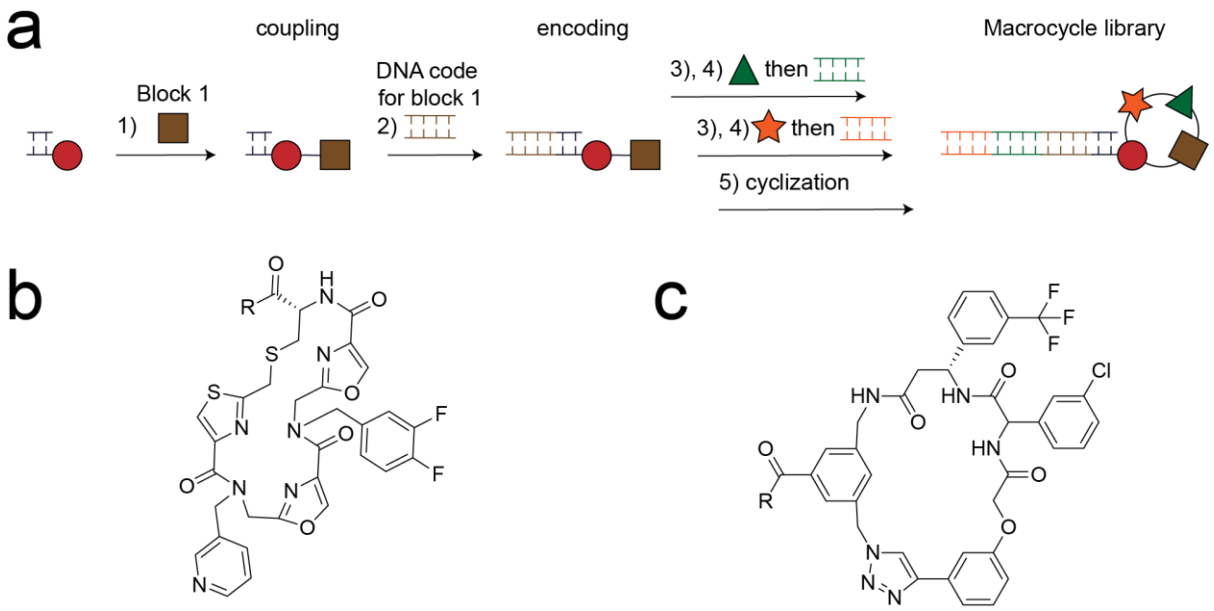


Figure 13. Generation of macrocyclic compounds using DNA-recorded library. (a) DNA-recorded library general principle. Building blocks are sequentially added, followed by their encoding by unique DNA fragments. (b) Nanomolar macrocyclic inhibitor (K_D around 30 nM) found using a DNA-recorded macrocycle library against streptavidin. R = Resin-linker. (c) Micromolar macrocyclic inhibitor (K_D of 1.9 μM) found by using a DNA-recorded macrocycle library against human serum albumin. R = Fluorescein-linker.

1.3.6 Combinatorial synthesis and screening of crude macrocycle libraries

Our lab proposed several approaches involving the combinatorial synthesis of macrocycle libraries with a focus on sub-kDa macrocycles, having the potential to be cell-permeable and orally available. The HTS of the crude mixtures of each macrocycle, being a unique combination of building blocks involved in the synthetic strategy, was then performed using traditional biochemical assays. Three combinatorial synthesis approaches were developed and are discussed below.

First-generation involves the cyclization of *m* di-, tri- or tetrapeptide with *n* bis-electrophile linkers allowing the generation of *m* × *n* macrocycles (Figure 14a).[104] The peptides contain

one thiol-containing amino acid located at the C-terminus, and a free amine from the N-terminal amino acid, allowing their purification by ether precipitation. Other amino acids are designed to be diverse by bringing various side chains or different backbone structures.[104] Nanomolar inhibitors of trypsin and thrombin were generated from libraries of 1,176 and 1,284 macrocycles, respectively (Figure 14b & 14c).

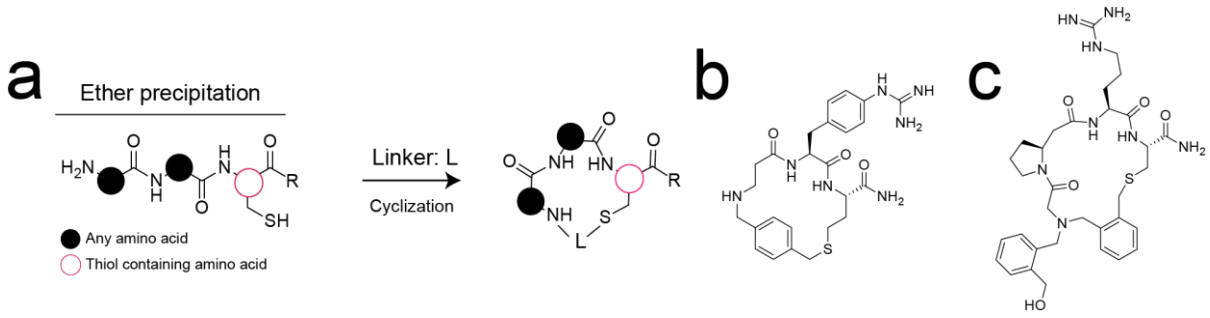


Figure 14. First-generation strategy to synthesize libraries of macrocyclic compounds. (a) General principle of the first-generation strategy. Ether precipitated peptides are cyclized with several linkers L. R = GGW. (b) Nanomolar inhibitor (K_i = 280 nM) discovered against trypsin using the first generation strategy. (c) Nanomolar inhibitor (K_i = 59 nM) discovered against thrombin using the first-generation strategy. It corresponds to a side product.

A second-generation was then developed to increase the chemical diversity of the synthesized macrocycle library. It is characterized by the step-wise addition and reaction of *m* linear peptides with *n* primary amines and *o* linkers, allowing the generation of $m \times n \times o$ macrocycles.[105] In this methodology, the linear peptides are composed of an N-terminal bromoacetamide, one or several amino acids, and one C-terminal thiol-containing amino acid with a removable side-chain protecting group. All linear peptides are HPLC-purified. The first diversification involves the reaction of the primary amines with the bromoacetamide to generate a secondary amine. Then, the thiol protecting group is released by reduction. The resulting linear peptides containing a free thiol and a secondary amine react with different bis-electrophile linkers used for cyclization (Figure 15a). This methodology was used to synthesize 3,780 macrocycles tested against thrombin to improve a macrocyclic hit previously obtained with the first-generation strategy. It allowed an improvement in binding by 5.5-fold against thrombin (Figure 15b).[105]

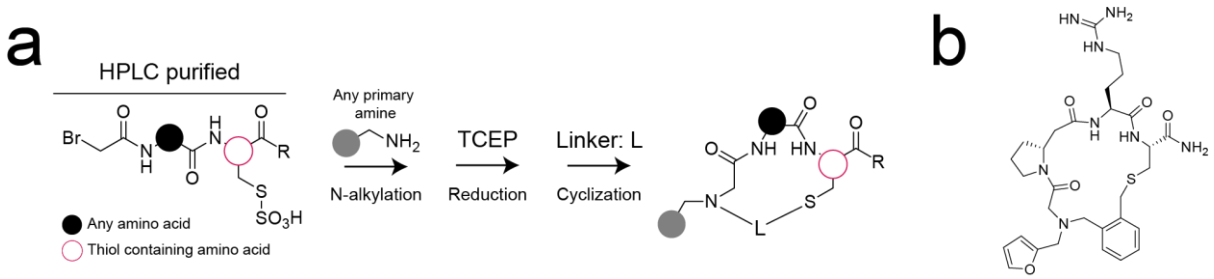


Figure 15. Second-generation strategy to synthesize libraries of macrocyclic compounds. (a) General principle of the second-generation strategy. HPLC-purified peptide containing N-terminal bromoacetamide and a C-terminal protected thiol are first diversified with primary amines, followed by the removal of the thiol protecting group by reduction and cyclized with bis-electrophile linkers L. R = GW. (b) Nanomolar inhibitor (K_i = 10.6 nM) discovered against thrombin using the second-generation strategy. It was used to improve a thrombin inhibitor previously discovered (see Figure 14c).

A third-generation synthetic strategy was developed to increase the synthesis throughput by removing any purification steps required in the preparation of the linear peptides. It involves the synthesis of *m* linear peptides containing thiol group at the N and C-terminus, the latter being immobilized on a thiol resin through a disulfide bond (Habeshian *et al.*, not published at this time). After synthesis, all amino acid protecting groups are removed in acidic conditions while the peptides stay attached to the resin (Figure 16a). From there, two options are possible to generate macrocycles. First, *m* high-quality disulfide macrocycles can be obtained in solution by cyclitive release using basic conditions, such as using DABCO (1,4-diazabicyclo[2.2.2]octane) in DMSO (Habeshian *et al.*, not published at this time). If one of the amino acids is a diamino acid, a peripheral free amine is generated after acidic cleavage of the protecting group. From there, the *m* disulfide macrocycles can be diversified with *n* carboxylic acids yielding *m* × *n* macrocycles (Option 1 in Figure 16 a). Alternatively, *m* high-quality dithiol linear peptides can be obtained in solution by reducing the disulfide bond attaching them to the resin using 1,4-butanedithiol (BDT) as reducing agents in DMF (Bognár *et al*., not published at this time). Then, these *m* linear peptides can be cyclized with *n* linkers yielding *m* × *n* macrocycles (Option 2 in Figure 16a). The former strategy (Option 1 in Figure 16a) was used to generate nanomolar macrocyclic inhibitors of thrombin and MDM2/p53 from libraries of 4,608 and 19,968 macrocycles respectively (Figure 16b & 16c) (Habeshian *et al.*, not published at this time). The later strategy (presented as option 2 in Figure 16a) is rather new. Therefore, no examples can be presented. A combination of both options (Option 1 and 2, Figure 16a), in which *m* linear peptides containing thiol group at the N and C-terminus and a free amine, cyclized with *n* bis-electrophile linkers followed by peripheral diversification with *o* carboxylic acids, yielding *m* × *n* × *o* macrocycles, was also developed. However, it hasn't been extensively used due to the resulting crude mixture complexity.

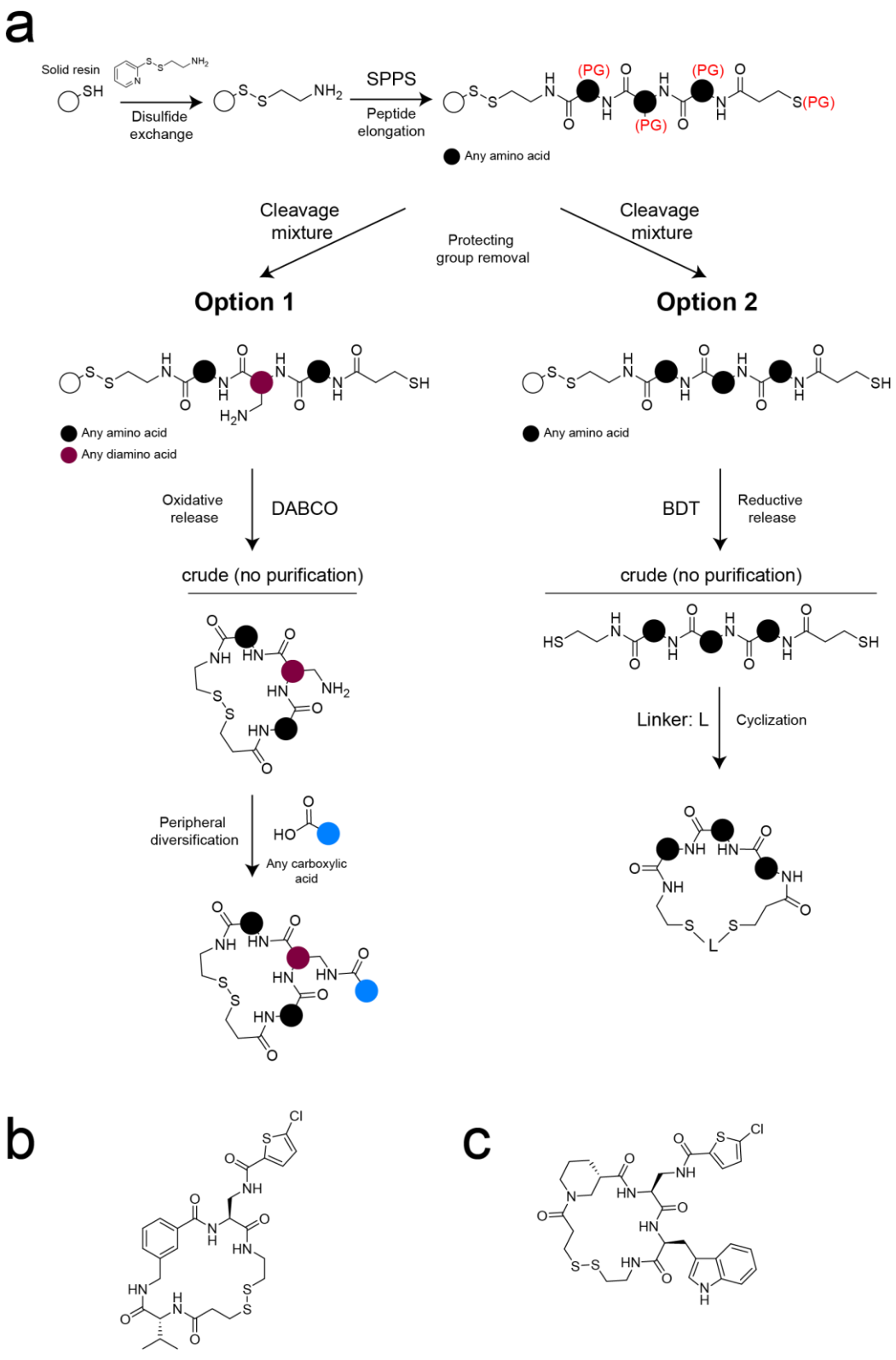


Figure 16. Third-generation strategy to synthesize libraries of macrocyclic compounds. (a) Peptides containing a thiol at the N- and C-terminus, the later immobilized on a thiol-containing resin through a disulfide bond, are synthesized on resin prior on-resin protecting group. Option 1: oxidative release of disulfide macrocycle followed by their diversification with carboxylic acids (in blue). Option 2: reductive release of the peptides followed by their cyclization with bis-electrophile linkers L. (b) Nanomolar thrombin inhibitor (K_i = 44 nM) discovered using the third-generation strategy option 1. (c) Nanomolar MDM2/p53 inhibitor (K_i = 650 nM) discovered using the third generation strategy option 1.

1.4 Assays for screening inhibitors of protein-protein interactions

1.4.1 General considerations

In HTS of individual compounds, compatible assays mimicking the targeted PPIs are required. They should offer a strong signal to be miniaturized in low volume assay plates, such as in 384-well or 1,536-well plates. These assays should be robust and reliable, in which both controls, representing the minimum and maximum of the signal, should be stable. Hence, the complexity of the assay should be minimized. In addition, the assay window, characterized by the difference in signal between the positive and negative controls, should be sufficient to have a clear differentiation between bound and inhibited PPIs. Sufficient assay windows and robust signals are essential parameters for the clear identification of any potential hits. Moreover, as large libraries of compounds are being tested, the assay should allow a rapid readout of the interaction of interest. End-point mix-and-read assays, which avoid multiple incubation/wash cycles, afford the fastest throughput to screen compound libraries against PPIs.[106] These include fluorescence polarization-based, FRET-based, TR-FRET-based, and bead-based assays (like AlphaScreen). As HTS of individual compounds is usually an expensive approach, the cost-per-well to monitor PPIs can also be a factor to consider when choosing an assay format.

1.4.2 Fluorescence polarization-based assays

Fluorescence polarization (FP) is a convenient approach to monitor PPIs. It traditionally involves one protein and a fluorescent probe mimicking the protein partner. FP assays principle is based on the observation that when a fluorescently labeled molecule is excited by polarized light, it emits light with a different degree of polarization depending on its bound state (Figure 17).[107] This difference in polarization is linked to the degree of freedom of the fluorescent probe. When the probe is free in solution, it has a high molecular tumbling. Therefore, upon excitation with linear polarized light, the resulting parallel and perpendicular emission will be similar, and no polarization will be observed. Upon binding to a large molecule (like a protein), which rotates at a slower rate, the degree of freedom of the fluorescent probe will decrease. Hence, once excited with linear polarized light, the parallel and perpendicular emissions will be different, resulting in an increase in polarization.[107] The PPI disruption by an inhibitor will result in the release of the probe in solution leading to a decrease of polarization.

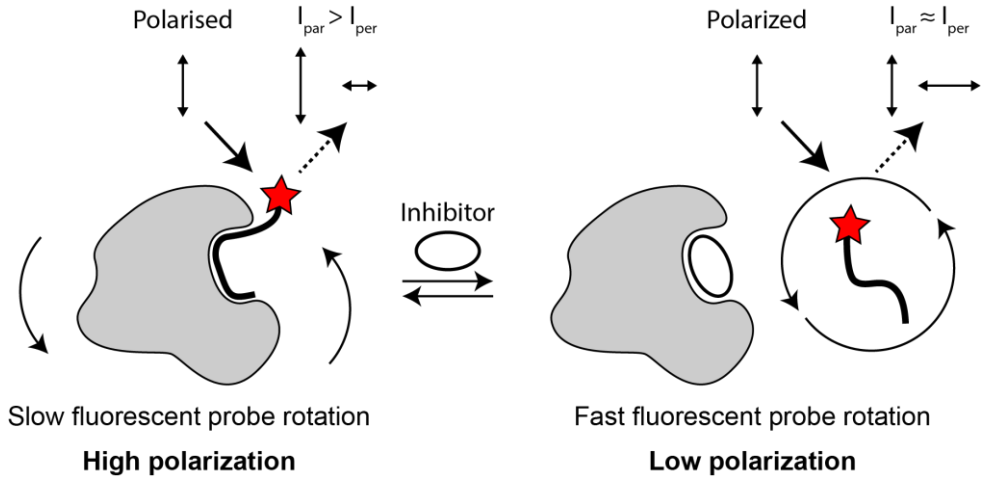


Figure 17. Fluorescence polarization (FP) assay principle. Upon binding to a large molecule, a fluorescent probe excited by polarized light would show a difference between the emitted parallel (I_{par}) and perpendicular (I_{per}) fluorescence intensities resulting in a polarization of the signal.

As peptides derive from the protein region involved in the binding to its protein partner, it is often used as a fluorescent probe, using a fluorophore covalently attached to the peptide. [108] For readout measurement, this assay requires a specific FP module linked to the absorption and emission of the probe used. FP assay total cost per well is usually quite low, making it an accessible assay format for PPI monitoring.[109] Despite being limited to systems whose size difference is large enough, FP assays proved practical in monitoring PPIs.[110]

1.4.3 FRET-based assays

Föster (or Fluorescence) resonance energy transfer (FRET)-based assays can be used to monitor PPIs. It traditionally involved one protein being labeled with a donor molecule and the protein partner labeled with an acceptor molecule. Upon binding of the two proteins, an increase in the FRET signal can be observed. FRET is a phenomenon based on the non-radiative energy transfer between an acceptor and donor molecules (Figure 18a). For the donor and the acceptor to be compatible FRET pairs, the donor's emission spectrum has to overlap with the absorption spectrum of an acceptor molecule, traditionally being a fluorophore (Figure 18b).[111] As a result, the acceptor molecule emits red-shifted fluorescent light, while the fluorescent emission of the donor decreases. The FRET signal is then defined by the ratio between the acceptor's and the donor's emitted light. FRET efficiency strongly depends on the distance between the donor and acceptor and is usually limited from 1 to 10 nm.[112] Thus, when one labeled protein (donor) binds to another one (acceptor), an increase in the FRET signal can be monitored (Figure 18c). If one of these proteins is competed off by an inhibitor, the distance between the two proteins will increase, resulting in a decrease of FRET signal.

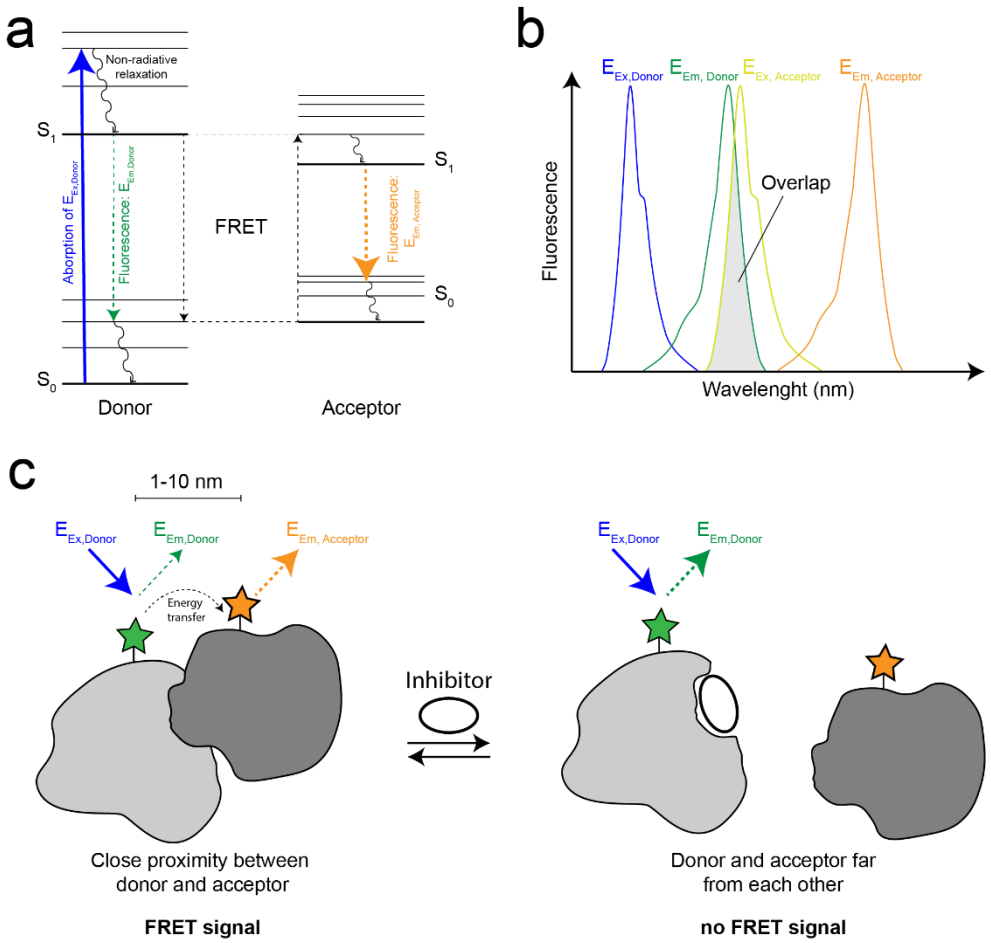


Figure 18. FRET principle to monitor PPI. (a) Representation of Jablonski diagram. A donor molecule is excited to a higher energy level (blue arrow). Non-radiative relaxation shifts the energy downhill and causes a redshift of final fluorescence (green arrow). In close proximity, some energy is transferred from the donor to the acceptor molecule, which will then emit red-shifted light (orange arrow). (b) Excitation and emission spectrums of representative donor and acceptor molecules. For FRET to happen, the emission of the donor molecule should overlap with the excitation of the acceptor molecule. (c) Application of FRET to monitor PPI (in gray and dark gray). In close proximity, the FRET signal can be observed. Upon inhibition, the interaction splits apart, resulting in the decrease of the FRET signal**.**

FRET pair labeling is often performed using fluorescent fusion proteins (such as CFP/YFP).[113] It can also be performed using fluorescently labeled antibodies targeting orthogonal tags (such as GST, His_6, or FLAG-tag), protein covalently modified, or a combination of these systems.[114] Peptides could also be used to mimic one of the two proteins, in which case the labeling is performed by covalent linkage of the donor/acceptor.

The FRET-based assay requires a fluorescent readout supported by most plate readers. Therefore, it is usually relatively low cost and affordable. Moreover, the FRET signal is quite robust, making it an interesting choice for monitoring PPIs. However, one downside is regarding the rather high excitation light required, which can excite the biological reagents or tested compounds and lead to potential false positives readout.[109] For this reason, an

improved format called time-resolved FRET (TR-FRET) was developed to limit these effects and has often been privileged over FRET for HTS against PPIs (see section 1.4.4).

1.4.4 TR-FRET-based assays

Time resolve FRET (TR-FRET) can be used to monitor PPIs. The principle of the assays is similar to the one for FRET discussed in section 1.4.3 but was designed to limit some false positive readouts that can be observed with FRET due to unspecific excitation of tested compounds or biological components. The PPI monitoring principle is the same as FRET, in which an increase of FRET signal is observed upon binding of one protein labeled with a FRET-donor to another protein labeled with a FRET-acceptor. Conversely, a decrease in the FRET signal is observed when an inhibitor can compete off one of the two proteins. The difference resides in the donor, europium or terbium chelates, characterized by exceptionally long fluorescent decay.[115] This characteristic enables reading the FRET signal after a time delay (usually several μs), which allows the fluorescent decay of unspecific excited molecules, typically with ns half-lives (Figure 19).[116] Thus, only specific proximity-based excitation of the acceptor will be observed during the measurement window.

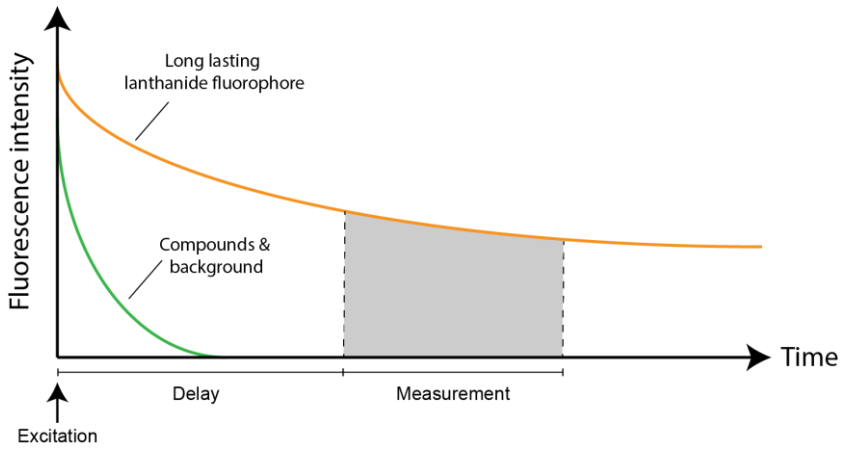


Figure 19. TR-FRET delay principle. Representation of the fluorescence intensity of long-lasting lanthanide fluorophores compared to compounds and background fluorescence. Measurement of the FRET signal is performed only after a delay allowing the decay of undesired fluorescence.

Common TR-FRET pairs involve Europium or Terbium chelates being excited at 340 nm, which emits at 620 nm, and red fluorescent dyes (such as Cy5), which emits at around 665 nm. Additionally, Terbium chelates can also be paired with green acceptors such as fluorescein or GFP due to an emission peak centered at 490 nm. The labeling system is similar to FRET, in which the protein or peptide of interest are labeled with the fluorescent pair either

covalently or using orthogonal tags that can be recognized by labeled ligands (such as antibodies, Protein-A, or Biotin).[117]

The TR-FRET-based assays require a specific module that allows time-resolved methods. Despite being usually more cost-intensive than FP-based or FRET-based assay, TR-FRET is often used to screen compound libraries against PPIs due to its reduced sensitivity toward background fluorescence and because of its robustness.

1.4.5 ALPHAScreen-based assays

Amplified luminescent proximity homogeneous assay (ALPHA) from Perkin-Elmer is a bead-based proximity assay that allows monitoring PPIs. It has its origins in a diagnostic assay termed LOCI (Luminescent Oxygen Channeling Immunoassay). The interaction of one protein bound to a donor bead interacting with a protein bound to an acceptor bead would generate a readout signal (Figure 20). The donor bead is composed of a photosensitizing phtalocyanine, which can be excited with low energy red-shifted light (680 nm). As a result, electronically excited singlet oxygen molecules are released and can diffuse in solution to a radius up to 200 nm. Acceptor beads are composed of Thioxene, Anthracene, and Rubrene. Upon interaction with the excited singlet oxygen, thioxene releases energy, which is transferred to anthracene followed by rubrene. Upon excitation, rubrene emits higher energy light at 520-620 nm.[118] Upon disruption of the PPIs by an inhibitor, the distance between the donor and acceptor beads would increase, resulting in a decrease of the signal. One donor bead is able to release up to 60,000 singlet oxygen molecules. Therefore, the signal of one biological molecule bound to it is significantly amplified, resulting in high sensitivity.

The labeling of two proteins by either the donor or the acceptor beads is performed through their binding to orthogonal tags. Hence, the two proteins need to have different recognition sites such as GST, His_6, FLAG-tag, or Fc that can be orthogonally recognized by the donor and acceptor beads.

ALPHAScreen-based assays require an ALPHAScreen enabled plate reader.[109] Despite being characterized by a high cost per well, the assay's robustness, high signal to background, and large proximity tolerance make the ALPHAScreen-based assay an interesting choice to monitor PPIs.

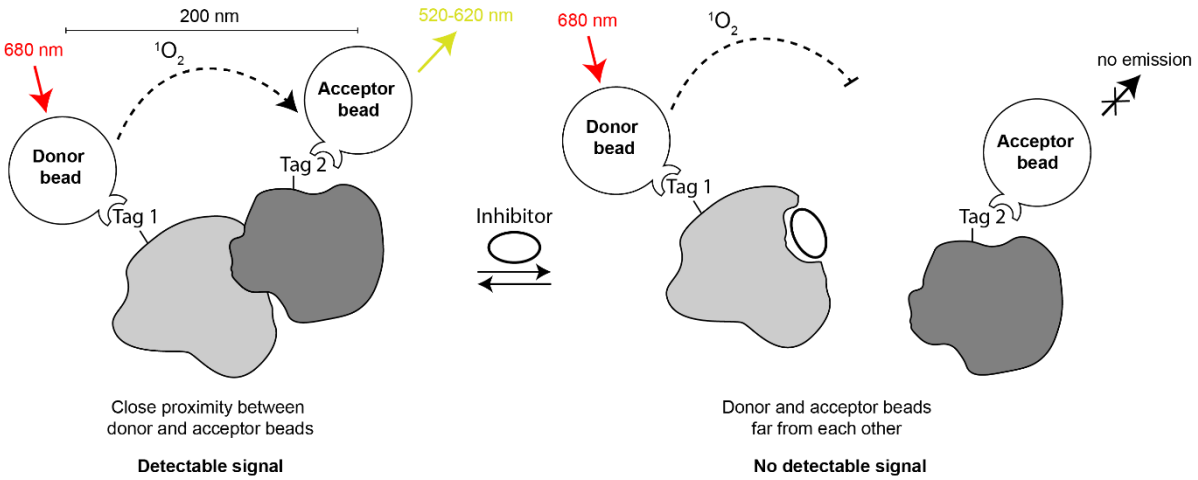


Figure 20. ALPHAScreen to monitor PPIs. Donor and acceptor beads bind to the two proteins (in gray and dark gray) to orthogonal recognition sites (Tag 1 and 2). Upon excitation by low-energy light (680 nm), the donor bead releases singlet oxygen. If the acceptor bead is in proximity (up to 200 nm), it emits higher energy light at 520 to 620 nm. Upon displacement of the PPI by an inhibitor, no signal is observed.

1.5 References

(1) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* **2002**.

(2) Loregian, A.; Palù, G. Disruption of protein-protein interactions: Towards new targets for chemotherapy. *Journal of Cellular Physiology*. **2005**.

(3) Nero, T. L.; Morton, C. J.; Holien, J. K.; Wielens, J.; Parker, M. W. Oncogenic protein interfaces: Small molecules, big challenges. *Nature Reviews Cancer*. **2014**.

(4) Yu, W.; Li, L.; Zheng, F.; Yang, W.; Zhao, S.; Tian, C.; Yin, W.; Chen, Y.; Guo, W.; Zou, L.; Deng, W. β-Catenin Cooperates with CREB Binding Protein to Promote the Growth of Tumor Cells. *Cell. Physiol. Biochem.* **2017**.

(5) Cory, S.; Adams, J. M. Killing cancer cells by flipping the Bcl-2/Bax switch. *Cancer Cell*. **2005**.

(6) Byrd, J. C.; Kipps, T. J.; Flinn, I. W.; Cooper, M.; Odenike, O.; Bendiske, J.; Rediske, J.; Bilic, S.; Dey, J.; Baeck, J.; O'Brien, S. Phase i study of the anti-CD40 humanized monoclonal antibody lucatumumab (HCD122) in relapsed chronic lymphocytic leukemia. *Leuk. Lymphoma* **2012**.

(7) Socinski, M. A.; Jotte, R. M.; Cappuzzo, F.; Orlandi, F.; Stroyakovskiy, D.; Nogami, N.; Rodríguez-Abreu, D.; Moro-Sibilot, D.; Thomas, C. A.; Barlesi, F.; Finley, G.; Kelsch, C.; Lee, A.; Coleman, S.; Deng, Y.; Shen, Y.; Kowanetz, M.; Lopez-Chavez, A.; Sandler, A.; Reck, M. Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. *N. Engl. J. Med.* **2018**.

(8) Dorr, P.; Westby, M.; Dobbs, S.; Griffin, P.; Irvine, B.; Macartney, M.; Mori, J.; Rickett, G.; Smith-Burchnell, C.; Napier, C.; Webster, R.; Armour, D.; Price, D.; Stammen, B.; Wood, A.; Perros, M. Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob. Agents Chemother.* **2005**.

(9) Dawidowski, M.; Emmanouilidis, L.; Kalel, V. C.; Tripsianes, K.; Schorpp, K.; Hadian, K.; Kaiser, M.; Mäser, P.; Kolonko, M.; Tanghe, S.; Rodriguez, A.; Schliebs, W.; Erdmann, R.; Sattler, M.; Popowicz, G. M. Inhibitors of PEX14 disrupt protein import into glycosomes and kill Trypanosoma parasites. *Science.* **2017**.

(10) Thompson, T. B.; Chaggar, P.; Kuhl, E.; Goriely, A. Protein-protein interactions in neurodegenerative diseases: A conspiracy theory. *PLoS Comput. Biol.* **2020**.

(11) Blazer, L. L.; Neubig, R. R. Small molecule protein-protein interaction inhibitors as CNS therapeutic agents: Current progress and future hurdles. *Neuropsychopharmacology*. **2009**.

(12) Lavecchia, A.; Cerchia, C. Recent advances in developing PCSK9 inhibitors for lipid-lowering therapy. *Future Medicinal Chemistry*. **2019**.

(13) Anand, P.; Brown, J. D.; Lin, C. Y.; Qi, J.; Zhang, R.; Artero, P. C.; Alaiti, M. A.; Bullard, J.; Alazem, K.; Margulies, K. B.; Cappola, T. P.; Lemieux, M.; Plutzky, J.; Bradner, J. E.; Haldar, S. M. XBET bromodomains mediate transcriptional pause release in heart failure. *Cell* **2013**.

(14) Madson, K. L.; Moore, T. L.; Lawrence, J. M.; Osborn, T. G. Cytokine levels in serum and synovial fluid of patients with juvenile rheumatoid arthritis. *J. Rheumatol.* **1994**.

(15) Palladino, M. A.; Bahjat, F. R.; Theodorakis, E. A.; Moldawer, L. L. Anti-TNF-α therapies: The next generation. *Nature Reviews Drug Discovery*. **2003**.

(16) Liu, S.; Desharnais, J.; Sahasrabudhe, P. V.; Jin, P.; Li, W.; Oates, B. D.; Shanker, S.; Banker, M. E.; Chrunyk, B. A.; Song, X.; Feng, X.; Griffor, M.; Jimenez, J.; Chen, G.; Tumelty, D.; Bhat, A.; Bradshaw, C. W.; Woodnutt, G.; Lappe, R. W.; Thorarensen, A.; Qiu, X.; Withka, J. M.; Wood, L. D. Inhibiting complex IL-17A and IL-17RA interactions with a linear peptide. *Sci. Rep.* **2016**.

(17) Zinzalla, G.; Thurston, D. E. Targeting protein-protein interactions for therapeutic intervention: A challenge for the future. *Future Medicinal Chemistry*. **2009**.

(18) Stumpf, M. P. H.; Thorne, T.; De Silva, E.; Stewart, R.; Hyeong, J. A.; Lappe, M.; Wiuf, C. Estimating the size of the human interactome. *Proc. Natl. Acad. Sci. U. S. A.* **2008**.

(19) London, N.; Raveh, B.; Schueler-Furman, O. Druggable protein-protein interactions - from hot spots to hot segments. *Current Opinion in Chemical Biology*. **2013**.

(20) Taguchi, K.; Yamamoto, M. The keap1–nrf2 system as a molecular target of cancer treatment. *Cancers*. **2021**.

(21) Yap, J. L.; Wang, H.; Hu, A.; Chauhan, J.; Jung, K. Y.; Gharavi, R. B.; Prochownik, E. V.; Fletcher, S. Pharmacophore identification of c-Myc inhibitor 10074-G5. *Bioorganic Med. Chem. Lett.* **2013**.

(22) White, P. W.; Faucher, A. M.; Goudreau, N. Small molecule inhibitors of the human papillomavirus E1-E2 interaction. *Curr. Top. Microbiol. Immunol.* **2010**.

(23) Perez, V. L.; Pflugfelder, S. C.; Zhang, S.; Shojaei, A.; Haque, R. Lifitegrast, a Novel Integrin Antagonist for Treatment of Dry Eye Disease. *Ocular Surface*. **2016**.

(24) Jain, A. K.; Barton, M. C. Bromodomain Histone Readers and Cancer. *Journal of Molecular Biology*. **2017**.

(25) Jones, S.; Thornton, J. M. Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*. **1996**.

(26) Scott, D. E.; Bayly, A. R.; Abell, C.; Skidmore, J. Small molecules, big targets: Drug discovery faces the protein-protein interaction challenge. *Nature Reviews Drug Discovery*. **2016**.

(27) Jubb, H.; Blundell, T. L.; Ascher, D. B. Flexibility and small pockets at protein-protein interfaces: New insights into druggability. *Progress in Biophysics and Molecular Biology*. **2015**.

(28) Smith, M. C.; Gestwicki, J. E. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert reviews in molecular medicine*. **2012**.

(29) Bogan, A. A.; Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**.

(30) Ran, X.; Gestwicki, J. E. Inhibitors of protein–protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Current Opinion in Chemical Biology*. **2018**.

(31) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*. **1997**.

(32) Fuller, J. C.; Burgoyne, N. J.; Jackson, R. M. Predicting druggable binding sites at the protein-protein interface. *Drug Discovery Today*. **2009**.

(33) Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wang, W. State-of-the-art strategies for targeting protein-protein interactions by small-molecule inhibitors. *Chemical Society Reviews*. **2015**.

(34) Jesus Perez de Vega, M.; Martin-Martinez, M.; Gonzalez-Muniz, R. Modulation of Protein-Protein Interactions by Stabilizing/Mimicking Protein Secondary Structure Elements. *Curr. Top. Med. Chem.* **2006**.

(35) Cunningham, A. D.; Qvit, N.; Mochly-Rosen, D. Peptides and peptidomimetics as regulators of protein–protein interactions. *Current Opinion in Structural Biology*. **2017**.

(36) Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science.* **2004**.

(37) Cai, Q.; Sun, H.; Peng, Y.; Lu, J.; Nikolovska-Coleska, Z.; McEachern, D.; Liu, L.; Qiu, S.; Yang, C. Y.; Miller, R.; Yi, H.; Zhang, T.; Sun, D.; Kang, S.; Guo, M.; Leopold, L.; Yang, D.; Wang, S. A potent and orally active antagonist (SM-406/AT-406) of multiple inhibitor of apoptosis proteins (IAPs) in clinical development for cancer treatment. *J. Med. Chem.* **2011**.

(38) Nieddu, E.; Pasa, S. Interfering with Protein-Protein Contact: Molecular Interaction Maps and Peptide Modulators. *Curr. Top. Med. Chem.* **2006**.

(39) Lehmann, C.; Friess, T.; Birzele, F.; Kiialainen, A.; Dangl, M. Superior anti-tumor activity of the MDM2 antagonist idasanutlin and the Bcl-2 inhibitor venetoclax in p53 wild-type acute myeloid leukemia models. *J. Hematol. Oncol.* **2016**.

(40) West, A. C.; Martin, B. P.; Andrews, D. A.; Hogg, S. J.; Banerjee, A.; Grigoriadis, G.; Johnstone, R. W.; Shortt, J. The SMAC mimetic, LCL-161, reduces survival in aggressive MYC-driven lymphoma while promoting susceptibility to endotoxic shock. *Oncogenesis* **2016**.

(41) Ward, G. A.; Lewis, E. J.; Ahn, J. S.; Johnson, C. N.; Lyons, J. F.; Martins, V.; Munck, J. M.; Rich, S. J.; Smyth, T.; Thompson, N. T.; Williams, P. A.; Wilsher, N. E.; Wallis, N. G.; Chessari, G. ASTX660, a novel non-peptidomimetic antagonist of cIAP1/2 and XIAP, potently induces TNFα-dependent apoptosis in cancer cell lines and inhibits tumor growth. *Mol. Cancer Ther.* **2018**.

(42) Musielak, B.; Kocik, J.; Skalniak, L.; Magiera-Mularz, K.; Sala, D.; Czub, M.; Stec, M.; Siedlar, M.; Holak, T. A.; Plewka, J. CA-170 - A Potent Small-Molecule PD-L1 Inhibitor or Not? *Molecules* **2019**.

(43) Carvajal, L. A.; Ben Neriah, D.; Senecal, A.; Benard, L.; Thiruthuvanathan, V.; Yatsenko, T.; Narayanagari, S. R.; Wheat, J. C.; Todorova, T. I.; Mitchell, K.; Kenworthy, C.; Guerlavais, V.; Annis, D. A.; Bartholdy, B.; Will, B.; Anampa, J. D.; Mantzaris, I.; Aivado, M.; Singer, R. H.; Coleman, R. A.; Verma, A.; Steidl, U. Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. *Sci. Transl. Med.* **2018**.

(44) Cheng, X.; Lee, T.-Y.; Ledet, G.; Zemade, G.; Tovera, J.; Campbell, R.; Purro, N.; Annamalai, T.; Masjedizadeh, M.; Liu, D.; Nawabi, R.; Kanwar, B. 751 Safety, Tolerability, and Pharmacokinetics of PTG-200, an Oral GI-Restricted Peptide Antagonist of IL-23 Receptor, in Normal Healthy Volunteers. *Am. J. Gastroenterol.* **2019**.

(45) Dirix, L. Y.; Takacs, I.; Jerusalem, G.; Nikolinakos, P.; Arkenau, H. T.; Forero-Torres,

A.; Boccia, R.; Lippman, M. E.; Somer, R.; Smakal, M.; Emens, L. A.; Hrinczenko, B.; Edenfield, W.; Gurtler, J.; von Heydebreck, A.; Grote, H. J.; Chin, K.; Hamilton, E. P. Avelumab, an anti-PD-L1 antibody, in patients with locally advanced or metastatic breast cancer: A phase 1b JAVELIN solid tumor study. *Breast Cancer Res. Treat.* **2018**.

(46) Yudin, A. K. Macrocycles: Lessons from the distant past, recent developments, and future directions. *Chemical Science*. **2015**.

(47) Iqbal, E. S.; Hartman, M. C. T. Shaping molecular diversity. *Nat. Chem.* **2018**.

(48) Giordanetto, F.; Kihlberg, J. Macrocyclic drugs and clinical candidates: What can medicinal chemists learn from their properties? *J. Med. Chem.* **2014**.

(49) Mallinson, J.; Collins, I. Macrocycles in new drug discovery. *Future Medicinal Chemistry*. **2012**.

(50) Claveria-Gimeno, R.; Vega, S.; Abian, O.; Velazquez-Campoy, A. A look at ligand binding thermodynamics in drug discovery. *Expert Opinion on Drug Discovery*. **2017**.

(51) Dinsmore, C. J.; Bogusky, M. J.; Culberson, J. C.; Bergman, J. M.; Homnick, C. F.; Zartman, C. B.; Mosser, S. D.; Schaber, M. D.; Robinson, R. G.; Koblan, K. S.; Huber, H. E.; Graham, S. L.; Hartman, G. D.; Huff, J. R.; Williams, T. M. Conformational restriction of flexible ligands guided by the transferred NOE experiment: Potent macrocyclic inhibitors of farnesyltransferase. *Journal of the American Chemical Society*. **2001**.

(52) Sund, C.; Belda, O.; Wiktelius, D.; Sahlberg, C.; Vrang, L.; Sedig, S.; Hamelink, E.; Henderson, I.; Agback, T.; Jansson, K.; Borkakoti, N.; Derbyshire, D.; Eneroth, A.; Samuelsson, B. Design and synthesis of potent macrocyclic renin inhibitors. *Bioorganic Med. Chem. Lett.* **2011**.

(53) Cherney, R. J.; Wang, L.; Meyer, D. T.; Xue, C. B.; Wasserman, Z. R.; Hardman, K. D.; Welch, P. K.; Covington, M. B.; Copeland, R. A.; Arner, E. C.; DeGrado, W. F.; Decicco, C. P. Macrocyclic amino carboxylates as selective MMP-8 inhibitors. *J. Med. Chem.* **1998**.

(54) Vinogradov, A. A.; Yin, Y.; Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *Journal of the American Chemical Society*. **2019**.

(55) Bockus, A. T.; Lexa, K. W.; Pye, C. R.; Kalgutkar, A. S.; Gardner, J. W.; Hund, K. C. R.; Hewitt, W. M.; Schwochert, J. A.; Glassey, E.; Price, D. A.; Mathiowetz, A. M.; Liras, S.; Jacobson, M. P.; Lokey, R. S. Probing the Physicochemical Boundaries of Cell Permeability and Oral Bioavailability in Lipophilic Macrocycles Inspired by Natural

Products. *J. Med. Chem.* **2015**.

(56) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral druggable space beyond the rule of 5: Insights from drugs and clinical candidates. *Chemistry and Biology*. **2014**.

(57) Appavoo, S. D.; Huh, S.; Diaz, D. B.; Yudin, A. K. Conformational Control of Macrocycles by Remote Structural Modification. *Chemical Reviews*. **2019**.

(58) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The exploration of macrocycles for drug discovery - An underexploited structural class. *Nature Reviews Drug Discovery*. **2008**.

(59) Schlünzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **2001**.

(60) Campbell, E. A.; Korzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S. A. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* **2001**.

(61) Bemer-Melchior, P.; Bryskier, A.; Drugeon, H. B. Comparison of the in vitro activities of rifapentine and rifampicin against Mycobacterium tuberculosis complex. *J. Antimicrob. Chemother.* **2000**.

(62) Bernard, L.; Vaudaux, P.; Vuagnat, A.; Stern, R.; Rohner, P.; Pittet, D.; Schrenzel, J.; Hoffmeyer, P. Effect of Vancomycin Therapy for Osteomyelitis on Colonization by Methicillin-Resistant Staphylococcus aureus : Lack of Emergence of Glycopeptide Resistance . *Infect. Control Hosp. Epidemiol.* **2003**.

(63) Loll, P. J.; Bevivino, A. E.; Korty, B. D.; Axelsen, P. H. Simultaneous recognition of a carboxylate-containing ligand and an intramolecular surrogate ligand in the crystal structure of an asymmetric vancomycin dimer. *J. Am. Chem. Soc.* **1997**.

(64) Schreiber, S. L.; Crabtree, G. R. The mechanism of action of cyclosporin A and FK506. *Immunology Today*. **1992**.

(65) Cohen, D. J.; Loertscher, R.; Rubin, M. F.; Tilney, N. L.; Carpenter, C. B.; Strom, T. B. Cyclosporine: A new immunosuppressive agent for organ transplantation. *Annals of Internal Medicine*. **1984**.

(66) Liu, J.; Farmer, J. D.; Lane, W. S.; Friedman, J.; Weissman, I.; Schreiber, S. L. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **1991**.

(67) Ganguly, A.; Cabral, F.; Yang, H.; Patel, K. D. Peloruside A is a microtubule-stabilizing

agent with exceptional anti-migratory properties in human endothelial cells. *Oncoscience* **2015**.

(68) Bode, C. J.; Gupta, M. L.; Reiff, E. A.; Suprenant, K. A.; Georg, G. I.; Himes, R. H. Epothilone and paclitaxel: Unexpected differences in promoting the assembly and stabilization of yeast microtubules. *Biochemistry* **2002**.

(69) Hart, S.; Goh, K. C.; Novotny-Diermayr, V.; Hu, C. Y.; Hentze, H.; Tan, Y. C.; Madan, B.; Amalini, C.; Loh, Y. K.; Ong, L. C.; William, A. D.; Lee, A.; Poulsen, A.; Jayaraman, R.; Ong, K. H.; Ethirajulu, K.; Dymock, B. W.; Wood, J. W. SB1518, a novel macrocyclic pyrimidine-based JAK2 inhibitor for the treatment of myeloid and lymphoid malignancies. *Leukemia* **2011**.

(70) William, A. D.; Lee, A. C. H.; Goh, K. C.; Blanchard, S.; Poulsen, A.; Teo, E. L.; Nagaraj, H.; Lee, C. P.; Wang, H.; Williams, M.; Sun, E. T.; Hu, C.; Jayaraman, R.; Pasha, M. K.; Ethirajulu, K.; Wood, J. M.; Dymock, B. W. Discovery of kinase spectrum selective macrocycle (16E)-14-methyl-20-oxa-5, 7,14,26-tetraazatetracyclo[19.3.1.1(2,6).1(8,12)]heptacosa-1(25),2(26),3,5,8(27) ,9,11,16,21,23-decaene (SB1317/TG02), a potent inhibitor of cyclin dependent kinases (CDKs), Janus. *J. Med. Chem.* **2012**.

(71) Lin, T. I.; Lenz, O.; Fanning, G.; Verbinnen, T.; Delouvroy, F.; Scholliers, A.; Vermeiren, K.; Rosenquist, Â.; Edlund, M.; Samuelsson, B.; Vrang, L.; De Kock, H.; Wigerinck, P.; Raboisson, P.; Simmen, K. In vitro activity and preclinical profile of TMC435350, a potent hepatitis C virus protease inhibitory. *Antimicrob. Agents Chemother.* **2009**.

(72) McCauley, J. A.; McIntyre, C. J.; Rudd, M. T.; Nguyen, K. T.; Romano, J. J.; Butcher, J. W.; Gilbert, K. F.; Bush, K. J.; Holloway, M. K.; Swestock, J.; Wan, B. L.; Carroll, S. S.; Dimuzio, J. M.; Graham, D. J.; Ludmerer, S. W.; Mao, S. S.; Stahlhut, M. W.; Fandozzi, C. M.; Trainor, N.; Olsen, D. B.; Vacca, J. P.; Liverton, N. J. Discovery of vaniprevir (MK-7009), a macrocyclic hepatitis C virus NS3/4a protease inhibitor. *J. Med. Chem.* **2010**.

(73) Macrocycle Library Screening Compounds – ChemBridge https://www.chembridge.com/screening_libraries/targeted_libraries/macrocycle-screening-library/index.php (accessed Jul 1, **2021**).

(74) Macrocycles – Asinex.com http://www.asinex.com/libraries-macrocyclic-5-html/ (accessed Jul 1, **2021**).

(75) Macrocycles - Enamine https://enamine.net/compound-collections/macrocycles (accessed Jul 1, **2021**).

(76) Macrocycle discovery platform – Polyphor https://www.polyphor.com/macrocycle-discovery-platform/ (accessed Jul 1, **2021**).

(77) Macrocycles Library – ChemDiv https://www.chemdiv.com/catalog/focused-and-targeted-libraries/macrocycles-library/ (accessed Jan 7, **2022**).

(78) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierskit, W. M.; Knappt, R. J. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **1991**.

(79) Morimoto, J.; Kodadek, T. Synthesis of a large library of macrocyclic peptides containing multiple and diverse N-alkylated residues. *Mol. Biosyst.* **2015**.

(80) Wu, X.; Upadhyaya, P.; Villalona-Calero, M. A.; Briesewitz, R.; Pei, D. Inhibition of Ras-effector interactions by cyclic peptides. *Medchemcomm* **2013**.

(81) Thakkar, A.; Cohen, A. S.; Connolly, M. D.; Zuckermann, R. N.; Pei, D. High-throughput sequencing of peptoids and peptide-peptoid hybrids by partial edman degradation and mass spectrometry. *J. Comb. Chem.* **2009**.

(82) Wang, P.; Arabaci, G.; Pei, D. Rapid sequencing of library-derived peptides by partial edman degradation and mass spectrometry. *J. Comb. Chem.* **2001**.

(83) Gao, Y.; Kodadek, T. Direct comparison of linear and macrocyclic compound libraries as a source of protein ligands. *ACS Comb. Sci.* **2015**.

(84) Joo, S. H.; Xiao, Q.; Ling, Y.; Gopishetty, B.; Pei, D. High-throughput sequence determination of cyclic peptide library members by partial Edman degradation/mass spectrometry. *J. Am. Chem. Soc.* **2006**.

(85) Lian, W.; Upadhyaya, P.; Rhodes, C. A.; Liu, Y.; Pei, D. Screening bicyclic peptide libraries for protein-protein interaction inhibitors: Discovery of a tumor necrosis factor-α Antagonist. *J. Am. Chem. Soc.* **2013**.

(86) Koivunen, E.; Arap, W.; Rajotte, D.; Lahdenranta, J.; Pasqualini, R. Identification of receptor ligands with phage display peptide libraries. *J. Nucl. Med.* **1999**.

(87) DeLano, W. L.; Ultsch, M. H.; De Vos, A. M.; Wells, J. A. Convergent solutions to binding at a protein-protein interface. *Science.* **2000**.

(88) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **2009**.

(89) Kale, S. S.; Villequey, C.; Kong, X. D.; Zorzi, A.; Deyle, K.; Heinis, C. Cyclization of peptides with two chemical bridges affords large scaffold diversities. *Nat. Chem.* **2018**.

(90) Krumpe, L. R. H.; Atkinson, A. J.; Smythers, G. W.; Kandel, A.; Schumacher, K. M.; McMahon, J. B.; Makowski, L.; Mori, T. T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. *Proteomics* **2006**.

(91) Roberts, R. W.; Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1997**.

(92) Goto, Y.; Katoh, T.; Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **2011**.

(93) Guillen Schlippe, Y. V.; Hartman, M. C. T.; Josephson, K.; Szostak, J. W. In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J. Am. Chem. Soc.* **2012**.

(94) Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H. Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol.* **2011**.

(95) Wang, H.; Liu, R. Advantages of mRNA display selections over other selection techniques for investigation of protein-protein interactions. *Expert Review of Proteomics*. **2011**.

(96) Plais, L.; Scheuermann, J. Macrocyclic DNA-encoded chemical libraries: a historical perspective. *RSC Chem. Biol.* **2022**.

(97) Gartner, Z. J.; Kanan, M. W.; Liu, D. R. Multistep small-molecule synthesis programmed by DNA templates. *J. Am. Chem. Soc.* **2002**.

(98) Seigal, B. A.; Connors, W. H.; Fraley, A.; Borzilleri, R. M.; Carter, P. H.; Emanuel, S. L.; Fargnoli, J.; Kim, K.; Lei, M.; Naglich, J. G.; Pokross, M. E.; Posy, S. L.; Shen, H.; Surti, N.; Talbott, R.; Zhang, Y.; Terrett, N. K. The discovery of macrocyclic XIAP antagonists from a DNA-programmed chemistry library, and their optimization to give lead compounds with in vivo antitumor activity. *J. Med. Chem.* **2015**.

(99) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-generation DNA-templated macrocycle libraries for the discovery of bioactive small molecules. *Nat. Chem.* **2018**.

(100) Li, Y.; de Luca, R.; Cazzamalli, S.; Pretto, F.; Bajic, D.; Scheuermann, J.; Neri, D. Versatile protein recognition by the encoded display of multiple chemical elements on a constant macrocyclic scaffold. *Nat. Chem.* **2018**.

(101) Paciaroni, N. G.; Ndungu, J. M.; Kodadek, T. Solid-phase synthesis of DNA-encoded libraries: Via an “aldehyde explosion” strategy. *Chem. Commun.* **2020**.

(102) Koesema, E.; Roy, A.; Paciaroni, N. G.; Kodadek, T. Synthesis and Screening of A DNA-Encoded Library of Non-Peptidic Macrocycles. *ChemRxiv* **2020**.

(103) Onda, Y.; Bassi, G.; Elsayed, A.; Ulrich, F.; Oehler, S.; Plais, L.; Scheuermann, J.; Neri, D. A DNA-Encoded Chemical Library Based on Peptide Macrocycles. *Chem. - A Eur. J.* **2021**.

(104) Kale, S. S.; Bergeron-Brlek, M.; Wu, Y.; Kumar, M. G.; Pham, M. V.; Bortoli, J.; Vesin, J.; Kong, X. D.; Franco Machado, J.; Deyle, K.; Gonschorek, P.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Thiol-to-amine cyclization reaction enables screening of large libraries of macrocyclic compounds and the generation of sub-kilodalton ligands. *Sci. Adv.* **2019**.

(105) Mothukuri, G. K.; Kale, S. S.; Stenbratt, C. L.; Zorzi, A.; Vesin, J.; Bortoli Chapalay, J.; Deyle, K.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Macrocycle synthesis strategy based on step-wise “adding and reacting” three components enables screening of large combinatorial libraries. *Chem. Sci.* **2020**.

(106) Arkin, M. R.; Glicksman, M. A.; Fu, H.; Havel, J. J.; Du, Y. Inhibition of Protein-Protein Interactions : Non-Cellular Assay Formats Assay Guidance Manual. In *Assay Guidance Manual [Internet]*; **2012**.

(107) Moerke, N. J. Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding. *Curr. Protoc. Chem. Biol.* **2009**.

(108) Hall, M. D.; Yasgar, A.; Peryea, T.; Braisted, J. C.; Jadhav, A.; Simeonov, A.; Coussens, N. P. Fluorescence polarization assays in high-throughput screening and drug discovery: A review. *Methods and Applications in Fluorescence*. **2016**.

(109) Milroy, L. G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. Modulators of protein-protein interactions. *Chemical Reviews*. **2014**.

(110) Kenny, C. H.; Ding, W.; Kelleher, K.; Benard, S.; Dushin, E. G.; Sutherland, A. G.; Mosyak, L.; Kriz, R.; Ellestad, G. Development of a fluorescence polarization assay to screen for inhibitors of the FtsZ/ZipA interaction. *Anal. Biochem.* **2003**.

(111) Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **1948**.

(112) Osad’Ko, I. S. Dependence of FRET efficiency on distance in single donor-acceptor pairs. *J. Chem. Phys.* **2015**.

(113) Song, Y.; Madahar, V.; Liao, J. Development of FRET assay into quantitative and high-throughput screening technology platforms for protein-protein interactions. *Ann. Biomed. Eng.* **2011**.

(114) Lichlyter, D. J.; Grant, S. A.; Soykan, O. Development of a novel FRET immunosensor technique. *Biosens. Bioelectron.* **2003**.

(115) Degorce, F.; Card, A.; Soh, S.; Trinquet, E.; Knapik, G. P.; Xie, B. HTRF: A technology tailored for drug discovery - A review of theoretical aspects and recent applications. *Current Chemical Genomics*. **2009**.

(116) Berezin, M. Y.; Achilefu, S. Fluorescence lifetime measurements and biological imaging. *Chem. Rev.* **2010**.

(117) Du, Y.; Fu, R. W.; Lou, B.; Zhao, J.; Qui, M.; Khuri, F. R.; Fu, H. A time-resolved fluorescence resonance energy transfer assay for high-throughput screening of 14-3-3 protein-protein interaction inhibitors. *Assay Drug Dev. Technol.* **2013**.

(118) Eglen, R. M.; Reisine, T.; Roby, P.; Rouleau, N.; Illy, C.; Bossé, R.; Bielefeld, M. The Use of AlphaScreen Technology in HTS: Current Status. *Curr. Chem. Genomics* **2008**.

2. Aim of the thesis

Protein-protein interactions (PPIs) are an important therapeutic target class due to their key role in many diseases, but the discovery of inhibitors has been challenging for many of them, especially when the interaction of interest is intracellular, and the inhibitor requires to be membrane permeable. Macrocycles proved to be a promising drug modality due to their capacity to target protein-protein interactions while having the potential to be cell-permeable or even orally available. However, their discovery was hampered by the lack of large macrocycle libraries to screen.

To address this gap, the goals of my Ph.D. thesis were to develop novel strategies to generate large libraries of macrocycles, to establish assays for screening them, and to identify macrocycle-based ligands of therapeutically relevant PPIs.

The first step towards this end was to establish a more efficient and economical approach to access combinatorial libraries of macrocycles synthesized in microwell plates. To achieve this, I aimed at applying acoustic droplet ejection (ADE) – a contact-less liquid transfer technology – to macrocycle synthesis strategies previously developed in our lab that were dependent on reagent transfer by disposable pipetting tips. This was expected to drastically reduce the reagent amount required for macrocycle synthesis and screening, and eliminate the need for plastic tips, both major cost drivers of the available macrocycle synthesis methods.

The second step in reaching my goals was to screen the synthesized macrocycle libraries against therapeutically relevant PPI targets, with the aim of finding ligands and potential lead compounds for drug development. Towards this end, I planned to establish reported assays or to develop new assays that are suited for high-throughput screening of crude macrocycle compound libraries against important PPI targets.

3. Picomole scale synthesis and screening of macrocyclic compound libraries by acoustic liquid transfer

3.1 Work contribution

The chapter 3 of this thesis is based on the published manuscript "Picomole scale synthesis and screening of macrocyclic compound libraries by acoustic liquid transfer". The following authors contributed:

Gontran Sangouard,[1] Alessandro Zorzi,[1] Yuteng Wu,[1] Edouard Ehret,[1] Mischa Schüttel,[1] Sangram Kale,[1] Cristina Dìaz Perlas,[1] Jonathan Vesin,[2] Julien Bortoli Chapalay,[2] Gerardo Turcatti,[2] and Christian Heinis[1]*

[1]Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

[2]Biomolecular Screening Facility, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

*Corresponding author. Email: christian.heinis@epfl.ch

Author contributions: Prof. Christian Heinis, Dr. Alessandro Zorzi, and Dr. Yuteng Wu, conceptualized the synthesis of macrocyclic compound libraries by acoustic liquid transfer and performed tests to assess the feasibility. I optimized the application at picomole scale of one synthetic strategy previously developed, expressed MDM2, synthesized all precursors for the macrocycle libraries, performed all the screening and hit validation. Edouard Ehret helped with characterizing hits. Jonathan Vesin and Julien Bortoli Chapalay, under the supervision of Prof. Gerardo Turcatti, helped in the robotic liquid transfer for the macrocycle library synthesis. Mischa Schüttel helped in demonstrating the proper mixing of the droplets by centrifugation. Dr. Sangram Kale developed an approach to synthesize 2-pyridylthio cysteamine. Dr. Cristina Dìaz Perlas helped in the SPR of some of the macrocyclic hits. Together with Prof. Christian Heinis, I prepared all the figures and wrote the first draft of the manuscript, and contributed to editing.

3.2 Abstract

Macrocyclic compounds are an attractive class of therapeutic ligands against challenging targets such as protein-protein interactions. However, the development of macrocycles as drugs is hindered by the lack of large combinatorial macrocyclic libraries, which are cumbersome, expensive, and time consuming to make, screen, and deconvolute. Here, we established a strategy for synthesizing and screening combinatorial libraries on a picomolar scale using acoustic droplet ejection to combine building blocks at nanoliter volumes, which reduced reaction volumes, reagent consumption, and synthesis time. As a proof-of-concept, we assembled a 2,700-member target-focused macrocyclic library that we could subsequently assay in the same microtiter synthesis plates, saving the need for additional transfers and deconvolution schemes. We screened the library against the MDM2-p53 protein-protein interaction and generated micromolar and sub-micromolar inhibitors. Our work synthesizing combinatorial macrocycle libraries at the picomole scale using acoustic liquid transfer provides a general strategy towards macrocycle ligand development.

3.3 Introduction

Macrocyclic compounds are of great interest to the pharmaceutical industry due to their ability to bind to challenging targets such as proteins with flat surfaces or protein-protein interactions, for which it has been difficult or impossible to generate ligands based on classical small molecules. Of particular interest are membrane permeable macrocyclic compounds that can reach intracellular targets, which requires that they are small (< 1 kDa) and have a limited polar surface area (< 200 Å^2).[1] However, the development of macrocyclic ligands within these parameters is hindered by the small size of commercially available libraries (typically around 10,000 molecules or smaller)[2–4] and the lack of efficient synthetic methodologies to generate target-focused libraries. Impressive progress to access larger macrocyclic compound libraries was made in recent years using DNA-encoding technologies,[5–8] but the synthesis of combinatorial macrocycle libraries based on split and pool methods and subsequent deconvolution can be challenging due to accumulation of side products in the many sequential coupling steps needed in macrocycle compounds synthesis. To access combinatorial libraries comprising ten-thousands of macrocycles while limiting the synthesis and deconvolution complexity, our laboratory has recently established an approach for library synthesis in microwell plates by the sequential "adding and reacting" of building blocks.[9] From an 8,988-member macrocycle library, we identified nanomolar inhibitors of protease targets. While this approach showed great promise for the development of macrocyclic ligands, we were limited by the high costs of pipetting tips used for liquid transfer and the mg-scale amounts of peptide needed as building blocks for synthesis.

Combinatorial library synthesis can potentially be miniaturized by performing reactions in nanoliter (nL) volumes. Previously, Dreher, Cernak, and co-workers used mosquito nL dispensing, an approach in which droplets are transferred by thin pins, to combine reagents and test many conditions for complex small-molecule synthesis.[10] This high-throughput nanomole-scale approach was subsequently used with an affinity-selection mass spectrometry bioassay to test 3,114 reaction conditions for 345 compounds.[11] Dömling and co-workers used acoustic droplet ejection (ADE) technology for the transfer of reagents in nL volumes and pioneered both, the scouting of chemical reactions and the combinatorial synthesis of drug-like scaffolds by ADE[12,13], and they recently synthesized and screened a library of several hundred small molecules at a 500 nmol scale.[14] With ADE, acoustic sound is used for contact-free liquid transfer in nL volumes from wells of a source plate to wells of an inverted destination plate positioned above.[15] The ADE technology has been broadly applied for the high-throughput screening of small-molecule libraries in the pharmaceutical industry, in which small molecules are transferred to assay reagents in microtiter plates.[15] Zhang et al.

have combined ADE with mass spectrometry for efficient reaction analysis.[16] To our knowledge, ADE has not been used to synthesize combinatorial compound libraries at a picomole scale for subsequent sampling of target engagement in high-throughput screens.

Herein, we applied acoustic dispensing to synthesize large combinatorial macrocyclic compound libraries by step-wise transfer of building blocks to 384-well microtiter plates, as illustrated in Figure 1a. In brief, *m* building blocks (in blue) were mixed with *n* building blocks (in yellow). Subsequently, *o* building blocks (in green) were added to cyclize the linear products to yield $m \times n \times o$ macrocyclic products, with each combination in a discrete well. We envisioned that the efficiency of coupling and cyclization reactions would yield macrocyclic compounds as the main product that can be screened without purification, simply by adding assay reagents to the synthesis plate.

3.4 Results and discussion

3.4.1 Validation of proper droplet merging

We first assessed the efficiency of reagent mixing in nL-sized droplets as transferred by acoustic sound using colored liquids and a standard acoustic transfer device (ECHO 650 from Labcyte/Beckman Coulter) (Figure 1b). The transfer of water was problematic for volumes smaller than 40 nL due to rapid air-drying (not shown). Transfers of reagents in DMSO or in DMSO-H_2O mixtures were efficient and prevented drying, even in droplets as small as 2.5 nL. We transferred reagents of two different colors, blue Coomassie and yellow fluorescein, in 2.5-, 5-, 10-, 20-, 40-, 80- or 160-nL volumes and found that the merged droplets immediately turned green, which indicated efficient mixing of the compounds in the droplets and solvent (Figure 1b).

We endeavored to use rather low reaction volumes in the double-digit nL range as this would limit the number of acoustic pulses (liquids are transferred in pulses of 2.5-nL volumes) and generate high reagent concentrations for efficient building-block coupling. However, we were concerned that at small volumes, the droplets of two reagents might occasionally not meet. The transfer of 40 nL volumes of two compounds to all wells of a 384-well plate revealed that droplets of this size merge in almost all but not all cases (Figure 1c & Figure S1). Further optimization of the transfer protocols helped somewhat, but in the end, we solved this by introducing a centrifugation step that ensured robust droplet merging (Figure 1c & Figure S2).

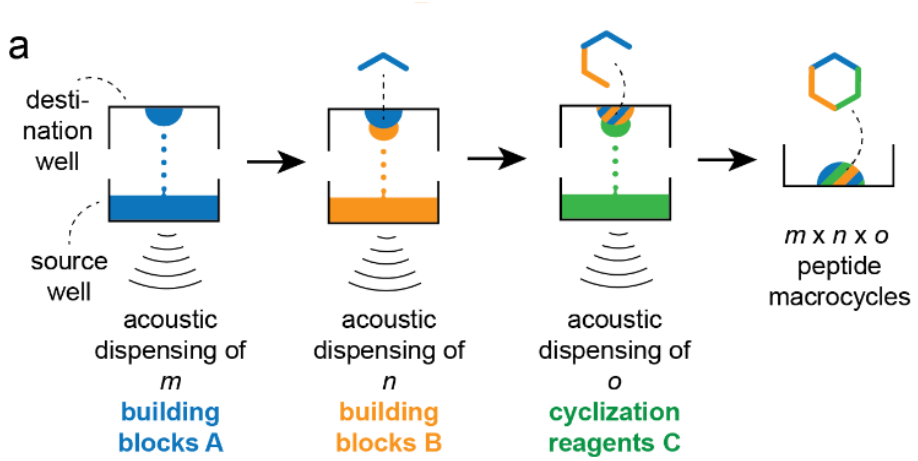


Figure 1. Combinatorial macrocyclic compound synthesis by acoustic droplet ejection. (a) Building blocks are transferred step-wise from source to destination wells in microtiter plates. Acoustic waves (indicated schematically) applied to the lower source wells transfer 2.5 nL droplets to the inverted destination well above. The three different building blocks are indicated by color. The product macrocycle is contained in the blue-yellow-green droplet. Figure 1 continues next page.

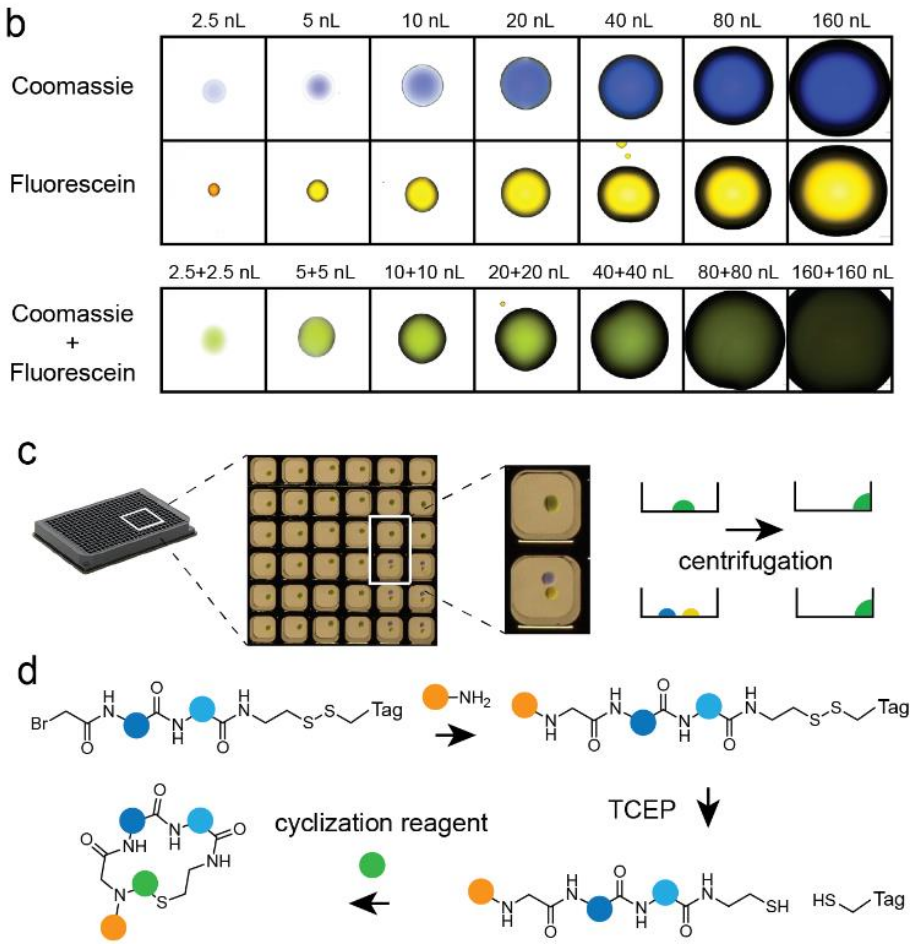


Figure 1 continued. (b) Microscopy images (10× amplification) of DMSO containing Coomassie (blue), fluorescein (yellow) or both (green). Solutions were transferred as 2.5-nL pulses to generate droplets of 2.5 to 160 nL. (c) Droplet merging in a 384-well microtiter plate. Two enlarged wells demonstrated merged (upper) or unmerged (lower) droplets. Droplets in all wells are merged by centrifugation of the plate. (d) Macrocycles are synthesized using a three-component strategy. Colored spheres represent varied positions in the combinatorial library. The cysteine thiol group is here protected by a disulphide-bridged short peptide (Tag).

3.4.2 Validation of macrocycle synthesis at picomole scale

We tested the acoustic dispensing strategy using a recently developed macrocycle synthesis approach (Figure 1d) in which *m* N_{α}-bromoacetyl-activated linear peptides were reacted with *n* primary amines, followed by a cyclization with *o* linkers to yield *m* × *n* × *o* macrocycles.[17] We had previously used this approach to affinity mature a macrocyclic thrombin inhibitor (P2, K_i = 42 nM) through combinatorial synthesis of thousands of variants of it. However, we had not applied the strategy to the synthesis or screening of random macrocyclic compound libraries, nor to any target other than thrombin. We first tested whether the three consecutive reactions, previously performed in larger reaction volumes of 20 µL (alkylation), 96 µL (reduction), and 100 µL (cyclization), were possible in 80 nL to 4 µL volumes, which were up to 250-fold smaller volumes, as shown in Figure 2a. We performed the three consecutive reactions with the three

building blocks shown in Figure 2a: i) a BrAc-activated model peptide (protected at the thiol group by a disulfide conjugated short peptide tag that helped also for its purification; Figure S3), ii) an example primary amine (2-furanmethanamine), and iii) cyclization reagents such as divinyl sulfone (**1**) (Figure 2a). The alkylation and disulfide-reduction reactions succeeded in the small volumes with over 95% yields (Figure 2b; desired products in red). The subsequent cyclization reactions with the seven tested linker reagents (Figure 2c) had yields ranging from more than 60% (reagents **1**-**3**, 61, 66, and 69%), to low or almost no product (reagents **4**-**7**, 2, 0.2, 26, 0.8%; Figure S4). We considered the yields of reagents **1**-**3** as sufficiently good for library synthesis and screens, and thus proceeded with them.

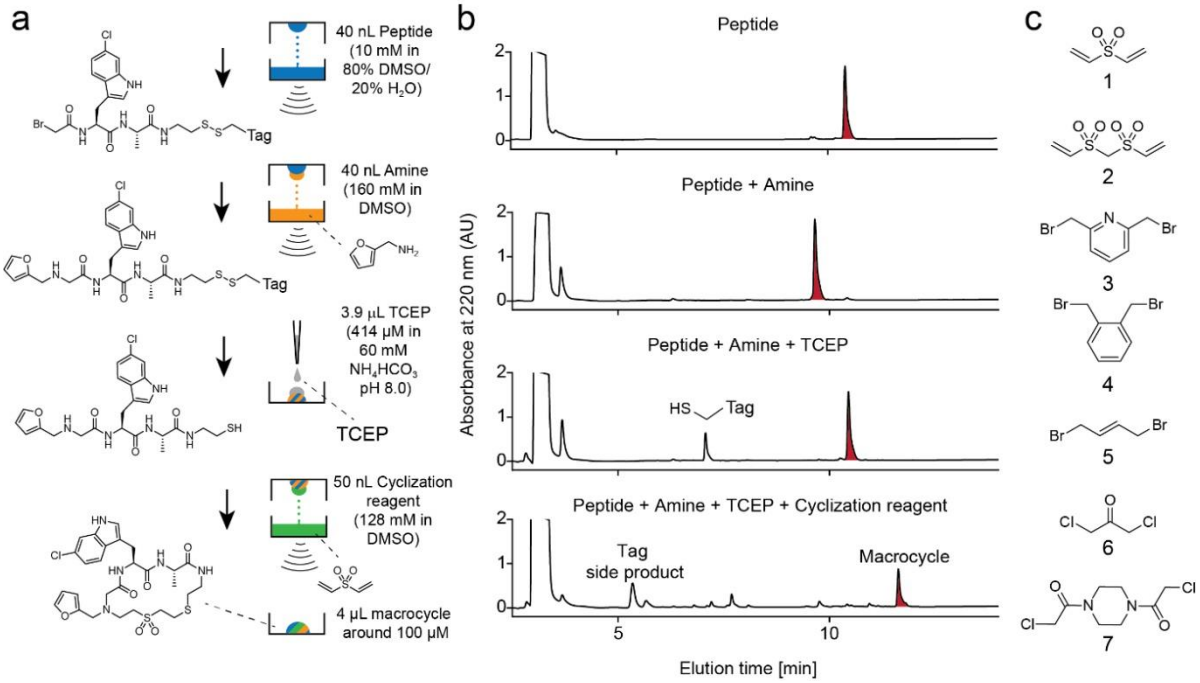


Figure 2. Macrocycle synthesis by acoustic liquid transfer. (a) Representative chemical structures of reagents, intermediates, and the final macrocyclic product. The peptide (blue), amine (yellow), and cyclization reagent (green) are transferred by acoustic sound. TCEP (grey) is transferred by a bulk pipette tip dispenser. (b) HPLC chromatograms of reactions analyzed at various steps during the synthesis. Peaks of desired products are highlighted in red. (c) Variations of bis-electrophile cyclization reagents used in macrocycle synthesis.

3.4.3 Screening of a focused library against MDM2:p53

We next tested whether the macrocycle libraries synthesized by acoustic dispensing were adequate for identifying ligands against a target of interest. As a proof-of-concept, we chose the oncology-driving protein-protein interaction between p53 and MDM2 as a target. Overexpression of MDM2 inhibits the activity of the tumor repressor p53, and MDM2 binders that inhibit the MDM2-p53 interaction have been shown to suppress tumor growth.[18] To

increase the chances of finding hits, we tailored the library design to include the amino acid 6-chloro-tryptophan ($Trp^{6\text{-}Cl}$) in each macrocycle. $Trp^{6\text{-}Cl}$ is a building block that forms the key interaction with MDM2 in short linear and stapled peptides.[19] To ensure it was in every macrocycle, we designed the library with the scheme shown in Figure 3, indicating how $Trp^{6\text{-}Cl}$ was incorporated in either the C-terminal (Library 1) or the N-terminal region (Library 2). To make the complete macrocycle, $Trp^{6\text{-}Cl}$ was accompanied by one of 10 randomly chosen amino acids that differ strongly in their backbones (i.e. a-, b-, g- and e-amino acids; **8**-**17**; blue in Figure 3), one of 45 amine building blocks (**18**-**62**, yellow in Figure 3), and one of three cyclization linkers (**1**-**3**, shown in Figure 2c).

To generate the peptide building blocks in the library, we synthesized and purified 10 bromoacetamide-functionalized peptides of the format BrAc-$Trp^{6\text{-}Cl}$-Xaa-Cys*-S-S-Tag (Library 1; Xaa = any of 10 structurally diverse amino acids; Cys* = cysteamine) and 10 of the format BrAc-Xaa-$Trp^{6\text{-}Cl}$-Cys*-S-S-Tag (Library 2), all with a short, disulfide-linked peptide tag that protected the thiol group and facilitated the purification (Tag = mercaptopropionic acid-Ser-Gly-Arg-Tyr). The 2,700 macrocycles were assembled robotically in 384-well plates by acoustic dispensing using the reagents, volumes, concentrations, and reaction conditions as indicated in Figure 2a. We obtained the final macrocyclic elements in volumes of 4 µL and at 100 µM, the latter concentration hypothetically assuming quantitative conversion of linear peptides to macrocycles.

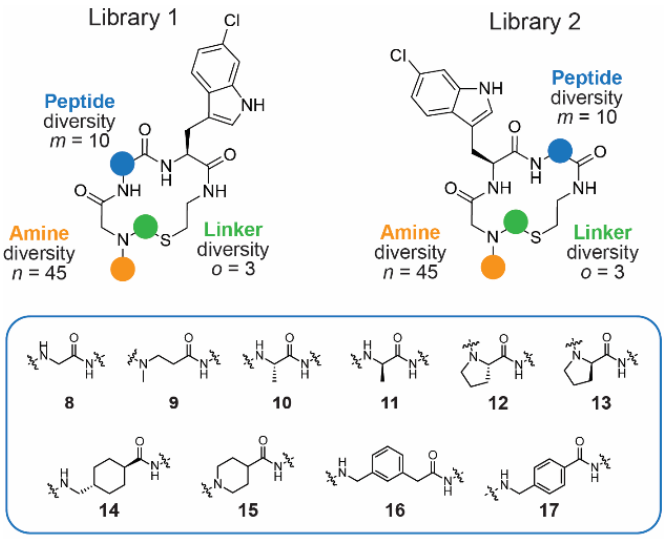


Figure 3. Macrocycle library design and building blocks. All macrocycles contain the amino acid $Trp^{6\text{-}Cl}$ in one of the two possible positions (C-terminal side: Library 1, N-terminal side: Library 2), a variable amino acid 8-17 (blue), an amine 18-62 (yellow), and a linker 1-3 (linkers are shown in Figure 2c). Figure 3 continues next page.

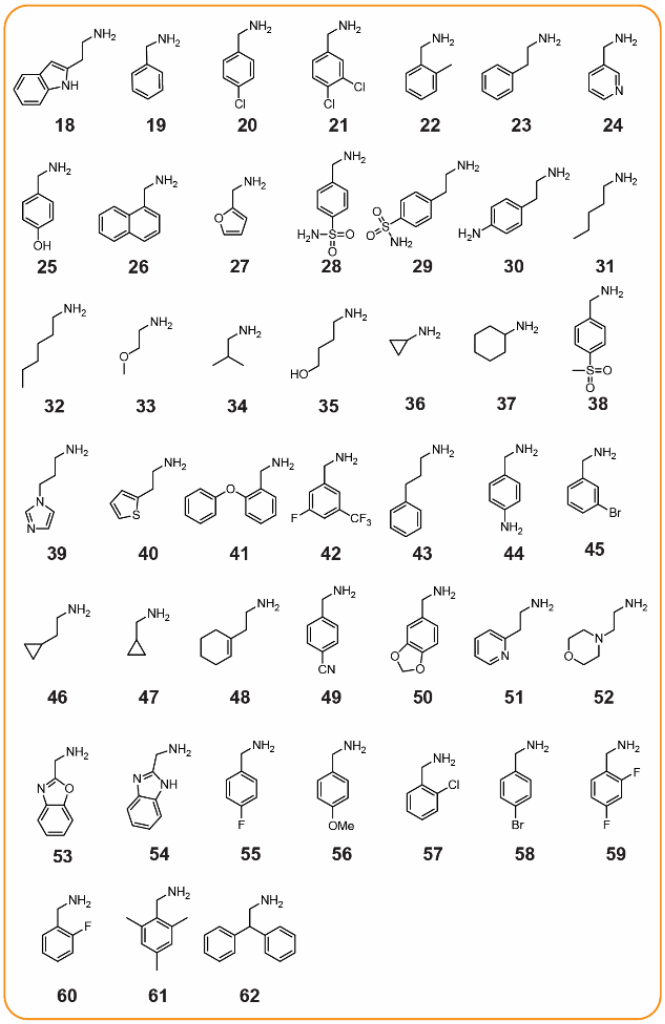


Figure 3 continued.

We screened the library for MDM2-p53 inhibitors by dispensing 11 µL of human MDM2 solution (as GST fusion; Figure S5; 1 µM in assay) and a fluorescent-labeled reporter peptide[20] (K_D = 0.5 µM; Figure S6; 50 nM in assay) to the 384-well synthesis plates containing the 2,700 macrocycles (26.7 µM in assay). Macrocycles that bound MDM2 displaced the reporter peptide and caused a decrease in fluorescence polarization, as shown in Figure 4a. The extent of reporter peptide displacement by the macrocycles is shown for Library 1 in the heat map in Figure 4b (see Library 2 in Figure S7).

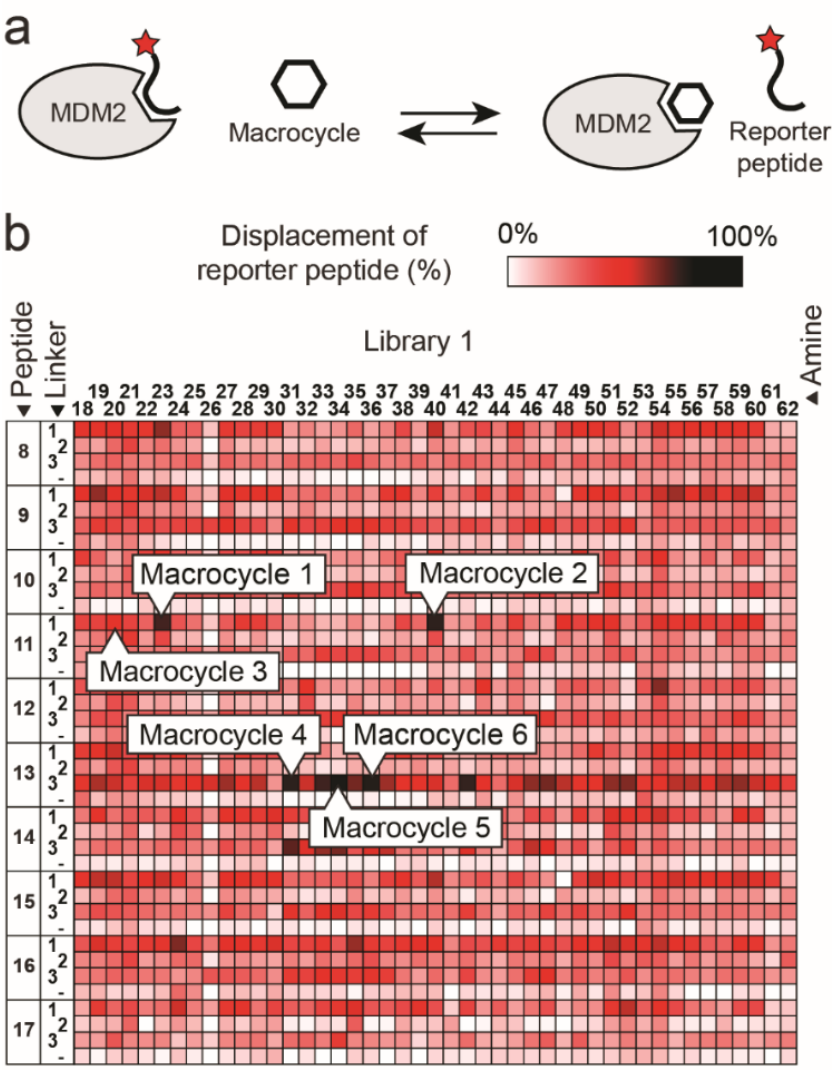


Figure 4. Screen for MDM2-p53 PPI inhibitors. (a) General schematic of the fluorescence polarization competition binding assay. (b) Binding of the 1,350 macrocycles of Library 1 to MDM2. The binding was assessed by displacement of a fluorescein-labeled MDM2-bound reporter peptide using fluorescence polarization. Darker colors indicate higher binding affinity. The building blocks used for library synthesis (shown in Figures 2c and 3) are indicated by number. Control reactions (linear peptides with no linker) are indicated by "-". Lead macrocycles (1 to 3) that were re-synthesized, purified, and characterized are indicated on the heat map. The screening result of library 2 is shown in Figure S7.

Several combinations of BrAc-peptide, amine, and cyclization reagent induced strong displacement of the reporter peptide, as indicated by the dark-red to black colors in the heat map. Control reactions in which no cyclization reagent was added showed weaker signals, which suggests that macrocyclization is necessary for strong binding activity.

To determine the reproducibility of the synthesis and screening assays and to verify that the activities observed during the screen were from the macrocyclic products, we studied macrocycles 1, 2 and 3 in greater depth (Figure 4b and Figure 5a). We repeated twice the synthesis and activity assay with the crude products which yielded the same results (Figure

5b). In addition, we repeated the reactions of all the 17 hits as well as of 16 randomly chosen reactions in triplicate, and found activities comparable to those in the screen (Figure S7d).

Separation of the reaction products by HPLC for an example reaction (macrocycle 1) showed that the desired macrocycle represented the main product (Figure 5c; fraction 9 highlighted in red), and that the fraction containing the macrocycle was active (Figure 5d and Figure S8). This result showed that the activity observed in the screen for this reaction resulted from the macrocycle and not a side product. A comparison of the activity of HPLC-purified macrocycles 1-3 showed that macrocycles 1 (IC_{50} = 2.7 ± 0.3 µM) and 2 (IC_{50} = 4.3 ± 1.5 µM) were more active than 3 (IC_{50} = 25 ± 2 µM), which was in line with the screening result assaying non-purified products (Figure 5a and 5e). From these results, we concluded that acoustic liquid transfer and combinatorial assembly of building blocks in nL volumes is suitable for the synthesis and function-based screening of combinatorial libraries.

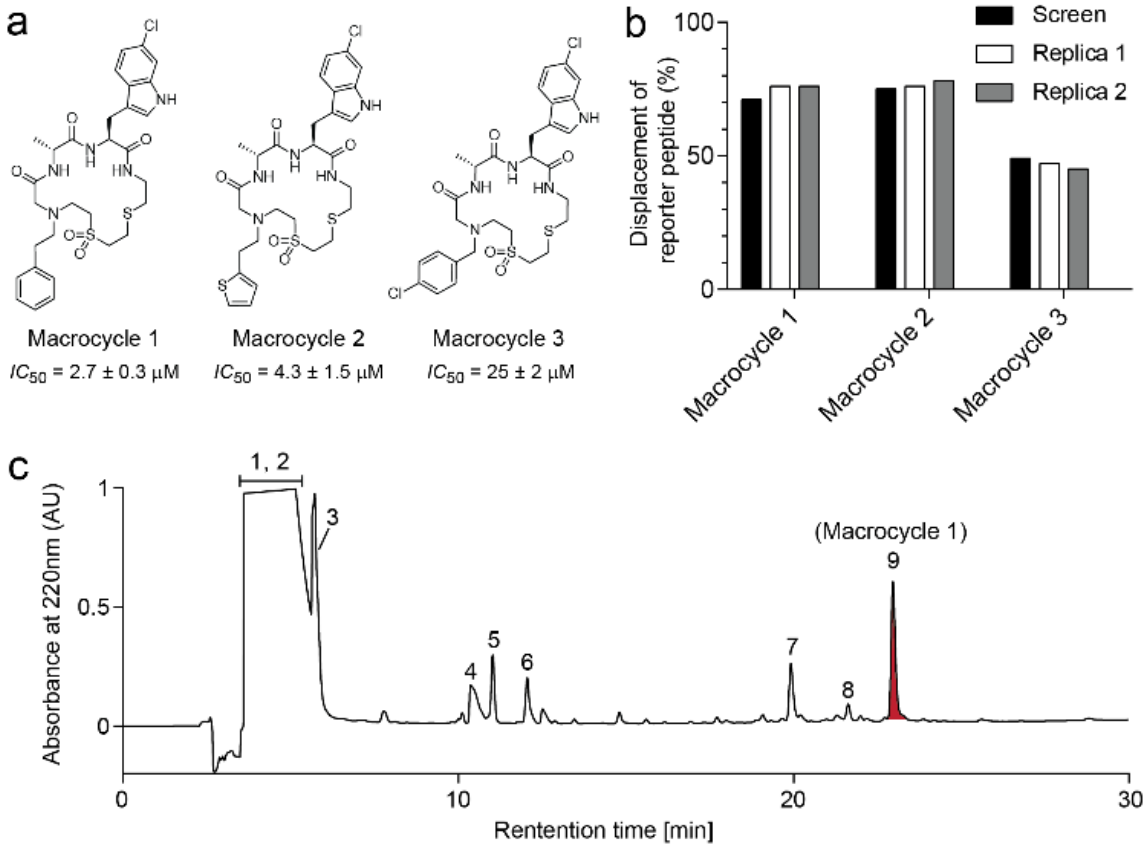


Figure 5. Characterization of hits from screen with Library 1. (a) Structures and IC_{50}s of macrocycles 1-3. Mean values and standard deviations (SDs) of three independent measurements are shown. (b) Macrocycles 1-3 were synthesized and assayed for MDM2 binding as under the screening conditions with little deviation in performance. (c) HPLC-chromatographic separation of the library well containing macrocycle 1. Collected peaks are indicated by numbers. Macrocycle 1 is the product in peak 9 (red). Figure 5 continues next page.

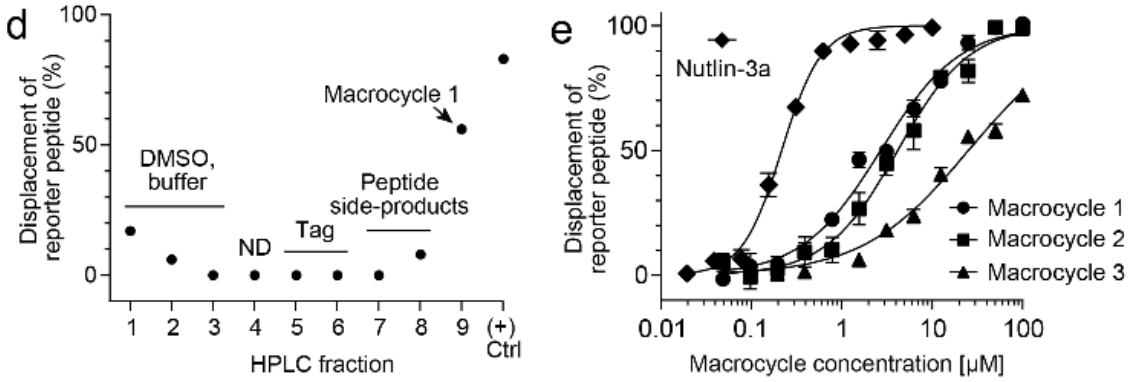


Figure 5 continued. (d) Activity of products in HPLC fractions determined by the MDM2 fluorescence polarization competition assay. Non-labeled reporter peptide was used as a positive control. (e) MDM2 binding of purified macrocycles 1-3 and a control MDM2 binder (nutlin 3a) determined by fluorescence polarization competition assay.

As further validation, we analyzed three additional hits from Library 1 (macrocycles 4, 5, & 6) and 11 hits from Library 2 (Figure S9, S10, & S11). In the reactions for macrocycles 4-6, we found a linear side product that was the most active species and we thus did not further analyze these three macrocycles. The observation that a side product can occasionally be more active than the desired product was previously found in libraries that we synthesized by pipetting,[9] and underscored the importance of hit verification using purified products.

Table 1. Binding affinities of purified macrocycles for MDM2

	Macrocycle	Linker	IC_{50} (μM)
Library 1	1	1	2.7 ± 0.3
	2	1	4.3 ± 1.5
	3	1	25.0 ± 2
	4	3	N.D.
	5	3	N.D.
	6	3	N.D.
Library 2	7	1	14.2 ± 1.9
	8	1	1.3 ± 0.8
	9	1	0.6 ± 0.2
	10	1	2.4 ± 0.9
	11	1	6.9 ± 1.8
	12	2	7.5 ± 0.6
	13	2	8.2 ± 1.1
	14	2	13.1 ± 2.7
	15	2	7.6 ± 2.0
	16	2	3.7 ± 0.5
	17	2	9.8 ± 4.2

IC_{50}s were determined by the MDM2 fluorescence polarization competition assay. Mean values and SDs of three independent measurements are indicated. N.D.: IC_{50}s were not determined as the most active species in the reactions of the screen were side products based on linear peptides.

From Library 2, all 11 hits studied were active macrocycles. Three of them were more potent than the two high-performing hits of Library 1, with the greatest inhibition achieved by macrocycle 9 (IC_{50} = 0.6 ± 0.2 µM), as shown in Table 1. We finally tested the binding of the best two macrocycles by surface plasmon resonance (SPR), which confirmed target-specific binding and dissociation constants in the low micromolar range (macrocycle 8: 11 µM, macrocycle 9: 5.5 µM) (Figure S12).

3.5 Conclusion

In summary, we have developed an efficient approach for the synthesis of combinatorial macrocycle libraries at a picomole scale by sequential transfer of three building blocks in nL-droplets using acoustic liquid dispensing. While ADE has previously been used for combinatorial reaction scouting and screening at a nanomole scale, our work is a *de novo* example of combinatorial library generation and screening at a picomole scale, illustrated with macrocycles. For classical HTS approaches, library compounds are often produced in milligram quantities, corresponding to a micromole scale, which is a million fold more than in the herein presented picomole scale synthesis and screening approach. We have demonstrated the efficiency of combinatorial library synthesis with ADE and identified low- to sub-micromolar inhibitors of the MDM2-p53 oncology target with target-focused macrocycles that contain $Trp^{6\text{-}Cl}$ to drive affinity for MDM2. While we synthesized and screened in 384-well plates, the ADE device used herein is also compatible with 3,072-well plates. In principle, macrocycle synthesis and screening can be conducted at even higher throughput and smaller scale, which will further reduce reagent use, material consumption, time per synthesis and increase library size and diversity. There are additional advantages of the ADE-based library approach, such as the omission of pipetting tips, which reduces the cost by around 0.3 cents per macrocycle and thus a large sum for libraries comprising thousands of macrocycles. Additionally, the innovative use of microwell plates for overall screening and sampling eliminates the need for large library storage and only requires preservation of the building blocks, which use a much smaller storage space. Taken together, this work provides a general strategy for the rapid combinatorial synthesis of libraries at the picomole scale to generate macrocyclic ligands with potential therapeutic applications.

3.6 Materials and methods

Merging reagents by acoustic liquid transfer

For testing the mixing of reagents and merging of droplets, liquids were transferred using an ECHO 550 or ECHO 650 device from Labcyte. All liquids were transferred in droplets of 2.5 nL using settings for DMSO or aqueous solvent transfer. For testing the mixing of transferred compounds, liquids saturated in fluorescein or Coomassie Brilliant Blue in DMSO were transferred from a 384-well ECHO low dead volume source plate (Labcyte 384LDV, Cat. # LP-0200) to a 96-well destination plate (Greiner Bio-One, 96 well, clear, flat bottom, polystyrene microplate). The droplets were visualized and imaged with a microscope (10×, EVOS FL Auto imaging, ThermoFisher Scientific).

For testing the transfer and merging of 40 nL droplets in microtiter plates, saturated solutions of fluorescein (80% [v/v] DMSO/ 20% [v/v] mQ water) and Coomassie Brilliant Blue (DMSO) were transferred from a 384-well ECHO low dead volume source plate to a 384-well low volume destination plate (ThermoFischer NUNC™ 384 shallow well std height plate, non-sterile, black, Cat. # 264705) using settings for DMSO transfer. The droplets were transferred in the following order: first 40 nL of fluorescein followed by 40 nL Coomassie Brilliant Blue. For enforcing merging of non-merged droplets, the plate was centrifuged at 2,000 × g (force measured at 11.5 cm from the centrifuge axis) corresponding to around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge and a Sorvall swing-out rotor. The droplets were imaged using a digital camera (Olympus OMD EM-5, 16 megapixel, camera with an Olympus M. Zuiko Premium 60/2.8 ED Macro lens and settings of 60 mm, F/13, 1/6 s, ISO 200).

Synthesis of thiol-protected and N^{α}-bromoacetyl-functionalized peptides

Peptides of the format BrAc-Xaa-Yaa-Cys*-S-S-Tag (Xaa and Yaa are the variable amino acids, Cys* = cysteamine. Tag = mercaptopropionic acid-Ser-Gly-Arg-Tyr) were synthesized on solid phase using Rink amide MBHA resin (50 μmol scale, loading of 0.3 mmol/g) in 5 mL syringe reactors using an automated peptide synthesizer (Intavis, MultiPep RSi). In all washing steps, 2 mL of DMF was used and the resin shaken for 1 min at 400 rpm.

The resin was swollen by adding three times 2.4 mL of DMF for 10 min. The amino acids (4.2-fold equiv., 0.22 M final conc.), the coupling agent HATU (3.9 equiv., 0.21 M final conc.) and the base NMM (10 equiv., 0.53 M final conc.), all in DMF, were mixed, 5 μL of NMP added to improve the solubility, incubated for one minute, transferred to the solid phase, and shaken for 45 min at 400 rpm. All Fmoc amino acids and the Trt-protected mercaptopropionic acid were coupled twice. The resin was washed 7 times with 2 mL DMF. Fmoc protecting groups

were removed by performing two cycles of 0.8 mL 20% (v/v) piperidine in DMF for 5 min. After conjugation of mercaptopropionic acid, the resin was washed three times with 600 μL DCM.

The trityl protecting groups were removed by incubating the resin twice for 15 min at 400 rpm with 4 mL of 10% (v/v) TFA, 2.5% (v/v) TIS and 87.5% (v/v) DCM. The resin was washed 5 times with 3 mL DCM. The trityl-deprotection step and the subsequent washing steps were performed on a TFA-compatible cleavage station.

The disulfide bridge was established by incubating the resin twice with 3 equiv. 2-pyridylthio cysteamine (1 mL, 0.15 M) in 30% (v/v) MeOH, 69.9% (v/v) DCM and 0.1% (v/v) AcOH for one hour at 400 rpm. 2-pyridylthio cysteamine was prepared as described below. The resin was washed twice with 3 mL 50% (v/v) DCM/ 50% (v/v) MeOH and twice with 3 mL DCM. For ensuring deprotonation of the cysteamine amino group, the resin was incubated with 200 μL 4 M NMM in DMF for 1 min and washed with DMF. The two variable amino acids were introduced using the coupling conditions described above.

The bromoacetyl group was introduced by incubating the resin three times with 5 equiv. bromoacetic acid (0.6 mL, 0.5 M) and 5 equiv. *N,N'*-diisopropylcarbodiimide (0.6 mL, 0.5 M) in DMF for 30 min at 400 rpm. The resin was washed three times with 2 mL DMF followed by two times with 0.6 mL of DCM.

The protecting groups were removed and the peptides cleaved from the resin by incubating in 5 mL of 95% (v/v) TFA, 2.5% (v/v) mQ water and 2.5% (v/v) TIS for 90 min at 400 rpm. The cleavage was performed on a TFA-compatible cleavage station. Around 4 mL of the TFA-cleavage solution was evaporated under a flow of nitrogen. The peptides were precipitated in 50 mL cold diethyl ether for 1 hr at -20°C and centrifuged at 3,400 × g (around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 10 min at 4°C. The supernatant was discarded.

For HPLC-purification, the solid was dissolved in 10% (v/v) ACN, 89.9% (v/v) mQ water and 0.1% (v/v) TFA and run over a reverse phase C18 column (Waters Sunfire™, 10 μm, 100 Å, 19 × 250 mm) using a linear gradient of solvent B (99.9% [v/v] ACN, 0.1% [v/v] TFA) over solvent A (99.9% [v/v] mQ water, 0.1% [v/v] TFA) from 10 to 34% in 34 min at a flow rate of 20 mL/min. The mass of the purified peptides was determined by electrospray ionization mass spectrometry using a Shimadzu-2020 single quadrupole LC-MS instrument. Fractions with the desired peptides were lyophilized. White powders were obtained and considered to be TFA salts.

Synthesis of 2-pyridylthio cysteamine

Cysteamine (1 equiv. in 30 mL MeOH, 0.39 M), 2,2'-dipyridyldisulfide (2 equiv., in 70 mL MeOH, 0.77 M) and AcOH (2 ml) were incubated overnight at RT under agitation. After removing the solvent by rotation evaporation, the 2-pyridylthio cysteamine product was precipitated by addition of 50 mL of cold diethyl ether, incubated for 30 min at -20°C and pelleted by centrifugation at 3,400 × g for 5 min at 4°C (around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor). The product was further purified in 7 cycles of dissolving the precipitate in 5 mL of MeOH and precipitation with 45 mL of cold diethyl ether and centrifugation. After the final centrifugation step, the white powder was dried by lyophilization. The resulting product was considered to be a chloride salt.

LC-MS analysis

Peptides and macrocycles were analyzed by a UHPLC and single quadrupole MS system (Shimadzu-2020). Samples were applied on a reversed-phase C18 column (Phenomenex Kinetex®, 2.6 μm, C18, 100 Å column, 50 × 2.1 mm) using a linear gradient of solvent B (99.95% [v/v] ACN and 0.05% [v/v] formic acid) over solvent A (99.95 % [v/v] mQ water and 0.05%[v/v] formic acid) from 0 to 100 % in 10 min at a flow rate of 1 ml/min using positive and negative modes for mass analysis.

Analytical HPLC

The purity of the peptides and macrocycles was analyzed by reversed-phase analytical HPLC (Agilent Technologies, 1260 Infinity). Fractions eluted from preparative HPLC (100 μL) were run over a reversed-phase C18 column (ZORBAX 300SB-C18, 5 μm, 4.6 mm × 20 mm) using a linear gradient of solvent B (99.9% [v/v] ACN, 0.1 % [v/v] TFA) over solvent A (94.9% [v/v] mQ water, 5% [v/v] ACN, 0.1% [v/v] TFA) from 0 to 100% in 15 min at a flow rate of 1 mL/min.

Macrocycle synthesis by acoustic liquid transfer

Acoustic liquid transfer was performed on an ECHO 550 from Labcyte/Beckman Coulter at the Biomolecular Screening Facility (BSF) of EPFL. All transfers were performed using DMSO transfer settings. ECHO-qualified 384-well low dead volume plates were used as source plates for peptides and amines. ECHO-qualified 384-well polypropylene plates (Labcyte 384 PP, Cat. # PP-0200) were used as source plates for linker reagents. Low volume 384-well plates (ThermoFischer NUNC™ 384 shallow well std height plates non-sterile, black, no lid; Cat. # 264705) were used as destination plates.

The building blocks were prepared in the following solvents and at the following concentrations, and transferred to 384-well source plates. N^{α}-bromoacetyl peptide linear precursors was prepared at 10 mM in 80% (v/v) DMSO and 20% (v/v) mQ water, and 10 μL were transferred to a 384-well low dead volume plate. Amines were prepared at 160 mM in DMSO and 10 μL were transferred to a 384-well low dead volume plate. Primary amines purchased as HCl salts were supplemented with NaOH (200 mM final concentration; 400 mM for amine **54**). Cyclization reagents were prepared at 128 mM in DMSO and 50 μL were transferred to an ECHO-qualified 384-well plate.

For the N-alkylation reactions, 40 nL of peptide (1 equiv.) and 40 nL of amine (16 equiv.) were transferred in droplets of 2.5 nL. The plates were sealed with a plastic foil, centrifuged for 2 min at RT to ensure complete droplet merging (130 × g, corresponding to 1,000 rpm in a Sigma 2-6 centrifuge with a 11121 rotor) and incubated for 1 hr at 37°C.

For the removal of the thiol protecting group (the disulfide-linked tag peptide), 3.87 μL of buffer containing 60 mM NH_4HCO_3, pH 8.0, and 414 μM TCEP (4 equiv.) were transferred from a 384-well source plate with a liquid handling workstation (Caliper Sciclone ALH3000) using a 384 pipetting head and 25 μL disposable plastic tips. The plates were sealed with a plastic lid, centrifuged to ensure proper mixing (130 × g, 2 min, as above), and incubated 1 h at 37°C.

For peptide cyclization, 50 nL of the bis-electrophile linker reagents (16 equiv.) were transferred to the destination plates. The plates were sealed with a plastic lid, centrifuged as described above, and incubated 2 hrs at 37°C. The reactions had volumes of 4 μL and contained 100 μM peptide, 100 μM Tag, 1.6 mM primary amines, 400 μM TCEP and 1.6 mM linker. The final solvent was 60 mM NH_4HCO_3 pH 8.0 containing 3.05% (v/v) DMSO.

For HPLC and mass spectrometric analysis of test reactions performed with peptide BrAc-$Trp^{6\text{-}Cl}$-Ala-Cys*-S-S-Tag, 2-furanmethanamine, and linker reagents **1** to **7**, identical reactions were performed at the same scale as described above in 24 independent wells, the products pooled, and analyzed by analytical HPLC (85 μL of pooled sample) or LC-MS (10 μL of pooled sample) as described above. Yields and quantities of side products were determined from peak analysis by LC-MS at 220 nm considering area under the peaks of peptide derived products.

Cloning of MDM2 expression vector

The following DNA coding for the amino acids 17 to 125 of human MDM2 (hMDM2; in bold), containing also suitable DNA restriction sites (underline) for cloning into the bacterial

expression vector pGEX-4T3 and expression as GST fusion, and sequences for primer annealing (italics), was ordered from Eurofins.

5'-
*ATTCTATGCGGCCCAGC*GGATCC**AGCCAGATTCCGGCAAGCGAACAAGAAACCCTGG**
TTCGTC
CGAAACCGCTGCTGCTGAAACTGCTGAAAAGCGTTGGTGCACAGAAAGATACCTATA
CCATGAAAGAGGTGCTGTTTTATCTGGGCCAGTATATTATGACCAAACGCCTGTATGA
TGAGAAACAGCAGCATATTGTGTATTGCAGCAATGATCTGCTGGGTGACCTGTTTGGT
GTTCCGAGCTTTAGCGTTAAAGAACACCGTAAAATCTATACCATGATCTATCGTAATC
TGGTGGTGGTTAATCAGCAAGAAAGCAGCGATAGCGGCACCAGCGTTAGCGAAAAC
TAAGAATTC*GGTTCTGGCGCTGAAA*-3'

The DNA was amplified by PCR using Phusion polymerase using the forward primer 5'-ATTCTATGCGGCCCAGC-3' and the reverse primer 3'-GGTTCTGGCGCTGAAA-5', and the DNA purified with a PCR purification kit (QIAquick® MinElute ® PCR Purification Kit, Qiagen).

The hMDM2 DNA (3.7 μg) and vector pGEX-4T3 (9.3 μg) were digested using *BamH*I (4 μL, 20,000 U/ml) and *EcoR*I (2 μL, 20,000 U/ml) in *EcoR*I buffer for 2 hrs at 37°C in a volume of 100 μL. The digested vector was additionally treated with FastAP phosphatase (13 μL, 1 U/μL) for 10 min at 37°C to prevent circularization. The cleaved PCR product and vector DNA were electrophoresed on an agarose gel using 1% (w/v) high quality agarose and TAE buffer containing 0.5 μg/mL of ethidium bromide and 1 mM guanosine. DNA products with the desired size were cut out using a scalpel under UV light (254 nm, 70% intensity) and purified by a gel extraction kit (QIAquick® Gel Extraction Kit, Qiagen). Concentrations were determined by measuring the absorption at 260 nm (NanoDrop 8000, Thermo Scientific).

Cleaved hMDM2 gene (443 ng) and pGEX-4T3 vector (446 ng) were ligated using T4 ligase (2 μL, 400,000 U/ml) in T4 ligase buffer in a volume of 21.2 μL for 4 hrs at RT. Completion of the ligation was analyzed by agarose electrophoresis and purified with a PCR purification kit.

The ligated vector (1 μL) was electroporated into 100 μL electrocompetent DH5$^{\alpha}$ *E.coli* cells using 1 mm electroporation cuvettes (EP-201, Cellprojects) and an electroporation device (MicroPulser, BIO-RAD). After electroporation, 1 mL of LB (Luria Broth) medium was immediately added to the cuvette, transferred to a sterile 14 mL tube, and incubated for 1 hr at 37°C under agitation (250 rpm). 100 mL of the culture were plated on LB agar plates containing ampicillin (100 μg/mL). After incubation overnight at 37°C, single colonies were

picked with sterile plastic tips to inoculate 5 mL of LB containing 100 μg/mL ampicillin in 50 mL Falcon tubes. After overnight incubation at 37°C and agitation at 250 rpm, cells were pelleted by centrifugation at 10,000 × g (around 11,000 rpm on an Eppendorf Centrifuge 5418) for 5 min at 4°C. Plasmid DNA was isolated using a plasmid miniprep kit (NucleoSpin® Plasmid Miniprep Kit, Qiagen). The DNA was sequenced (Microsynth AG) and a clone with the desired sequence was identified.

Sequence of hMDM2-GST

The DNA sequence was confirmed by Sanger sequencing using the primer Gex-5-for (5'-CCAGCAAGTATATAGCATGG-3') as is shown in Figure S5a. Expression in *E.coli* yielded the GST-hMDM2 fusion protein shown in Figure S5b.

Recombinant expression of human GST-hMDM2

The vector for expressing GST-hMDM2 was inserted into electrocompetent *E.coli* BL21 (DE3) cells by electroporation and plated on agar plates containing ampicillin (100 μg/mL). A 5 mL culture containing LB medium and ampicillin (100 μg/mL) was inoculated with cells of a single colony and incubated overnight at 37°C with agitation at 250 rpm. A culture flask containing 1 L LB medium and ampicillin (100 μg/mL) was inoculated with 5 ml cells of the pre-culture and incubated at 37°C under agitation at 250 rpm until OD_{600} reached 0.8 (around 4 h). The culture was cooled on ice for 10 min and protein expression induced by addition of IPTG to a final concentration of 1 mM. The culture was incubated for 5 hrs at 25°C under agitation at 250 rpm. The culture was centrifuged for 5 min 5,500 × g (around 5,000 rpm on a Thermo Scientific SORVALL RC BIOS centrifuge with a Fiberlite F8-6x1000y rotor) at 4°C. The supernatant was discarded, the pellet transferred to a 50 mL Falcon tube, flash-frozen in liquid nitrogen, and stored at -80°C.

The pellet was thawed on ice, suspended in cold lysis buffer containing 10 mM Na_2HPO_4, 1.8 mM $NaH_2PO_4*H_2O$, 137 mM NaCl, 2.7 mM KCl and 4 mM DTT, pH 7.4, supplemented with lysozyme (100 μg/mL), DNaseI (13 U/ml) and PMSF (0.2 mM). After 30 min incubation on ice, the cells were sonicated (SONICS Vibra cell; Method: Pulse 15 s on/ 45 s off, Amp 35%, 10 min) and centrifuged for 1 hr at 15,000 × g (around 10,000 rpm on a Beckman Coulter™ Allegra™ 25R using a TA-10-250 rotor) at 4°C. The supernatant was collected in a 50 mL Falcon tube and run over two connected 1 mL glutathione sepharose columns (GSTrap FF, GE Life Sciences) using a peristatic pump (BIORAD Econo Pump) at a flow rate of 1 mL/min. After washing with 10 column volumes of cold washing buffer (10 mM Na_2HPO_4, 1.8 mM $NaH_2PO_4*H_2O$, 137 mM NaCl, 2.7 mM KCl and 4 mM DTT, pH 7.4), the protein was eluted

with 5 column volumes of elution buffer (50 mM Tris-HCl, pH 8.0, 4 mM DTT and 20 mM reduced glutathione). Protein in the fractions was denatured and reduced with loading buffer (5-fold loading buffer containing 2% [w/v] SDS, 25% [v/v] glycerol, 0.05% [w/v] bromophenol blue, 60 mM Tris-HCl pH 6.8 supplemented with pure β-Mercaptoethanol to a final concentration of 10 mM prior use), analyzed by SDS-PAGE using a 12% (w/v) RunBlue SDS protein gel and standard SDS running buffer (40 mM tricine, 60 mM TRIS, 0.1% (w/v) SDS, 2.5 mM sodium meta bisulfite, pH 8.4). The proteins in the gel were stained for 2 hrs with 0.1% (w/v) Coomassie R-250 in 50% (v/v) methanol, 10% (v/v) glacial acetic acid and 40% (v/v) mQ water, and de-stained overnight using de-staining solution containing 50% (v/v) mQ water, 40% (v/v) methanol and 10% (v/v) acetic acid. The fractions containing pure fusion protein were combined, the volume reduced by concentration to around 2 mL, and buffer exchanged with cold PBS by 5 cycles of adding 20 mL PBS and centrifugational concentration at 3,400 × g (around 4,000 rpm with a Thermo Scientific Multifuge 3L-R centrifuge and a Sorval 75006445rotor) at 4°C with a Macrostep® Advance Centifugal Device having a 10,000 MWCO.

The protein concentration was determined by measuring absorbance at 280 nm ($MW_{GST\text{-}MDM2}$ = 38.8 kDa, $\varepsilon_{280\,nm}$ = 53,290 $M^{-1}cm^{-1}$). Protein purity was assessed by SDS-PAGE as described above. Around 80 aliquots of 50 μL with a concentration of 1.39 mg/mL were prepared, flash-frozen with liquid nitrogen, and stored at -80°C.

Synthesis of fluorescein-labeled MDM2-binding peptide

A peptide with the amino acids 15-29 of p53 (SQETFSDLWKLLPEN) labeled with fluorescein at the N-terminus spaced by a linker (GSGS) was synthesized as follows. The peptide GSGSSQETFSDLWKLLPEN-NH_2 was synthesized by standard Fmoc solid-phase chemistry at a scale of 25 μmol on Rink amide MBHA resin (loading: 0.3 mmol/g). In the last step of synthesis, 5(6)-carboxyfluorescein was coupled to the N-terminus as described before. The peptide was cleaved, precipitated with diethyl ether, and purified by RP-HPLC.

Measuring binding affinity of reporter peptide by fluorescence polarization

The binding of the reporter peptide was determined by incubating the peptide (50 nM final concentration) with 2-fold dilutions of GST-hMDM2 (form 25 to 0.01 μM final concentration) in assay buffer (10 mM Na_2HPO_4, 1.8 mM NaH_2PO_4*H_2O, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 0.01% [v/v] Tween-20) in a 384 well plate (ThermoFischer NUNC™ 384 shallow well std height plates non-sterile, black), incubation for 10 min at RT, and measuring fluorescence anisotropy (Infinite F2000 pro TECAN, E_{ex}= 485 nm, E_{em} = 535 nm, 25°C). The G-factor was calculated using 15 μL of fluorescein at 50 nM in assay buffer as a reference (minimum of polarization),

and assay buffer as a blank. The K_D was calculated using Graphpad Prism 5 using the following formula (1):

$$y = a + (b - a)\frac{(K_d + x + P) - \sqrt{(K_D + x + P)^2 - 4xP}}{2P} \qquad (1)$$

Where y is the anisotropy, a the probe signal in absence of fluorescent ligand (blank), b probe signal in the presence of saturating concentrations of ligand (reference), x and P the protein and probe concentrations respectively.

Establishing fluorescence polarization competition assay

For establishing the MDM2 fluorescence polarization competition assay, we synthesized the MDM2-binding peptide described above but without fluorescein (Ac-GSGSSQETFSDLWKLLPEN-NH_2) to use it as a control that was expected to displace the fluorescent probe (control peptide). Fluorescent reporter peptide (50 nM final concentration) bound to around 70% by GST-MDM2 (1 μM final concentration) were incubated with 2-fold dilutions of control peptide (20 to 0.02 μM final concentration) in assay buffer in a 384 well plate, incubated for 30 min, and fluorescence anisotropy measured as described above. The fluorescent reporter peptide and the control peptide were initially dissolved in DMSO, which added a small volume of DMSO. All wells were therefore complemented with DMSO to reach a final concentration of 1%.

The data were normalized considering negative control (no control peptide) and reporter probe alone (no protein) as 0 and 100% displacement of the reporter peptide, respectively. Sigmoidal curves were fitted to the data using Graphpad Prism 5 software and the following dose-response equation (2):

$$y = \frac{100}{1 + 10^{(\log IC_{50} - x)p}} \qquad (2)$$

Where y = displacement of reporter probe (%), x = peptide concentration and p = Hill slope. IC_{50} indicates the functional strength of the inhibitor (concentration resulting in 50 % displacement of reporter probe) and were derived from the fitted curve.

Screening macrocycle compound library

Macrocycles were synthesized in 384 well plates using ADE exactly as described above. To eight dedicated wells of each 384 well plates, 4 μl of positive control peptide (37.5 μM) was added manually. To eight other wells, only reaction buffer, TCEP and DMSO were dispensed as negative control. The other wells were filled combinatorially with the reagents to assemble

the macrocycles, wherein one or two peptides were fit into one 384 well plate, along with all amines and linkers.

GST-hMDM2 and the fluorescent reporter probe were diluted using screening buffer (100 mM Na_2HPO_4, 18.5 mM $NaH_2PO_4.H_2O$, 137 mM NaCl, 2.7 mM KCl + 0.01% [v/v] Tween-20, pH 7.4) and mixed. The dilutions were chosen so that concentrations of 1.36 μM and 68 nM, respectively, were reached. After incubation for 10 minutes, 11 μl of the reagent mix was added to the each well of the assay plates using a bulk microplate dispenser (MultiFlo, BioTek) and a 1 μl cassette. Final volumes were 15 μl/well containing 1 μM GST-hMDM2, 50 nM fluorescent reporter peptide, and 26.7 μM macrocycle (assuming 100% yield). Final DMSO content was 0.8% (v/v). Final concentration of positive control was 10 μM. A plastic lid was manually added and the plates were incubated for 30 min at RT. The fluorescence anisotropy was recorded as described above.

Calculating percent of reporter peptide displacement

The raw anisotropy data was converted into % of reporter peptide displacement using the following equation wherein A is anisotropy:

$$\text{Displacement of reporter peptide (\%)} = 100 \times \frac{\text{A (negative control)} - \text{A (macrocycle reaction)}}{\text{A (negative control)} - \text{A (positive control)}} \qquad (3)$$

The A (negative control) is the anisotropy in absence of a competitor. The A (positive control) is the anisotropy in absence of MDM2. On each 384-well plate, eight wells were dedicated for each one of the two controls and mean values were calculated and used in the equation above. For presenting the data in the heat maps shown in Figure 4b and Figure S7, the values (% of reporter peptide displacement) were converted into colors applying the color gradient shown in Figure 4b and using Excel.

Hit confirmation

Macrocycle hits showing strong displacement of reporter peptide were synthesized at a 10 nmol scale, which was 25-fold higher than in the screen. All reagents were transferred by pipetting to an Eppendorf tube which was centrifuged after each transfer. The following synthesis describes the procedures for the 10 nmol scale. The N_α-bromoacetyl peptides (1 μl, 10 mM, 80% [v/v] DMSO and 20% [v/v] mQ: H_2O) were mixed with primary amines (1 μl, 160 mM in DMSO) and incubated at 37°C for 1 h. A volume of 96.7 μl TCEP (60 mM NH_4HCO_3, 414 μM TCEP, pH 8.0) was added and incubated at 37°C for 1 h for deprotecting the thiol group. The cyclization was performed by adding bis-electrophile reagent (1.25 μl, 128 mM in

DMSO) and incubating for 2 h at 37°C. Negative controls were performed using the solvent and buffer only following the same protocol. The binding of the reaction products to hMDM2 was measured using the fluorescence polarization competition assay as described above.

Identification of active compounds in reactions

Macrocycle were synthesized at a 50 nmol scale essentially as described for the 10 nmol scale above, which was 125-fold higher than in the screen. The products obtained were separated on a reverse phase C18 column (Nova-Pak® C18, 6 µm, 7.8 x 300mm) using UHPLC (Thermo Scientific, UltiMate 3000, *Flow cell*: SST, 13 µL, 10 mm, 120 bar) and a linear gradient of solvent B (99.9% [v/v] ACN and 0.1% [v/v] TFA) over solvent A (99.9% [v/v] mQ:H_2O and 0.1% [v/v] TFA) from 0 to 80% in 30 min at a flow rate of 4 ml/min. Fractions of peaks were lyophilized and the white powders dissolved in 12.5 µL DMSO followed by addition of 238 µL of mQ: H_2O. The activity of the fractions was tested using the fluorescence polarization competition assay as described above.

Preparative macrocycle synthesis

Macrocycles were produced at a 25 µmol scale without the tag on cysteamine 4-methoxytrityl resin (Novabiochem®, loading: 0.92 mmole/g) and the N-alkylation step was performed on solid phase. Fmoc solid phase synthesis and bromoacetylation were performed as described above. Primary amines were incubated with the resin (10 equiv. in DMF, 0.5 mL, 0.5 M) for 1 h under agitation. The resin was washed three times with 3 ml DMF and twice with 3 ml of DCM. Cleavage and ether precipitation were performed as described above. The resulting white powder was solubilized in a small volume of DMSO, complemented with buffer (60 mM NH_4HCO_3, pH 8.0), and cyclized with 4 equiv. of bis-electrophile reagent. After 2 h incubation at 37°C, the mixture was lyophilized. The resulting solid powder was solubilized in 10% (v/v) DMSO, 20% (v/v) ACN, 69.9% (v/v) mQ: H_2O and 0.1% (v/v) TFA, and purified by RP on a C18 column (Nova-Pak® C18, 6 µm, 7.8 x 300mm) by UHPLC (Thermo Scientific, UltiMate 3000, *Flow cell:* PEEK, 0.7 µL, 0.4 mm, 100 bar) using a linear gradient of solvent B (99.9% [v/v] ACN and 0.1% [v/v] TFA) over solvent A (99.9% [v/v] mQ: H_2O and 0.1% [v/v] TFA) from 20 to 70% in 30 min at a flow rate of 4 ml/min. After lyophilization the macrocycles were dissolved in DMSO and adjusted to a concentration of 10 mM. The mass of the purified peptides was determined by electrospray ionization MS (ESI-MS) as described above.

Measuring macrocycle binding by SPR

The binding affinity of macrocycles for immobilized MDM2 was measured by surface plasmon resonance (SPR). The experiments were performed using a BiacoreTM 8K instrument (GE Healthcare). GST-hMDM2 fusion protein (10 μg/mL) was dissolved in 10 mM MES buffer (pH 6.0) and immobilized on a CM5 series S chip by a standard amine coupling method in running buffer (10 mM Na_2HPO_4, 1.85 mM $NaH_2PO_4 \cdot H_2O$, 137 mM NaCl, 2.7 mM KCl and 0.005% (v/v) Tween-20). The immobilization level was around 3000 resonance units (RUs). The reference cell was treated the same way without GST-hMDM2. For the measurement of binding kinetics and dissociation constants, five serial dilutions (4-fold) of macrocycles were prepared in running buffer containing 0.4% (v/v) DMSO and analyzed in single cycle kinetics mode with the contact and dissociation times of 60 s and 120 s, respectively.

3.7 Supplementary information

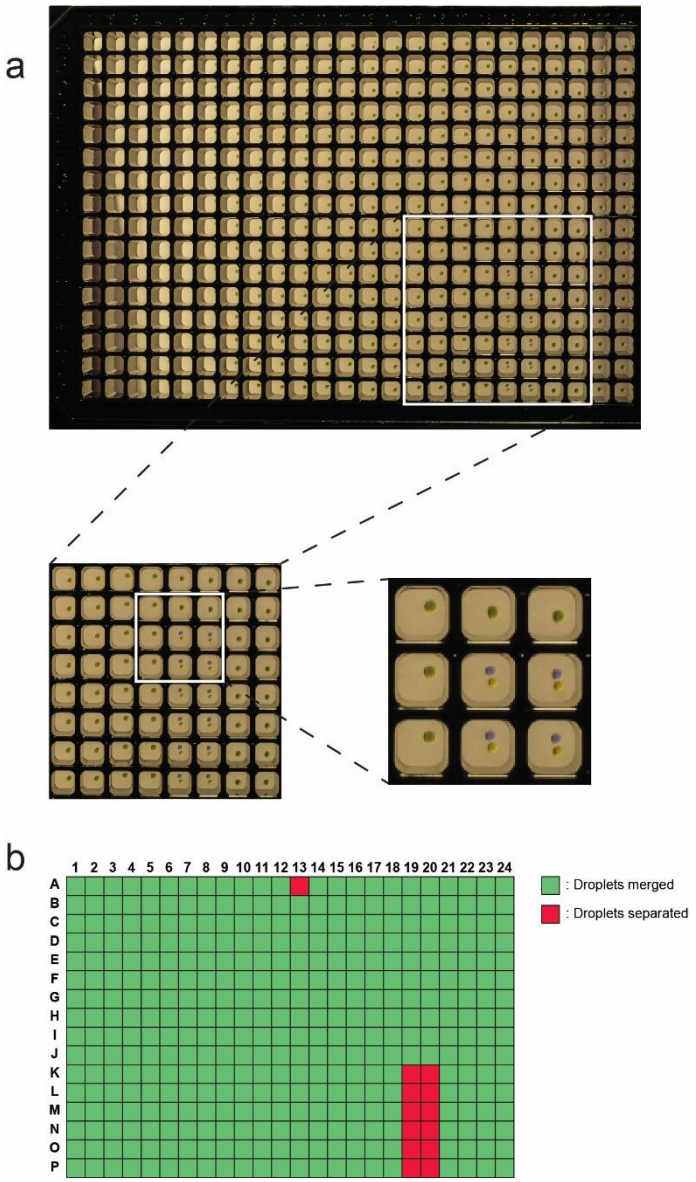


Figure S1. Merging of droplets transferred by acoustic dispensing in a 384-well microtiter plate. (a) Photo of a microtiter plate to which 20 nL Coomassie and fluorescein droplets were transferred. Below is shown an enlarged region in which the two droplets did not merge in four of the nine wells. (b) Indication of wells in which droplets were merged (green) or not merged (red).

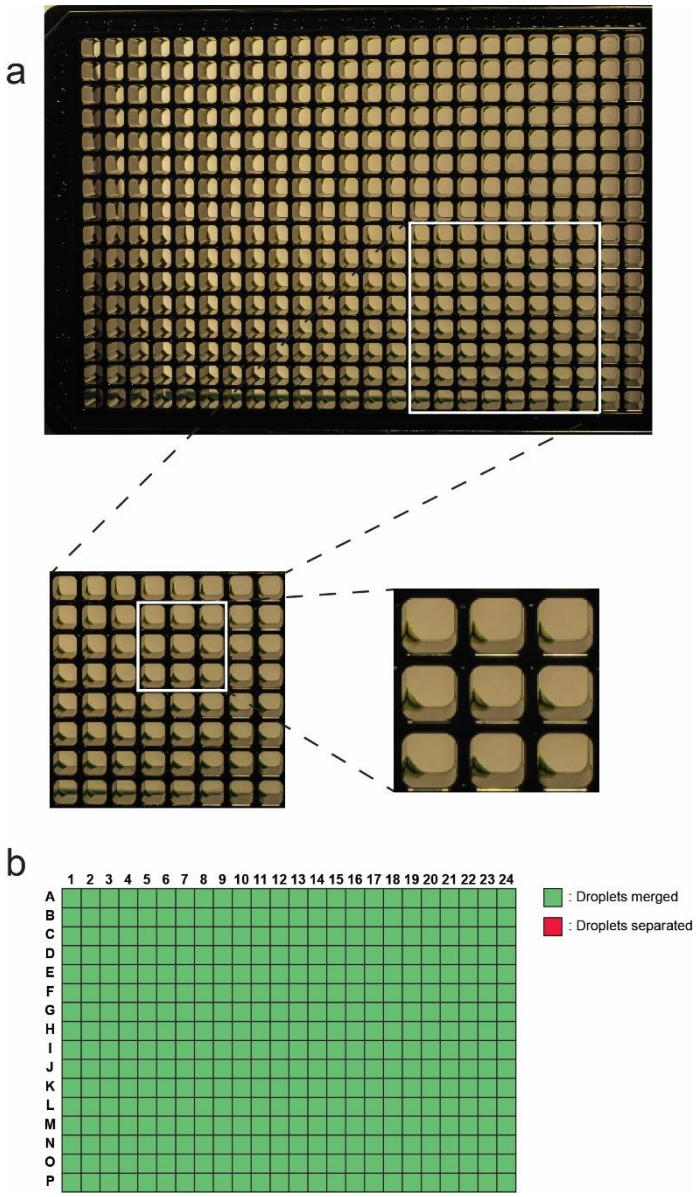


Figure S2. Enforcing droplet meeting in a microtiter plate by centrifugation. (a) Photo of the microtiter plate shown in Figure S1 after centrifugation and enlarged regions below. The merged liquids are adhered to the lower left corner of the wells. (b) In all wells, the droplets were merged (green).

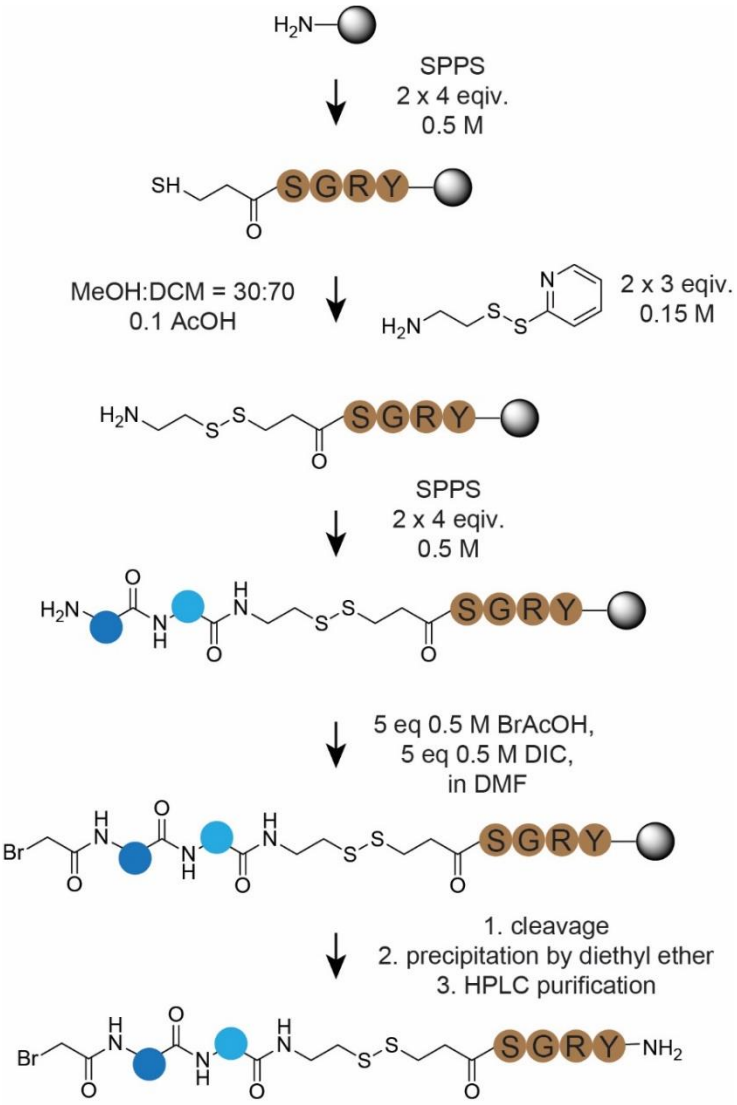


Figure S3. Synthesis of thiol-protected and bromoacetamide-functionalized peptides. Peptides were synthesized on solid phase using rink amide resin (50 μmol scale). They contain a C-terminal mercaptopropionic acid-Ser-Gly-Arg-Tyr tag that serves two functions, one facilitating purification of the peptides by precipitation with diethylether, and one protecting the peptides' thiol group by a disulfide bridge.

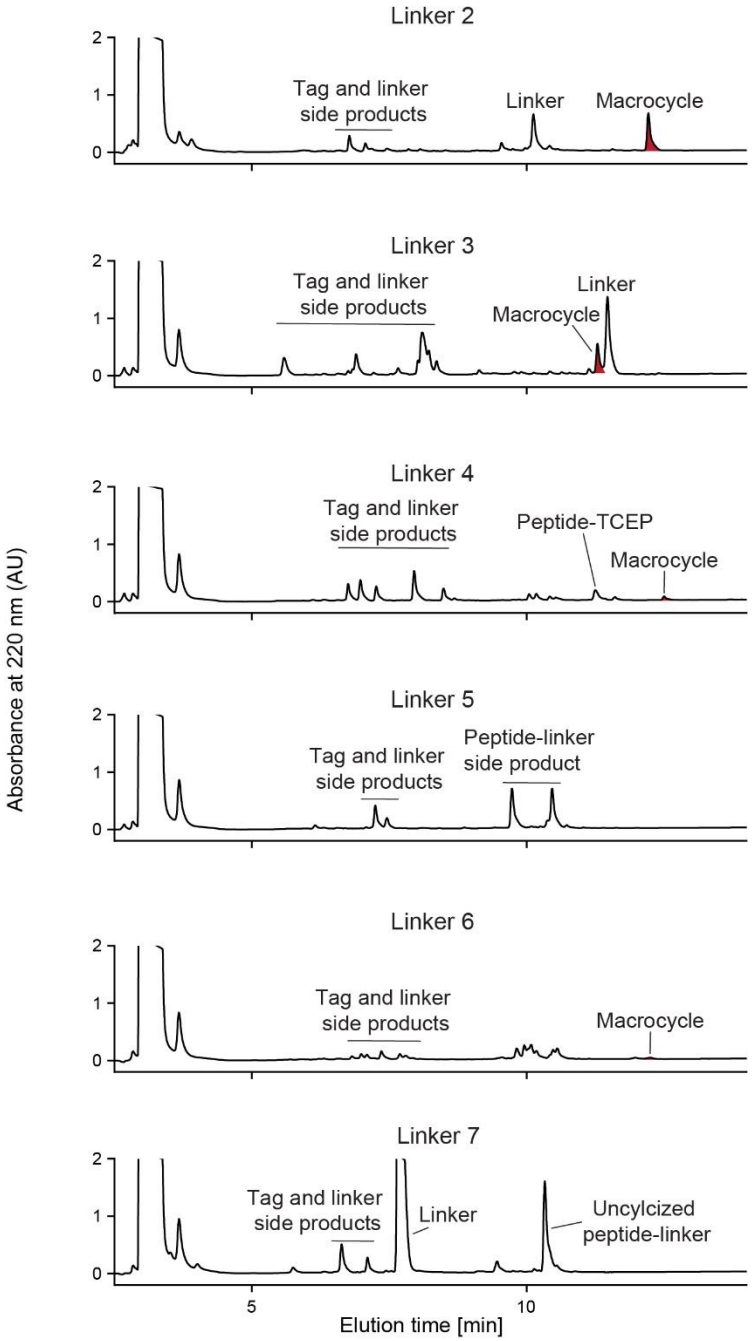


Figure S4. Thiol-to-amine macrocyclization by cyclization reagents 2-7. A BrAc-activated model peptide (Figure 2a) was reacted with the example primary amine 2-furanmethanamine, and cyclized by the bis-electrophile reagents **2**-**7** (indicated as "linker"). Reagents were transferred by ADE as shown in Figure 2a.

DNA sequence:

5'-
ATCCTCCAAAATCGGATCTGGTTCCGCGTGGATCC**AGCCAGATTCCGGCAAGCGAACAAGAAA**

CCCTGGTTCGTCCGAAACCGCTGCTGCTGAAACTGCTGAAAAGCGTTGGTGCACAGAAAGATA
CCTATACCATGAAAGAGGTGCTGTTTTATCTGGGCCAGTATATTATGACCAAACGCCTGTATGA
TGAGAAACAGCAGCATATTGTGTATTGCAGCAATGATCTGCTGGGTGACCTGTTTGGTGTTCCG
AGCTTTAGCGTTAAAGAACACCGTAAAATCTATACCATGATCTATCGTAATCTGGTGGTGGTTA
ATCAGCAAGAAAGCAGCGATAGCGGCACCAGCGTTAGCGAAAACTAAGAATTCCCGGGTCGA
CTCGAGCGGCCGCATCGTGACTGACTGACGATCTGCCTCGCGCGTTTCGGTGATGACGGTGAA
AACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGC
AGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCA
GTCACGTAGCGATAGCGGAGTGTATAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTT
TATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTG
CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAAC
CCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCC
TTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA
AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAA
GATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATG
TGGCGCGGTATTATCCCGTGTTGACGCCSGGGCAAGAGCAACTCGGTCGCCGCATACACAAGT-
3'

Protein sequence:

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQS

MAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRL

CHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQG

*WQATFGGGDHPPKSD*LVPRGS**SQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYLGQYI**

MTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIYRNLVVVNQQESSDSGTSVSEN

Figure S5. DNA and protein sequence of GST-hMDM2. The region of the expression vector that was confirmed by Sanger sequencing is shown. In bold is the gene coding for hMDM2 (17-125 aa). The cleavage sites of the restriction enzymes BamHI and EcoRI are underlined. The stop codon is highlighted in red. Amino acid sequence of the GST-hMDM2 fusion protein. The amino acids of GST are shown in italics and those of hMDM2 (17-125 aa) in bold.

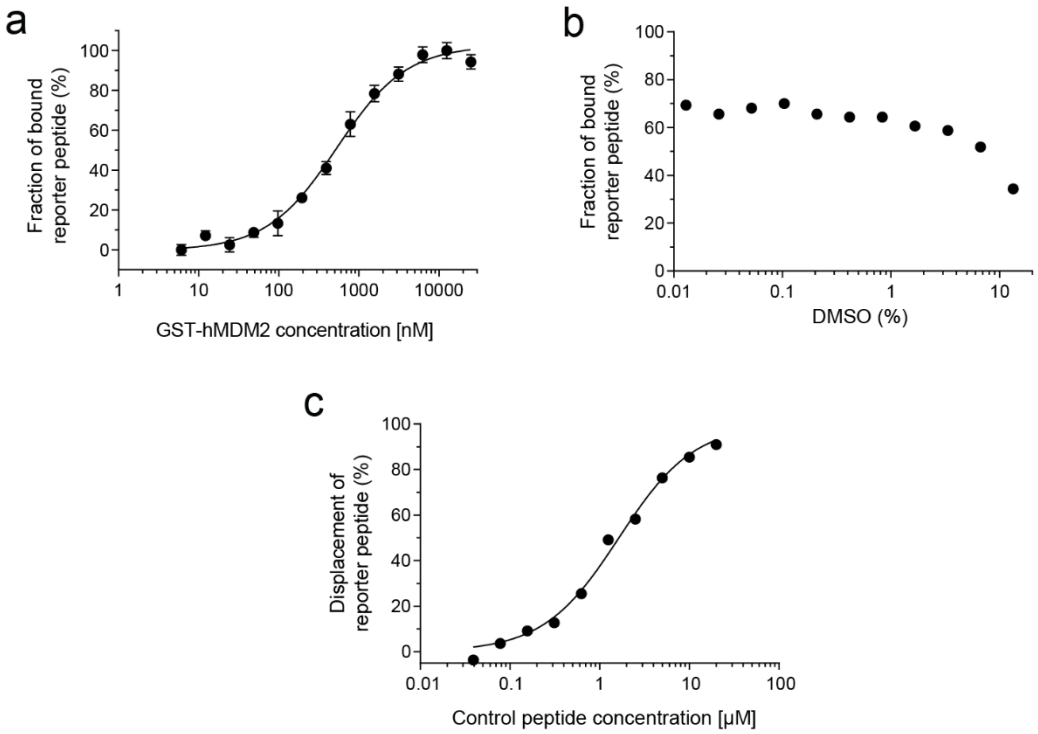


Figure S6. Fluorescence polarization assay development. (a) Binding of the fluorescent reporter peptide 5(6)-FAM-GSGSSQETFSDLWKLLPEN-NH_2 to GST-hMDM2 measured by fluorescence anisotropy. Means and standard deviations of three measurements are shown. (b) Effect of different DMSO concentrations on the binding of GST-hMDM2 (1 µM) to the fluorescent reporter peptide (50 nM) in the fluorescence polarization assay. Average values of two measurements are shown. (c) Fluorescence polarization competition assay. The displacement of the fluorescent reporter peptide (50 nM) from GST-hMDM2 (1 µM) at increasing concentration of the non-fluorescent control peptide was quantified by measuring fluorescence anisotropy.

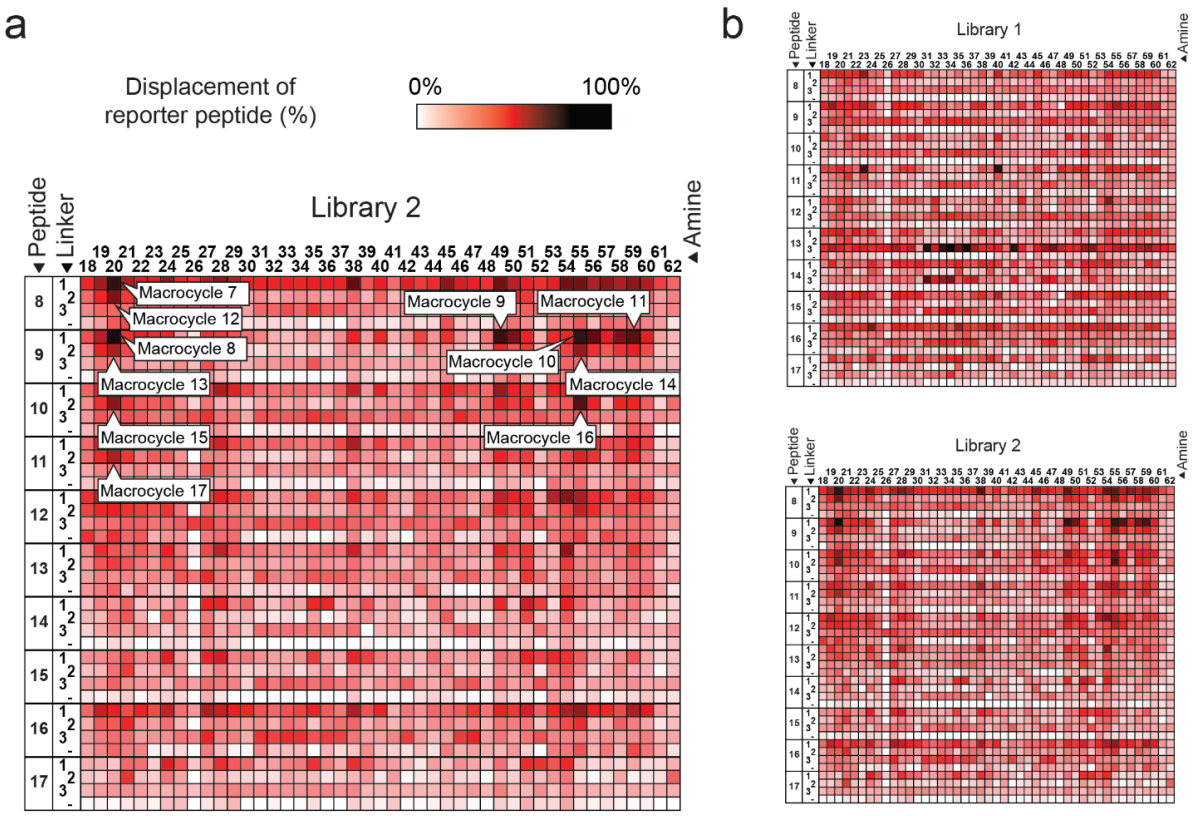
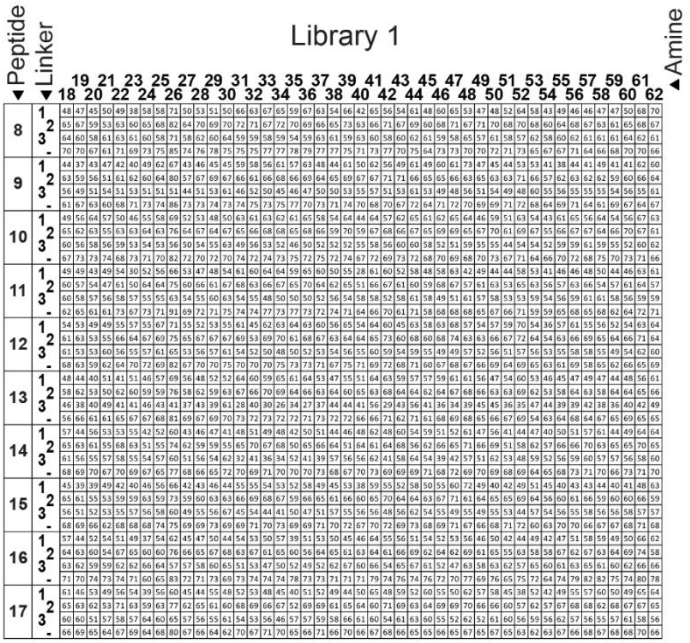


Figure S7. Screen for MDM2-p53 PPI inhibitors. (a) Binding of the 1,350 macrocycles of Library 2 to MDM2. The binding was assessed by displacement of a fluorescein-labeled MDM2-bound reporter peptide using fluorescence polarization. Darker colors indicate higher binding affinity. The building blocks used for library synthesis (shown in Figures 2c and 3) are indicated by number. Control reactions (linear peptides with no linker) are indicated by "-". Macrocycles that were re-synthesized, purified, and characterized are indicated on the heat map. (b) Heat map from screens with Library 1 and 2 without annotations. (c) Raw data of screen. The values indicate the measured anisotropy. The mean anisotropy of the negative controls (no competitor) and the positive controls (reporter peptide only) were 77 and 13, respectively. (d) Repetition of the screen for the 17 hit compounds and 16 randomly chosen reactions. The randomly chosen reactions are from Library 2 and are indicated with a black arrow in the top left panel. The bar graphs show the % of displaced reporter peptide for the screen and the three replicates. Figure S7 continues next page.

C



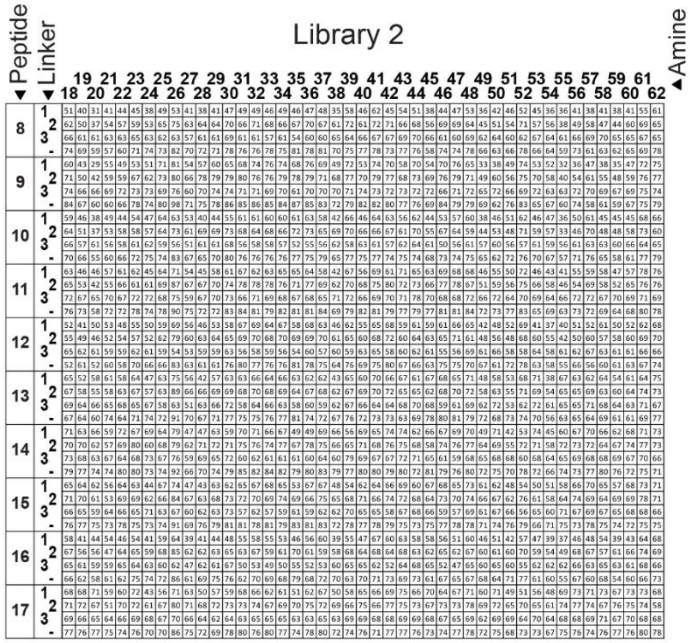


Figure S7 continued. Figure S7 continues next page.

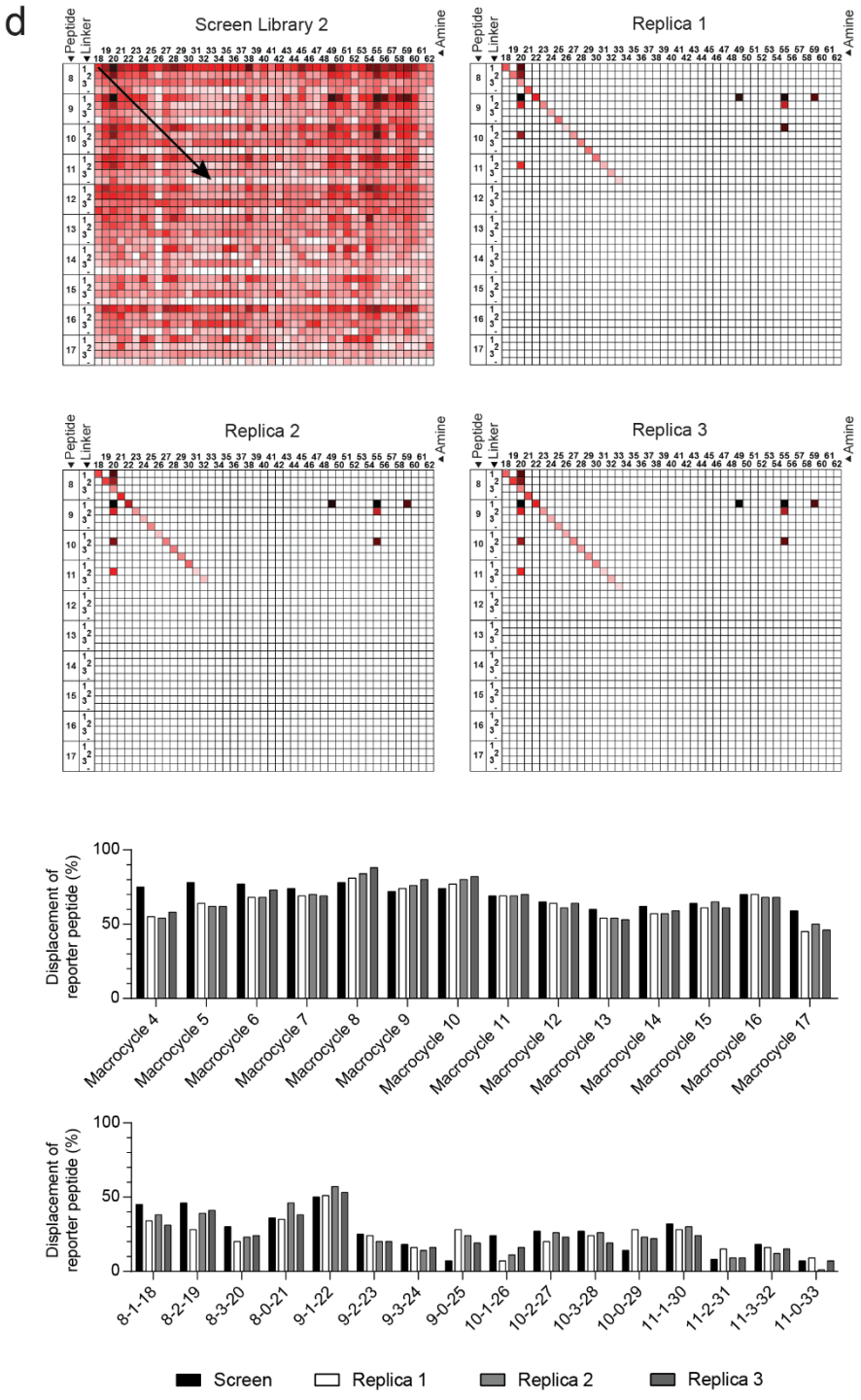


Figure S7 continued.

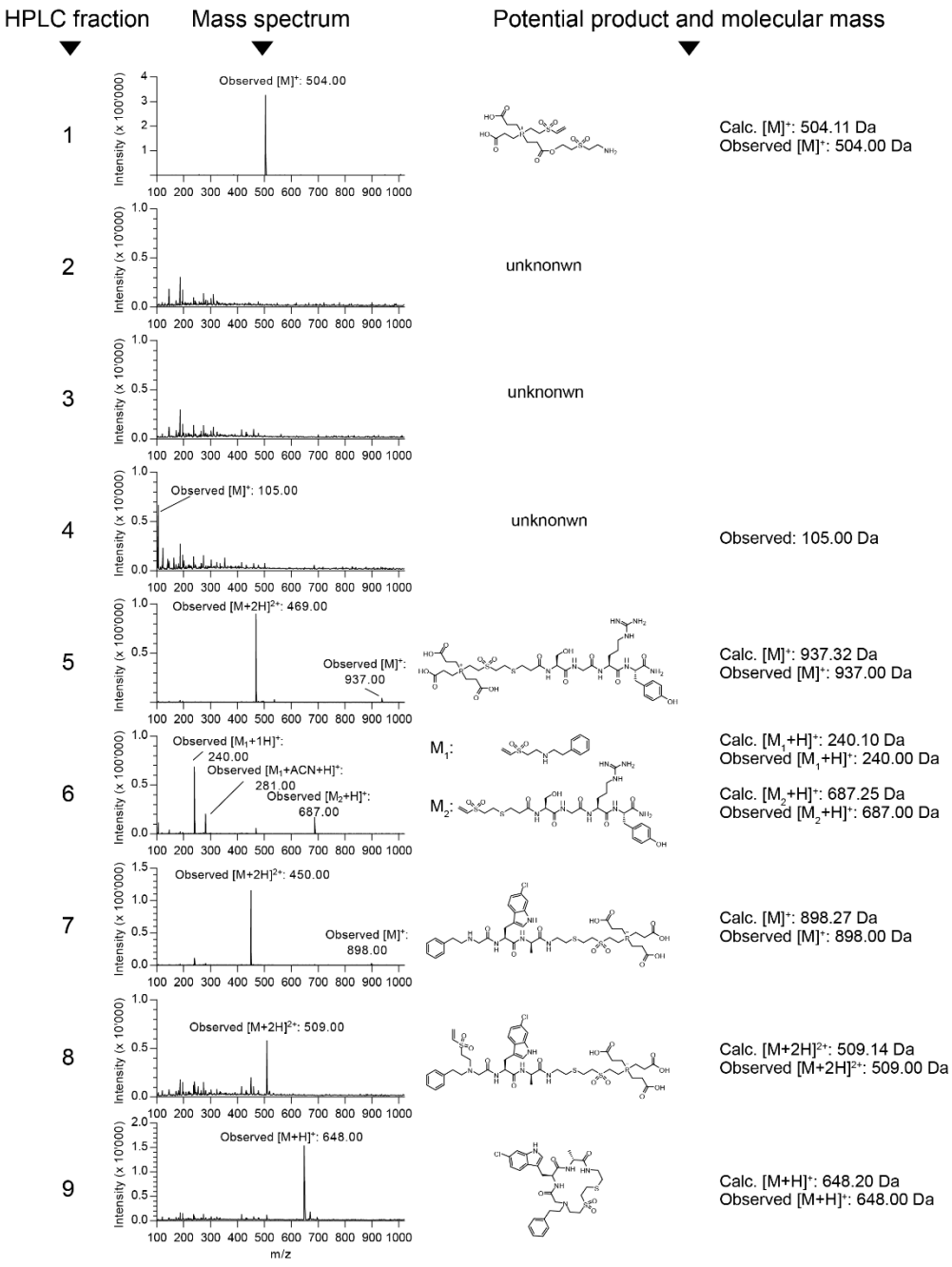


Figure S8. Side-products identified in the reaction mix of macrocycle 1. Macrocycle 1 was synthesized at a 50 nmol scale using the same "mixing and reaction" procedure that was used to prepare the library, the reaction run over a C18 column by RP-HPLC, and molecules that eluted as peaks shown in Figure 5c analyzed by ESI-MS. Peak 9 contains the macrocycle 1.

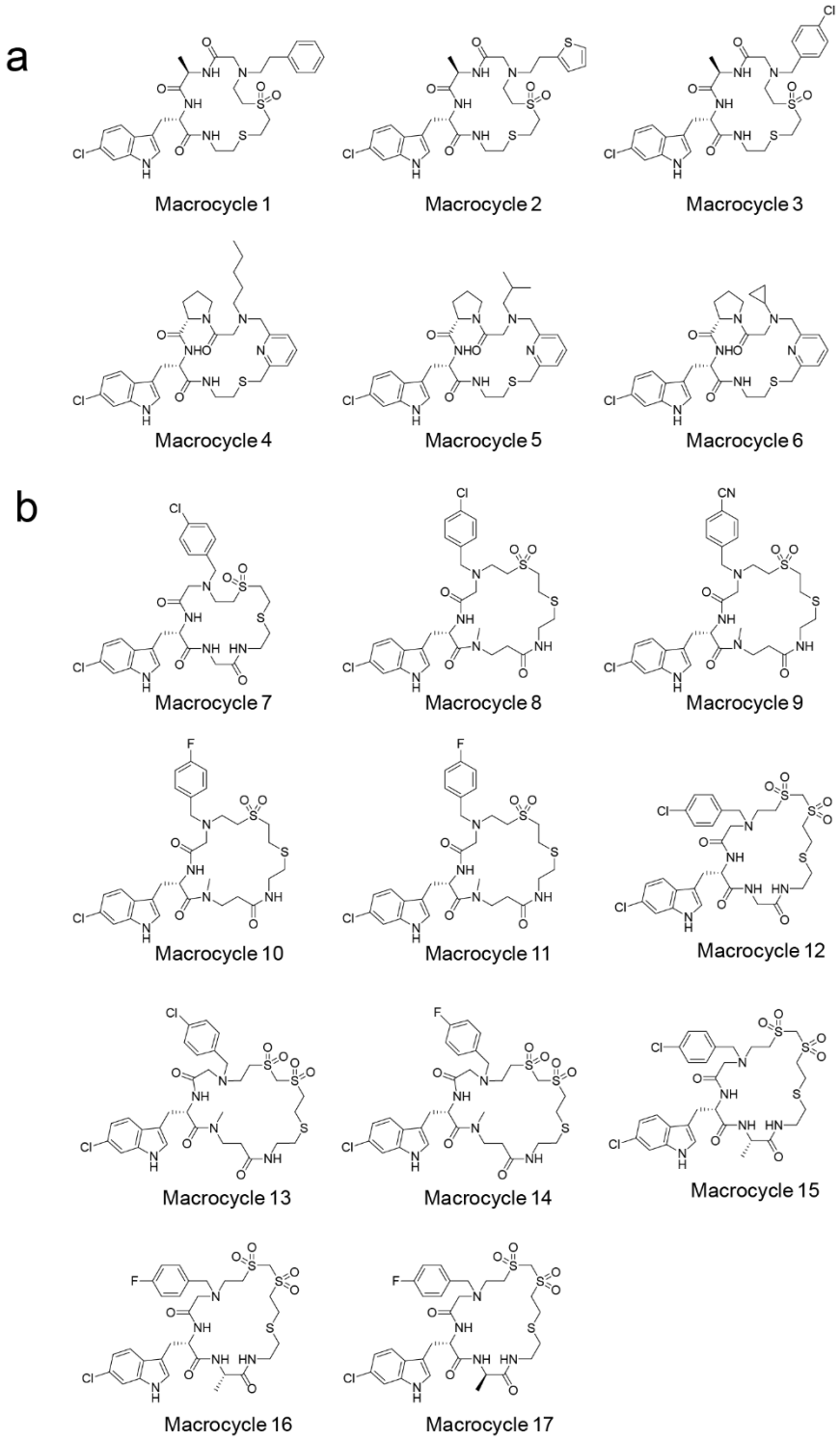


Figure S9. Chemical structures of macrocycles. Hits from Library 1 (a) and Library 2 (b).

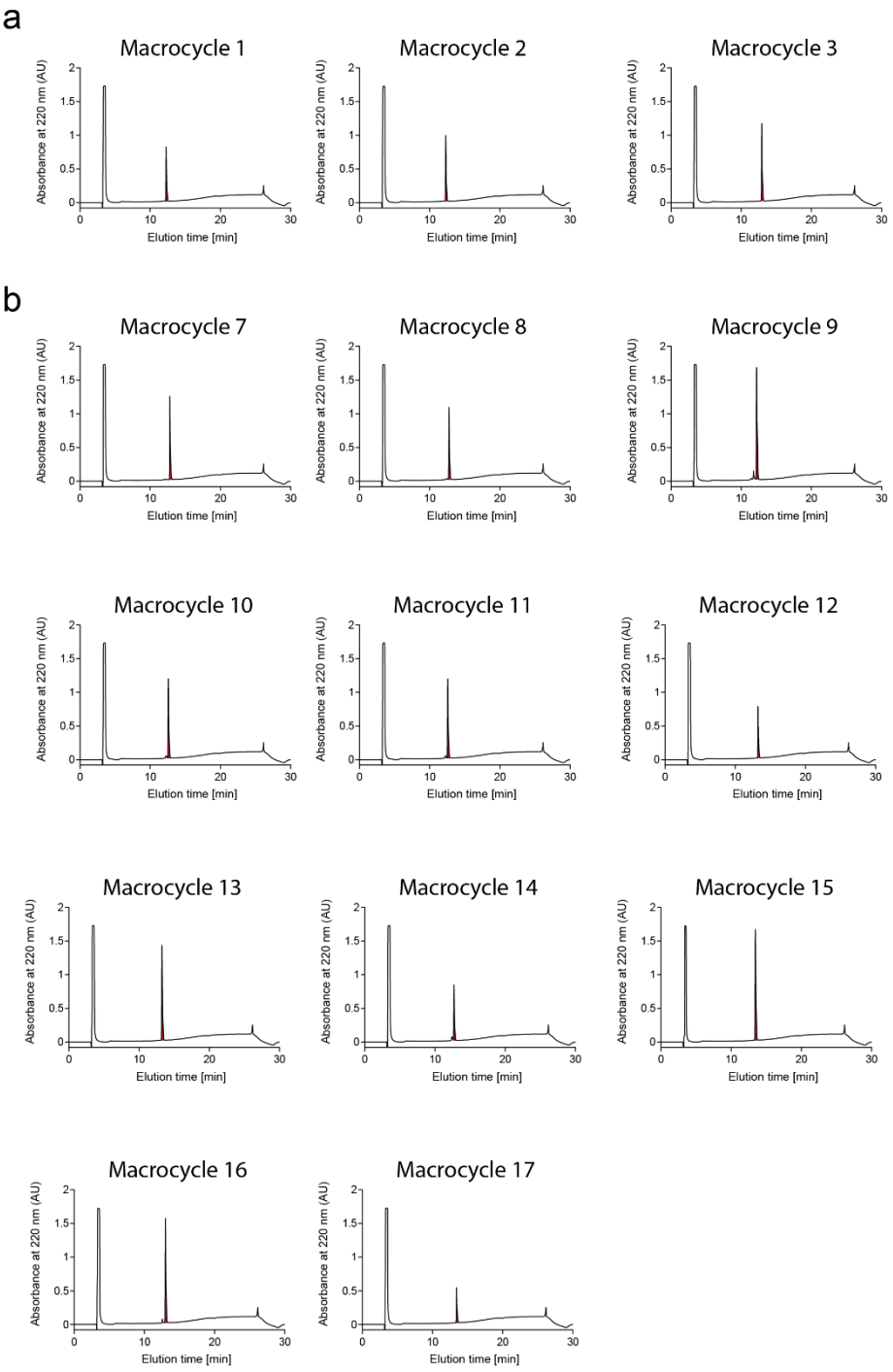


Figure S10. HPLC analysis of macrocycles. Hits from Library 1 (a) and Library 2 (b).

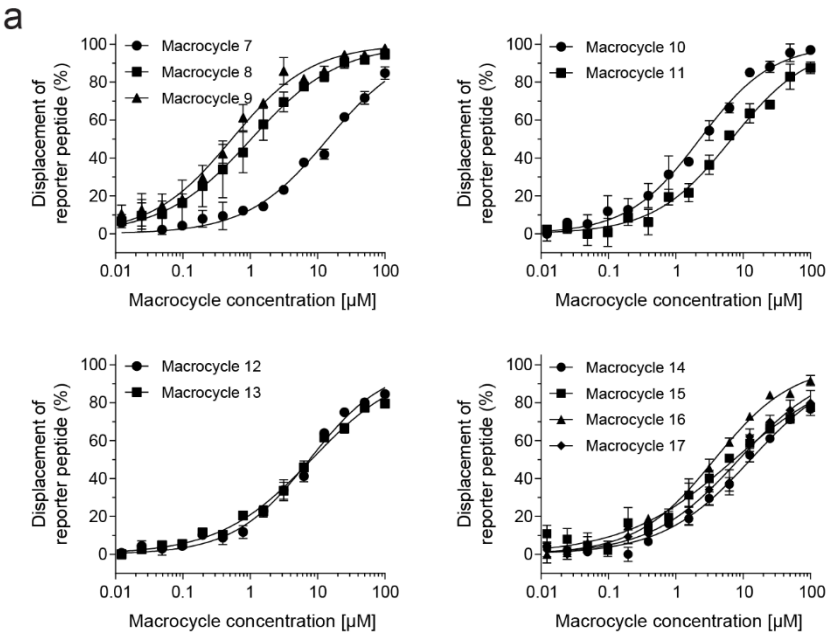


Figure S11. Binding of purified macrocycles to MDM2. The binding to GST-hMDM2 was measured with the fluorescence competition assay. Mean values and standard deviations of three independent measurements are shown. (a) Displacement of reporter peptide in % calculated based on the change in fluorescence polarization, as described in the materials and methods. (b) Raw data of fluorescence polarization competition assay. The extent of polarization is indicated as anisotropy.

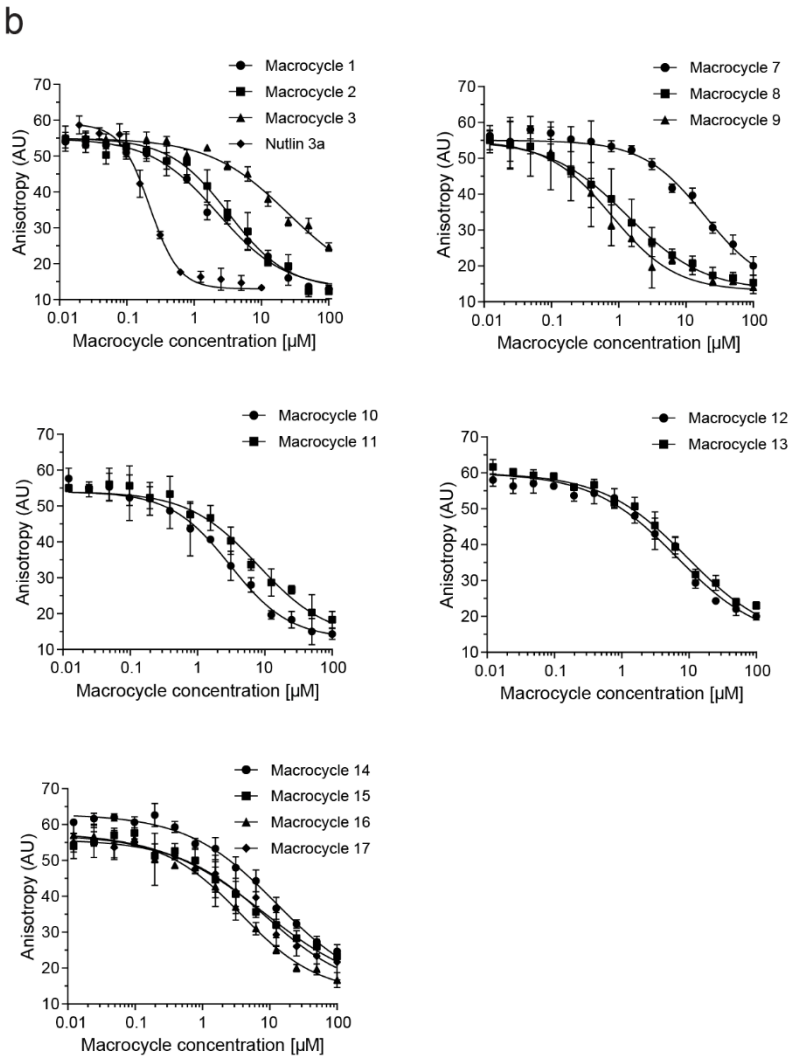


Figure S11 continued.

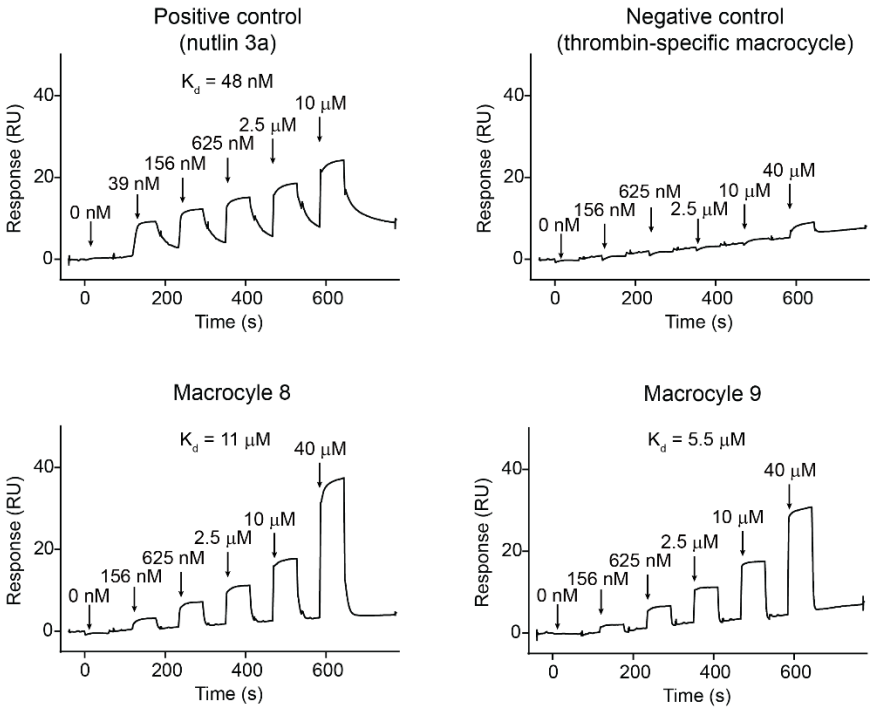


Figure S12. Measurement of macrocycle binding affinity by SPR. MDM2 was immobilized and macrocycles were run over the chip at the indicated concentrations. Single cycle SPR sonograms of the best two macrocycles and two controls are shown. The small molecule MDM2 binder nutlin 3a was used as a positive control. A macrocycle with similar size and format, binding to a different target (thrombin), was used as a negative control.

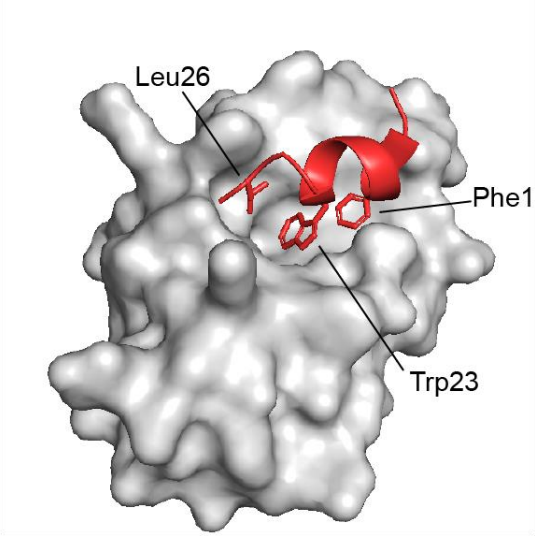


Figure S13. Structure of MDM2:p53. X-ray structure (PDB entry 4HFZ) of p53 peptide (aa 15-29 in red) bound to MDM2 (aa 17-125). Highlighed are the amino acids being key to the interaction.

3.8 References

(1) Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Lee, M. D.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W. D.; Duvall, J. R.; Kihlberg, J. Structural and conformational determinants of macrocycle cell permeability. *Nat. Chem. Biol.* **2016**.

(2) "Asinex macrocycle library," can be found under http://www.asinex.com/libraries-macrocyclic-5-html/, **n.d.**

(3) "Chembridge macrocycle library," can be found under https://www.chembridge.com/screening_libraries/targeted_libraries/macrocycle-screening-library/index.php, **n.d.**

(4) Ermert, P. Design, Properties and Recent Application of Macrocycles in Medicinal Chemistry. *Chim. Int. J. Chem.* **2017**.

(5) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-generation DNA-templated macrocycle libraries for the discovery of bioactive small molecules. *Nat. Chem.* **2018**.

(6) Li, Y.; De Luca, R.; Cazzamalli, S.; Pretto, F.; Bajic, D.; Scheuermann, J.; Neri, D. Versatile protein recognition by the encoded display of multiple chemical elements on a constant macrocyclic scaffold. *Nat. Chem.* **2018**.

(7) Stress, C. J.; Sauter, B.; Schneider, L. A.; Sharpe, T.; Gillingham, D. A DNA- Encoded Chemical Library Incorporating Elements of Natural Macrocycles. *Angew. Chemie Int. Ed.* **2019**.

(8) Roy, A.; Koesema, E.; Kodadek, T. High- Throughput Quality Control Assay for the Solid- Phase Synthesis of DNA- Encoded Libraries of Macrocycles**. *Angew. Chemie Int. Ed.* **2021**.

(9) Kale, S. S.; Bergeron-Brlek, M.; Wu, Y.; Kumar, M. G.; Pham, M. V.; Bortoli, J.; Vesin, J.; Kong, X. D.; Franco Machado, J.; Deyle, K.; Gonschorek, P.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Thiol-to-amine cyclization reaction enables screening of large libraries of macrocyclic compounds and the generation of sub-kilodalton ligands. *Sci. Adv.* **2019**.

(10) Buitrago Santanilla, A.; Regalado, E. L.; Pereira, T.; Shevlin, M.; Bateman, K.; Campeau, L.-C.; Schneeweis, J.; Berritt, S.; Shi, Z.-C.; Nantermet, P.; Liu, Y.; Helmy, R.; Welch, C. J.; Vachal, P.; Davies, I. W.; Cernak, T.; Dreher, S. D. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science.* **2015**.

(11) Gesmundo, N. J.; Sauvagnat, B.; Curran, P. J.; Richards, M. P.; Andrews, C. L.; Dandliker, P. J.; Cernak, T. Nanoscale synthesis and affinity ranking. *Nature* **2018**.

(12) Shaabani, S.; Xu, R.; Ahmadianmoghaddam, M.; Gao, L.; Stahorsky, M.; Olechno, J.; Ellson, R.; Kossenjans, M.; Helan, V.; Dömling, A. Automated and accelerated synthesis of indole derivatives on a nano-scale. *Green Chem.* **2019**.

(13) Wang, Y.; Shaabani, S.; Ahmadianmoghaddam, M.; Gao, L.; Xu, R.; Kurpiewska, K.; Kalinowska-Tluscik, J.; Olechno, J.; Ellson, R.; Kossenjans, M.; Helan, V.; Groves, M.; Dömling, A. Acoustic Droplet Ejection Enabled Automated Reaction Scouting. *ACS Cent. Sci.* **2019**.

(14) Gao, K.; Shaabani, S.; Xu, R.; Zarganes-Tzitzikas, T.; Gao, L.; Ahmadianmoghaddam, M.; Groves, M. R.; Dömling, A. Nanoscale, automated, high throughput synthesis and screening for the accelerated discovery of protein modifiers. *RSC Med. Chem.* **2021**.

(15) Hadimioglu, B.; Stearns, R.; Ellson, R. Moving Liquids with Sound. *J. Lab. Autom.* **2016**.

(16) DiRico, K. J.; Hua, W.; Liu, C.; Tucker, J. W.; Ratnayake, A. S.; Flanagan, M. E.; Troutman, M. D.; Noe, M. C.; Zhang, H. Ultra-High-Throughput Acoustic Droplet Ejection-Open Port Interface-Mass Spectrometry for Parallel Medicinal Chemistry. *ACS Med. Chem. Lett.* **2020**.

(17) Mothukuri, G. K.; Kale, S. S.; Stenbratt, C. L.; Zorzi, A.; Vesin, J.; Bortoli Chapalay, J.; Deyle, K.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Macrocycle synthesis strategy based on step-wise “adding and reacting” three components enables screening of large combinatorial libraries. *Chem. Sci.* **2020**.

(18) Chène, P. Inhibiting the p53–MDM2 interaction: an important target for cancer therapy. *Nat. Rev. Cancer* **2003**.

(19) García-Echeverría, C.; Chène, P.; Blommers, M. J. J.; Furet, P. Discovery of Potent Antagonists of the Interaction between Human Double Minute 2 and Tumor Suppressor p53. *J. Med. Chem.* **2000**.

(20) Zhang, R.; Mayhood, T.; Lipari, P.; Wang, Y.; Durkin, J.; Syto, R.; Gesell, J.; McNemar, C.; Windsor, W. Fluorescence polarization assay and inhibitor design for MDM2/p53 interaction. *Anal. Biochem.* **2004**.

4. A high-throughput screening assay for developing IL-23 receptor antagonists

4.1 Work contribution

The chapter 4 of this thesis is based on the manuscript draft "A high-throughput screening assay for developing IL-23 receptor antagonists". The following authors contributed:

Gontran Sangouard,[1] Maurice Philippe Biedermann,[1] Edward Will,[1] Soraya Quinche,[2] Durrer Laurence,[2] Florence Pojer[2] and Christian Heinis[1]*

[1]Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

[2]Protein Production and Structure Core Facility, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

*Corresponding author. Email: christian.heinis@epfl.ch

Author contributions: Together with Prof. Christian Heinis, I conceptualized the development of the high-throughput screening assay against IL-23/IL-23R protein-protein interaction and the macrocycle library design. I performed the plasmid preparation and protein purification, the optimization of the FP assay, all the macrocycle library synthesis and screening, the entire development of the TR-FRET assay, the hit validation including the active impurity synthesis and test. Maurice Philippe Biedermann contributed in the development of the assay based on fluorescence polarization and in the design for the expression of the IL-23R. Edward Will helped in identifying the active side-products. Soraya Quiche and Durrer Laurence, under the supervision of Dr. Florence Pojer, performed all the transient expressions of IL-23R. Together with Prof. Christian Heinis, I prepared all the figures and wrote the manuscript.

4.2 Abstract

Interleukin 23 (IL-23) is a proinflammatory cytokine involved in several autoimmune inflammatory diseases such as psoriasis and inflammatory bowel disease. Monoclonal antibodies binding IL-23 or its receptor IL-23R, including the approved drugs ustekinumab, guselkumab, tildrakizumab and risankizumab have validated inhibition of the IL-23:IL-23R protein-protein interaction (PPI) as a therapeutic strategy. Towards the development of an orally available drug against this target, we have set out to establish a robust high-throughput screening assay for identifying molecules that inhibit the PPI. In a first attempt, we tested a screening assay based on fluorescence polarization (FP) competition that turned out to be prone to artifacts resulting from compounds interfering with the assay. In a second attempt, we established a robust assay based on time-resolved FRET (TR-FRET) that proved efficient for screening binders of IL-23R. Application of the assay to a pilot-scale library of 3,072 cyclic peptides allowed identification of hits with low micromolar affinity, the most potent inhibitor having an IC_{50} of 1.3 ± 0.2 μM. The assay is robust and suited for the high-throughput screening of large compounds libraries. The learnings from developing and testing the two different assay types can be applied for establishing high-throughput assays for other PPI targets.

4.3 Introduction

IL-23 is a member of the IL-12 cytokine family, which are messenger proteins produced by activated antigen-presenting cells such as dendritic cells and macrophages, and have a key role in triggering the immune response.[1] The IL-12 cytokine family is characterized by their heterodimeric nature cross-utilizing different cytokine subunits and sharing signaling receptors, while having different influence on the immunological balance. IL-23 is composed of the IL-23p19 subunit that is disulfide-linked to the IL-12p40 subunit, common to IL-12, and signals via complex of the IL-23 receptor (IL-23R) and IL-12 receptor subunit β1 (IL-12Rβ1).[2,3] The most important biological function of IL-23 is to amplify the proliferation of IL-17 secreting $CD4^+$ T helper cells (Th17 cells).[4]

IL-23 was identified as a key mediator in many autoimmune inflammatory diseases such as encephalomyelitis,[5] psoriasis,[6] and inflammatory bowel disease (IBD).[7] Additionally, IL-23 and IL-23R proved to be crucial in tumor incidence and growth.[8]

Several drugs targeting IL-23 were developed for the treatment of psoriasis, psoriatic arthritis, Crohn's disease and ulcerative colitis all being based on monoclonal antibodies.[9,10] The antibody ustekinumab, binds the IL-23p40 subunit and thus blocks both, signaling of IL-23 and IL-12, and the antibodies risankizumab, guselkumab and tildrakizumab bind IL-23p19 subunit and thus block only signaling via the IL-23R. Monoclonal antibodies were also developed against the receptor, IL-23R, inhibiting the binding of IL-23p19, such as AS2762900-00 that showed therapeutic effects in a pre-clinical model of psoriasis.[11,12]

Oral drugs targeting the IL-23:IL-23R PPI do not exist, despite obvious advantages such as more convenience to patients, more flexible dosing, and a smaller application barrier than injectable treatments, for example for treating milder forms of the diseases. Protagonist Therapeutics and Janssen Biotech have developed peptide-based antagonists of IL-23R for the oral treatment of IBD. A first generation peptidic IL-23R antagonist, PTG-200 is in Phase 2 and the second generation peptides PN-235 and PN-232 are tested in Phase 1.[13] Overall, the development of small molecules inhibitors appears challenging due to the lack of rather shallow binding sites at the IL-23:IL23R PPI. Our laboratory has recently screened cyclic peptide libraries against IL-23R, which identified a hot spot binding region exactly where the receptor is binding IL-23.[14] Peptides isolated shared strong consensus sequences including a conserved tryptophan residues that was needed for binding, suggesting that small molecules or macrocyclic compounds with tryptophan (derivative) as anchor residues could potentially be developed.

Towards the screening of small molecules inhibiting the IL-23:IL-23R PPI, a cell-based HTS screening was recently developed and applied to screen 24,000 compounds.[15] The assay, using an IL-23-responsive cell-based luciferase reporter system and engineered cells showed a robust performance with a signal-to-background >7-fold and Z′ > 0.5. However, more than 1,000 hits were identified due to toxic effect on the cells or off-target effects. Herein, we aimed at establishing a HTS assay that is assaying directly the IL-23:IL-23R PPI in a cell free context. Towards this end, we developed a HTS assay to mimic and monitor the interaction between IL-23 p19 subunit and its cognate receptor IL-23R, and performed a pilot screen of 3,072 macrocycles. In first attempt, we tried to develop an assay based on FP, which turned out to be not suited for our libraries. A second assay based on TR-FRET was suited and allowed screening of the pilot macrocycle compound library.

4.4 Results and discussion

4.4.1 Screening assay based on fluorescence polarization

In a previous work, we had used phage display to develop cyclic peptides that bind with high affinity to IL-23R. All peptides isolated shared an amino acid motif that resembled the sequence of an α-helix in IL-23 that forms the main contact with IL-23R, suggesting that the peptides mimic the helix (Figure 1a) (Dìaz Perlas *et al*, not published at this time). Competition experiments indeed showed that the peptides inhibit the IL-23/IL-23R interaction. Herein, we proposed to use such a peptide for developing a screening assay based on FP that works as follows. A peptide labeled with a fluorophore is incubated with an excess of IL-23R so that most of the peptide is bound. Binding to the receptor reduces the rotation of the peptide and thus also of the fluorophore, which can be measured by FP. Compounds binding to IL-23R and inhibiting the interaction of receptor and peptide can be detected by a reduction of the fluorescence polarization, as illustrated in Figure 1b. The assay would allow screening for compounds that bind exactly where the peptide binds, which is an ideal side for inhibiting the IL-23R:IL-23 PPI. In addition, the finding that all phage-selected peptides are directed to this binding suggests that this region is a hotspot and thus a surface of IL-23R where ligands may be more easily developed to than other surface regions.

For the FP-based screening assay, we required the N-terminal immunoglobulin domain D1 of IL-23R that interacts with IL-23. Trauger and co-workers reported that the extracellular region of IL-23R, formed by the domains D1-D2-D3, can be efficiently expressed as fusion with an antibody Fc fragment in mammalian cells.[16] We cloned a vector for transiently expressing this region with Fc (the fusion protein termed in the following IL23-R-Fc), and additionally also for D1-Fc, and performed test expressions in mammalian cells (Figure S1). While the single domain construct showed poor yields, IL23-R-Fc was efficiently expressed at around 15 mg/L in HEK cells. We produced the fusion protein at a larger scale and obtained 7.8 mg per liter culture after purification. For binding around 50% of the fluorescence peptide, the receptor protein had to be used at a concentration that equals the dissociation constant (around 80 nM). Assuming a screening assay volume of 10 μl, a library of around 150,000 compounds could theoretically be screened with the amount of protein produced.

As a fluorophore-labeled reporter binding to IL-23R, we used the peptide L<u>C</u>TWVDYWL<u>C</u>H-NH_2 derived from the phage display selections described above, that was cyclized by reacting two cysteine side chains (underlined) with the linker reagent divinylsulfone (Dìaz Perlas *et al*, not published at this time). We synthesized the peptide with a fluorescein moiety at either the N-terminus (conjugated to the N-terminal amino group; peptide 1) or the C-terminus

(conjugated with via the side chain of an appended lysine; peptide 2) (Figure 1c). Binding tests by FP showed that both peptides were binding with affinities in the expected 80 nM range (Figure 1d). However, the two peptides displayed unexpected high FP values already in absence of receptor, which we speculated could potentially result from the hydrophobic nature of the peptides that may stick to the microwell plates, or precipitates partially, both leading to a fraction of peptide that is polarized.

In order to improve the solubility of the peptide, we appended multiple charged amino acids, either four arginines or four glutamic acids, and either at the N-terminal end, the C-terminal end, or distributed to both ends (Figure 1c). All peptides showed improved solubility, and for the peptides containing four glutamic acids, the FP of the unbound peptides showed values typically found for fluorescein-labeled peptides (Figure 1e & Figure S2). For the subsequent experiments, we chose to work with the peptide that presented the largest change of FP and a strong binding affinity, being fluorescent peptide 6 (sequence: fluorescein-EELCTWVDYWLCHEE-NH_2; cyclized by divinylsulfone). The peptide showed a dissociation constant of 26 ± 3 nM, which was in the same range as the affinity of the peptides without the additional glutamic acids, indicating that the extensions did not affect the receptor interaction. We further tested the compatibility of the assay with DMSO, that often occurs in HTS assays due to addition of compounds from stocks in DMSO, and found that the solvent is tolerated up to 4% (v/v) and thus at a high concentration (Figure S3). Finally, we tested if the peptide 6 can be displaced, using a non-labeled peptide binding to the same site (HCTWMDFWRCS; cyclized by divinylsulfone; K_D = 200 nM), and if the displacement can be monitored by FP, which was the case (Figure 1f).

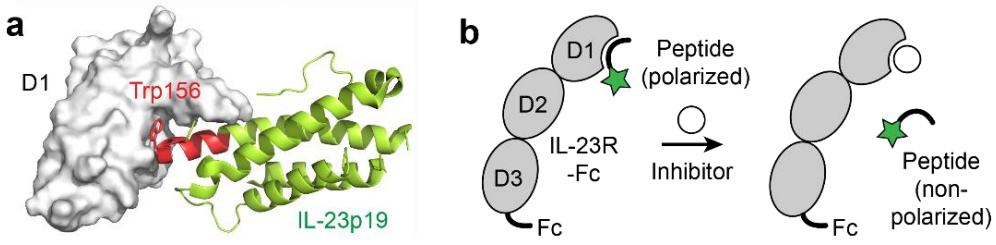


Figure 1. Development of fluorescence polarization competition assay. (a) X-ray structure of IL-23p19/IL-12p40 bound to IL-23R (PDB entry 5MZV) with only D1 of IL-23R (grey) and IL-23p19 (green) shown. The α-helix forming the main contact with the receptor is shown in red and the Trp156 is indicated. (b) Principle of screening assay based on fluorescence polarization. Figure 1 continues next page.

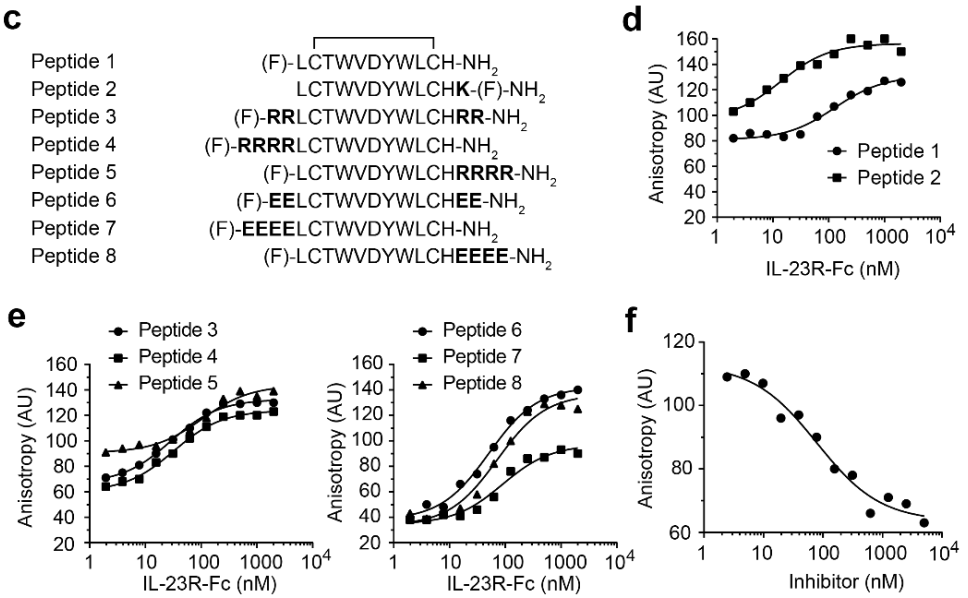


Figure 1 continued. (c) Sequences of fluorescein-labeled cyclic peptides. Fluorescein is indicated as (F). (d) and (e) Fluorescence polarization of peptides 1 and 2 (d) or 3 to 8 (e) binding to IL-23R. The peptides were applied at a concentration of 10 nM. (f) Fluorescence polarization competition assay. The displacement of peptide 6 (10 nM) from IL-23R-Fc (60 nM) at increasing concentrations of non-labeled peptide (sequence: HCTWMDFWRCS; cyclized by divinylsulfone) was quantified by measuring fluorescence anisotropy.

4.4.2 Screening test compounds using the FP assay

In order to test the FP-based assay in a screen with test compounds, we synthesized 70 small cyclic peptides that had peptide sequences similar to those of peptides identified in the aforementioned phage display selections against IL-23R. Specifically, we synthesized peptides comprising between three and seven amino acids, a cysteine or homocycteine residue at the N-terminus or a middle position, and mercaptoethylamine (Mea) at the C-terminus (Figure 2a), and cyclized the peptides with bis-electrohilic reagents via the two thiol groups (Figure 2b). The peptides contained all a tryptophan residue that is key for binding to IL-23R. In many of the peptides, we inserted two consecutive aromatic amino acids downstream of the key tryptophan. Additionally, we inserted amino acids that were also abundant in the phage-selected peptides such as serine, threonine, glutamine and glutamate. We synthesized the peptides on solid-phase using an approach that yields peptides of high purity without the need of a chromatographic purifications step (Habeshian *et al*, Bognár *et al*. both not published at this time), cyclized them at a concentration of 1 mM using a 4-fold molar excess of the five bis-electrophile reagents, quenched excess of electrophiles with 16-fold excess of 2-mercaptoethanol.

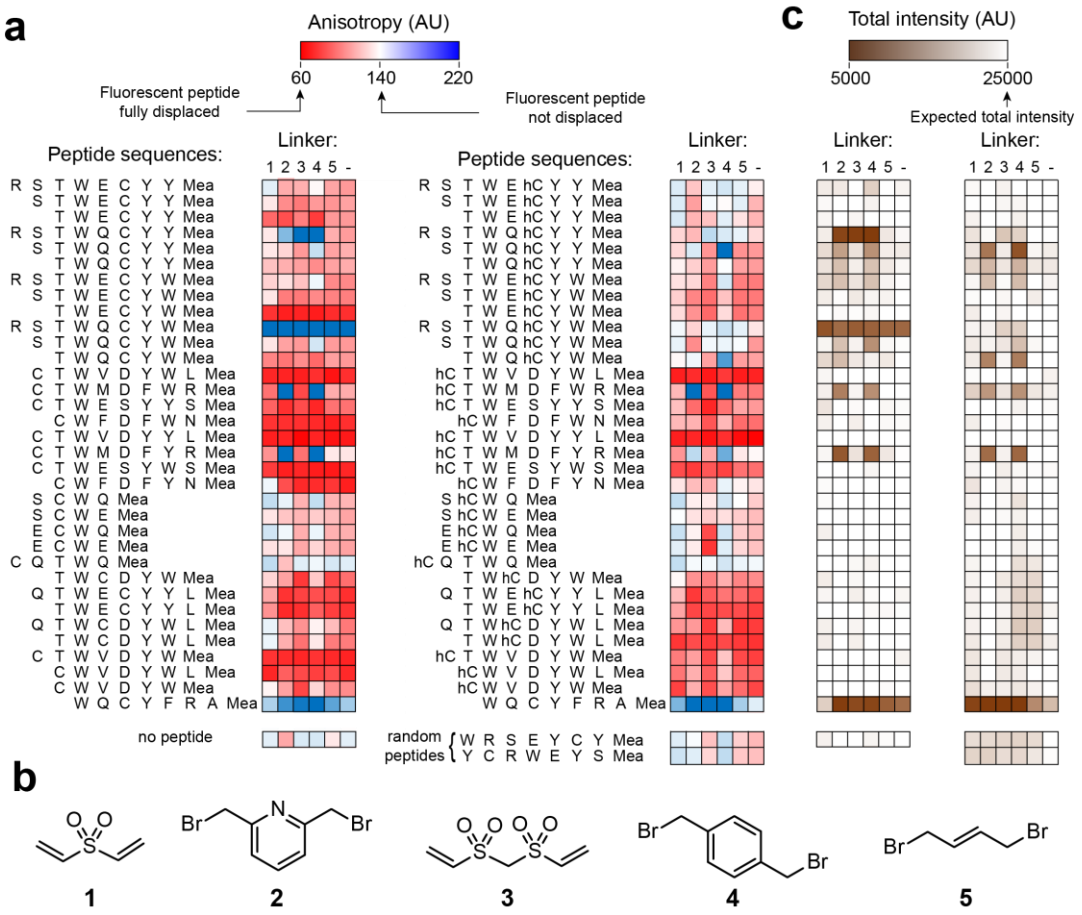


Figure 2. Testing the fluorescence polarization-based assay with potential IL-23R binders. (a) A panel of 82 linear peptides with sequences inspired by phage display-selected binders of IL-23R were cyclized combinatorially via their two thiol groups with the five bis-electrophile reagents shown in panel (b). The first 40 peptides have the same sequences as the following sequences except for a cysteine (C) that is replaced by a homocysteine (hC). The last two peptides are random peptide sequences having similar amino acids as the other peptides. All peptides contain mercaptoethylamine (Mea) at the C-terminal end, which provided the second thiol group for peptide cyclization. Displacement of the peptide 6 by the peptides was detected in 384-microwell plates by measuring the anisotropy and is indicated in color. White = no displacement, red = full displacement, blue = unexpected increase of anisotropy. (b) Five bis-electrophile reagents. (c) For all wells, the total fluorescence intensity was measured and is shown in colors from white (expected intensity) to brown (unexpected low fluorescence intensity).

We screened the products without purification as follows. We transferred the crude reactions into 384-well plates (150 nl at 1 mM in DMSO) and added 15 μl of IL-23R-Fc (60 nM final conc.) and the fluorescent probe (10 nM final conc.), resulting in a final cyclic peptide concentration of around 10 μM. Given that the peptides were shorter and had different sequences than those identified by phage display, we expected that only some of them bind IL-23R and most likely with lower affinity. Several of the peptides decreased the anisotropy from a value of around 140 (no inhibitor control; in arbitrary units, AU) to values of around 60 which is expected for full displacement of the fluorescent probe (inhibitor control; red values in Figure 2a), suggesting binding to IL-23R (Figure 2). For some of the wells, we observed unexpected strong increases in anisotropy values, as high as 220 units (blue in Figure 2a)

which hinted to problems with the assay. We noted that all peptides showing a strong increase in anisotropy were containing a positive charge. In order to analyze the unexpected effects, we verified total fluorescence intensity in the screening plates, which revealed that peptides with particularly high anisotropy values showed much reduced fluorescence emission (change from 25,000 AU to 5,000 AU; shown in brown in Figure 2c). We speculated that the unexpected increase in anisotropy and the drop in total fluorescence emission is caused by peptides sticking to the wells or aggregating, and the negatively charged fluorescent probe binding to the positively charged peptides. The result showed a weakness of FP competition assay, being the sensitivity to precipitation effects induced by library compounds. An additional hint for problems with the assay came from two random peptide sequences (containing also a tryptophan), that we included as negative controls and showed some binding (last two sequences in Figure 2a). Given the large number of peptides that showed these undesired effects (increased anisotropy and decrease of total fluorescence), we concluded that the FP-based assay is not suited for screening libraries of cyclic peptides against IL-23R.

4.4.3 Screening assay based on TR-FRET

An assay format that is routinely used for screening compounds against PPIs is based on TR-FRET, measuring fluorescence-resonance transfer (FRET) between a donor and acceptor fluorophore linked to the interacting proteins, and reduction of the FRET signal by PPI inhibitors respectively. In TR-FRET, a fluorescence donor is used that transfers energy to an acceptor for a prolonged time period, allowing measuring emission of the acceptor fluorophore at a delayed time point, when the excitation light is turned off already and no scattering effects, direct emission of the donor, or fluorescence emission by the screened compounds, are detected. Herein, we were strongly interested in the TR-FRET because the assay format is compatible with the use of high concentrations of blocking agents such as albumin or detergents, that we felt are needed to prevent the above observed solubility and precipitation effects when screening hydrophobic compounds against IL-23R.

We conceived the assay format shown in Figure 3a, in which TR-FRET is used to measure binding of the extracellular region of IL-23R to the IL-23-mimicking peptide described above. The his-tagged fusion protein of IL-23R extracellular domains and antibody Fc fragment (IL-23R-Fc) is associated with a monoclonal anti-His antibody labeled with the lanthanide-based-fluorophore europium (Eu-AntiHis$_6$), and the peptide 6 is labeled with the acceptor fluorophore Cy5 instead of fluorescein (EELCTWVDYWLCHEE-PEG$_2$-progargylglycine(Cy5)-NH$_2$; peptide 9) (Figure 3a). Europium, excited at 340 nm, transfers energy to the closeby Cy5. Emission of light from Cy5 and europium are then measured with a time delay of around 60

μs at 665 nm and 620 nm respectively. Inhibition of the PPI is expected to reduce the emission of Cy5, and to some extent also to increase the emission of europium.

In order to find the optimal conditions for the assay, we tested TR-FRET at several concentration combinations of IL-23R-Fc (0, 5, 10, 20, 40, 80 nM) and the peptide 9 (0, 30, 60, 120, 250 and 500 nM) using Eu-AntiHis$_6$ at 1 nM. For all these tests, we applied 0.1% (w/v) of BSA as blocking agent, that we hoped would prevent in a later step precipitation or sticking caused by hydrophobic screening compounds. At all concentrations of the two binding partners, efficient TR-FRET was observed, with the Cy5 acceptor reaching an equal emission fluorescence intensity (measured at 665 nm) to the emission of the donor europium (measured at 620 nm) at the highest concentrations of protein and peptide (Figure 3b). The best signal to background ratios reached were around 30 and were found for IL-23R-Fc concentrations of 10 nM and above, and peptide 9 at 30 and 60 nM (Figure 3c). We considered concentrations of 10 nM IL-23R-Fc, 50 nM peptide 9, and 1 nM Eu-AntiHis$_6$ as ideal to reach a signal-to-background ratio of 30, and used these conditions in the subsequent experiments. Addition of the IL-23-binding peptide (HCTWMDFWRCS; cyclized by divinylsulfone;,K_D = 200 nM) to this assay reduced the TR-FRET signal in a dose-dependent manner and suggested that the assay format could be used for screening IL-23R:IL23 inhibitors (Figure 3d).

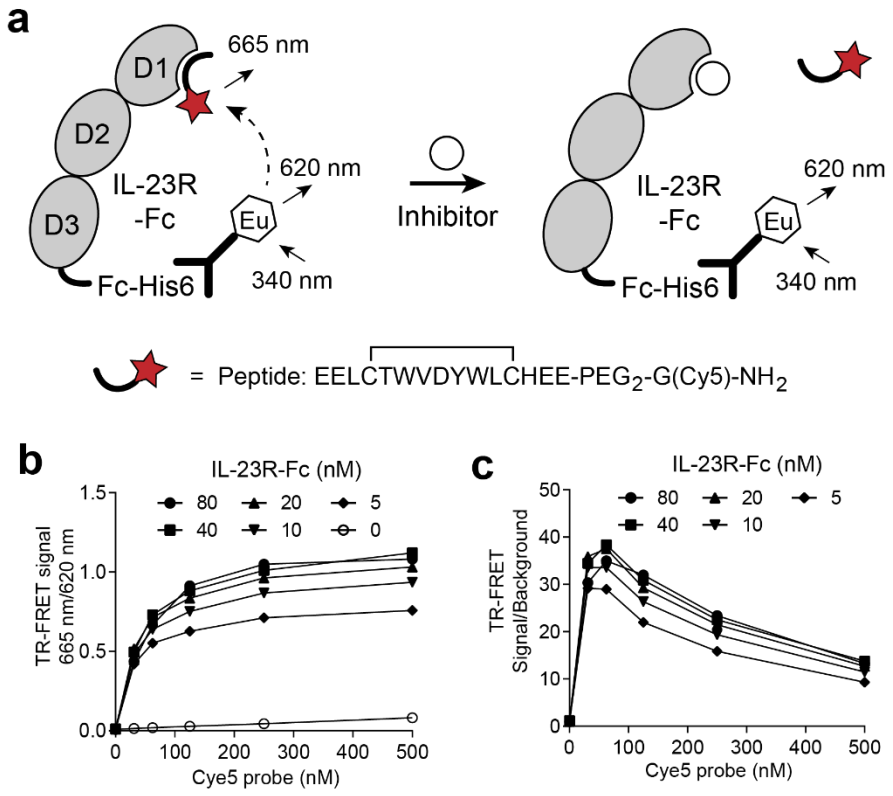


Figure 3. Development of assay based on TR-FRET. (a) Schematic depiction of assay principle. The sequence of the Cy5-labeled peptide is shown below (peptide 9). Figure 3 continues next page.

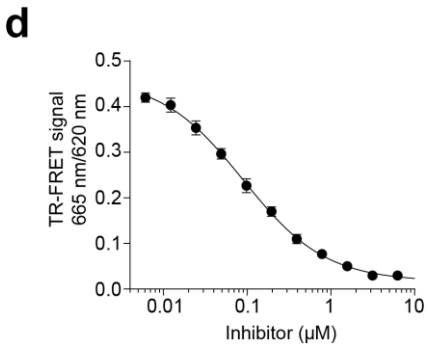


Figure 3 continued. (b) Testing of different concentration combinations of IL-23R and peptide 9 using Eu-AntiHis$_6$ at 1 nM. The TR-FRET signal is indicated as a ratio of Cy5 emission (665 nm) and europium emission (620 nm). (c) Signal-to-background from experiment in (c) wherein background is the signal measured in absence of IL-23R. (d) Competition assay based on TR-FRET. Eu-AntiHis$_6$ (1 nM), IL-23R (10 nM) and peptide 9 (50 nM) were incubated with the indicated concentrations of the non-labelled peptide (sequence: HCTWMDFWRCS; cyclized by divinylsulfone). Mean values and SDs are shown.

4.4.4 Screening test compounds using the TR-FRET assay

We assessed the performance and reliability of the TR-FRET assay by screening the same 80 small cyclic peptides as with the FP assay. We transferred the crude reactions into 384-well plates (200 nl at 1 mM in DMSO) and added sequentially 10 μl of IL-23R-Fc (10 nM final conc.) and 9.8 μl of a mixture containing the Eu-AntiHis$_6$ (1 nM final conc.) and the Cy5-labelled peptide 9 (50 nM final conc.). The screen showed results that we could expect based on the peptide sequences, namely that the longer peptides bind to the receptor while shorter ones do not (Figure 4a). Pleasingly, the two negative control peptides (random sequences, last two) did not show any binding. Peptides differing in only one amino acid, namely cysteine (C) replaced by homocycsteine (hC) showed very similar activities, which was further assuring that the assay works well. As an additional control, we excited directly Cy5 at 620 nm and measured emission at 665 nm, to see if the fluorescence signal was stable, or if it was also affected by some of the peptide sequences (signals shown in brown) (Figure 4b). The vast majority of wells showed a signal around 47,000 AU seen for the control, which showed that presence of the blocking agent albumin suppressed unspecific interactions of the fluorescent peptide. We concluded that the TR-FRET-based assay is robust and should be suited for screening compounds against IL-23R.

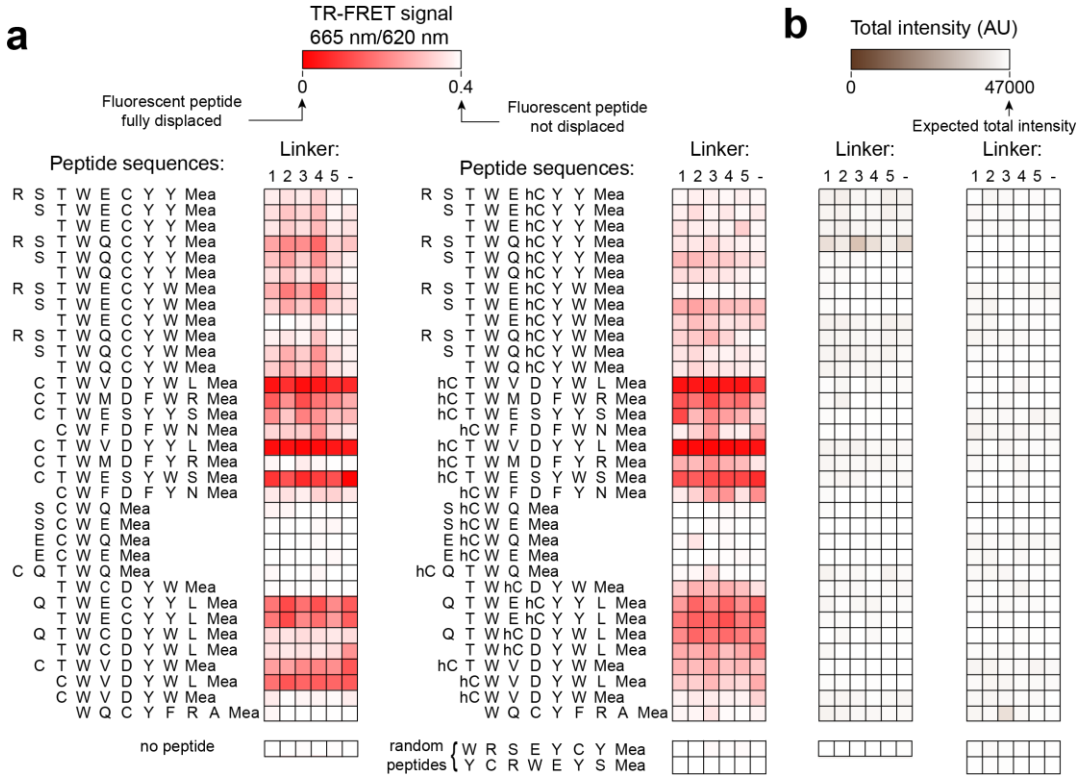


Figure 4. Testing the TR-FRET assay with potential IL-23R binders. (a) The assay was tested using the same cyclic peptides as shown in Figure 2. Displacement of the peptide 9 by the peptides was detected in 384-microwell plates by measuring a decrease in TR-FRET signal, assessed by a decrease of fluorescence emission of Cy5 (665 nM) and an increase of fluorescence emission of europium (620 nm). White = no displacement, red = full displacement. (b) The fluorescence intensity of Cy5 (independent of the FRET donor europium) was measured by exciting the probe at 620 nm and measuring emission at 665 nm. White color indicates the expected intensity and brown unexpected reduced intensity.

4.4.5 Screening a library of 3,072 cyclic peptides

We finally applied the established TR-FRET assay to screen a large library of cyclic peptides against IL-23R. We synthesized the library described in Figure 5, that contains peptides of 3-, 4- and 5-amino acids that all contain at both ends thiol groups so that they could be cyclized by incubation with bis-electrophile reagents. The peptides were synthesized on solid phase in 96-well plates using the same method as described for the pilot-scale library above. The thiol groups were introduced through the building blocks mercaptopropionic acid (Mpa, at the N-terminus) and mercaptoethylamine (Mea, at the C-terminus). To increase the chance of finding ligands of IL-23R, we tailored the library, choosing amino acids that were enriched in the phage display-selected peptides described above, as follows. All peptides were chosen to contain a tryptophan and a tyrosine residue at random positions, and additionally, the random amino acids (X) shown in Figure 5b. For the random amino acids (X), we used β-, γ- and δ-amino acids that alter the backbone, as well as α-amino acids containing polar, charged, aliphatic and aromatic side chains. Of the 165,660 peptide sequences that were possible to

form combinatorially, we randomly synthesized 384 (Figure S4). All peptides were cyclized by reaction with the seven bis-electrophile reagents shown in Figure 5c, and one control reaction without linker yielding 3,072 cyclic peptides. LC-MS analysis of eight randomly picked peptides showed that the synthesis was good and suggested that the library quality was sufficiently good for screening (Figure S5).

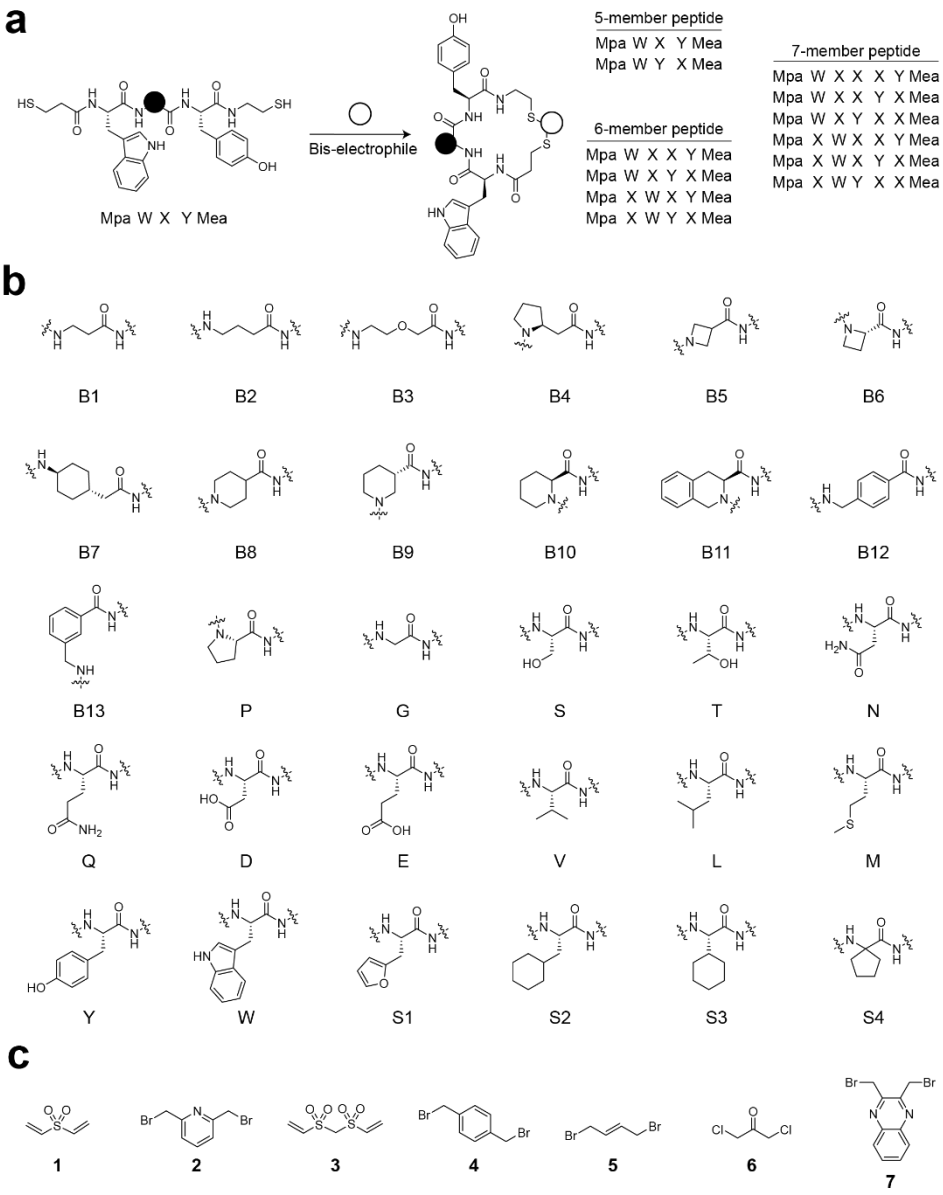


Figure 5. Design of pilot-scale cyclic peptide library. (a) Cyclization reaction shown for an example peptide and formats of linear peptides used for generating the library. The black dot indicates a random amino acid and the white dot indicates a bis-electrophile reagent / cyclization linker. All linear peptides synthesized contain two thiol groups, that are introduced through the building blocks mercaptopropionic acid (Mpa; at N-terminus) and mercaptoethylamine (Mea; at C-terminus). The peptides contain three, four or five random amino acids (X), of which one is always tryptophan (W) and one tyrosine (Y). (b) Random amino acids used in position (X). (c) Bis-electrophile reagents used for peptide cyclization.

We screened the library using the TR-FRET assay, as described for the pilot-scale library above. Several cyclization reactions showed strong inhibition, as seen by dark red color in Figure 6.

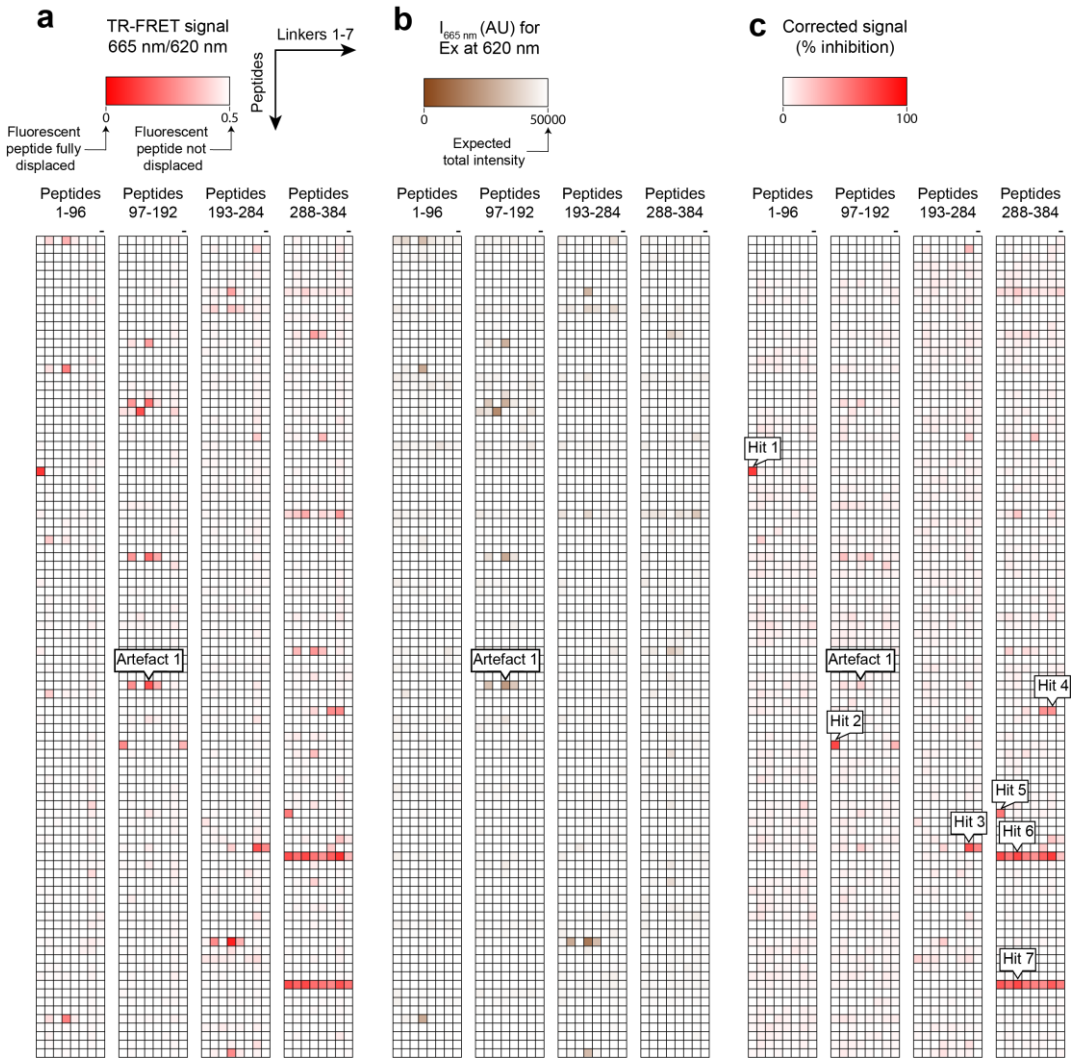


Figure 6. Screening pilot-scale library against IL-23R using TR-FRET-based assay. (a) Displacement of Cy5-labelled peptide 9 by 3,072 cyclic peptides (crude reactions), measured by TR-FRET in 384-well plates. The ratio of fluorescence intensity of Cy5 (665 nM) and fluorescence intensity of europium (620 nm) is shown. White = no displacement, red = full displacement. (b) Identification of artifacts due to effect of cyclic peptides onto Cy5-labelled peptide. The fluorescence intensity of Cy5 (independent of the FRET donor europium) was measured by exciting the probe at 620 nm and measuring emission at 665 nm. White color indicates the expected intensity, and brown unexpected reduced intensity. (c) Inhibition of IL-23R indicated in %. The inhibition was calculated using the data of panel (a) and (b) as described in the materials and methods.

In order to exclude that the activities were resulting from precipitation-triggered effects on the fluorescent peptide, we excited Cy5 and recorded emission (brown colors in Figure 6). This test revealed that some peptides were affected by a change in the Cy5 fluorescence, but that most of the hits seen were likely based on a reduced TR-FRET signal resulting from inhibition

of the IL-23R-Fc:peptide 9 interaction. We mathematically subtracted the effect of cyclic peptides on the Cy5-labeled peptide 9, and represented the activity of the cyclic peptides as % inhibition of the PPI.

4.4.6 Hit characterization

To assess the reproducibility of the TR-FRET assay, the screen, and the mathematical signal correction, we repeated the screen with the hits indicated in Figure 6 (hit 1 to 7). In this repetition, we included additional peptides as controls, such as one peptide that affected the Cy5 signal (artifact 1) and eight randomly chosen peptides that showed no activity in the initial screen (inactive 1 to 8). We found comparable activities for nearly all replicates as shown in Figure S6). Only hit 5 did not reproduce and was therefore considered as a false hit.

We chose seven of the cyclization reactions, indicated as "hits" in Figure 5, and analyzed the cyclization reactions by LC-MS to see if the desired products were present. Analysis of hits 1, 3, 4, 6 and 7 showed efficient but not complete conversion into the desired cyclic peptide, and for hits 2 and 5 mostly non-cyclized or disulfide-cyclized peptide (Figure S7). In order to see if the activities observed in the TR-FRET are based on the cyclic peptide products, we re-synthesized peptides of the hits 1, 3, 4, 6 and 7, purified the cyclic peptides by HPLC, and tested their activities using the TR-FRET based assays. The most active cyclic peptide showed an IC_{50} in the medium micromolar range (hit 4, 58 μM), which was a surprisingly low activity, given that the peptide concentration in the HTS was 10 μM and strong effects were seen by TR-FRET. This suggested that the activities found in the screen were due to side products (Figure S8). In order to identify the active side-products, we run the crude reactions over a reversed-phase column, collected 20 fractions, lyophilized them, and tested the activity of all fractions (Figure 7a & 7b, Figure S9). This experiment confirmed that for several hits, the most active species were not the desired cyclic peptides, but side products that eluted at different times. The hits 6 and 7 showed particularly active side products that we identified by high resolution MS as N-terminal thiol tBu-capped peptide single alkylated at the C-terminal thiol by the bis-electrophile linker and quenched on the remaining reactive site by 2-mercaptoethanol, and that competed with the IL-23R-Fc/peptide interaction with an IC_{50} of 1.3 ± 0.2 μM and 7.6 ± 0.3 μM respectively (Figure 7c & 7d). Identification of a highly active IL-23R ligand showed that the newly developed TR-FRET assay is suited for the HTS of compound libraries against this important therapeutic targets.

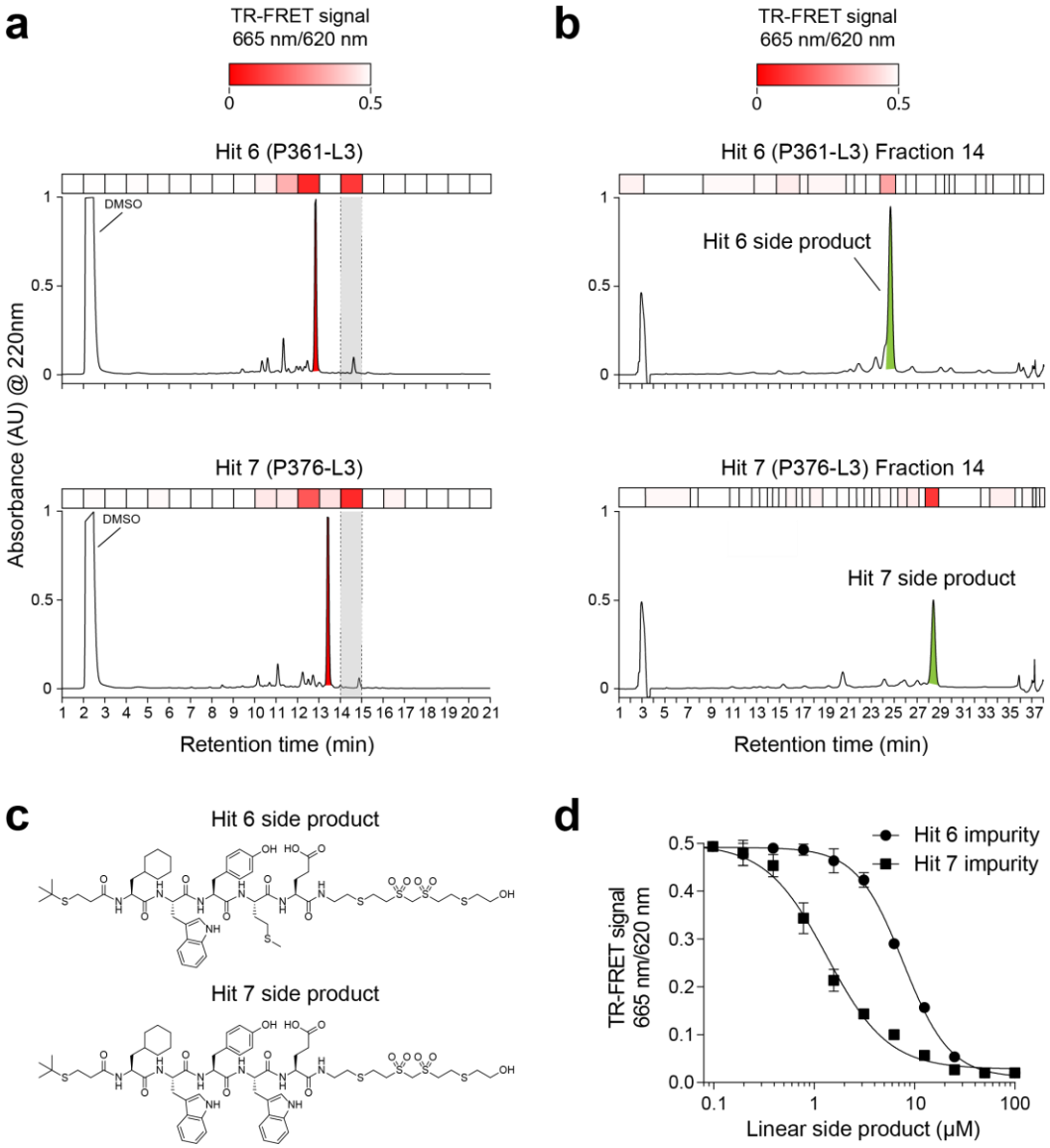


Figure 7. Identification of active species. (a) Competition assay of all species in the crude mixture of hit 6 and 7 (separated in 20 different fractions) against the TR-FRET assay. Highlighted in red is the macrocycle. Highlighted in grey is the fraction chosen (Fraction 14) to identify the active specie. (b) Competition assay of all species in the selected fraction of hit 6 and 7 highlighted in (a) against the TR-FRET assay. In green is the identified active specie. (c) Proposed structures of hit 6 and hit 7 side products. (d) Competition assay of pure hit 6 impurity and hit 7 impurity against the TR-FRET assay.

4.5 Conclusion

We have developed a TR-FRET-based assay for the high-throughput screening of compounds inhibiting the IL-23:IL23R PPI. The *in vitro* assay measuring the interaction of IL-23R and a synthetic peptide is designed to identify exclusively compounds that bind to a hotspot region of IL-23R forming the main contact with IL-23, and thus inhibitors of the PPI. Application of the HTS assay to a pilot-scale library of 3,072 crude macrocyclic products led to the identification of single-digit micromolar inhibitors of the IL-23:IL23R PPI. Taken together, this study provides a robust and accessible *in vitro* assay to screen large compound libraries against IL-23R and to identify inhibitors of the IL-23:IL-23R PPI required towards the development of orally available drugs for important diseases.

4.6 Materials and methods

Cloning expression vector for IL-23R extracellular domain

The following DNA sequences coding for IL-23R amino acids 24 to 353 (IL-23R extracellular region) or amino acids 24 to 125 (D1) (in bold), CD33 signal peptide (italic and underlined), IgG1 Fc domain (in italics), and His_6-tag (CATCACCATCACCATCAC) were synthesized and inserted into plasmid pUC57-Kan by GENEWIZ using DNA restriction sites *Xba*I and *Xho*I (underlined). The synthetic genes were sub-cloned into the eukaryotic expression vector pEXPR-IBA42 using the restriction enzymes *Xba*I and *Xho*I. The IL-23R amino acids were counted considering the endogenous signal peptide for secretion accounting for the amino acids 1 to 23.

Gene corresponding to IL-23R extracellular region-Fc fusion (IL-23R-Fc)

5'-

TCTAGACCCACC*ATGCCGCTGCTGCTACTGCTGCCCCTGCTGTGGGCAGGGGCCCTG*
*GCT***GGAATTACAAATATAAACTGCTCTGGCCACATCTGGGTAGAACCAGCCACAATTT**
TTAAGATGGGTATGAATATCTCTATATATTGCCAAGCAGCAATTAAGAACTGCCAACC
AAGGAAACTTCATTTTTATAAAAATGGCATCAAAGAAAGATTTCAAATCACAAGGATT
AATAAAACAACAGCTCGGCTTTGGTATAAAAACTTTCTGGAACCACATGCTTCTATGT
ACTGCACTGCTGAATGTCCCAAACATTTTCAAGAGACACTGATATGTGGAAAAGACAT
TTCTTCTGGATATCCGCCAGATATTCCTGATGAAGTAACCTGTGTCATTTATGAATATT
CAGGCAACATGACTTGCACCTGGAATGCTGGGAAGCTCACCTACATAGACACAAAAT
ACGTGGTACATGTGAAGAGTTTAGAGACAGAAGAAGAGCAACAGTATCTCACCTCAA
GCTATATTAACATCTCCACTGATTCATTACAAGGTGGCAAGAAGTACTTGGTTTGGGT
CCAAGCAGCAAACGCACTAGGCATGGAAGAGTCAAAACAACTGCAAATTCACCTGG
ATGATATAGTGATACCTTCTGCAGCCGTCATTTCCAGGGCTGAGACTATAAATGCTAC
AGTGCCCAAGACCATAATTTATTGGGATAGTCAAACAACAATTGAAAAGGTTTCCTGT
GAAATGAGATACAAGGCTACAACAAACCAAACTTGGAATGTTAAAGAATTTGACACC
AATTTTACATATGTGCAACAGTCAGAATTCTACTTGGAGCCAAACATTAAGTACGTATT
TCAAGTGAGATGTCAAGAAACAGGCAAAAGGTACTGGCAGCCTTGGAGTTCACTGTT
TTTTCATAAAACACCTGAAACAGTTCCCCAGGTCACATCAAAAGCATTCCAACATGAC
ACATGGAATTCTGGGCTAACAGTTGCTTCCATCTCTACAGGGCACCTTACTTCTGACA
ACAGAGGAGAC*GAGCCCAAATCTCAGGACAAAACTCACACATGCCCACCGTGCCCAG*
CACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACAC
CCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGA

AGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAG
ACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACC
GTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAG
CCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACC
ACAGGTGTACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAAT
GGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCC
TTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCT
TCTCATGCTCCGTGATGCACGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTC
*CCTGTCTCCGGGT*CATCACCATCACCATCACTGATAACTCGAG-3'

Gene corresponding to D1-Fc

5'-
TCTAGACCCACC*ATGCCGCTGCTGCTACTGCTGCCCCTGCTGTGGGCAGGGGCCCTG*
*GCT***GGAATTACAAATATAAACTGCTCTGGCCACATCTGGGTAGAACCAGCCACAATTT**
TTAAGATGGGTATGAATATCTCTATATATTGCCAAGCAGCAATTAAGAACTGCCAACC
AAGGAAACTTCATTTTTATAAAAATGGCATCAAAGAAAGATTTCAAATCACAAGGATT
AATAAAACAACAGCTCGGCTTTGGTATAAAAACTTTCTGGAACCACATGCTTCTATGT
ACTGCACTGCTGAATGTCCCAAACATTTTCAAGAGACACTGATATGTGGAAAAGACAT
TTCTTCTGGATATCCGCCA*GAGCCCAAATCTCAGGACAAAACTCACACATGCCCACCGT*
GCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAA
GGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAG
CCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAAT
GCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTC
CTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCA
ACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCG
AGAACCACAGGTGTACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGT
CAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGA
CGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGG
GAACGTCTTCTCATGCTCCGTGATGCACGAGGCTCTGCACAACCACTACACGCAGAAG
*AGCCTCTCCCTGTCTCCGGGT*CATCACCATCACCATCACTGATAACTCGAG-3'

The ordered plasmid were suspended in autoclaved mQ water and electroporated (1 μl at 400 ng/ml) into 50 μl electrocompetent XL1-blue E.coli cells using 1 mm electroporation cuvettes

(EP-201, Cellprojects) and an electroporation device (MicroPulser, BIO-RAD). Directly after the electroporation pulse, 1 ml of warm Luria Broth media (LB) was added to the cuvette and the mixture transferred to a sterile 14 ml tube before incubation at 37°C for 1 hr under agitation (250 rpm). A volume of 50 μl of the culture were plated on LB agar containing kanamycin (50 μg/ml) and incubated at 37°C overnight. Single colonies were picked to inoculate 10 ml of LB-containing 50 μg/ml kanamycin in 50 ml Falcon tubes and incubated at 37°C overnight under agitation (250 rpm). The cells were pelleted by centrifugation at 10,000 × g (around 11,000 rpm on an Eppendorf Centrifuge 5418) for 5 min at room temperature (RT). The plasmids were isolated using a plasmid miniprep kit (NucleoSpin® plasmid miniprep kit, Qiagen).

The plasmids (around 8 μg) and the vector pEXPR-IBA42 (around 8 μg) were digested using *Xba*I (4 μl,10,000 U/ml) and *Xho*I (2 μl, 10,000 U/ml) in CutSmart™ Buffer (BioLabs Inc.) in a final volume of 100 μl for 2 hrs at 37°C. The digested vector was additionally treated with FastAP phosphatase (8 μl, 1 U/ml) for 10 min at 37°C. The digested plasmids and vector were electrophoresed on a high quality agarose gel 1% (w/v) containing 0.5 μg/ml of ethidium bromide and 1 mM guanosine and TAE buffer. The DNA bands corresponding to the genes and vector were extracted using a scalpel under UV light (254 nm, 70% intensity) and extracted by a gel extraction kit (QIAquick® gel extraction kit, Qiagen). The concentration of the purified DNA was determined by measuring the absorption at 260 nm (NanoDrop 8000, Thermo Scientific).

The genes (around 500 ng) and the vector pEXPR-IBA42 (around 500 ng) were ligated in a volume of 20 μl for 4 hrs at RT using T4 ligase (2 μl, 400,000 U/ml) and T4 ligase buffer. Completion of the ligation reaction was assessed by agarose electrophoresis. The ligation mixtures were purified (NucleoSpin® plasmid miniprep kit, Qiagen) and concentrations determined by measuring the absorption at 260 nm. The ligated vectors were electroporated (1 μl at 500 ng/μL) into 50 μl electrocompetent XL1-blue E.coli cells as described above. A volume of 50 μl containing electroporated cells in LB media were plated on agar LB plates containing ampicillin (100 μg/ml) and incubated overnight at 37°C. Single colonies were selected to inoculate 10 ml of LB containing 100 μg/ml of ampicillin in a 50 ml Falcon tube and grown overnight at 37°C under agitation (250 rpm). The cells were pelleted by centrifugation at 10,000 × g (around 11,000 rpm on an Eppendorf Centrifuge 5418) for 5 min at 4°C and plasmid DNA was isolated using a plasmid miniprep kit (NucleoSpin® plasmid miniprep kit, Qiagen). Plasmid containing the correct DNA as verified by Sanger sequencing using the forward primer 5'-GAGAACCCACTGCTTACTGGC-3' and reverse primer 5'-TAGAAGGCACAGTCGAGG-3'.

Mammalian transient test expression of IL-23R-Fc and D1-Fc

Vector pEXPR-IBA42 containing the IL-23R-Fc or D1-Fc gene were transiently transfected into HEK293E or Expi-CHO-S cells. Transfection in HEK293E cells was performed by incubating the vector containing the desired gene (15 μg) and 30 μl of PEI-MAX at 1 mg/ml, into $10 \cdot 10^6$ cells in 0.5 ml of RPMI medium (Gibco, GlutaMAX, cat. 61870-010) + 0.1% pluronic F-68 for 90 min at 37°C under agitation. A volume of 9.5 ml of Excell293 medium (Sigma, cat. 14571C) containing 4 mM glutamine was then added and 5 ml of the resulting solution were transferred in an another tube and supplemented with valproic acid sodium salt (vpa) to a final concentration of 3.75 mM, leaving the other 5 ml without any vpa. The two tubes were then agitated for 7 days at 37°C. The cells were removed by centrifugation leaving only the expression medium with expressed protein. Transfection in Expi-CHO-S cells (Gibco, cat. A29127) was performed by incubating the vector containing the desired gene (30 μg) and 150 μl of PEI-MAX at 1 mg/ml, into $50 \cdot 10^6$ cells in 10 ml of ProCHO5 medium (Lonza, cat. BELN12-766Q) containing 4 mM glutamine for 60 min at 31°C under agitation. A volume of 5 ml of the resulting solution was transferred in another tube and supplemented with DMSO to reach a final concentration of 2% (v/v), leaving the other 5 ml without any DMSO. The two tubes were then agitated for 7 days at 31°C. The cells were then removed by centrifugation leaving only the expression medium with expressed protein. Samples of 150 μl of each expression media were concentrated to around 20 μl by centrifugation at 10,000 × g (around 11,000 rpm on an Eppendorf Centrifuge 5418) using a 3 kDa MWCO Nanosep® centrifugal filter. The samples were analyzed by SDS-PAGE using a 12% (w/v) RunBlue SDS protein gel and standard SDS running buffer (40 mM tricine, 60 mM TRIS, 0.1% [w/v] SDS, 2.5 mM sodium metabisulfite, pH 8.4). The protein in the gel were stained using staining solution (0.1% [w/v] Coomassie R-250 in 50% [v/v] methanol, 10% [v/v] glacial acetic acid and 40% [v/v] mQ water) for 2 hrs. Destaining was performed using destaining solution (50% [v/v] mQ water, 40% [v/v] methanol and 10% [v/v] acetic acid) overnight. Three of the 5 ml test expressions that showed protein expression (D1-Fc and IL-23R-Fc in HEK293 with vpa, and D1-Fc in CHO with DMSO) were run over a 1 ml His-trap column (HisTrap™ excel, GE Healthcare) to purify His-tagged protein. After washing with 10 column volumes of wash buffer (20 mM Na_2HPO_4, 3.7 mM $NaH_2PO_4.H_2O$, 500 mM NaCl, 10 mM Imidazole, pH 7.4), the protein was eluted with 5 column volumes of elution buffer (20 mM Na_2HPO_4, 3.7 mM $NaH_2PO_4.H_2O$, 500 mM NaCl, 500 mM Imidazole, pH 7.4) and analyzed by SDS-PAGE. The proteins were concentrated to 2 ml using a 10,000 MWCO Macrostep® Advance Centrifugational Device and centrifuging at 3,400 × g (around 4,000 rpm with a Thermo Scientific Multifuge 3L-R with a Sorval 75006445 rotor) at 4°C. The concentration of imidazole was reduced in 5 cycles of adding 20 ml of cold wash buffer and centrifugational concentration to 2 ml. A final round of centrifugational concentration

was performed to reduce the volumes to 100 μl and the samples were analyzed by SDS-PAGE. The protein concentration was determined based on the absorbance at 280 nm (For IL-23R-Fc: MW = 64,713 Da, ε = 111,730 $M^{-1} \cdot cm^{-1}$. For D1-Fc: MW = 38,584 Da, ε = 53,860 $M^{-1} \cdot cm^{-1}$, NanoDrop 8000, Thermo Scientific).

Large scale expression of IL-23R-Fc in mammalian cells

A 10 ml pre-culture of LB medium containing 100 μg/ml of ampicillin in a 50 ml falcon tube was inoculated with XL1 *E. coli cells* containing pEXPR-IBA42 with the IL-23R-Fc gene and incubated overnight at 37°C under agitation (250 rpm). A culture flask containing 1 L of LB medium with ampicillin (100 μg/ml) was then inoculated with the entire pre-culture and incubated overnight at 37°C under agitation (250 rpm). The culture was then centrifuged at 5,500 × g (around 5,000 rpm using a Thermo Scientific SORVALL RC BIOS centrifuge with a Fiberlite™ F8-6x1000y rotor) for 5 min at 4°C. The supernatant was discarded and the plasmid DNA was purified from the resulting pellet using a maxiprep kit (Macherey-Nagel, NucleoBond® Xtra Maxi). The plasmid was sterilized by precipitation in isopropanol followed by a 70% ethanol wash and air-drying close to a flame. IL-23R-Fc was produced in HEK293 cells by transient expression as follows. The vector pEXPR-IBA42 containing IL-23R-Fc gene (1.5 mg) was transfected into HEK293 cells in 1 L media containing valproic acid as described previously. After 1 week expression, the media containing the secreted protein was run over two connected 1 ml His-trap columns (HisTrap™ excel, GE Healthcare) overnight at 4°C using a peristatic pump (BIORAD Econo Pump) at a constant flow rate of 0.7 ml/min. After washing with 20 column volumes of wash buffer, the protein was eluted with 10 column volumes of elution buffer and collected in fractions of 5 ml. Samples of the fractions were analyzed by SDS-PAGE. The fractions containing IL-23R-Fc were combined and the volume reduced to around 2 ml using a 10,000 MWCO Macrostep® Advance Centrifugational Device and centrifuging at 3,400 x g (around 4,000 rpm with a Thermo Scientific Multifuge 3L-R with a Sorval 75006445 rotor) at 4°C. The concentration of imidazole was reduced as described above and the protein concentration determined by measuring absorbance.

Synthesis of fluorescent peptides for FP assay

Peptides were synthesized by Fmoc chemistry using Rink amide MBHA (50 μmol scale, loading of 0.3 mmol/g) in 5 ml syringes on an automated peptide synthesizer (Intavis, MultiPep RSi). The resin was initially washed and swollen by adding three times 2.4 ml of DMF for 10 min. The amino acids were coupled as follows: amino acid (4.2 equiv., 0.22 M final), coupling agent HATU (3.9 equiv., 0.21 M final), the base *N*-methyl-morpholine (NMM; 10 equiv., 0.53

M final), all in a total volume of 460 μl DMF, and 5 μl of *N*-methyl-2-pyrrolidone (NMP) were mixed and incubated for 1 min, transferred to the resin, and shaken for 45 min at 400 rpm. Each coupling step was done twice followed by seven cycles of washing as follows. A volume of 2 ml of DMF was added, shaken for 1 min (400 rpm) and removed by applying vacuum for 1 min. Two cycles of capping were performed by adding 0.8 ml of 5% (v/v) Ac_2O, 6% (v/v) 2,6-lutidine in DMF for 5 min. After seven washing cycles, the Fmoc protecting groups were removed in two cycles of adding 0.8 ml of 20% (v/v) piperidine in DMF and incubation for 5 min.

The peptides were labelled on solid phase with fluorescein, either at the N-terminus by coupling 5(6)-carboxyfluorescein to the N-terminal amino group, or to the C-terminus by coupling 5(6)-carboxyfluorescein to the ε-amino group of a lysine introduced at the C-terminus. For the labelling at the C-terminus, Fmoc-Lys(Dde)-OH was introduced at the C-terminus. The Dde protecting group was selectively deprotected before 5(6)-carboxyfluorescein coupling by incubating the resin four times with 2 ml of 2% (v/v) hydrazine in DMF for 3 min. As hydrazine also cleaves Fmoc groups, the N-terminus was protected by a Boc protecting group.

The protecting groups were removed and the peptides cleaved from the resin using 5 ml of 90% (v/v) TFA, 2.5% (v/v) thioanisole, 2.5% (v/v) 1,2-ethanedithiol, 2.5% (v/v) mQ water, 2.5% (w/v) phenol for 4 hrs at RT under agitation (200 rpm). After transferring into a 50 ml Falcon tube, around 4 ml of TFA was evaporated under a constant flow of nitrogen. Peptides were then precipitated in 50 ml cold diethylether for 20 min at -20°C, and centrifuged at 3,400 × g (around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 10 min at 4°C. Solid yellow pellets were obtained and air-dried for 20 min while the supernatants were discarded.

For the peptide cyclization, the crude peptides were dissolved in 4% (v/v) DMSO, 20% (v/v) ACN and 66% (v/v) aqueous buffer (60 mM NH_4HCO_3, pH 8.0), and cyclization reagent divinylsulfone added (5 μl at 9.96 M). The reaction mixture was incubated at 30°C for 2 hrs, and then lyophilized.

For HPLC-purification, the crude peptides were dissolved in 10% (v/v) DMSO, 30% (v/v) ACN, 59.9% (v/v) mQ water, and 0.1% (v/v) TFA, and run on a reversed-phase C18 column (Waters Sunfire™, 10 μm, 100 Å, 19 × 250 mm) with a linear gradient of solvent B (99.9% [v/v] ACN, 0.1% [v/v] TFA) over solvent A (99.9% [v/v] mQ water, 0.1% [v/v] TFA) from 30 to 60% in 30 min at a flow rate of 20 ml/min. The masses of the purified peptides were confirmed by electrospray ionization mass spectrometry using a single quadrupole MS instrument described below. Fractions with the desired mass were combined and lyophilized. Yellow powders were obtained and considered as TFA salts. Purity was assessed by analytical HPLC.

LC-MS analysis

Linear and cyclic peptides were analyzed with an ultra HPLC (UHPLC) and single quadrupole MS system (Shimadzu-2020). Samples were always prepared as a mixture of ACN and mQ water, and applied on a reversed-phase C18 column (Phenomenex Kinetex®, 2.6 μm, C18, 100 Å column, 50 × 2.1 mm) with a linear gradient of solvent B (99.95% [v/v] ACN and 0.05% [v/v] formic acid), over solvent A (99.95% [v/v] mQ water and 0.05% [v/v] formic acid) starting from 10 to 100% at a constant flow rate of 1 ml/min.

Analytical HPLC

The purity of linear and cyclic peptides were analyzed with reversed-phase analytical HPLC (Agilent Technologies, 1260 Infinity). All fractions were prepared as a mixture of ACN and mQ water containing 0.1% (v/v) TFA, and run against a reversed-phase C18 column (ZORBAX 300SB-C18, 5 μm, 4.6 mm x 20 mm) using a linear gradient of solvent B (99.9% [v/v] ACN, 0.1% [v/v] TFA) over solvent A (94.9% [v/v] mQ water, 5% [v/v] ACN, 0.1% [v/v] TFA) starting from 0 to 100% at a constant flow rate of 1 ml/min. Purities were analyzed based on absorbance peaks at 220 nm.

Fluorescence polarization assay

The binding of fluorescent peptides were assessed by incubating the fluorescent peptide (10 nM final concentration) with 2-fold dilutions of IL-23R (IL-23R-Fc or D1-Fc) in assay buffer (27 mM TRIS, 137 mM NaCl, 2.7 mM KCl pH 7.4, 0.01% [v/v] Tween-20) in a 384 well plate (ThermoFischer NUNC™ 384 shallow well, std height plate, non-sterile, black) for 10 min at RT. The fluorescence anisotropy was measured (Infinite F200 pro TECAN, E_{ex} = 485 nm, E_{em} = 535 nm, 25°C) using 10 nM 5(6)-carboxyfluorescein in assay buffer as a reference (minimum of polarization) and assay buffer alone as a blank for the calculation of the G-factor. The K_D was obtained from fitted data using Graphpad Prism 5 and the following formula (1):

$$y = a + (b - a)\frac{(K_D + x + P) - \sqrt{(K_D + x + P)^2 - 4xP}}{2P} \qquad (1)$$

Where y is the anisotropy measured, a and b are the signal in absence or with fluorescent ligand respectively, x is the protein concentration and P is the probe concentration.

Fluorescence polarization competition assay

The control peptide HCTWMDFWRCS-NH_2 cyclized with divinylsulfone was used to test the displacement of peptide 6 (10 nM final concentration) from IL-23R-Fc (60 nM final

concentration; leading to 75% of probe bound). The competition experiments were performed in assay buffer (27 mM TRIS, 137 mM NaCl, 2.7 mM KCl pH 7.4, 0.01% [v/v] Tween-20) and a 384-well plate. The peptide 6/ IL-23R-Fc complex was incubated with 2-fold dilutions of the control peptide (5 μM to 3 nM final concentration), incubated for 30 min at RT, and the fluorescence anisotropy was measured as described above. As inhibitor peptide and peptide 6 stocks were in DMSO, the final concentration of DMSO in the assay was 1% (v/v).

The IC_{50} indicating the inhibitor concentration resulting in 50% displacement of the fluorescent peptide was obtained from fitted data using Graphpad Prism 5 and the following equation (2):

$$y = \frac{100}{1+10^{(\log IC_{50}-x)p}} \qquad (2)$$

With y being displacement of reporter peptide (%), x the peptide concentration and p the Hill slope.

Synthesis of 2-pyridylthio cysteamine·HCl

Cysteamine (1 equiv. in 30 ml MeOH, 0.39 M), 2,2'-dipyridyldisulfide (2 equiv., in 70 ml MeOH, 0.77 M) and AcOH (2 ml) were stirred overnight at RT under agitation in a 250 ml balloon. Reaction volume was then decreased by rotation evaporation to around 15 ml and split between 8 different 50 ml Falcon tubes, complemented to around 5 ml with MeOH, and desired 2-pyridylthio cysteamine was precipitated by adding 45 mL of RT diethyl ether on top. After 30 min at -20°C, the precipitate was obtained by centrifugation at 3,400 × g (around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 5 min at 4°C and discarding the supernatant. The precipitate was further purified in 6 cycles by dissolving in 5 ml of MeOH followed by precipitation and centrifugation as described above. After a final centrifugation step, the white powder was dried by lyophilisation. The product was considered to be a chloride salt.

Preparation of polystyrene-S-S-cysteamine resin

Into six 20 ml syringes was added 534 mg of resin (Rapp Polymere A SH resin, 200-400 mesh, 0.82 mmol/g loading), corresponding to 0.44 mmol scale per syringe, and 2.64 mmol scale in total. The resin of each syringe was washed with 15 ml of DCM, then swelled in 15 ml of 30% (v/v) MeOH/70% (v/v) DCM for 20 min. Previously synthesized 2-pyridylthio cysteamine·HCl (2.05 g, 3.5 equiv.) was dissolved in 21 ml of MeOH, then 49.44 ml DCM and 1.94 ml DIPEA (11.15 mmol, 4.2 equiv.) were added. A volume of 12.1 ml of this solution was pulled into each syringe, and the syringes were shaken (250 rpm) at RT for 3 hrs. After this time, the solutions were discarded, and the resins were washed with 2 × 20 ml 30% (v/v) MeOH/70% (v/v) DCM

then 2 × 20 ml DMF. The resins were then combined into two syringes and washed with 9 ml of 1.2 M DIPEA solution in DMF for 5 min to ensure complete deprotonation of all amines. This solution was discarded, and the resin was washed with 2 × 20 ml ml DMF and 3 × 20 ml DCM, then kept under vacuum overnight to yield a free-flowing powder.

Cyclic peptide library synthesis

Cyclic peptides were synthesized by cyclizing linear peptides via two thiol groups using bis-electrophile reagents. The linear peptides were synthesized in 96-well synthesis plates (standard 96-well filter plate, 0.7 ml/well, 10 μm PE frit) on polystyrene-S-S-cysteamine resin (5 μmol scale, 0.82 mmol/g loading) in using an automated peptide synthesizer (Intavis, Multipep RSi). All amino acids, including the N-terminal building block S-trityl-mercaptopropionic acid, were coupled following standard Fmoc-chemistry as described above.

For protecting groups removal, the bottom of the 96-well synthesis plate was sealed by pressing the plate onto a soft 6 mm thick ethylene-vinyl acetate foam pad (Rayher Hobby GmbH, 78 263 01), and the resin of each well was incubated with around 600 μl of 95% (v/v) TFA, 2.5% (v/v) mQ water and 2.5% (v/v) TIS. The plates were covered with a polypropylene adhesive seal, then weighed down by placing a weight (around 2 kg) on top to ensure that no leakage occurred. After 3 hrs, the synthesis plates were placed onto 2 ml deep-well plates, and the TFA mixture was removed by filtration on a TFA-compatible cleavage station. The wells were washed 3 times with around 600 μl of DCM, then allowed to air dry for 30 min.

The peptides were released from the resin by reduction of the disulfide bond linking the peptide to the resin. Plates were pressed into foam pads as described above to plug the openings. A volume of 200 μl 150 mM 1,4-butanedithiol (BDT, 4 equiv.), 150 mM triethylamine (TEA, 4 equiv.) in DMF was added to each well and incubated for 5 hrs. The synthesis plates were covered with PP plastic seals to prevent reagent evaporation. The 96-well synthesis plates were then transferred to 96-deep well plate recipient (Biotix, deep well plate, 96-round well, 1.2ml) and centrifuged at 480 × g (around 1,500 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 2 min at RT. The reductive release of the peptides was repeated once applying again a volume of 200 μl. A volume of 61 μl of 10% (v/v) TFA in water (8 equiv.) was added to each well of the plate. The solvent, BDT and TEA were removed using a SpeedVac (RVC 2-33 CDplus, 3 hrs, 30°C, 1750 rpm, 0.1 mbar, cooling trap = -85°C) and a plate compatible module (TMM rotor, cat. 124700). The dried peptides were either powder or oil. A volume of 50 μl DMSO was added to each well, the plates sonicated for fully dissolving the peptides, and the peptide transferred to ECHO-qualified 384-well PP plates

(Labcyte 384 PP, Cat. # PP-0200). The plates were centrifuged at 480 × g (around 1,500 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 2 min at RT. The purity of the peptides was analyzed by LC-MS.

The concentrations of the peptides were determined by measuring the concentrations of sulfhydryl groups using the Ellman's assay. Volumes of 23.97 μl of 10% (v/v) DMSO, 150 mM NH_4HCO_3, pH 8.0 and 6 μl of 10 mM 5,5'-dithio-bis-[2-nitrobenzoic acid] (20 equiv. if 100% yield for 5 μmol scale) in 10% (v/v) DMSO, 150 mM NH_4HCO_3, pH 8.0 were transferred to a 384 well plate (384 Greiner PS black F-bottom uClear MB) using a bulk dispenser (Certus Flex, Gyger). Volumes of 30 nl of the reduced peptides were transferred directly to the 384 well plate using acoustic droplet ejection (ECHO 650, Labcyte, DMSO settings). The plate was then manually sealed with a plastic foil and centrifuged at 480 × g (around 1,500 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 2 min at RT. The peptide concentration was determined by measuring the absorbance at 412 nm (Tecan Infinite M200 PRO) comparing it to a calibration curve. Peptides having concentrations above 20 mM were diluted to 20 mM by adding DMSO using the bulk dispenser.

The peptides were cyclized by transferring 1,250 nl of reduced peptide (20 mM in DMSO, 1 equiv.) into wells of ECHO-qualified 384-well PP plates using acoustic droplet ejection (ECHO 650, Labcyte, DMSO settings), and dispensing 25 μl of bis-electrophile reagents dissolved at 4 mM (4 equiv.) in 60 mM NH_4HCO_3 prepared in 50% (v/v) ACN/50% (v/v) mQ water pH 8.0 using the bulk dispenser. The plates were manually sealed with a PP plastic lid and centrifuged at 480 × g (around 1,500 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 2 min at RT. After 2 hrs incubation at RT, unreacted cyclization reagent was quenched by adding 12.5 μl of 32 mM β-mercaptoethanol (BME; 16 equiv.) in 60 mM NH_4HCO_3 50% (v/v) ACN/50% (v/v) mQ water pH 8.0 using a bulk dispenser, manually sealed and centrifuged as described above. After 2 hrs incubation at RT, the mixture was evaporated directly in the plate using the SpeedVac (10 hrs, 40°C, 1,750 rpm, 0.1 mbar, cooling trap = -85°C). The cyclic peptide pellets/oils were dissolved in 25 μl DMSO using a bulk dispenser to reach a final concentration of 1 mM. The purity of the cyclic peptides was analyzed by LC-MS.

Screening cyclic peptides with the FP assay

The synthesized cyclic peptides at a concentration of 1 mM in DMSO were stored in ECHO-qualified 384-well PP plates and transferred into low volume 384-well plates (ThermoFischer NUNC™ 384 shallow well std height plates non-sterile, black, no lid; Cat. # 264705) using acoustic droplet ejection (ECHO 650, Labcyte, DMSO transfer settings). To sixteen dedicated wells to be used as negative (maximum of anisotropy) and positive controls (minimum of

anisotropy), 150 nl of DMSO were transferred. To all the other wells, 150 nl of cyclic peptide at 1 mM were transferred. IL-23R-Fc (60 nM final concentration) and the fluorescent peptide 6 (10 nM final concentration) were diluted using screening buffer (33 mM TRIS, 137 mM NaCl, 2.7 mM KCl + 0.01% [v/v] Tween-20, pH 7.4) and mixed. After 10 min incubation at RT, 15 μl of this reagent mix was added to each well of the assay plates (except positive control) using a bulk microplate dispenser (Certus Flex, Gyger). For the positive controls, the same procedure was applied but without protein in the solution. The concentration of cyclic peptides was 10 μM. Final DMSO content was 1% (v/v). A plastic lid was manually added and the plates were incubated for 30 min at RT. The fluorescence anisotropy was recorded as described above.

Probe synthesis for TR-FRET assay

Peptide 9 (sequence: EELCTWVDYWLCHEE-PEG_2-Propargylglycine-NH_2) was synthesized by Fmoc-chemistry using Rink amide MBHA resin (50 μmol scale, loading of 0.3 mmol/g) loaded in a 5 ml syringe reactor on an automated peptide synthesizer (Intavis, MultiPep RSi). The synthesis, deprotection and cleavage, ether precipitation and HPLC-purification were done as described above. Cy5 was coupled to the purified peptide by click chemistry as follows. Sulfo-Cy5-N_3 and the peptide were prepared as 10 mM DMSO stocks. $CuSO_4$ and tris(benzyltriazolylmethyl)amine (TBTA) were prepared at 100 mM in mQ water. Sodium ascorbate was prepared at 500 mM in mQ water. Prior to the reaction, $CuSO_4$ and TBTA at 100 mM were mixed in a 1:1 ratio yielding a mixture containing both species at 50 mM. A characteristic blue color attesting of the proper coordination of TBTA on Cu^{2+} was obtained. The reaction was performed by mixing 28.82 μl of peptide at 10 mM (1 equiv.) and 28.82 μl of Sulfo-Cy5-N_3 at 10 mM (1 equiv.) with 5.76 μl of the $CuSO_4$/TBTA mixture at 50 mM (1 equiv.). The reaction was initiated by adding 8.65 μl of sodium ascorbate at 500 mM (15 equiv.). After 1 hr incubation at RT, completion of the reaction was monitored by LC-MS. The mixture was purified by reversed-phase on a C18 column (Nova-Pak® C18, 6 μm, 7.8 x 300mm) by UHPLC (Thermo Scientific, UltiMate 3000, Flow cell: PEEK, 0.7 μl, 0.4 mm, 100 bar) using a linear gradient of solvent B (99.9% [v/v] ACN and 0.1% [v/v] TFA) over solvent A (99.9% [v/v] mQ: H_2O and 0.1% [v/v] TFA) from 10 to 60% in 30 min at a flow rate of 4 ml/min. Purification of the product immediately after the one hour reaction was important to prevent side reactions. The HPLC-fraction with the desired peptide product was identified by electrospray ionization MS (ESI-MS) as described above. After lyophilisation, a solid powder of peptide TFA salt was obtained.

Establishing the TR-FRET assay

Different combinations of concentrations of the Cy5-labeled peptide 9 and of the protein IL-23R-Fc were tested in a TR-FRET assay as follows. Volumes of 7.5 μl of 2-fold dilutions of Cy5-labeled peptide 9 (from 1,333 to 83 nM, or no peptide) and 7.5 μl of 2-fold dilutions of IL-23R-Fc (from 213 to 13 nM, or no protein) in TR-FRET assay buffer (50 mM TRIS, 50 mM NaCl, 0.1% [w/v] BSA, 0.01% [v/v] Tween-20, pH 7.4, filtered with a 0.2 μm filter) were added to a 384-well plate. After 10 min incubation at RT (protected from light), 5 μl of Eu-AntiHis$_6$ (LANCE-Eu-W1024 Anti 6xHis, PerkinElmer) prepared at 4 nM in assay buffer was added. The final assay volume was 20 μl containing peptide 9 at final concentrations ranging from 500 to 31 nM or no peptide, IL-23R-Fc at final concentrations ranging from 80 to 5 nM or no protein, and Eu-AntiHis$_6$ at a final concentration of 1 nM. A plastic lid was manually added and the plate was centrifuged at 480 × g (around 1,500 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 2 min at RT. After 30 min incubation at RT, the TR-FRET signal was measured using a plate reader (PheraStar, BMG) exciting at 337 nm (TRF laser, 5 flashes) and measuring with a delay of 60 μs the emission at both, 620 nm (from europium) and 665 nm (from Cy5) with a measurement window of 400 μs. TR-FRET signal was calculated from the ratio between the emission at 665 nm over the emission at 620 nm (equation 3).

$$TR\ FRET = \frac{F_{665\ nm}}{F_{620\ nm}} \qquad (3)$$

The signal-to-background (S/B) was calculated from the ratio between the TR-FRET signal observed at a defined protein, probe and europium concentrations, over the TR-FRET signal observed in the same condition but without protein (equation 4).

$$S/B = \frac{TR\ FRET_{protein/probe/europium}}{TR\ FRET_{probe/europium}} \qquad (4)$$

Competition binding assay by TR-FRET

Cy5 labelled peptide 9 (50 nM final concentration), IL-23R-Fc (10 nM final concentration) and Eu-AntiHis$_6$ (1 nM final concentration) were incubated with 2-fold dilutions of a non-labeled IL-23R-binding peptide (control peptide; 50 to 0.02 μM final concentrations) in TR-FRET assay buffer in a 384 well plate (20 μl final volume), centrifuged, incubated for 30 min, and TR-FRET signal measured as described above. The Cy5-labelled peptide 9 and the control peptide were initially prepared in DMSO, which added a small volume of DMSO. Therefore, all wells were complemented with DMSO to reach a final concentration of 1% (v/v) DMSO. Sigmoidal curves

were fitted to the data using Graphpad Prism 5 software and the dose-response equation (equation 2).

Screening cyclic peptides with the TR-FRET assay

The library of cyclic peptides was synthesized as described above and stored in ECHO-qualified 384-well PP plates at 1 mM in DMSO. The peptides were transferred to low volume 384-well plates (ThermoFischer NUNC™ 384 shallow well std height plates non-sterile, black, no lid; Cat. # 264705) using acoustic droplet ejection (ECHO 650, Labcyte, DMSO transfer settings). To sixteen dedicated wells to be used as negative (maximum of TR-FRET signal) and positive controls (minimum of TR-FRET signal), 200 nl of DMSO were transferred. To all the other wells, 200 nl of cyclic peptide at 1 mM were transferred. IL-23R-Fc in TR-FRET assay buffer (10 µl, 10 nM final concentration) was added to each well using the bulk microplate dispenser (Certus Flex, Gyger). A mixture of Cy5 labelled peptide 9 (50 nM final concentration) and Eu-AntiHis$_6$ (1 nM final concentration) diluted in TR-FRET assay buffer, in a volume of 9.8 µl, were added to the wells using the bulk microplate dispenser. Final DMSO content was 1% (v/v). A plastic lid was manually added and the plates were centrifuged at 1,100 × g (around 1,500 rpm with SIGMA 4-16KS with a swing-out rotor 11660) for 2 min at RT. After 30 min incubation at RT (protected from light), the TR-FRET signal was recorded as described above.

Correction of signal

The potential artifact effects (such as quenching) of the tested cyclic peptides on the fluorescent probe were assessed by recording the fluorescent intensity observed at 665 nm (Cy5 emission) of the same plate upon direct excitation at 620 nm with a plate reader (Tecan Infinite M200 PRO). The fluorescent intensity signals observed were normalized considering the average of the negative and positive controls as 100%. If the compounds do not have any influence on the Cy5-labelled peptide, the measured fluorescent intensity should not be impacted (stay at 100%). Any effect on the resulting Cy5 emission signal will have a direct impact on the recorded TR-FRET signal. TR-FRET signal recorded was then normalized considering the average of the negative and positive controls as 100 and 0% respectively. In order to distinguish real hits from false hits due to artifact effects on the fluorescent probe from the cyclic peptides, we calculated a corrected signal based on the difference between the normalized fluorescence intensity upon excitation at 620 nm, and the corresponding normalized TR-FRET signal observed (equation 5). The corrected signal was expressed as the percentage of inhibition observed.

$$Corrected\ signal\ (\%\ inhibition) =$$

$$Normalized\ I_{665\ nm} upon\ Ex.at\ 620\ nm\ (\%) - Normalized\ TR\ FRET\ signal\ (\%) \quad (5)$$

Synthesis of cyclic peptides at a preparative scale

Peptides were synthesized at a 50 μmol scale in 5 ml syringe reactors essentially as described for the 5 μmol scale using 10-fold larger volumes. The products obtained were separated on a reversed-phase C18 column (XBridge® Peptide BEH C18, 5 μm, 300 Å, 10 mm × 250 mm) using semi-prep HPLC (Waters Sunfire™, 10 μm, 100 Å, 19 × 250 mm) over solvent A (99.9% [v/v] mQ water, 0.1% [v/v] TFA) from 20 to 70% in 30 min at a flow rate of 8 ml/min. After lyophilisation, the cyclic peptides were dissolved in DMSO to a concentration of 10 mM. The masses of the purified cyclic peptides were confirmed by LC-MS.

Identification of active compounds in reactions

A small fraction (5%; corresponding to around a 2.5 μmol scale synthesis) of the cyclic peptides produced at a preparative scale were separated on a reversed-phase C18 column (Nova-Pak® C18, 6 μm, 7.8 x 300mm) by UHPLC (Thermo Scientific, UltiMate 3000, Flow cell: PEEK, 0.7 μl, 0.4 mm, 100 bar) using a linear gradient of solvent B (99.9% [v/v] ACN and 0.1% [v/v] TFA) over solvent A (99.9% [v/v] mQ waterand 0.1% [v/v] TFA) from 20 to 80% in 20 min at a flow rate of 4 ml/min collecting fractions of 1 min per tube. The solvent in the fractions in plastic tubes was evaporated using a SpeedVac (RVC 2-33 CDplus,10 hrs, 40°C, 1,750 rpm, 0.1 mbar, cooling trap = -85°C) with a tube compatible module (TMM rotor, cat. 124627). The powder/oil obtained in the tubes was dissolved in a large volume of 60 μl of DMSO and transferred to an ECHO-qualified 384-well PP plate. The DMSO in the wells was again evaporated using the SpeedVac (RVC 2-33 CDplus, 3 hrs, 30°C, 1,750 rpm, 0.1 mbar, cooling trap = -85°C) and a plate compatible module (TMM rotor, cat. 124700). The resulting powder/oil was dissolved in 20 μl of DMSO and the plate was centrifuged at 1,100 x g (around 2,500 rpm with SIGMA 4-16KS with a swing-out rotor 11660) for 2 min at RT. The plate was turned and centrifuged again to make sure that all cyclic peptides was dissolved in DMSO. Volumes of 200 nl of each fraction was transferred to wells of a 384-well plate using acoustic dispensing. The activity of the fractions was tested using the TR-FRET assay as described above. Fractions showing activity were analyzed by LC-MS in order to identify the active species.

Preparation of 3-(tert-butylsulfanyl)propionic acid (tBu-Mpa)

In a 250 ml balloon placed in an ice bath, a volume of 50 ml of mQ: water was added followed by the dropwise addition of 80 ml H_2SO_4. Volumes of 40 ml tert-butanol (8.9 equiv.), then 4.1 ml 3-mercaptopropionic acid (1 equiv.) were added drop-by-drop. After 10 min incubation in the ice bath, the reaction mixture was warmed to 30°C using a water bath, and incubated for 3 hrs at 30°C. The resulting solution was transferred in a separatory funnel and the aqueous phase was washed with 3 × 25 ml Et_2O. Compound was then extracted by 4 × 50 ml DCM, all organic phase fractions were combined, and washed with 50 ml mQ water followed by 2 × 50 ml brine solution. Anhydrous sodium sulfate was then added in the organic phase until this one was fully dried as observed by the formation of larger crystal clumps, later removed by filtration. The solvent was then removed by rotation evaporation until a yellow oil was obtained. This oil was then solubilized in 10 ml 50% (v/v) ACN/ 50% mQ water, frozen in liquid nitrogen and lyophilized. Purity of the compound was assessed by ^{1}H-NMR.

Synthesis of linear impurities (tBu-Mpa-peptide-linker-BME)

Identified linear impurities were synthesized by reacting linear peptides via one C-terminal thiol with bis-electrophile reagent and quenching solution, while the N-terminal thiol was capped with tBu preventing its reactivity. Linear peptides were synthesized in 5 ml syringes on polystyrene-S-S-cysteamine resin (50 μmol scale, 0.82 mmol/g loading) using the automated peptide synthesizer. All amino acids, including the synthesized N-terminal building block *3*-(tert-butylsulfanyl)propionic acid, were coupled following standard Fmoc-chemistry as described above. Peptide deprotection, release, cyclization protocol and quenching, and purification were performed as described before. The purity of the linear species were analyzed by analytical HPLC.

4.7 Supplementary information

Supplementary Results

Identification of side-products by mass spectrometry

In order to identify the active species in fractions of hit 6, we performed a second round of HPLC separation for the active fractions (12 and 14), followed by LC-MS for mass determination (Figure S10). As hit 6 showed strong structural similarity with hit 7, we also selected its active fractions (12 and 14) and followed the same protocol. Based on the high-resolution masses, the species were identified as linear side products that contained a tBu group, which, we suspected, was capped by the N-terminal thiol during on-resin peptide deprotection. Such a capping prevented the cyclization and lead to the monoalkylation of the C-terminal thiol, followed by addition of β-mercaptoethanol on the remaining linker reactive site during quenching (Figure S11a).

Supplementary Figures

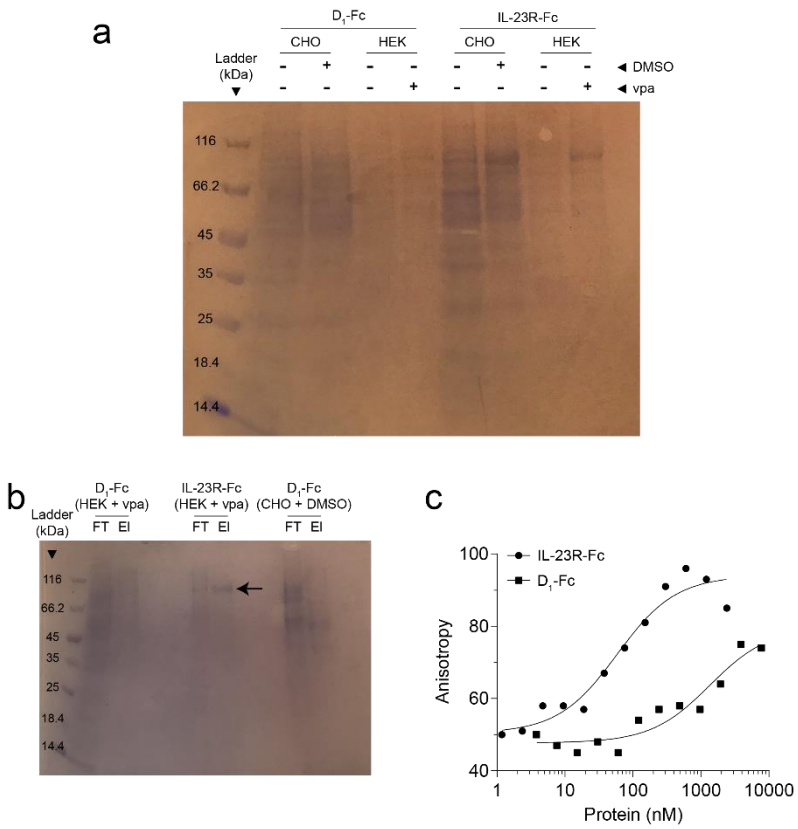


Figure S1. Mammalian transient test expression of D1-Fc or IL-23R-Fc. (a) SDS-PAGE of the expression media (concentrated 6.5-fold) corresponding to all conditions tested for small-scale transient expression of D1-Fc (aa 24 to 125) or IL-23R-Fc (aa 24 to 353) in mammalian cells (in CHO with or without DMSO, and HEK with or without valproic acid). Protein bands can not be directly compared to the ladder due to protein glycosylation affecting the molecular weight and retention time of the protein. (b) SDS-PAGE of flow through (FT) and elution (El) fractions (concentrated 10-fold) after His-column purification of three conditions. The arrow shows the desired protein. (c) Binding of fluorescent reporter peptide 6 (10 nM) against IL-23R-Fc or D1-Fc obtained from purification of small-scale expression in HEK+vpa and CHO+DMSO respectively (see (b)), and measured by fluorescence anisotropy.

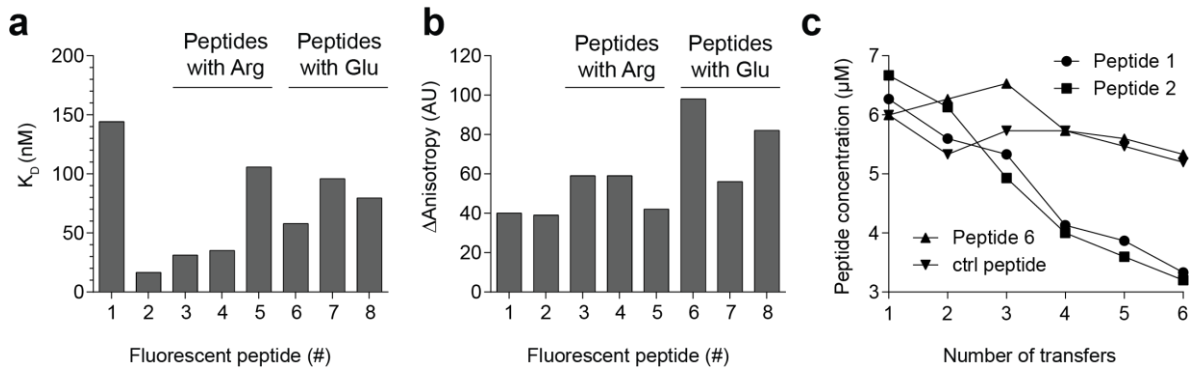


Figure S2. Comparison of IL-23R-binding peptides containing charged amino acids. (a) Dissociation constants of all fluorescent probes tested against IL-23R-Fc. All sequences are in Figure 1c. (b) Assay window for all fluorescent probes tested. (c) Testing the sticking of peptides 1, 2, 6 and ctrl to plastic. The peptides were transferred several times to new tubes, each time using new pipetting tips. Loss of peptide was monitored by nanodrop through a decrease in observed concentration. Ctrl peptide is a fluorescent peptide used in another fluorescence polarization assay.

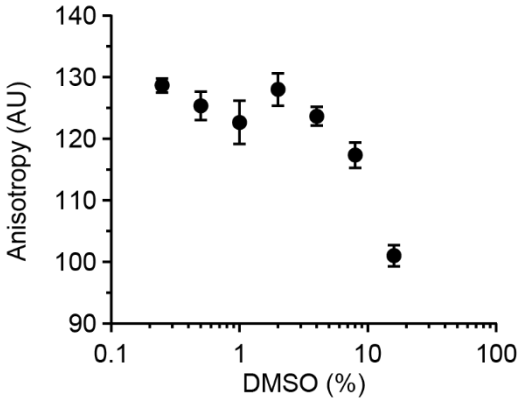


Figure S3. Effect of different DMSO concentrations on the binding of the fluorescent peptide. Peptide 6 (10 nM) was incubated with IL-23R-Fc (60 nM; leading to 75% bound peptide) and increasing amounts of DMSO were added. Means and standard deviation of three measurements are shown.

P#	Sequence	P#	Sequence	P#	Sequence	P#	Sequence	P#	Sequence
P1	Mpa W M Y Mea	P78	Mpa W S2 Y S1 Mea	P155	Mpa W W Y B4 Mea	P232	Mpa W S4 T Y S4 Mea	P309	Mpa Q W S T Y Mea
P2	Mpa W B5 Y Mea	P79	Mpa W L Y P Mea	P156	Mpa B10 W Y S2 Mea	P233	Mpa W B3 B12 Y S Mea	P310	Mpa B1 W B11 B5 Y Mea
P3	Mpa W B2 Y Mea	P80	Mpa W P Y Q Mea	P157	Mpa P W Y B9 Mea	P234	Mpa W B4 N Y B6 Mea	P311	Mpa B13 W S2 B6 Y Mea
P4	Mpa W B9 Y Mea	P81	Mpa W B8 Y S1 Mea	P158	Mpa B6 W Y V Mea	P235	Mpa W E B13 Y S Mea	P312	Mpa B11 W B1 Q Y Mea
P5	Mpa W T Y Mea	P82	Mpa W E Y B7 Mea	P159	Mpa B9 W Y G Mea	P236	Mpa W B9 B12 Y M Mea	P313	Mpa B6 W M Y E Mea
P6	Mpa W S Y Mea	P83	Mpa W B9 Y B11 Mea	P160	Mpa P W Y N Mea	P237	Mpa W B13 D Y B2 Mea	P314	Mpa E W B1 Y Y Mea
P7	Mpa W B4 Y Mea	P84	Mpa W B4 Y W Mea	P161	Mpa B11 W Y T Mea	P238	Mpa W S B2 Y D Mea	P315	Mpa S1 W P Y S2 Mea
P8	Mpa W V Y Mea	P85	Mpa W L Y E Mea	P162	Mpa B3 W Y S Mea	P239	Mpa W S3 B11 Y S1 Mea	P316	Mpa T W S Y B11 Mea
P9	Mpa W N Y Mea	P86	Mpa W B3 Y B3 Mea	P163	Mpa B3 W Y B5 Mea	P240	Mpa W Y B7 Y L Mea	P317	Mpa S W B3 Y B2 Mea
P10	Mpa W E Y Mea	P87	Mpa W S1 Y Y Mea	P164	Mpa S3 W Y M Mea	P241	Mpa W B12 Y B11 B8 Mea	P318	Mpa B3 W B8 Y S2 Mea
P11	Mpa W D Y Mea	P88	Mpa W E Y B1 Mea	P165	Mpa L W Y L Mea	P242	Mpa W S2 Y M B6 Mea	P319	Mpa L W S Y W Mea
P12	Mpa W Y Y Mea	P89	Mpa W B1 Y B13 Mea	P166	Mpa B7 W Y B13 Mea	P243	Mpa W Q Y S B7 Mea	P320	Mpa B2 W B10 Y B6 Mea
P13	Mpa W Y T Mea	P90	Mpa W N Y B4 Mea	P167	Mpa S W Y L Mea	P244	Mpa W D Y B1 N Mea	P321	Mpa B10 W B7 Y B8 Mea
P14	Mpa W Y B2 Mea	P91	Mpa W V Y B13 Mea	P168	Mpa V W Y B1 Mea	P245	Mpa W B11 Y L N Mea	P322	Mpa S2 W L Y B4 Mea
P15	Mpa W Y S2 Mea	P92	Mpa W G Y M Mea	P169	Mpa W B2 B4 B1 Y Mea	P246	Mpa W P Y S2 B8 Mea	P323	Mpa B1 W B12 Y S4 Mea
P16	Mpa W Y M Mea	P93	Mpa W T Y T Mea	P170	Mpa W B7 M L Y Mea	P247	Mpa W B6 Y B1 S2 Mea	P324	Mpa S2 W L Y W Mea
P17	Mpa W Y B11 Mea	P94	Mpa W W Y S2 Mea	P171	Mpa W L B9 B11 Y Mea	P248	Mpa W B1 Y B4 Y Mea	P325	Mpa B12 W B4 Y B3 Mea
P18	Mpa W Y Q Mea	P95	Mpa W B13 Y S4 Mea	P172	Mpa W B11 B11 S4 Y Mea	P249	Mpa W V Y S3 B2 Mea	P326	Mpa D W B4 Y B5 Mea
P19	Mpa W Y W Mea	P96	Mpa W Q Y G Mea	P173	Mpa W M B6 D Y Mea	P250	Mpa W S Y B7 V Mea	P327	Mpa B4 W G Y T Mea
P20	Mpa W Y B4 Mea	P97	Mpa B5 W B3 Y Mea	P174	Mpa W W G S3 Y Mea	P251	Mpa W S1 Y S1 B9 Mea	P328	Mpa Y W B11 Y P Mea
P21	Mpa W Y B6 Mea	P98	Mpa D W B3 Y Mea	P175	Mpa W B10 B3 B13 Y Mea	P252	Mpa W B11 Y W E Mea	P329	Mpa B9 W W Y B12 Mea
P22	Mpa W Y D Mea	P99	Mpa B10 W D Y Mea	P176	Mpa W S Q B6 Y Mea	P253	Mpa W B9 Y D S Mea	P330	Mpa V W Q Y B2 Mea
P23	Mpa W Y B9 Mea	P100	Mpa L W E Y Mea	P177	Mpa W G B4 T Y Mea	P254	Mpa W B3 Y D B6 Mea	P331	Mpa Q W S3 Y B1 Mea
P24	Mpa W Y P Mea	P101	Mpa E W S4 Y Mea	P178	Mpa W B8 P B12 Y Mea	P255	Mpa W B3 Y V T Mea	P332	Mpa S3 W S1 Y B10 Mea
P25	Mpa W P S3 Y Mea	P102	Mpa B13 W N Y Mea	P179	Mpa W B7 S1 B6 Y Mea	P256	Mpa W B4 Y B10 T Mea	P333	Mpa N W S2 Y B13 Mea
P26	Mpa W B10 B5 Y Mea	P103	Mpa V W V Y Mea	P180	Mpa W B12 B7 B1 Y Mea	P257	Mpa W B5 Y B6 Q Mea	P334	Mpa B7 W S2 Y S1 Mea
P27	Mpa W W S4 Y Mea	P104	Mpa B7 W T Y Mea	P181	Mpa W Y N B9 Y Mea	P258	Mpa W B10 Y B9 B9 Mea	P335	Mpa S1 W N Y B10 Mea
P28	Mpa W L S1 Y Mea	P105	Mpa D W S4 Y Mea	P182	Mpa W B5 B1 V Y Mea	P259	Mpa W L Y W P Mea	P336	Mpa S3 W B2 Y T Mea
P29	Mpa W G V Y Mea	P106	Mpa B2 W S1 Y Mea	P183	Mpa W Y B2 B5 Y Mea	P260	Mpa W S3 Y B3 S3 Mea	P337	Mpa B9 W M Y Y Mea
P30	Mpa W W V Y Mea	P107	Mpa B11 W B7 Y Mea	P184	Mpa W E L S2 Y Mea	P261	Mpa W S3 Y P B5 Mea	P338	Mpa B13 W B2 Y D Mea
P31	Mpa W B9 B9 Y Mea	P108	Mpa S3 W S2 Y Mea	P185	Mpa W B1 V S4 Y Mea	P262	Mpa W S4 Y S3 S2 Mea	P339	Mpa B6 W S1 Y G Mea
P32	Mpa W S3 B10 Y Mea	P109	Mpa B12 W M Y Mea	P186	Mpa W G B11 G Y Mea	P263	Mpa W W Y S1 B12 Mea	P340	Mpa G W B6 Y M Mea
P33	Mpa W B7 B1 Y Mea	P110	Mpa Y W B6 Y Mea	P187	Mpa W E D B7 Y Mea	P264	Mpa W S1 Y N D Mea	P341	Mpa B5 W B5 Y S4 Mea
P34	Mpa W B8 B1 Y Mea	P111	Mpa B3 W S2 Y Mea	P188	Mpa W S S4 B13 Y Mea	P265	Mpa W P Y S4 B5 Mea	P342	Mpa B8 W B9 Y V Mea
P35	Mpa W B13 N Y Mea	P112	Mpa Y W B4 Y Mea	P189	Mpa W B5 E S1 Y Mea	P266	Mpa W B5 Y P Q Mea	P343	Mpa S4 W G Y S3 Mea
P36	Mpa W M B7 Y Mea	P113	Mpa S W G Y Mea	P190	Mpa W N S T Y Mea	P267	Mpa W G Y B8 B10 Mea	P344	Mpa P W W Y D Mea
P37	Mpa W Y B11 Y Mea	P114	Mpa P W Q Y Mea	P191	Mpa W V S2 B4 Y Mea	P268	Mpa W S4 Y S4 V Mea	P345	Mpa W W B8 Y G Mea
P38	Mpa W B12 L Y Mea	P115	Mpa M W B13 Y Mea	P192	Mpa W S4 B5 B12 Y Mea	P269	Mpa W B13 Y S L Mea	P346	Mpa B11 W T Y B9 Mea
P39	Mpa W B10 W Y Mea	P116	Mpa M W T Y Mea	P193	Mpa W B6 P B3 Y Mea	P270	Mpa W W Y B11 B11 Mea	P347	Mpa M W B3 Y V Mea
P40	Mpa W N B8 Y Mea	P117	Mpa B10 W B10 Y Mea	P194	Mpa W S2 L B2 Y Mea	P271	Mpa W B7 Y E B13 Mea	P348	Mpa B8 W B11 Y Q Mea
P41	Mpa W T N Y Mea	P118	Mpa B3 W B11 Y Mea	P195	Mpa W D B2 V Y Mea	P272	Mpa W B12 Y Y B11 Mea	P349	Mpa M W Y B6 B10 Mea
P42	Mpa W Q B2 Y Mea	P119	Mpa W W B1 Y Mea	P196	Mpa W B6 B8 D Y Mea	P273	Mpa W B6 Y B7 B12 Mea	P350	Mpa B1 W Y B5 S Mea
P43	Mpa W B4 P Y Mea	P120	Mpa P W P Y Mea	P197	Mpa W Q S2 Q Y Mea	P274	Mpa W B7 Y B12 G Mea	P351	Mpa G W Y P B8 Mea
P44	Mpa W N Y Y Mea	P121	Mpa B2 W Y Y Mea	P198	Mpa W B2 E M Y Mea	P275	Mpa W N Y Y M Mea	P352	Mpa W W Y S1 S1 Mea
P45	Mpa W B5 E Y Mea	P122	Mpa Q W S3 Y Mea	P199	Mpa W N M B2 Y Mea	P276	Mpa W M Y B13 B4 Mea	P353	Mpa B13 W Y T W Mea
P46	Mpa W S2 L Y Mea	P123	Mpa W W B12 Y Mea	P200	Mpa W S1 V Y Y Mea	P277	Mpa V W S3 D Y Mea	P354	Mpa S4 W Y B4 B1 Mea
P47	Mpa W B11 B8 Y Mea	P124	Mpa N W B6 Y Mea	P201	Mpa W S1 W M Y Mea	P278	Mpa B4 W B9 B4 Y Mea	P355	Mpa M W Y N B8 Mea
P48	Mpa W B12 S2 Y Mea	P125	Mpa G W D Y Mea	P202	Mpa W B12 S4 B3 Y Mea	P279	Mpa P W B8 T Y Mea	P356	Mpa T W Y P G Mea
P49	Mpa W S4 B3 Y Mea	P126	Mpa B1 W E Y Mea	P203	Mpa W B3 B12 B9 Y Mea	P280	Mpa B6 W G S1 Y Mea	P357	Mpa B4 W Y B6 L Mea
P50	Mpa W B1 B7 Y Mea	P127	Mpa L W B7 Y Mea	P204	Mpa W S2 T P Y Mea	P281	Mpa S2 W B6 S1 Y Mea	P358	Mpa B2 W Y Q B12 Mea
P51	Mpa W S2 P Y Mea	P128	Mpa B12 W L Y Mea	P205	Mpa W B3 B2 Y S1 Mea	P282	Mpa N W B2 W Y Mea	P359	Mpa D W Y V B12 Mea
P52	Mpa W D B2 Y Mea	P129	Mpa B8 W S1 Y Mea	P206	Mpa W W B3 Y Q Mea	P283	Mpa E W Q B9 Y Mea	P360	Mpa B12 W Y B9 D Mea
P53	Mpa W B3 B5 Y Mea	P130	Mpa B8 W B10 Y Mea	P207	Mpa W T S2 Y N Mea	P284	Mpa B8 W S2 B6 Y Mea	P361	Mpa S2 W Y M E Mea
P54	Mpa W M S4 Y Mea	P131	Mpa S2 W B13 Y Mea	P208	Mpa W B7 S Y B7 Mea	P285	Mpa B10 W B7 S3 Y Mea	P362	Mpa G W Y Y M Mea
P55	Mpa W B8 B9 Y Mea	P132	Mpa B11 W G Y Mea	P209	Mpa W B11 B5 Y N Mea	P286	Mpa D W B4 S4 Y Mea	P363	Mpa E W Y G V Mea
P56	Mpa W B2 B6 Y Mea	P133	Mpa D W Y B8 Mea	P210	Mpa W B9 S1 Y S4 Mea	P287	Mpa S4 W E S4 Y Mea	P364	Mpa S1 W Y B5 M Mea
P57	Mpa W B5 G Y Mea	P134	Mpa B5 W Y M Mea	P211	Mpa W L B4 Y W Mea	P288	Mpa S3 W B6 M Y Mea	P365	Mpa S1 W Y V S3 Mea
P58	Mpa W S E Y Mea	P135	Mpa S4 W Y E Mea	P212	Mpa W B5 S3 Y B7 Mea	P289	Mpa B7 W B12 B2 Y Mea	P366	Mpa B6 W Y B4 B11 Mea
P59	Mpa W B1 B12 Y Mea	P136	Mpa B10 W Y B13 Mea	P213	Mpa W B4 S4 Y B13 Mea	P290	Mpa S3 W W S2 Y Mea	P367	Mpa V W Y M L Mea
P60	Mpa W V D Y Mea	P137	Mpa D W Y Y Mea	P214	Mpa W L M Y S2 Mea	P291	Mpa Q W L L Y Mea	P368	Mpa B9 W Y L N Mea
P61	Mpa W B12 Y T Mea	P138	Mpa N W Y Q Mea	P215	Mpa W B2 W Y B4 Mea	P292	Mpa B4 W S4 B2 Y Mea	P369	Mpa B7 W Y S2 P Mea
P62	Mpa W M Y L Mea	P139	Mpa Y W Y B2 Mea	P216	Mpa W B12 S4 Y B11 Mea	P293	Mpa B6 W W B7 Y Mea	P370	Mpa B13 W Y B12 B5 Mea
P63	Mpa W P Y B9 Mea	P140	Mpa T W Y B3 Mea	P217	Mpa W N B1 Y B1 Mea	P294	Mpa B8 W T G Y Mea	P371	Mpa B5 W Y S B1 Mea
P64	Mpa W Q Y S3 Mea	P141	Mpa E W Y Q Mea	P218	Mpa W S1 B4 Y V Mea	P295	Mpa B9 W B10 L Y Mea	P372	Mpa B1 W Y L S3 Mea
P65	Mpa W B7 Y Y Mea	P142	Mpa S2 W Y S4 Mea	P219	Mpa W B8 G Y B3 Mea	P296	Mpa P W D E Y Mea	P373	Mpa B12 W Y S3 B2 Mea
P66	Mpa W B3 Y B5 Mea	P143	Mpa B2 W Y S2 Mea	P220	Mpa W B11 B11 Y M Mea	P297	Mpa W W B13 B4 Y Mea	P374	Mpa T W Y B7 B7 Mea
P67	Mpa W B13 Y Q Mea	P144	Mpa Q W Y S3 Mea	P221	Mpa W V B10 Y W Mea	P298	Mpa B10 W B12 B11 Y Mea	P375	Mpa Y W Y B1 G Mea
P68	Mpa W W Y B12 Mea	P145	Mpa S1 W Y B12 Mea	P222	Mpa W B6 V Y P Mea	P299	Mpa Y W S4 E Y Mea	P376	Mpa S2 W Y W E Mea
P69	Mpa W B5 Y B6 Mea	P146	Mpa B1 W Y B12 Mea	P223	Mpa W V B8 Y Q Mea	P300	Mpa B9 W M B12 Y Mea	P377	Mpa B10 W Y S4 B9 Mea
P70	Mpa W B4 Y B12 Mea	P147	Mpa Y W Y S1 Mea	P224	Mpa W E B5 Y Y Mea	P301	Mpa G W V V Y Mea	P378	Mpa B3 W Y B13 B7 Mea
P71	Mpa W S Y V Mea	P148	Mpa S W Y B10 Mea	P225	Mpa W Q S1 Y G Mea	P302	Mpa S4 W B8 V Y Mea	P379	Mpa B6 W Y G B2 Mea
P72	Mpa W S3 Y G Mea	P149	Mpa M W Y W Mea	P226	Mpa W Q B6 Y B8 Mea	P303	Mpa B2 W B3 S Y Mea	P380	Mpa B9 W Y W S4 Mea
P73	Mpa W B7 Y B8 Mea	P150	Mpa B6 W Y P Mea	P227	Mpa W S1 L Y S2 Mea	P304	Mpa B5 W T P Y Mea	P381	Mpa D W Y S4 B6 Mea
P74	Mpa W B9 Y N Mea	P151	Mpa B11 W Y Y Mea	P228	Mpa W B7 S2 Y B9 Mea	P305	Mpa S1 W B13 N Y Mea	P382	Mpa W W Y B11 V Mea
P75	Mpa W D Y B11 Mea	P152	Mpa B4 W Y B5 Mea	P229	Mpa W G B13 Y E Mea	P306	Mpa G W Y B7 Y Mea	P383	Mpa S W Y D T Mea
P76	Mpa W B2 Y B4 Mea	P153	Mpa M W Y P Mea	P230	Mpa W B2 E Y S3 Mea	P307	Mpa T W P B8 Y Mea	P384	Mpa B4 W Y B3 W Mea
P77	Mpa W S4 Y S3 Mea	P154	Mpa B1 W Y G Mea	P231	Mpa W B13 V Y B10 Mea	P308	Mpa W W P S3 Y Mea		

Figure S4. Peptides for IL-23R-tailored cyclic peptide library. The chemical structures of the building blocks are shown in Figure 5.

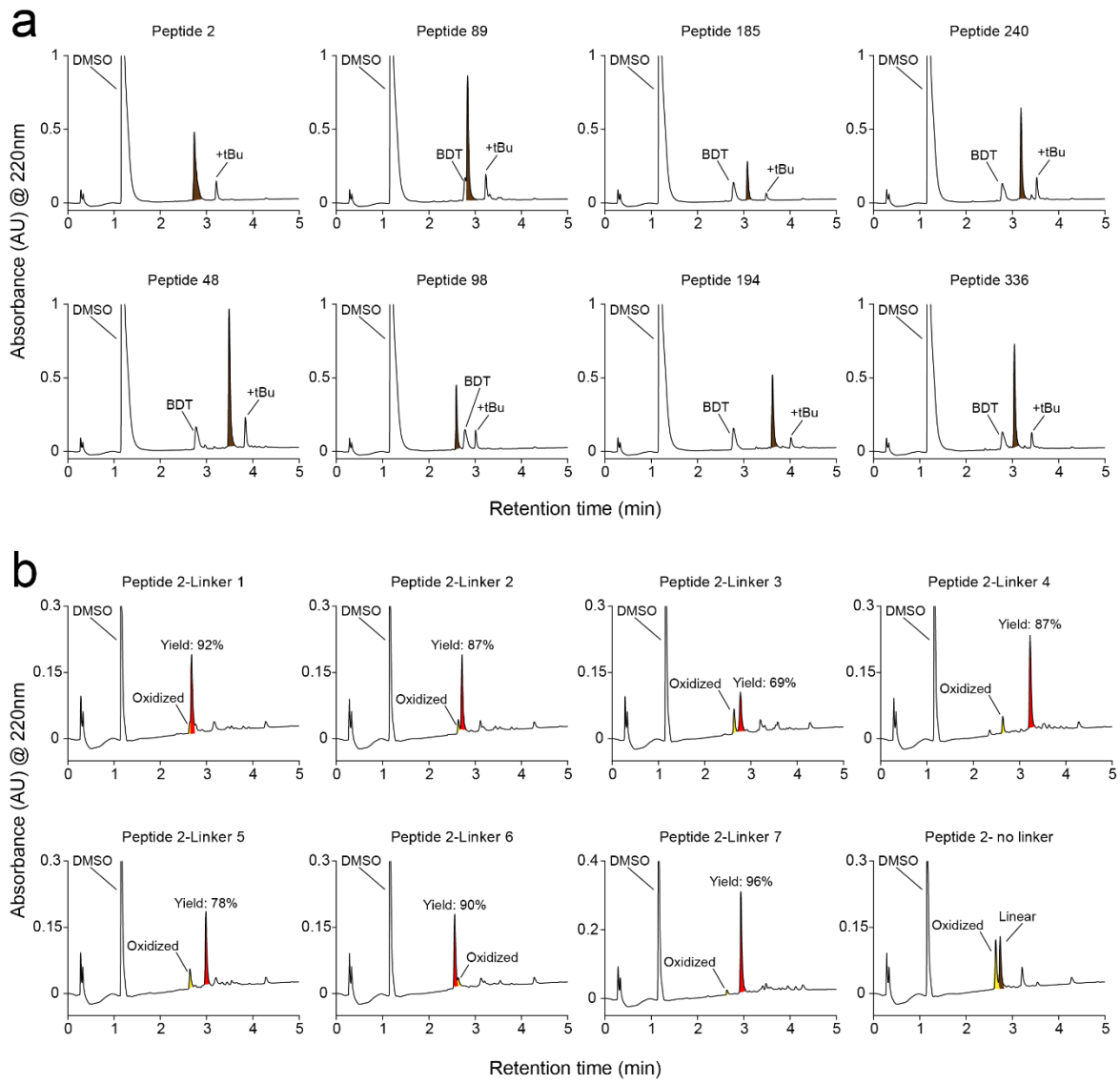


Figure S5. Quality control cyclic peptide library. (a) UV spectra of randomly chosen reduced peptides. Highlighted in brown are the desired species. (b) UV spectra of cyclic peptides obtained after cyclization. Highlighted in red are the desired species. Highlighted in yellow and brown are the oxidized and reduced peptides, respectively. Yields were calculated considering desired specie, oxidized and reduced peptides.

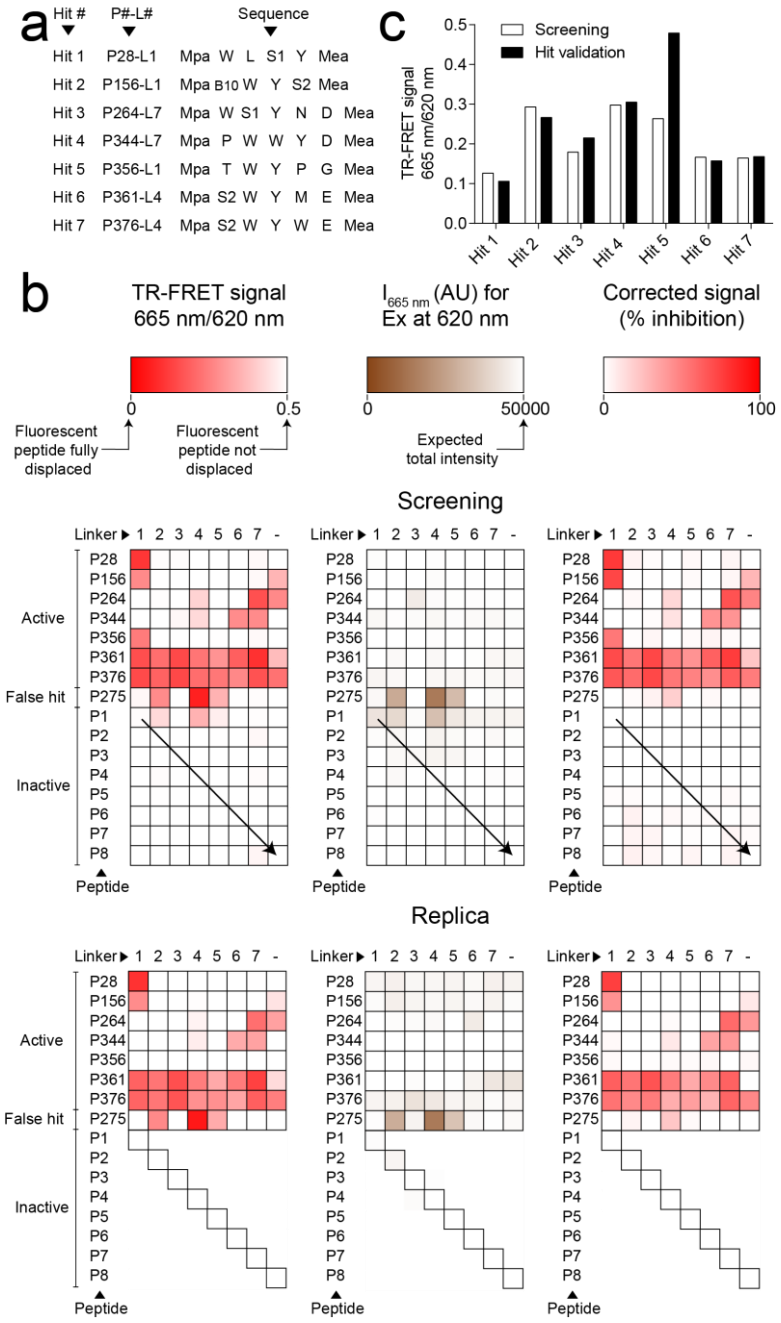


Figure S6. Reproducibility of hits in HTS. (a) Information concerning all the hits considered. (b) Replica of the hits considered. (c) Replica of active, false hit identification and inactive compound. The arrow represents the inactive hits randomly chosen as a replica.

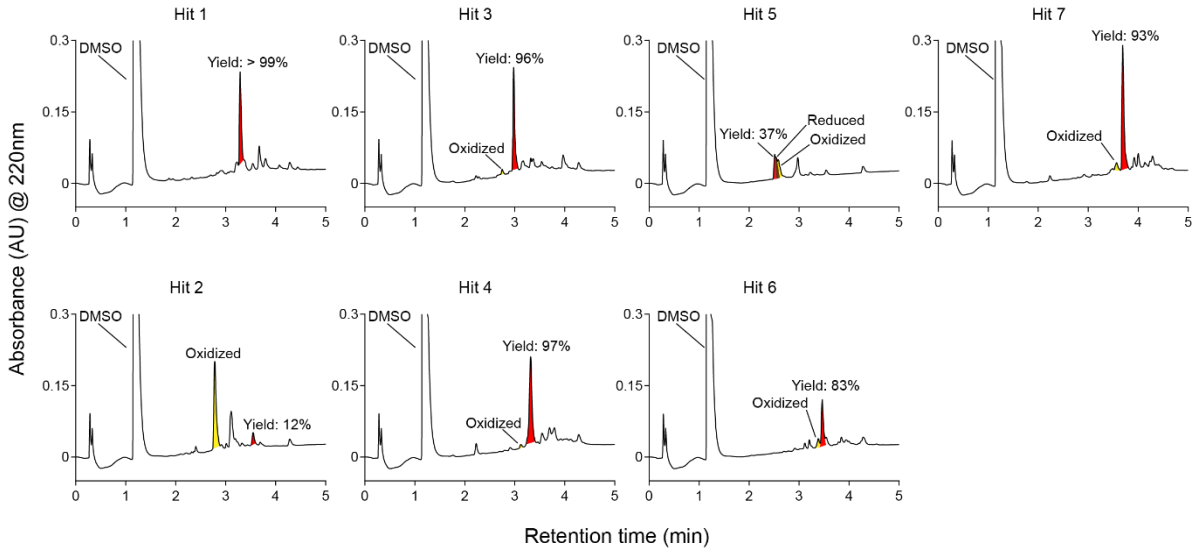


Figure S7. LCMS spectra of hits identified in screen. Highlighted in red are the desired species. Highlighted in yellow and brown are the oxidized and reduced peptides, respectively. Yields were calculated considering desired specie, oxidized and reduced peptides.

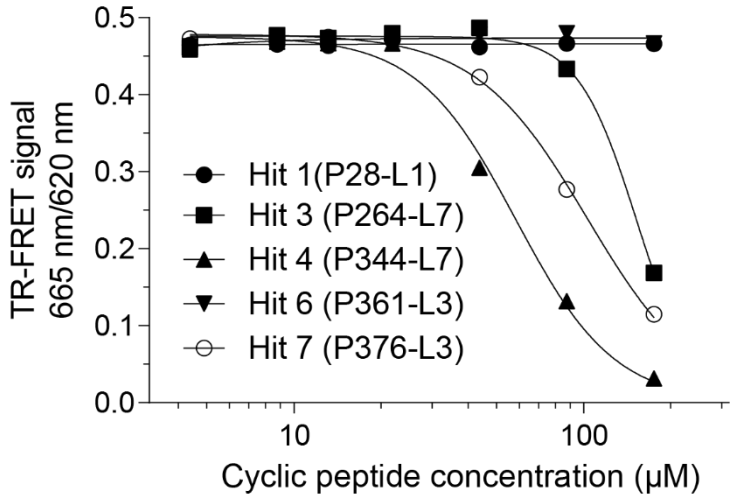


Figure S8. Binding of purified cyclic peptides to IL-23R. Competition assay of pure considered hits against the TR-FRET assay. Eu-AntiHis$_6$ (1 nM), IL-23R (10 nM) and peptide 9 (50 nM) were incubated with the indicated concentrations of the cyclic peptides.

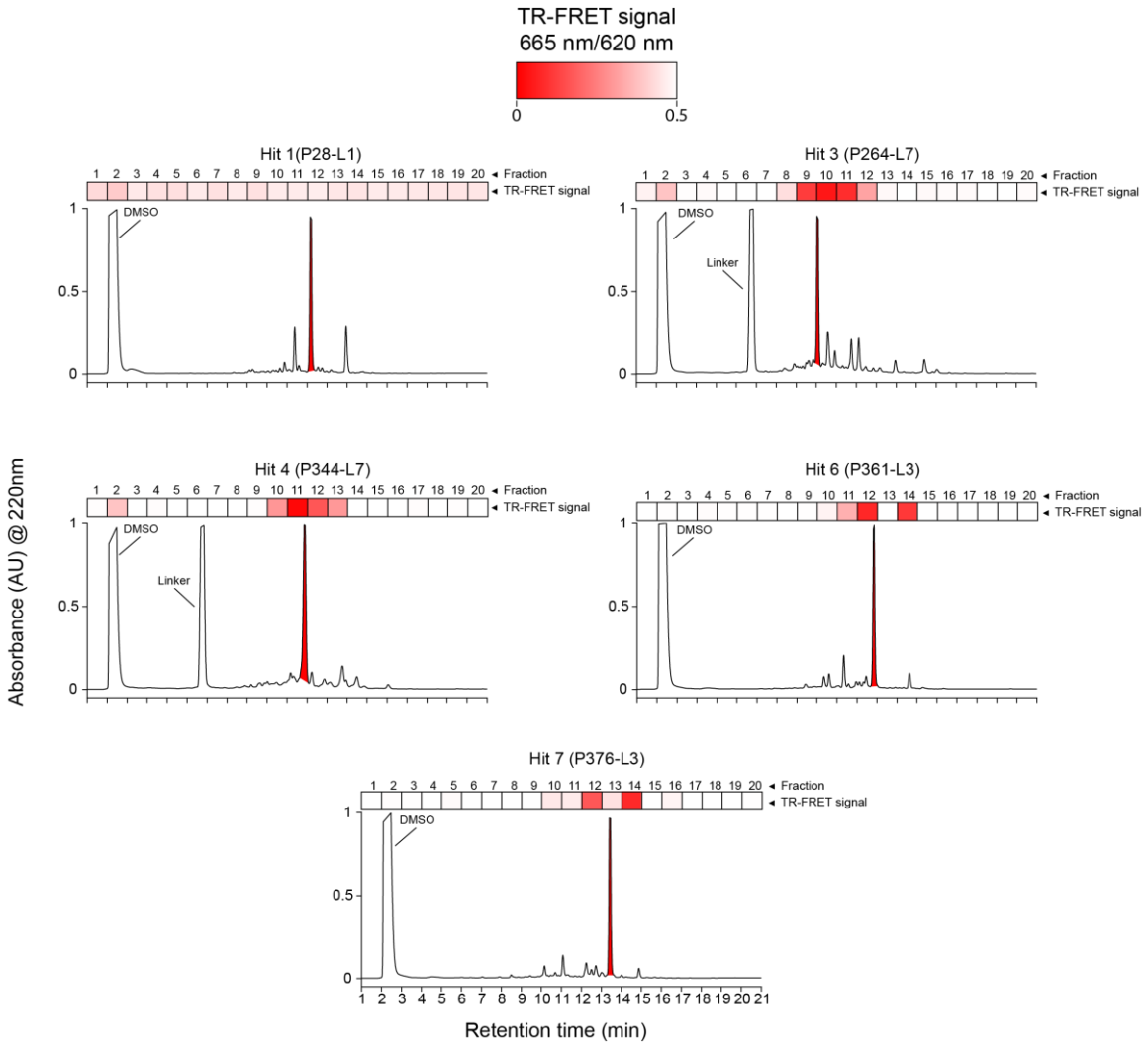


Figure S9. LCMS spectra of hits identified in screen. Competition assay of all species in the crude mixture of all considered hits (separated in 20 different fractions) against the TR-FRET assay. Highlighted in red are the desired cyclic peptides.

4.8 References

(1) Teng, M. W. L.; Bowman, E. P.; McElwee, J. J.; Smyth, M. J.; Casanova, J. L.; Cooper, A. M.; Cua, D. J. IL-12 and IL-23 cytokines: From discovery to targeted therapies for immune-mediated inflammatory diseases. *Nature Medicine*. **2015**.

(2) Oppmann, B.; Lesley, R.; Blom, B.; Timans, J. C.; Xu, Y.; Hunte, B.; Vega, F.; Yu, N.; Wang, J.; Singh, K.; Zonin, F.; Vaisberg, E.; Churakova, T.; Liu, M. ru; Gorman, D.; Wagner, J.; Zurawski, S.; Liu, Y. J.; Abrams, J. S.; Moore, K. W.; Rennick, D.; De Waal-Malefyt, R.; Hannum, C.; Bazan, J. F.; Kastelein, R. A. Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity* **2000**.

(3) Parham, C.; Chirica, M.; Timans, J.; Vaisberg, E.; Travis, M.; Cheung, J.; Pflanz, S.; Zhang, R.; Singh, K. P.; Vega, F.; To, W.; Wagner, J.; O'Farrell, A.-M.; McClanahan, T.; Zurawski, S.; Hannum, C.; Gorman, D.; Rennick, D. M.; Kastelein, R. A.; de Waal Malefyt, R.; Moore, K. W. A Receptor for the Heterodimeric Cytokine IL-23 Is Composed of IL-12Rβ1 and a Novel Cytokine Receptor Subunit, IL-23R. *J. Immunol.* **2002**.

(4) Langrish, C. L.; Chen, Y.; Blumenschein, W. M.; Mattson, J.; Basham, B.; Sedgwick, J. D.; McClanahan, T.; Kastelein, R. A.; Cua, D. J. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J. Exp. Med.* **2005**.

(5) Cua, D. J.; Sherlock, J.; Chen, Y.; Murphy, C. A.; Joyce, B.; Seymour, B.; Lucian, L.; To, W.; Kwan, S.; Churakova, T.; Zurawski, S.; Wiekowski, M.; Lira, S. A.; Gorman, D.; Kastelein, R. A.; Sedgwick, J. D. Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* **2003**.

(6) Lee, E.; Trepicchio, W. L.; Oestreicher, J. L.; Pittman, D.; Wang, F.; Chamian, F.; Dhodapkar, M.; Krueger, J. G. Increased Expression of Interleukin 23 p19 and p40 in Lesional Skin of Patients with Psoriasis Vulgaris. *J. Exp. Med.* **2004**.

(7) Yen, D.; Cheung, J.; Scheerens, H.; Poulet, F.; McClanahan, T.; Mckenzie, B.; Kleinschek, M. A.; Owyang, A.; Mattson, J.; Blumenschein, W.; Murphy, E.; Sathe, M.; Cua, D. J.; Kastelein, R. A.; Rennick, D. IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. *J. Clin. Invest.* **2006**.

(8) Langowski, J. L.; Zhang, X.; Wu, L.; Mattson, J. D.; Chen, T.; Smith, K.; Basham, B.; McClanahan, T.; Kastelein, R. A.; Oft, M. IL-23 promotes tumour incidence and growth. *Nature* **2006**.

(9) Yang, K.; Oak, A. S. W.; Elewski, B. E. Use of IL-23 Inhibitors for the Treatment of Plaque Psoriasis and Psoriatic Arthritis: A Comprehensive Review. *American Journal of Clinical Dermatology*. **2021**.

(10) Almradi, A.; Hanzel, J.; Sedano, R.; Parker, C. E.; Feagan, B. G.; Ma, C.; Jairath, V. Clinical Trials of IL-12/IL-23 Inhibitors in Inflammatory Bowel Disease. *BioDrugs* **2020**.

(11) Sasaki-Iwaoka, H.; Ohori, M.; Imasato, A.; Taguchi, K.; Minoura, K.; Saito, T.; Kushima, K.; Imamura, E.; Kubo, S.; Furukawa, S.; Morokata, T. Generation and characterization of a potent fully human monoclonal antibody against the interleukin-23 receptor. *Eur. J. Pharmacol.* **2018**.

(12) Sasaki-Iwaoka, H.; Taguchi, K.; Okada, Y.; Imamura, E.; Kubo, S.; Furukawa, S.; Morokata, T. AS2762900-00, a potent anti-human IL-23 receptor monoclonal antibody, prevents epidermal hyperplasia in a psoriatic human skin xenograft model. *Eur. J. Pharmacol.* **2019**.

(13) Protagonist Therapeutics to Present at the Jefferies Next Generation IBD Therapeutics Summit - Protagonist Therapeutics, Inc https://www.protagonist-inc.com/investors-media/press-releases/news-details/2021/Protagonist-Therapeutics-to-Present-at-the-Jefferies-Next-Generation-IBD-Therapeutics-Summit/default.aspx (accessed Jan 2, **2022**).

(14) Kong, X. D.; Moriya, J.; Carle, V.; Pojer, F.; Abriata, L. A.; Deyle, K.; Heinis, C. De novo development of proteolytically resistant therapeutic peptides for oral administration. *Nat. Biomed. Eng.* **2020**.

(15) Varghese, T. M.; Dudas, P. L.; Allen, S. J.; Schneeweis, J. E.; Finley, M. F. A. Optimization of a High-Throughput Cell-Based Screening Strategy to Identify Small-Molecule Inhibitors of IL-23 Signaling. *SLAS Discov.* **2021**.

(16) Suen, K. F.; Turner, M. S.; Gao, F.; Liu, B.; Althage, A.; Slavin, A.; Ou, W.; Zuo, E.; Eckart, M.; Ogawa, T.; Yamada, M.; Tuntland, T.; Harris, J. L.; Trauger, J. W. Transient expression of an IL-23R extracellular domain Fc fusion protein in CHO vs. HEK cells results in improved plasma exposure. *Protein Expr. Purif.* **2010**.

5. Screening macrocycle libraries against KRAS(G12D)

5.1 Work contribution

Together with Prof. Christian Heinis, I conceptualized the end-point assay against KRAS(G12D) and the macrocycle library design. I expressed and purified the protein, synthesized and purified the peptides for the assay, and established the assay. I also synthesized the entire macrocycle library and performed the screening and the hit validation. In addition, I made all the figures and wrote the chapter.

5.2 Abstract

KRAS mutations were found to be involved in around 20% of all human cancers, therefore offering promising drug targets in chemotherapy treatments. Of particular interest is KRAS(G12D), which represents the most frequent mutation observed in cancer. Despite more than 30 years of research, there are still no drugs in the market targeting this high hanging fruit, emphasizing the need of novel molecular scaffolds. Macrocycles could potentially fill this gap, as they can target challenging targets while having the potential to be cell permeable and bioavailable. Here, we present the synthesis and screening of 92,160 macrocycles against KRAS(G12D) using a novel assay based on fluorescence polarization. Unfortunately, no active molecules were found which further confirms the difficulty to tackle KRAS(G12D). However, our synthetic method keeps evolving allowing new chemical space to be reached, increasing the chances to get an active molecule and provide potential therapies for numerous incurable cancers.

5.3 Introduction

Rat sarcoma (RAS) protein families are small intracellular membrane-anchored GTPases and were shown to have key roles in cell proliferation signals.[1] They are expressed as KRAS, HRAS, and NRAS isoforms in human cells. The activity of RAS proteins is controlled by an OFF/ON molecular switch upon GDP or GTP bound states, respectively. These two states are mediated by guanine nucleotide exchange factors such as Son of Sevenless homolog 1 (SOS1), which promotes GDP exchange for GTP, leading to RAS activation and interaction with downstream signal molecules. RAS activation stops upon GTP hydrolysis via intrinsic GTPase activity and GTPase-accelerating proteins. RAS is known to act as downstream effectors of cell surface receptors, such as epidermal growth factor (EGF) receptor or fibroblast growth factor (FGF) receptor involved in numerous key cellular pathways driving cellular growth, differentiation, and apoptosis.[2]

RAS genetic alterations were found to be the most prevalent mutation among human cancers (around 30%), making them an attractive chemotherapy drug target.[3–5] From all the three isoforms, KRAS is the most frequently mutated (in around 20% of all cancers).[6] Targeting KRAS could represent numerous important therapeutic options, including for aggressive forms of cancer such as pancreatic ductal adenocarcinomas (PDAC), where it was found to be mutated at around 90%.[7] The most common alterations of KRAS are point mutations at Gly^{12} (in 80% of KRAS mutant malignancies), of which 41% are G12D, 28% G12V, and 14% G12C.[8] Therefore, KRAS(G12D) represents one of the most important chemotherapy drug targets for numerous intractable cancers.

Despite more than 30 years of efforts in targeting the RAS family, there are still no drugs currently on the markets. Due to its lack of reachable pockets for conventional small molecules, the RAS family was often considered undruggable.[9] However, some small molecules still proved to be able to target this class, such as the recently published KRAS(G12D) specific potent binders MRTX1133.[10] Other molecular formats such as peptides,[11–14] engineered protein antagonists,[15] or targeted covalent inhibitors,[16] have also shown promising results, but cellular permeability and bioavailability is often a challenge.

Macrocycles are a class of compounds known to be able to target flat and featureless interfaces.[17] Of particular interest are sub-KDa macrocycles with a limited polar surface area, which present potential membrane permeability and bioavailability.[18] In previous work, we have developed a synthetic methodology that allowed us to generate and screen crude reactions of large libraries of sub-KDa macrocycles. It involves the synthesis without purification of numerous disulfide-bridge macrocycles containing a free amine followed by their

peripheral diversification using carboxylic acids at the picomole scale (Figure S3 & S7a, Habeshian *et al.,* not published at this time).

Herein, we discuss the synthesis of 1,536 disulfide-based macrocycles followed by their diversification with 60 acids yielding 92,160 sub-kDa macrocycles. This macrocycle library was subsequently screened against KRAS(G12D) using a novel assay based on fluorescence polarization to find macrocyclic binders. As the disulfide-bond in the macrocycle is sensitive to reductive conditions, which would likely affect the macrocycle integrity *in vivo*, we envisioned replacing it during the hit optimization of any discovered binders.

5.4 Results and discussion

5.4.1 Screening assay based on fluorescence polarization

We considered developing an end-point assay based on fluorescence polarization. A 19-mer cyclic KRAS(G12D)-selective peptide inhibitor (KRpep-2d) formed via oxidation of two cysteine residues (sequence: Ac-RRRRCPLYISYDPVCRRRR-NH_2) was chosen as a potential probe.[13] It was shown to act as an allosteric modulator of KRAS(G12D), blocking its activation by the protein-protein interaction with SOS1, and exhibited selective inhibition against cancerous cell lines.[13] We envisioned that if a sub-kDa macrocyclic inhibitor could displace this probe, it would conserve a similar allosteric mechanism while having more drug-like properties. Based on a co-crystal structure involving KRpep-2d and KRAS(G12D), it was found that the cyclic moiety seems to be the key aspect for activity while the linear parts (all the arginines at the C- and N-terminus) were solvent-exposed.[19] Structural requirement investigation, including the deletion of all the arginine, confirmed this observation as important activity was conserved when only the cyclic part was involved.[20] Therefore, we decided to synthesize and test two probes involving the common cyclic sequence with fluorescein directly coupled at the N-terminus (probe 1: Fluorescein-RCPLYISYDPVC-NH_2) or C-terminus through the side chain of a first lysine residue (probe 2: Ac-CPLYISYDPVCR(Fluorescein)K-NH_2), both cyclized through oxidation of the two cysteine residues. One arginine was kept to facilitate ether precipitation during peptide synthesis as well as the resulting cyclic peptide solubility.

Both probes (final concentration: 10 nM) were tested by direct fluorescent polarization assay against KRAS(G12D) and showed binding affinity of 88 ± 12 nM and 150 ± 20 nM for probe 1 and 2, respectively (Figure 1a & Figure S2). Both binding affinities were in the expected range, validating the probe design (K_D = 51 nM by SPR for KRpep-2: Ac-RRCPLYISYDPVCRR-NH_2).[13] Probe 1 was selected for the assay development as it provided a larger assay window and had a better binding affinity (Figure S2).

Optimization of the assay was done by testing the DMSO tolerance, followed by the selectivity assessment, and finally, the reliability of the assay. All experiments were performed with probe 1 (10 nM final) and KRAS(G12D) at a concentration leading to 70% bound reporter probe (250 nM protein final). (Figure 1a). First, DMSO was titrated from 13.3% (v/v) to 0.01% (v/v) final concentration, against a mixture containing probe 1/KRAS(G12D). Some fluctuations were observed mainly after 2% (v/v) DMSO final, considered a maximum limit (Figure 1b). Then, the selectivity of the binding interaction was assessed by titrating a control peptide, which was the same peptide as the probe but acetylated, against the same probe 1/ KRAS(G12D) mixture

(Figure 1c). Complete displacement of the reporter probe was observed in a dose-dependent manner, validating the interaction specificity. Finally, the reliability of the assay was tested by performing 32 replicates of both negative (no inhibitor) and positive controls (no protein) (Figure 1d). A good reproducibility with little variations of both controls was observed. A Z' factor of 0.8 was obtained, validating the suitability of the assay for HTS.

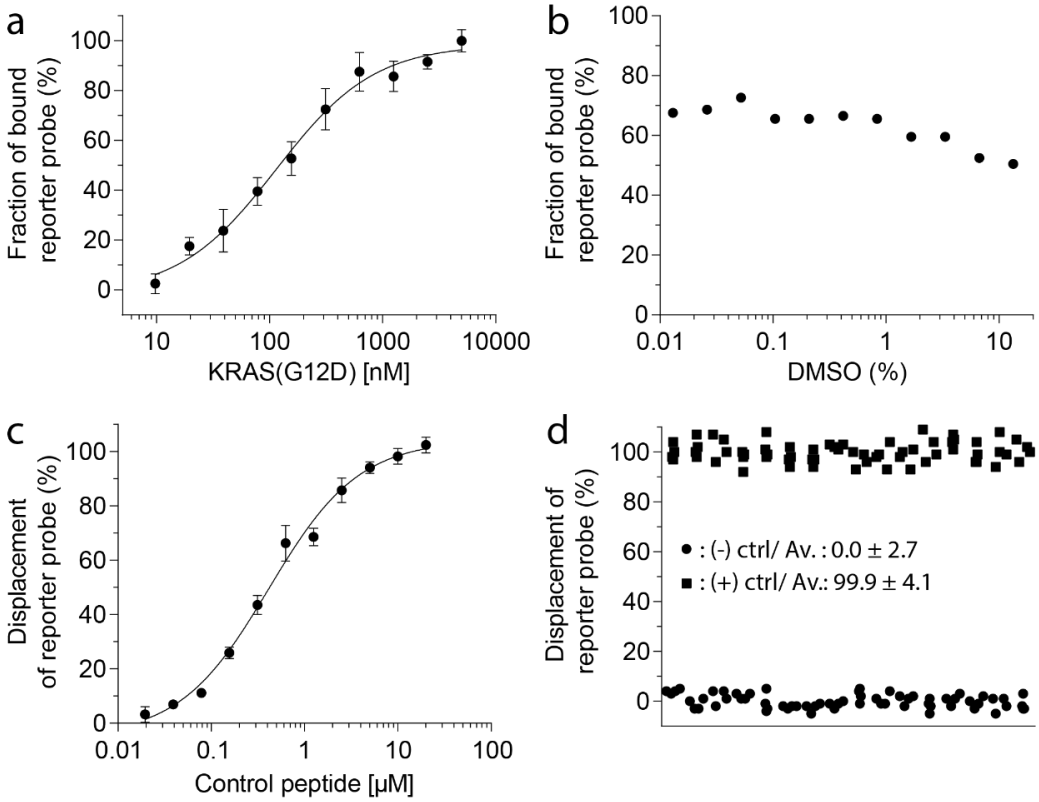


Figure 1. Fluorescence polarization assay establishment. (a) Binding of probe 1 to KRAS(G12D) measured by fluorescence anisotropy and normalized considering the minimum and the maximum as 0 and 100% fraction bound, respectively. Means and standard deviations of three measurements are shown. (b) Effect of different DMSO concentrations on the binding of the fluorescent reporter peptide (10 nM) to KRAS(G12D) at a concentration leading to 70% bound reporter probe (250 nM). (c) Fluorescence polarization competition assay. The displacement of reporter peptide (10 nM) from KRAS(G12D) (250 nM) at increasing concentrations of control peptide was quantified by measuring fluorescence anisotropy. Raw data were normalized considering negative control (no inhibitor) and positive control (no protein) as 0 and 100%, respectively. Means and standard deviations of three measurements are shown. (d) Reliability of the assay. Each control was reproduced 32 times. (-) ctrl (0%) represents the assay in the absence of inhibitor, and (+) ctrl (100 %) represents the assay without protein.

5.4.2 Macrocycle library design, synthesis and quality control

As the long-term goal of this project was to generate potential bioavailable KRAS(G12D) macrocyclic inhibitors, we decided to synthesize sub-kDa macrocycles containing no charges. In addition, we limited the number of amino acids with the chosen systematic sequence Mpa-X_1-X_2-X_3-cys (X: any amino acid) cyclized through the thiol oxidation (Figure 2a). One of these

amino acid positions was set to be a diamino acid required for the peripheral diversification strategy and selected among eight structurally varied options (**1**-**8**; brown in Figure 2a & 2c). Each diamino acid was equally represented (48 times per library). Then, to increase our chances of finding a potential hit, we decided to include one building block with properties known to be important in the cyclic peptide binding. KRpep-2d alanine scanning revealed that the key amino acids involved in the interaction against KRAS(G12D) are Leu^{7}, Ile^{9}, and Asp^{12}.[20] In addition, the synthesis of KRpep-2d derivatives showed that numerous non-canonical aliphatic amino acids could be used to replace Leu^{7} and Ile^{9} leading to substantial improvement in binding.[14] Therefore, in the attempt to mimic some of these key interactions, we decided to include at least one aliphatic side chain in each macrocycle, which was randomly selected among ten different one (**A**-**J**; purple in Figure 2a & 2c). Finally, the last amino acid was selected to cover a large set of chemical properties. Therefore, it was chosen to either vary the macrocycle backbone (**a**-**k**; black, Library 1), to be polar (**l-v**; yellow, Library 2), to be aromatic (**w-ah**; blue, Library 3), or finally aliphatic (**ai-as**; green Library 4) (Figure 2d). This last building block was randomly chosen for each macrocycle from its corresponding list of amino acids. To maximize the macrocycle chemical diversity, all the permutations of these three amino acids were covered equally in each library (64 times per library, Figure 2b), each library contained 384 macrocycles yielding 1,536 macrocycles (see all macrocycle sequences in Figure S4).

The overall quality of the synthesized macrocycle libraries was assessed directly after oxidative release by randomly selecting eight different macrocycles per library and analyzing them by LCMS (Figure S5 & Table S1). For libraries 1, 2, 3, and 4, purity averages of their respective eight samples were 80%, 79%, 81%, and 90% (purity average of 83% considering all the 32 samples analyzed LCMS). The main undesired species were dimers, and tBu-capped species, which we suspected were capped by the N-terminal thiol during group deprotection. Overall, the quality obtained corresponded to the expected one using this synthetic strategy and was considered sufficient for screening.

The libraries concentrations were then determined by measuring the absorbance of tryptophan-containing macrocycles at 280 nm using a Nanodrop (Figure S6). The average concentration obtained from these samples was considered representative of their library. For the libraries which didn't contain tryptophan as one of their amino acids (all except library 3), eight macrocycles were randomly chosen to have one of their amino acids substituted (sequences in Figure S4). For library 1, 2, 3, and 4, concentrations of 12.8 ± 5.2, 7.7 ± 3.6, 7.2 ± 2.6, and 6.4 ± 2.2 mM were measured respectively. The fluctuations were judged acceptable to pursue directly with the screen. The library concentrations were adjusted with

DMSO to 6.4 mM based on their average measured concentration before combinatorial diversification.

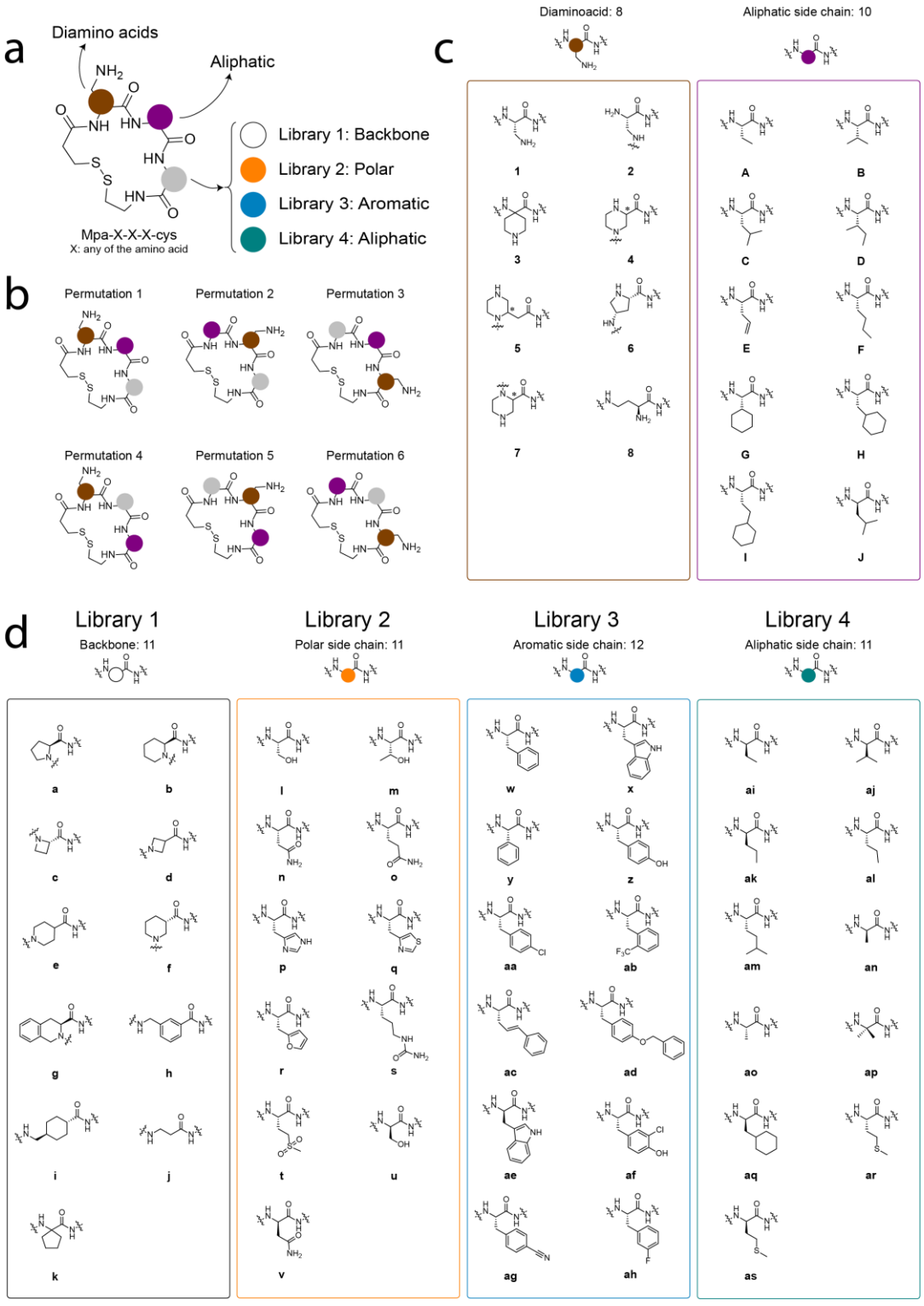


Figure 2. KRAS-oriented library design. (a) Macrocycle building blocks for all macrocycles. Each macrocycle was synthesized with one diamino acid (in brown) and one aliphatic side chain amino acid (in purple). Depending on the library, the last amino acid was chosen from the backbone (in blank), polar (in orange), aromatic (in blue), or aliphatic (in green) side chain class of amino acids (see 2 (d)). Mpa (N-terminus) and cys (C-terminus) mean 3-mercaptopropionic acid and cysteamine, respectively. (b) All permutations of the three building blocks. (c) Diamino acids (in brown) and aliphatic side chain

amino acids (in purple) used for all four libraries. A star (*) represents the stereocenter in the racemic building blocks. (d) Backbone (in blank), polar (in orange), aromatic (in blue), or aliphatic (in green) side-chain amino acids used for library 1, 2, 3, and 4, respectively.

To increase the chemical diversity of the 1,536 synthesized macrocycles, peripheral diversification of the free amine with carboxylic acids at picomole scale was performed following a method previously established in the lab (Figure S7a, Habeshian *et al.*, not published at this date). For this, 60 nL of macrocycles (6.4 mM in DMSO) and 60 nL of activated carboxylic acids (64 mM in DMSO) were transferred by acoustic droplet ejection directly into assay plates. In total, 60 acids were used for combinatorial diversification yielding 23,040 macrocycles per library and 92,160 macrocycles in total. The acids were selected to cover a large scope of chemical diversity containing aromatic (halogenated or not), heterocyclic, polar, or aliphatic elements (Figure S7b). After 5 hours of incubation at room temperature, the reactions were quenched by adding 4.88 µL TRIS buffer. We obtained the final macrocyclic element in 5 µL at 76.8 µM. For acylation control, we selected the same macrocycles picked during the quality control of each library and made them react with acids chosen randomly. The proper acylation of the macrocycles was monitored by pulling together five wells of the identical reaction and submitting it to LCMS for analysis. As the samples were too diluted to properly distinguish UV peaks, only qualitative analysis of the reaction could be performed comparing reacted against unreacted TIC peaks (Figure S8 & Table S2). As TIC is not quantitative due to the difference between ion flying capabilities, only high (> 80%) and low (< 10%) ratios were considered as high and low conversions, respectively. In the majority of the analyzed samples, high conversion was observed, therefore validating the overall library quality.

5.4.3 Screening of macrocycle library and hit validation

We next screened the synthesized macrocycle library for KRAS(G12D) inhibitors by dispensing 10 µL of a solution containing the fluorescent-labeled reporter peptide (Probe 1, 2 nM final) and KRAS(G12D) (250 nM final) to the 384-well synthesis plates containing the 92,160 macrocycles (25 µM final, 1.5% [v/v] DMSO). The extent of reporter peptide displacement by the macrocycles is shown in the heat map in Figure 3. Altogether, the assay proved to be robust as little fluctuation could be observed. Several combinations of acids and macrocycles suggested strong displacement of the reporter peptide, as indicated by the red colors in the heat map. Control reactions including acids alone showed that despite a37, no acids showed any disturbance to the assay. For this reason, all hits including a37 were not considered. Controls involving macrocycle alone showed little activity, suggesting that the

macrocycles without peripheral diversifications were not suitable to displace the reporter peptide.

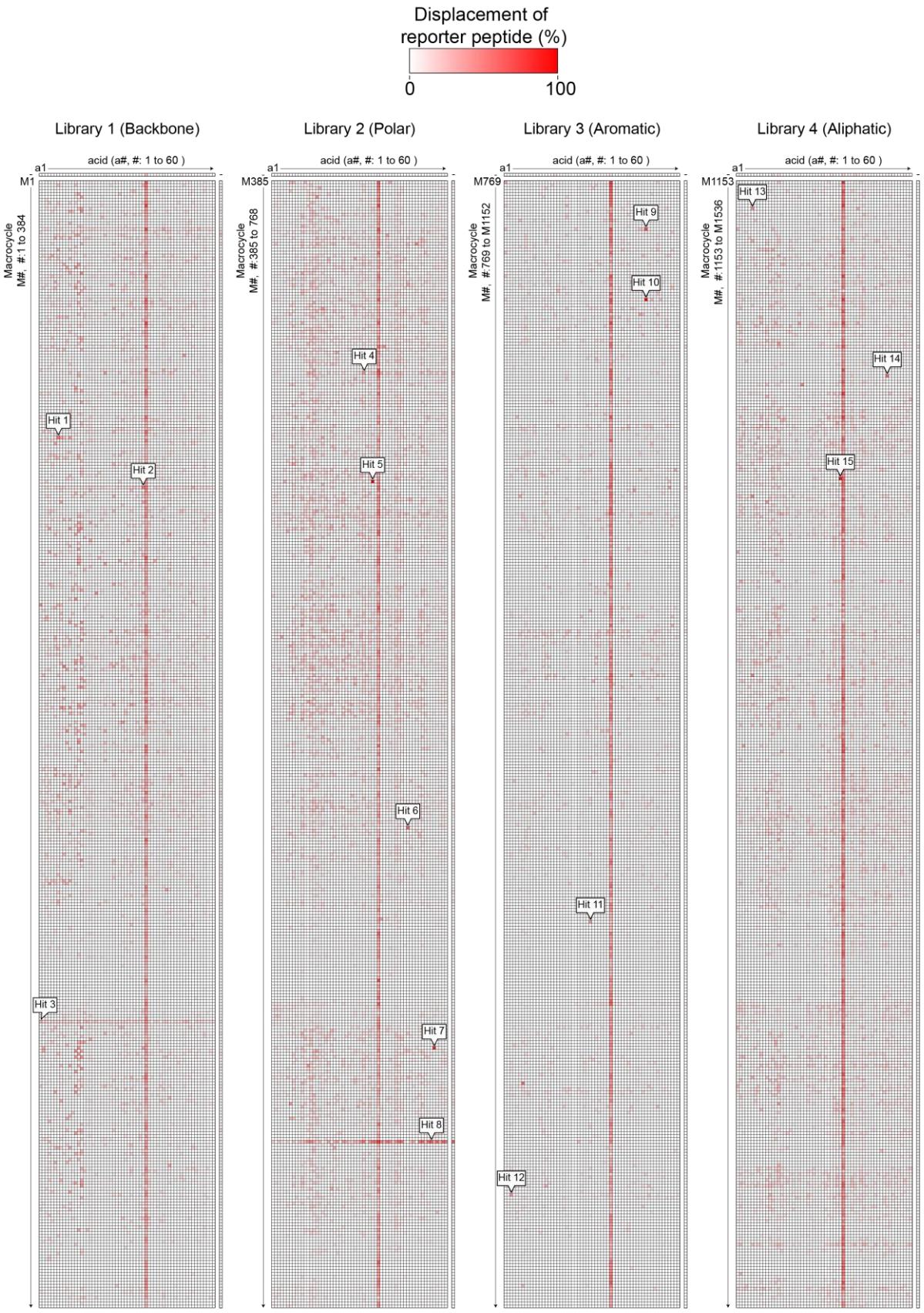


Figure 3. Screen for KRAS(G12D) inhibitor. Binding of 92,160 macrocycles to KRAS(G12D) assessed by displacement of a fluorescein-labeled KRAS(G12D) bound reporter peptide using fluorescence polarization. The raw data were normalized considering the control without inhibitor and the ones without protein as 0 and 100% displacement of reporter peptide. Controls without macrocycles or acids are represented with "-" horizontally and vertically, respectively.

To determine the reproducibility of the activity observed for some macrocycles, we decided to investigate further the most active hits (>50 % displacement of reporter peptide) or some macrocycles showing activity with several acids (Hit 1 to 15, Figure 3 & 4a). Analyzing the total fluorescence intensity observed in the FP assay during the screen can help recognize false hits.[21] Any drastic total intensity variation from the controls would suggest compound interference on the readout. Therefore, we represented and analyzed the total intensity for the entire screen (see heat map in Figure S9). All the hits considered showing extreme total intensities (around five times more than controls) were immediately considered as false positives (Hit 5, 7, 10, and 15; represented in black in the heat map, Figure S9). Worryingly, for most of the other remaining hits considered, some fluctuation of the total intensity could also be perceived.

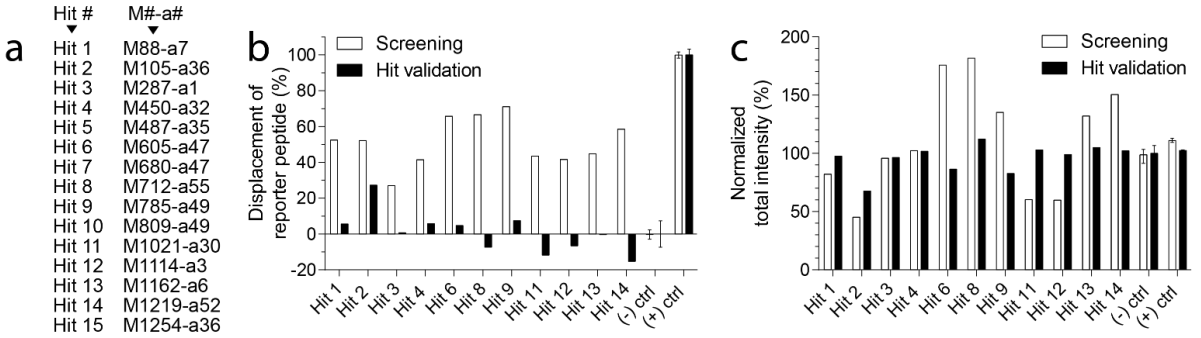


Figure 4. Hit reproducibility assessment. (a) All the hits considered with their corresponding macrocycle-acid combination. (b) Reproducibility of the hits against the developed fluorescence polarization assay. (c) Normalized total intensity observed in the fluorescence polarization assay. The total polarized fluorescence intensity was normalized, considering the average of all positive and negative controls as 100%.

Next, we repeated the hit with the corresponding macrocycles and acids and tested them following the same conditions as in the screen (Figure 4b). Unfortunately, no hits were reproduced. We also compared the total intensities observed during the screen and the hit validation by normalizing them, considering their average positive and negative controls total intensities as 100% (Figure 4c). Interestingly, the total intensities observed in the hit validation were not fluctuating as much as their corresponding values in the screening. This suggested that the fluctuations observed in the screen were most likely not due to compound interference but more likely because of some external factors such as bubbles present during the screen, which could be expected due to the presence of detergent.

5.5 Conclusion

In summary, we have successfully synthesized 92,160 macrocycles and applied them against KRAS(G12D) using a novel fluorescence polarization assay. Unfortunately, no inhibitors were found, further emphasizing the difficulty of tackling this target. The challenges faced, such as the poor hit reproducibility, are characteristic of the HTS complexity. Additionally, this work further validated the macrocycle synthesis strategy followed as high-quality products were obtained at every step. The lack of active molecule against KRAS(G12D) could be explained by improper molecular format or simply because more chemical diversity should be tested. Our macrocycle synthesis strategy keeps evolving, allowing us to expand further the chemical space that can be reached and increase the library size. Therefore, there is still hope in finding a drug-like macrocyclic molecule against KRAS(G12D) that could provide a therapy against a large spectrum of intractable cancers.

5.6 Material and methods

LC-MS analysis

Peptides and macrocycles were analyzed with an ultra HPLC (UHPLC) and single quadrupole MS system (Shimadzu-2020). Samples were dissolved in a mixture of acetonitrile (ACN) and mQ water before being applied on a reversed-phase C18 column (Phenomenex Kinetex®, 2.6 mm, C18, 100 Å column, 50 × 2.1 mm) with a linear gradient of solvent B (99.95% [v/v] ACN and 0.05% [v/v] formic acid) over A (99.95% [v/v] mQ water and 0.05% [v/v] formic acid) starting from 10 to 100% at a constant flow rate of 1 mL/min. Both positive and negative modes were used for mass analysis.

Cloning of KRAS(G12D) expression vector

Expression strategy was based on Gray *et al.* for the expression of KRAS(1-169) but inserting G12D for our application.[22] The following DNA coding for the amino acids 1 to 169 of KRAS(G12D) (in bold) and coding region for two N-terminal His_8-tags separated by a 28 amino acids linker followed by a TEV cleavage site and BamHI restriction site (in italic), was synthesized and inserted in the plasmid pUC-57Amp using suitable DNA restriction site (NcoI and XhoI; underlined), by GENEWIZ (gene-containing plasmid arbitrarily named pUC-57Amp-KRAS for simplicity). Desired protein was to be obtained by cloning the synthesized gene into the bacterial expression vector pET28-b.

5'-
CCATGGGTCACCATCACCATCATCACCATCATGGTGCTACCGGCTCTACCGCTGGCTC
TGGCACCGCTGGTTCTACCGGTGCTTCTGGCGCTTCCACCGGCGGTACCGGCGCTAC
*CCATCATCACCATCATCACCACCATGAAAACCTGTATTTTCAGGGCGGATCC***ATGACGG**
AATACAAACTGGTGGTGGTGGGTGCGGACGGCGTGGGTAAAAGTGCTCTGACGATC
CAACTGATCCAAAATCACTTCGTGGATGAATATGACCCGACCATTGAAGATTCTTACC
GTAAACAGGTGGTTATAGACGGCGAAACCTGCCTGCTGGATATCCTGGACACGGCCG
GTCAGGAAGAATATTCAGCTATGCGTGATCAATACATGCGCACCGGCGAAGGTTTTC
TGTGTGTGTTCGCGATAAACAACACGAAATCGTTCGAAGATATCCATCACTATCGTGA
ACAGATCAAACGCGTTAAAGATAGTGAAGACGTCCCGATGGTCCTGGTGGGCAACAA
ATGCGATCTGCCGTCCCGGACCGTTGATACGAAACAGGCGCAAGACCTGGCCCGCA
GCTACGGTATTCCGTTCATCGAAACCTCTGCAAAAACGCGCCAAGGCGTTGATGATG
CCTTCTACACCCTGGTTCGTGAAATCCGGAAACACAAAGAAAAATAACTCGAG-3'

The ordered plasmid was resuspended in autoclaved mQ water and electroporated (1 μL at 400 ng/mL) into 100 μL electrocompetent DH5$^{\alpha}$ *E.coli* cells using 1 mm electroporation cuvettes (EP-201, Cellprojects) and an electroporation device (MicroPulser, BIO-RAD). Immediately after electroporation, 1 mL of warm LB medium was added to the cuvette and transferred to a sterile 14 mL tube to be incubated for 1 hr at 37°C under agitation (250 rpm). 50 μL of the resulting culture were plated on LB agar plates containing ampicillin (100 μg/mL) and incubated overnight at 37°C. Single colonies obtained were selected to inoculate 10 mL of LB containing 100 μg/mL of ampicillin in 50 mL Falcon tubes. After overnight incubation at 37°C under agitation (250 rpm), cells were pelleted by centrifugation at 10,000 x g (around 11,000 rpm on an Eppendorf Centrifuge 5418) for 5 min at RT. pUC-57Amp-KRAS was isolated using a plasmid miniprep kit (NucleoSpin® Plasmid Miniprep Kit, Qiagen).

The obtained pUC-57Amp-KRAS (9.0 μg) and vector pET28-b (4.3 μg) were digested using *Nco*I (4 μL, 10,000 U/mL) and *Xho*I (4 μL, 10,000 U/mL) in *Xho*I/*Nco*I compatible buffer (50 mM TRIS, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT, pH 7.5) in a final volume of 100 μL. After 2 hrs incubation at 37°C, the digested vector was additionally treated with FastAP phosphatase (2.5 μL, 1 U/mL) for 10 min at 37°C. The cleaved plasmids were then purified by agarose electrophoresis using 1% (w/v) high-quality agarose and TAE buffer with 0.5 μg/mL of ethidium bromide and 1 mM guanosine. Desired KRAS containing gene and cleaved vector were cut out of the gel using a scalpel under UV light (254 nm, 70% intensity). Gel extraction kit (QIAquick® Gel Extraction Kit, Qiagen) was used for purification. Obtained concentrations were obtained by measuring the absorption at 260 nm (NanoDrop 8000, Thermo Scientific).

Ligation of cleaved KRAS containing gene (391 ng) and pET28-b vector (392 ng) was performed using T4 ligase (2 μL, 400,000 U/mL) in T4 ligase buffer in a volume of 22.1 μL for 4 hrs at RT. Completion of the ligation was evaluated by agarose electrophoresis. Purification of the ligation mixture was then performed (NucleoSpin® Plasmid Miniprep Kit, Qiagen).

The ligated vector containing KRAS gene (arbitrarily called pET28-b-KRAS) was electroporated (1 μL at 390 ng/μL) in 100 μL DH5$^{\alpha}$ *E.coli* cells using the same protocol described previously. Plating of the 50 μL containing electroporated cells in LB media was done on Kanamycin containing (50 μg/mL) LB agar plate. After incubation overnight at 37°C, single colonies were selected to inoculate 10 mL of LB containing 50 μg/mL Kanamycin in 50 mL Falcon tubes. After overnight incubation at 37°C under agitation (250 rpm), cells were pelleted by centrifugation at 10,000 x g (around 11,000 rpm on an Eppendorf Centrifuge 5418) for 5 min at 4°C. pET28-b-KRAS was then isolated using a plasmid miniprep kit (NucleoSpin® Plasmid Miniprep Kit, Qiagen) and submitted to sequencing to identify a colony containing the desired gene sequence.

Sequence of KRAS(G12D)

The DNA sequence was confirmed by Sanger sequencing using T7 primer (5'-TAATACGACTCACTATAGGG-3') as shown in Figure S1a. Expression in *E.coli* yielded the KRAS(G12D) protein shown in Figure S1b.

Recombinant expression of human KRAS(G12D)-GDP loaded

The confirmed vector pET28-b-KRAS was inserted in electrocompetent *E. coli* BL21(DE3) cells following the protocol described previously and plated on agar plates containing kanamycin (50 µg/mL). A 10 mL pre-culture containing LB medium and kanamycin (50 µg/mL) was inoculated with cells from a single colony and incubated overnight at 37°C under agitation (250 rpm). A culture flask containing 1 L of LB medium with kanamycin (50 µg/mL) was then inoculated with 10 mL of the pre-culture and incubated at 37°C under agitation (250 rpm). Once optical density at 600 nm (OD_{600}) reached 0.8 (takes around 3 hrs), the culture was cooled on ice for 10 min. Protein induction was done overnight at 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) final concentration and at 18°C under agitation (180 rpm). The culture was then centrifuged at 5,500 x g (around 5,000 rpm using a Thermo Scientific SORVALL RC BIOS centrifuge with a Fiberlite™ F8-6x1000y rotor) for 5 min at 4°C. The supernatant was discarded, and the resulting pellet was transferred to a 50 mL Falcon tube, flash-frozen in liquid nitrogen, and stored at -80°C prior purification.

The pellet was thawed on ice, resuspended in cold lysis buffer containing 50 mM TRIS, 300 mM NaCl ,pH 7.5 supplemented with lysozyme (200 µg/mL final), DNAseI (13 U/mL final) and finally with Phenylmethylsulfonyl fluoride (PMSF) (0.2 mM final, prepared as a stock of 200 mM in Isopropanol). After 30 min incubation on ice, the cells were sonicated (SONICS Vibra cell; Method: Pulse: 5 s on/5 s off, Amp 20%, 5 min) followed by centrifugation at 15,000 x g (around 10,000 rpm with a Beckman Coulter™ Allegra™ 25R using a TA-10-250 rotor) for 1 hr at 4°C. The supernatant was transferred into a new 50 mL Falcon tube and run over two connected 1 mL His-trap columns (HisTrap™ FF crude, GE Healthcare) using a peristaltic pump (BIORAD Econo Pump) at a flow rate of 1 mL/min. After washing with 10 column volumes of binding buffer (50 mM TRIS, 300 mM NaCl and 40 mM Imidazole, pH 7.5), the protein was eluted with 10 column volumes of elution buffer (50 mM TRIS, 300 mM NaCl, and 500 mM Imidazole, pH 7.5) and collected in fractions of 4 mL. SDS-PAGE using a 12% (w/v) RunBlue SDS protein gel and standard SDS running buffer (40 mM tricine, 60 mM TRIS, 0.1% (w/v) SDS, 2.5 mM sodium metabisulfite, pH 8.4) was used to verify protein quality as well as which fraction contained the desired protein. A sample of each fraction (20 µL) was denatured and reduced using 5-fold loading buffer (2% [w/v] SDS, 25% [v/v] glycerol, 0.05% [w/v]

bromophenol blue, 60 mM Tris-HCl pH 6.8 supplemented with pure b-mercaptoethanol to a final concentration of 10 mM). After running the gel for 70 min at 160 mV, the protein in the gel were stained using staining solution (0.1% [w/v] Coomassie R-250 in 50% [v/v] methanol, 10% [v/v] glacial acetic acid and 40% [v/v] mQ water) for 2 hrs. Destaining was performed using destaining solution (50% [v/v] mQ water, 40% [v/v] methanol and 10% [v/v] acetic acid) overnight. The fractions containing the desired protein were combined and the volume reduced to around 2 mL using a 10,000 MWCO Macrostep® Advance Centrifugational Device and centrifugating at 3,400 x g (around 4,000 rpm with a Thermo Scientific Multifuge 3L-R with a Sorval 75006445 rotor) at 4°C. Buffer exchange was performed using the same centrifugation based system by 5 cycles of adding 20 mL of cold 20 mM HEPES, 150 mM NaCl at pH 7.5 on top of the concentrated 2 mL.

Protein concentration was determined to be 3.1 mg/mL based on the absorbance at 280 nm ($MW_{KRAS(G12D)}$ = 24.82 kDa, $\varepsilon_{KRAS(G12D)}$ = 13,540 $M^{-1}.cm^{-1}$, NanoDrop 8000, Thermo Scientific).

Guanosine diphosphate (GDP) was loaded directly in the concentrated protein stock obtained by adding $MgCl_2$ (stock at 1 M in mQ water) and GDP (stock at 0.1 M in mQ water) to final concentration of 5 mM and 800 μM, respectively. After incubation at 4°C overnight, the GDP loaded protein was transferred in aliquots, flash-freezed with liquid nitrogen, and stored at -80°C.

Synthesis of fluorescein-labeled KRAS(G12D)-binding peptide

Two cyclic fluorescent peptides, probe 1 (Fluorescein-R<u>C</u>PLYISYDPV<u>C</u>-NH_2) and probe 2 (Ac-<u>C</u>PLYISYDPV<u>C</u>R(Fluorescein)K-NH_2), were synthesized as follows.

Peptides were synthesized on solid phase using traditional Fmoc-chemistry with Rink amide MBHA resin (50 μmol scale, loading of 0.3 mmol/g) in 5 mL syringe reactors loaded in an automated peptide synthesizer (Intavis, MultiPep RSi). In every washing step, 2 mL of N,N-dimethylformamide (DMF) were used under agitation at 400 rpm for 1 min.

Resin was initially swollen in three cycles of 2.4 mL DMF for 10 min. For each coupling, the amino acids (4.2-fold equiv., 0.22 M final), coupling agent HATU (3.9 equiv., 0.21 M final), the base N-methyl morpholine (NMM) (10 equiv, 0.53 M final), all in DMF, and 5 μL of N-methyl-2-pyrrolidone (NMP) were mixed for 1 min, transferred to the solid phase resin, and shaken for 45 min at 400 rpm. Each amino acid was coupled twice followed by seven cycles of wash with 2 mL DMF. Capping was then performed with 0.8 mL of caping mixture (5% [v/v] Acetic anhydride Ac_2O, 6% [v/v] 2,6-Lutidine in DMF) for 5 min followed by seven cycles of DMF

wash. Fmoc protecting groups were removed in two cycles of 0.8 mL of 20% (v/v) piperidine in DMF for 5 min.

In the last step of the synthesis, 5(6)-carboxyfluorescein (5(6)-FAM) was treated as a traditional amino acid and coupled as described above. For the first fluorescent probe, 5(6)-FAM was coupled directly at the N-terminus of the peptide. For the second fluorescent probe, prior to 5(6)-FAM coupling, N-terminal acylation using capping mixture was performed, followed by Lys(Dde) side chain deprotection in four cycles of 3 min using 4 mL of 2% (v/v) hydrazine in DMF followed by 5(6)-FAM coupling as described above. A final wash in two cycles of 2 mL DCM for 1 min under agitation (400 rpm) was performed.

The protecting groups were removed and the peptides cleaved from resin using 5 mL of 95% (v/v) TFA, 2.5% mQ water and 2.5% (v/v) triisopropylsilane (TIS) for 90 min under agitation (400 rpm). After transferring in a 50 mL Falcon tube using a TFA compatible station, around 4 mL of TFA-cleavage solution was evaporated under a constant flow of nitrogen. The peptides were then precipitated in 50 mL of cold diethylether for 1 hr at -20°C and centrifuged at 3,400 x g (around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 10 min at 4°C. Solid yellow pellets were obtained while the supernatants were discarded.

For the oxidation of the two cysteines side chains, the obtained solids were dissolved in 3 mL of dimethylsulfoxide (DMSO), 1 mL ACN and 2 mL of 1 M TRIS pH 8.5 and incubated overnight at room temperature (RT) protected from light. The resulting cyclization mixtures were diluted up to 20 mL with mQ water and lyophilized.

For HPLC-purification, the solids obtained were dissolved in 10% (v/v) DMSO, 30% (v/v) ACN, 59.9% (v/v) mQ water, 0.1% (v/v) TFA, and run on a reverse phase C18 column (Waters Sunfire™, 10 μm, 100 Å, 19 x 250 mm) with a linear gradient of solvent B (99.9% [v/v] ACN, 0.1% [v/v] TFA) over A (99.9% [v/v] mQ water, 0.1% [v/v] TFA) from 30 to 50% in 30 min at a flow rate of 20 mL/min. The mass of purified peptides was confirmed by electrospray ionization mass spectrometry using the Shimadzu-2020 single quadrupole LC-MS instrument. Fractions containing the desired mass were combined and lyophilized. Yellow powders were obtained and considered as TFA salts.

Measuring binding affinity of fluorescent peptides by direct fluorescence polarization assay

The binding of the two fluorescent peptides was assessed by incubating the peptides (10 nM final concentration) with 2-fold dilutions of KRAS(G12D) (starting from 5,000 to 10 nM final concentration) in assay buffer (10 mM Na_2HPO_4, 1.8 mM NaH_2PO_4*H_2O, 137 mM NaCl, 2.7 mM KCl, pH 7.4, 0.01% [v/v] Tween-20) in a 384 well plate (ThermoFischer NUNC™ 384

shallow well std height plates non-sterile, black). After 10 min incubation at RT, the fluorescence anisotropy was measured (Infinite F2000 pro TECAN, E_{ex} = 485 nm, E_{em} = 535 nm, 25°C). The G-factor was calculated using 10 nM 5(6)-FAM in assay buffer as a reference, and assay buffer alone as a blank. The K_D was calculated from fitted data using GraphPad Prism 5 and the following formula (1):

$$y = a + (b - a)\frac{(K_d + x + P) - \sqrt{(K_d + x + P)^2 - 4xP}}{2P} \qquad (1)$$

Where y is the anisotropy measured, a and b are the signal in absence or with fluorescent ligand respectively, x is the protein concentration, and P is the probe concentration.

Establishing fluorescence polarization competition assay

To assess probe specificity and displacement event of fluorescent probes bound to KRAS(G12D), we synthesized the same peptide described above without fluorescein (Ac-RCPLYISYDPVC-NH_2) to be used as a control peptide. Fluorescent peptide (10 nM final concentration) bound at around 70% to KRAS(G12D) (250 nM final concentration) were incubated with 2-fold dilutions of control peptide (20 to 0.02 µM final concentration) in assay buffer in the same 384 well plates described previously. After 30 min incubation at RT, the fluorescence anisotropy was measured as described above. All peptides (probe and control peptide) were initially prepared in DMSO. Therefore, to account for the volume of DMSO added, all wells were complemented with DMSO to reach a final concentration of 1%.

The data were normalized considering negative control (without control peptide) and reported probe alone (no protein) as 0 and 100% displacement of the reporter peptide, respectively. To calculate the IC_{50}, a sigmoidal curve was fitted to the data using Graphpad Prism 5 and the following equation (2):

$$y = \frac{100}{1 + 10^{(\log IC_{50} - x)p}} \qquad (2)$$

With y being displacement of reporter peptide (%), x the peptide concentration and p the Hill slope. IC_{50} indicates the concentration resulting in 50% displacement of the fluorescent peptide (relative to the functional strength of the inhibitor.

Z' factor calculation

The Z' factor was calculated using the following formula (3).

$$Z' = 1 - \frac{3*(\sigma_{pos} + \sigma_{neg})}{|\mu_{pos} + \mu_{neg}|} \qquad (3)$$

Where $\sigma_{pos}/\sigma_{neg}$ and μ_{pos}/μ_{neg} stand for standard deviation (σ) and average (μ) of positive (pos) and negative (neg) controls.

Synthesis of 2-pyridylthio cysteamine.HCl

Cysteamine (1 equiv. in 30 mL MeOH, 0.39 M), 2,2'-dipyridyldisulfide (2 equiv., in 70 mL MeOH, 0.77 M) and AcOH (2 ml) were incubated overnight at RT under agitation in a 250 mL balloon. After removing the solvent by rotation evaporation, the 2-pyridylthio cysteamine resulting product (around 15 mL left) was separated in 8 different 50 mL Falcon tubes and complemented with MeOH up to around 5 mL prior precipitation with 45 mL of RT diethyl ether. After 30 min of incubation at -20°C, the solid product was pelleted by centrifugation at 3,400 × g (around 4,000 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 5 min at 4°C. The solid was further purified in 6 cycles by dissolving the precipitate in 5 mL MeOH followed by the addition of 45 mL of RT diethyl ether, incubation at -20°C, and centrifugation. After the final centrifugation step, the white powder was dried by lyophilization. The resulting product was considered to be a chloride salt.

Preparation of cysteamine loaded polystyrene SH resin

Loading of polystyrene SH resin (Rapp Polymere A SA, 200-400 mesh, loading of 0.82 mmol/g) with cysteamine by disulfide exchange was prepared as follows.

Polystyrene SH resin (530 mg) was transferred to a 20 mL syringe, washed with 15 mL of dichloromethane (DCM) followed by swelling in 15 mL of 30% (v/v) MeOH/70 % (v/v) DCM for 20 min. 2-pyridylthio cysteamine.HCl (4.25 eq., 0.124 M final) was first dissolved in MeOH followed by the addition of DCM to reach 30% (v/v) MeOH/70 % (v/v) DCM. After adding N,N-diisopropylethylamine (DIPEA) (4.25 eq, 0.124 M final), the resulting mixture was transferred into the syringe and agitated (250 rpm) for 3 hrs at RT. Washes were then performed with 2 times 20 mL 30% (v/v) MeOH/70 % (v/v) DCM followed by 2 times 20 mL DMF. To ensure complete amine deprotonation, 5 min incubation with 9 mL 1.2M DIPEA in DMF were performed followed by 2 washes with 20 mL of DMF and 3 washed of 20 mL DCM. After being dried in the lyophilizer overnight, the resin was stored at 4°C.

Macrocycles library generation

Macrocycle libraries obtained from disulfide cyclized linear peptides were synthesized as follows.

Linear peptides were synthesized on cysteamine loaded polystyrene SH (5 µmol scale, loading considered as 0.82 mmol/g) in up to four 96 wells synthesis plates (Standard 96-Well Filter Plate, 0.7 mL/Well 10 µm PE Frit) at once using an automated peptide synthesizer (Intavis, Multipep RSi) and standard Fmoc-chemistry as described above. All linear peptides contained an N-Fmoc/N-Boc-protected diamino acid and a terminal S-trityl-mercaptopropionic acid.

Prior to removal of the protecting group, all the outlets of the 96 well synthesis plate were sealed at once using ethylene-vinyl acetate foam (sealing foam) and mechanical pressure to ensure proper sealing. The protecting groups of the peptides bound to the resin were deprotected in the synthesis plate with 600 µL per well of 95% (v/v) TFA, 2.5% (v/v) mQ water and 2.5% (v/v) TIS for 3 hrs, and a polypropylene (PP) plastic seal to prevent TFA evaporation. A wash in 3 cycles of 300 µL DCM was then performed followed by resin air drying for 30 min at RT. A new sealing foam was then used to close all the outlets of the 96 well plates. Macrocycles formation due to peptide cyclorelease through on-resin thiols oxidation was performed in the 96 well synthesis plate by adding 200 µL per well of 0.150 M DABCO in DMSO and incubating overnight at RT with all plates closed with an aluminium foil to minimize losses in case of defective sealing. The 96 well synthesis plates were then transferred in a 96-deep well plate recipient (Biotix, Deep Well Plate, 96-Round Well, 1.2mL) and centrifuged at 480 x g (around 1'500 rpm on a Thermo Scientific Multifuge 3L-R centrifuge with Sorval 75006445 rotor) for 2 min at RT. Macrocycle library concentrations were estimated by measuring the absorbance average at 280 nm of Trp or Tyr-containing macrocycles (ε_{Trp}: 5,500 $M^{-1}.cm^{-1}$/ε_{Tyr}: 1,490 $M^{-1}.cm^{-1}$, NanoDrop 8000, Thermo Scientific).

Quality control of the library was assessed by analyzing the UV-spectra and mass by LC-MS of some representative macrocycles (around 20 for a 384 macrocycles library). These were selected to cover as many amino acids involved in a synthesis as possible (up to 31 per synthesis). The macrocycle libraries (50 µL) were transferred into 384-well ECHO PP plate (Labcyte, 384-Well PP 2.0 Mictroplate, Echo® Qualified) and stored at -20°C.

Macrocycles peripheral diversification by acoustic liquid transfer

Macrocycles were combinatorially diversified through the coupling of their free amine with activated carboxylic acids. For library assembly, ECHO 650 device from Labcyte was used with DMSO settings and destination-centric methodology. Separate ECHO qualified 384-well PP plate (Labcyte™, 384-Well PP 2.0 Mictroplate, Echo® Qualified) were used as source plates for the macrocycles and the activated acids. Low volume 384-well plates

(ThermoFischer NUNC™ 384 shallow well std height plates non-sterile, black, no lid; Cat. # 264705) were used as destination plates.

All acids were prepared at 127.6 mM in a solution of DMSO containing 127.6 mM DABCO and mixed with HBTU at 127.6 mM in DMSO in equal amount to reach a final concentration of 63.8 mM for all components before being manually transferred (40 μL) in a 384-well ECHO PP plate. For combinatorial diversification, macrocycles (60 nL, 6.39 mM, 1 eq.) were transferred followed by carboxylic acid/HBTU/DABCO (60 nL, 63.8 mM, 10 eq.) in destination plates. To ensure droplet merging and proper coupling, all destination plates were manually sealed with a plastic lid, centrifuged at 1,100 x g (around 2,500 rpm with SIGMA 4-16KS with a swing-out rotor 11660) and incubated 5 hrs at RT. After dispensing 105 nL of DMSO with a bulk dispenser to ensure proper solubility, the activated acids were quenched by dispensing 4880 nL of 100 mM TRIS, 137 mM NaCl, 2.7 mM KCl, pH 7.4 in all plates using a bulk dispenser (Gyger, Certus Flex Liquid Dispenser) and by incubating overnight at RT. The reactions had a final volume of 5 μL containing 75 μM macrocycles and 767 μM of carboxylic acid/HBTU/ DABCO containing DMSO content of 4.5% (v/v).

For quality control, the products from five identical reactions were pooled and analyzed by LC-MS (10 μL of pooled sample). Qualitative conversion was estimated based on TIC-peaks.

Screening macrocycle compound library

Macrocycles were synthesized in 384 well plates exactly as described above. The entire library assembly and screening involving sources or destination plates swapping was done automatically using a robotic arm (Labcyte™,ACCESS™ station). To prevent DMSO hydration of the source plates, microclime lids (Labcyte, MicroClime ® Environmental Microplate Lid) containing DMSO were used. To prevent evaporation after adding quenching solution and after adding assay reagents in the destination plates, microclime lids (Labcyte™, MicroClime ® Environmental Microplate Lid) filled with mQ water were used. All lids were removed by air suction from the ACCESS™ station prior to use.

To sixteen dedicated wells of each 384 well plates, 120 nL of DMSO were acoustically transferred using ECHO650 followed by simultaneous dispensing of 105 nL of DMSO and 4775 nL of 100 mM TRIS, 137 mM NaCl, 2.7 mM KCl, pH 7.4 in all plates using a bulk dispenser (Certus Flex Liquid Dispenser) with eight wells to be used as negative and positive controls respectively. The other wells were filled combinatorically with the macrocycles and activated acids (one acid per plate) exactly as described above. One control plate undergoing the same synthetic process but including macrocycles alone or activated acids alone, the other component being replaced by the required amount in DMSO, was included for each screening.

KRAS(G12D) and the fluorescent probe were diluted using screening buffer (30 mM TRIS, 137 mM NaCl, 2.7 mM KCl, 0.015% Tween-20, pH 7.4) and mixed. The dilutions were chosen so that concentrations of 375 nM and 3 nM were reached, respectively. After 10 min, 10 µL of the reagent mix was added to each well of the plates using a bulk dispenser (Gyger, Certus Flex Liquid Dispenser). For positive control (8 wells), the same volume was added but of a mixture containing only fluorescent probe in screening buffer at a concentration of 3 nM. Final volume of 15 µL per well containing 250 nM KRAS(G12D) (except for the 8 wells of positive control, which didn't contain any protein), 2 nM fluorescent peptide, and 25 µM macrocycles were reached. A final of 1.5% (v/v) DMSO was reached. A plastic lid was manually added, and the plates were incubated for 30 min at RT. The fluorescence anisotropy was recorded using a plate reader inserted in the ACCESS™ station (BMG, PHERAstar® FSX, E_{ex} = 485 nm, E_{em} = 535 nm).

Normalization of fluorescence anisotropy into reporter peptide displacement

The raw anisotropy data were converted into reporter peptide displacement in % using the following equation (4):

$$\text{Displacement of reporter peptide } (\%) = 100 \times \frac{\text{A (negative control)} - \text{A (macrocycle reaction)}}{\text{A (negative control)} - \text{A (positive control)}} \quad (4)$$

Where A stands for anisotropy. A (negative control) corresponds to the mean value of the anisotropy in the absence of a competitor (8 wells per plate). The A (positive control) corresponds to the mean value of the anisotropy in the absence of KRAS(G12D) (8 wells).

5.7 Supplementary information

a. DNA sequence:

5'-
CTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGTCACCATCACC
ATCATCACCATCATGGTGCTACCGGCTCTACCGCTGGCTCTGGCACCGCTGGT
TCTACCGGTGCTTCTGGCGCTTCCACCGGCGGTACCGGCGCTACCCATCATCA
CCATCATCACCACCATGAAAACCTGTATTTTCAGGGCGGATCC**ATGACGGAAT**
ACAAACTGGTGGTGGTGGGTGCGGACGGCGTGGGTAAAAGTGCTCTGACGA
TCCAACTGATCCAAAATCACTTCGTGGATGAATATGACCCGACCATTGAAGA
TTCTTACCGTAAACAGGTGGTTATAGACGGCGAAACCTGCCTGCTGGATATC
CTGGACACGGCCGGTCAGGAAGAATATTCAGCTATGCGTGATCAATACATGC
GCACCGGCGAAGGTTTTCTGTGTGTTCGCGATAAACAACACGAAATCGTT
CGAAGATATCCATCACTATCGTGAACAGATCAAACGCGTTAAAGATAGTGAA
GACGTCCCGATGGTCCTGGTGGGCAACAAATGCGATCTGCCGTCCCGGACC
GTTGATACGAAACAGGCGCAAGACCTGGCCCGCAGCTACGGTATTCCGTTCA
TCGAAACCTCTGCAAAAACGCGCCAAGGCGTTGATGATGCCTTCTACACCCT
GGTTCGTGAAATCCGGAAACACAAAGAAAAATAACTCGAGCACCACCACCAC
CACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTC
TTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGAC
GCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGC
GTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCC
TTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGC
TCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTG
ATTAGGGTGATGGTTCACGTANNGGGCCATCGCCCTGATAAACGGTTTTCCCC
CTTGACGTTGGAG-3'

b. Protein sequence:

*MGHHHHHHHHGATGSTAGSGTAGSTGASGASTGGTGATHHHHHHHH*ENLYFQ

G*GS***MTEYKLVVVGADGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLL**

DILDTAGQEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSED

VPMVLVGNKCDLPSRTVDTKQAQDLARSYGIPFIETSAKTRQGVDDAFYTLVREI

RKHKEK

Figure S1. DNA and protein sequence of KRAS(G12D). (a) Region of the expression vector that was confirmed by Sanger sequencing. In bold is the gene coding for KRAS(G12D) (1-169 aa). The cleavage sites of restriction enzyme NcoI and XhoI are underlined. Stop codon is highlighted in red. (b) Final protein sequence expressed. In bold are the amino acids of KRAS(G12D) (1-169 aa). Amino acids containing the two His_8-Tag separated by a 28-mer amino acid linker is in italic followed by the TEV cleavable part.

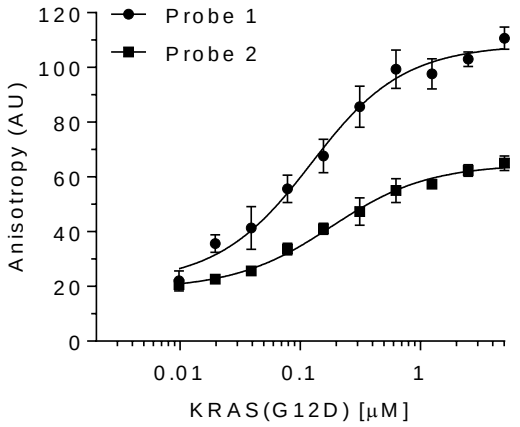


Figure S2. Raw data of probe 1 and probe 2 in direct fluorescence polarization assay. Binding of fluorescent reporter peptides probe 1 corresponding to 5(6)-FAM-R<u>C</u>PLYISYDPV<u>C</u>-NH_2 and probe 2 to Ac-<u>C</u>PLYISYDPV<u>C</u>RK(5(6)-FAM)-NH_2 to KRAS(G12D), both probes being cyclized through cysteines oxidation. Both probes final concentration was 10 nM. Means and standard deviation of three measurements are shown.

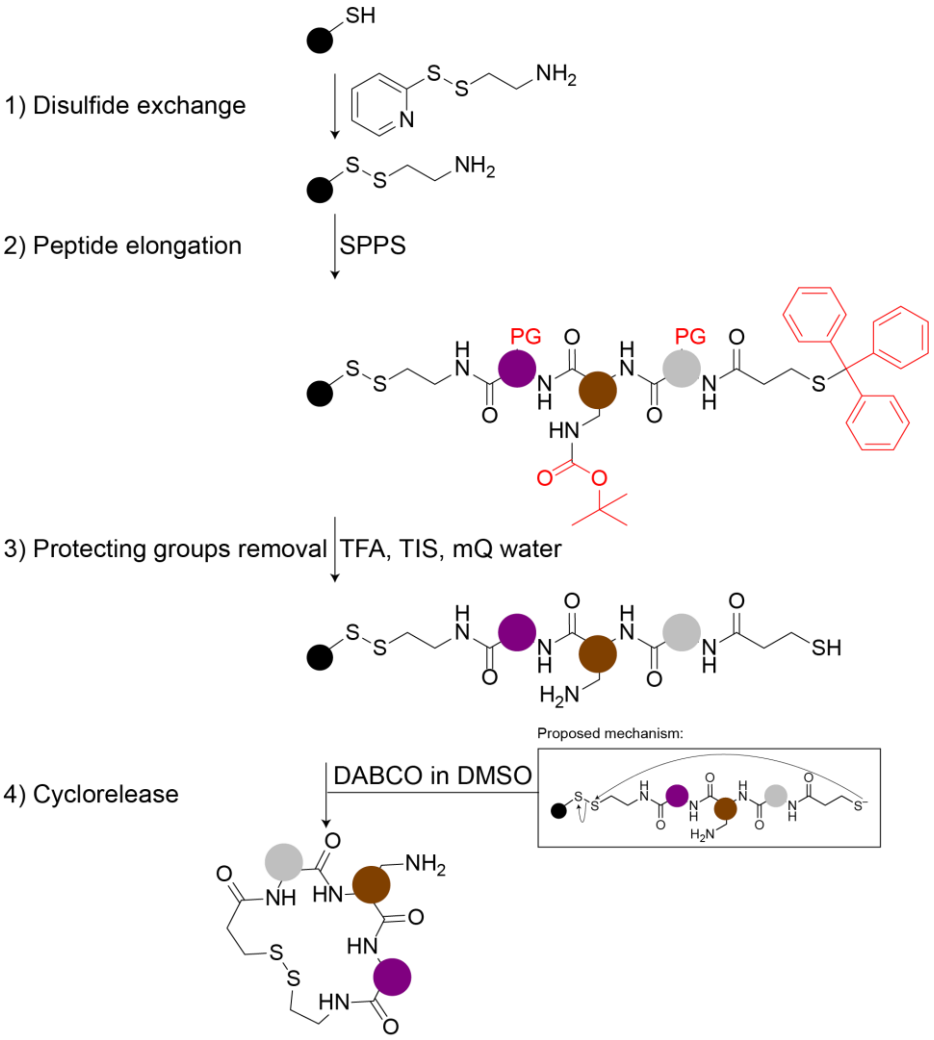


Figure S3. Ultra-parallel synthesis of high quality macrocycles without purification. Step-by-step methodology starting from the on resin disulfide exchange followed by peptide elongation, protecting group removal and thiol oxidative cyclization to yield high quality macrocycles. In purple and brown are the aliphatic side chain and diamino acid respectively. In gray is the last building block where properties vary depending on the library.

a) Library 1: Backbone

M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence
M1	Mpa-1-A-k-cys	M49	Mpa-7-J-j-cys	M97	Mpa-F-5-c-cys	M145	Mpa-c-3-A-cys	M193	Mpa-1-a-H-cys	M241	Mpa-7-d-E-cys	M289	Mpa-J-b-5-cys	M337	Mpa-b-J-3-cys
M2	Mpa-1-G-b-cys	M50	Mpa-7-I-h-cys	M98	Mpa-J-5-e-cys	M146	Mpa-W-3-F-cys	M194	Mpa-1-g-C-cys	M242	Mpa-7-a-D-cys	M290	Mpa-H-h-5-cys	M338	Mpa-i-E-3-cys
M3	Mpa-1-E-j-cys	M51	Mpa-7-B-d-cys	M99	Mpa-H-5-i-cys	M147	Mpa-b-3-J-cys	M195	Mpa-1-d-E-cys	M243	Mpa-7-c-A-cys	M291	Mpa-B-d-5-cys	M339	Mpa-j-G-3-cys
M4	Mpa-1-I-d-cys	M52	Mpa-7-D-b-cys	M100	Mpa-I-5-a-cys	M148	Mpa-g-3-B-cys	M196	Mpa-1-c-F-cys	M244	Mpa-7-g-H-cys	M292	Mpa-A-a-5-cys	M340	Mpa-f-I-3-cys
M5	Mpa-1-B-W-cys	M53	Mpa-7-A-a-cys	M101	Mpa-D-5-b-cys	M149	Mpa-a-3-E-cys	M197	Mpa-1-e-B-cys	M245	Mpa-7-j-C-cys	M293	Mpa-I-j-5-cys	M341	Mpa-h-H-3-cys
M6	Mpa-1-C-e-cys	M54	Mpa-7-H-f-cys	M102	Mpa-G-5-f-cys	M150	Mpa-j-3-H-cys	M198	Mpa-1-b-I-cys	M246	Mpa-7-h-I-cys	M294	Mpa-C-e-5-cys	M342	Mpa-c-C-3-cys
M7	Mpa-1-D-g-cys	M55	Mpa-7-G-e-cys	M103	Mpa-A-5-h-cys	M151	Mpa-i-3-C-cys	M199	Mpa-1-j-A-cys	M247	Mpa-7-e-G-cys	M295	Mpa-E-f-5-cys	M343	Mpa-e-B-3-cys
M8	Mpa-1-H-c-cys	M56	Mpa-7-F-c-cys	M104	Mpa-B-5-k-cys	M152	Mpa-k-3-D-cys	M200	Mpa-1-f-J-cys	M248	Mpa-7-b-J-cys	M296	Mpa-G-W-5-cys	M344	Mpa-g-D-3-cys
M9	Mpa-2-H-f-cys	M57	Mpa-8-G-b-cys	M105	Mpa-J-6-h-cys	M153	Mpa-f-4-I-cys	M201	Mpa-2-d-B-cys	M249	Mpa-8-g-C-cys	M297	Mpa-J-e-6-cys	M345	Mpa-j-A-4-cys
M10	Mpa-2-F-k-cys	M58	Mpa-8-A-f-cys	M106	Mpa-E-6-e-cys	M154	Mpa-e-4-F-cys	M202	Mpa-2-j-E-cys	M250	Mpa-8-j-B-cys	M298	Mpa-F-d-6-cys	M346	Mpa-c-F-4-cys
M11	Mpa-2-D-c-cys	M59	Mpa-8-J-h-cys	M107	Mpa-I-6-j-cys	M155	Mpa-i-4-A-cys	M203	Mpa-2-a-G-cys	M251	Mpa-8-f-J-cys	M299	Mpa-E-a-6-cys	M347	Mpa-h-H-4-cys
M12	Mpa-2-C-d-cys	M60	Mpa-8-H-g-cys	M108	Mpa-H-6-d-cys	M156	Mpa-b-4-G-cys	M204	Mpa-2-f-F-cys	M252	Mpa-8-i-I-cys	M300	Mpa-G-i-6-cys	M348	Mpa-b-D-4-cys
M13	Mpa-2-G-j-cys	M61	Mpa-8-F-c-cys	M109	Mpa-F-6-b-cys	M157	Mpa-c-4-H-cys	M205	Mpa-2-k-C-cys	M253	Mpa-8-h-D-cys	M301	Mpa-C-j-6-cys	M349	Mpa-a-I-4-cys
M14	Mpa-2-A-b-cys	M62	Mpa-8-D-j-cys	M110	Mpa-C-6-g-cys	M158	Mpa-d-4-C-cys	M206	Mpa-2-e-H-cys	M254	Mpa-8-e-G-cys	M302	Mpa-B-g-6-cys	M350	Mpa-k-J-4-cys
M15	Mpa-2-E-g-cys	M63	Mpa-8-B-k-cys	M111	Mpa-B-6-i-cys	M159	Mpa-g-4-J-cys	M207	Mpa-2-b-I-cys	M255	Mpa-8-a-H-cys	M303	Mpa-A-h-6-cys	M351	Mpa-g-G-4-cys
M16	Mpa-2-B-a-cys	M64	Mpa-8-E-i-cys	M112	Mpa-D-6-f-cys	M160	Mpa-h-4-B-cys	M208	Mpa-2-c-A-cys	M256	Mpa-8-d-F-cys	M304	Mpa-H-b-6-cys	M352	Mpa-f-B-4-cys
M17	Mpa-3-G-j-cys	M65	Mpa-E-1-a-cys	M113	Mpa-C-7-d-cys	M161	Mpa-i-5-F-cys	M209	Mpa-3-a-I-cys	M257	Mpa-B-i-1-cys	M305	Mpa-B-h-7-cys	M353	Mpa-h-E-5-cys
M18	Mpa-3-A-g-cys	M66	Mpa-I-1-k-cys	M114	Mpa-B-7-h-cys	M162	Mpa-b-5-H-cys	M210	Mpa-3-j-E-cys	M258	Mpa-G-g-1-cys	M306	Mpa-J-g-7-cys	M354	Mpa-c-H-5-cys
M19	Mpa-3-I-b-cys	M67	Mpa-B-1-g-cys	M115	Mpa-J-7-k-cys	M163	Mpa-f-5-E-cys	M211	Mpa-3-e-G-cys	M259	Mpa-H-e-1-cys	M307	Mpa-A-b-7-cys	M355	Mpa-d-I-5-cys
M20	Mpa-3-E-d-cys	M68	Mpa-F-1-i-cys	M116	Mpa-I-7-a-cys	M164	Mpa-k-5-B-cys	M212	Mpa-3-g-C-cys	M260	Mpa-J-k-1-cys	M308	Mpa-C-j-7-cys	M356	Mpa-j-C-5-cys
M21	Mpa-3-C-i-cys	M69	Mpa-A-1-e-cys	M117	Mpa-G-7-j-cys	M165	Mpa-h-5-I-cys	M213	Mpa-3-b-H-cys	M261	Mpa-E-h-1-cys	M309	Mpa-G-k-7-cys	M357	Mpa-a-B-5-cys
M22	Mpa-3-D-e-cys	M70	Mpa-C-1-j-cys	M118	Mpa-H-7-c-cys	M166	Mpa-c-5-J-cys	M214	Mpa-3-i-D-cys	M262	Mpa-D-f-1-cys	M310	Mpa-H-c-7-cys	M358	Mpa-i-J-5-cys
M23	Mpa-3-B-k-cys	M71	Mpa-G-1-c-cys	M119	Mpa-A-7-b-cys	M167	Mpa-d-5-A-cys	M215	Mpa-3-d-A-cys	M263	Mpa-A-b-1-cys	M311	Mpa-D-i-7-cys	M359	Mpa-b-A-5-cys
M24	Mpa-3-F-f-cys	M72	Mpa-D-1-f-cys	M120	Mpa-F-7-e-cys	M168	Mpa-a-5-D-cys	M216	Mpa-3-f-F-cys	M264	Mpa-I-c-1-cys	M312	Mpa-I-d-7-cys	M360	Mpa-k-G-5-cys
M25	Mpa-4-G-g-cys	M73	Mpa-W-2-h-cys	M121	Mpa-B-8-b-cys	M169	Mpa-f-6-A-cys	M217	Mpa-4-j-I-cys	M265	Mpa-E-j-2-cys	M313	Mpa-I-h-8-cys	M361	Mpa-b-C-6-cys
M26	Mpa-4-E-a-cys	M74	Mpa-I-2-c-cys	M122	Mpa-A-8-i-cys	M170	Mpa-h-6-B-cys	M218	Mpa-4-a-W-cys	M266	Mpa-G-a-2-cys	M314	Mpa-B-k-8-cys	M362	Mpa-a-H-6-cys
M27	Mpa-4-B-j-cys	M75	Mpa-B-2-g-cys	M123	Mpa-F-8-h-cys	M171	Mpa-b-6-F-cys	M219	Mpa-4-b-B-cys	M267	Mpa-J-i-2-cys	M315	Mpa-A-f-8-cys	M363	Mpa-d-W-6-cys
M28	Mpa-4-A-h-cys	M76	Mpa-D-2-d-cys	M124	Mpa-G-8-e-cys	M172	Mpa-g-6-C-cys	M220	Mpa-4-i-C-cys	M268	Mpa-I-c-2-cys	M316	Mpa-C-j-8-cys	M364	Mpa-c-F-6-cys
M29	Mpa-4-D-e-cys	M77	Mpa-J-2-j-cys	M125	Mpa-D-8-g-cys	M173	Mpa-k-6-I-cys	M221	Mpa-4-c-G-cys	M269	Mpa-A-g-2-cys	M317	Mpa-J-d-8-cys	M365	Mpa-h-B-6-cys
M30	Mpa-4-C-f-cys	M78	Mpa-H-2-k-cys	M126	Mpa-H-8-j-cys	M174	Mpa-d-6-H-cys	M222	Mpa-4-g-J-cys	M270	Mpa-C-k-2-cys	M318	Mpa-E-a-8-cys	M366	Mpa-j-D-6-cys
M31	Mpa-4-I-b-cys	M79	Mpa-E-2-b-cys	M127	Mpa-E-8-f-cys	M175	Mpa-j-6-J-cys	M223	Mpa-4-k-D-cys	M271	Mpa-F-d-2-cys	M319	Mpa-G-i-8-cys	M367	Mpa-e-I-6-cys
M32	Mpa-4-H-d-cys	M80	Mpa-G-2-i-cys	M128	Mpa-C-8-c-cys	M176	Mpa-e-6-G-cys	M224	Mpa-4-h-F-cys	M272	Mpa-H-e-2-cys	M320	Mpa-D-e-8-cys	M368	Mpa-k-J-6-cys
M33	Mpa-5-G-d-cys	M81	Mpa-H-3-c-cys	M129	Mpa-c-1-D-cys	M177	Mpa-b-7-D-cys	M225	Mpa-5-k-C-cys	M273	Mpa-I-g-3-cys	M321	Mpa-b-G-1-cys	M369	Mpa-e-H-7-cys
M34	Mpa-5-B-k-cys	M82	Mpa-F-3-a-cys	M130	Mpa-f-1-B-cys	M178	Mpa-d-7-B-cys	M226	Mpa-5-g-F-cys	M274	Mpa-C-d-3-cys	M322	Mpa-j-J-1-cys	M370	Mpa-k-E-7-cys
M35	Mpa-5-H-e-cys	M83	Mpa-J-3-d-cys	M131	Mpa-k-1-A-cys	M179	Mpa-c-7-I-cys	M227	Mpa-5-c-B-cys	M275	Mpa-F-i-3-cys	M323	Mpa-h-C-1-cys	M371	Mpa-W-B-7-cys
M36	Mpa-5-I-f-cys	M84	Mpa-D-3-f-cys	M132	Mpa-h-1-E-cys	M180	Mpa-g-7-F-cys	M228	Mpa-5-a-I-cys	M276	Mpa-G-e-3-cys	M324	Mpa-e-I-1-cys	M372	Mpa-c-D-7-cys
M37	Mpa-5-C-c-cys	M85	Mpa-I-3-b-cys	M133	Mpa-b-1-I-cys	M181	Mpa-j-7-E-cys	M229	Mpa-5-h-D-cys	M277	Mpa-D-c-3-cys	M325	Mpa-f-E-1-cys	M373	Mpa-b-I-7-cys
M38	Mpa-5-D-a-cys	M86	Mpa-E-3-i-cys	M134	Mpa-i-1-H-cys	M182	Mpa-f-7-A-cys	M230	Mpa-5-b-H-cys	M278	Mpa-E-f-3-cys	M326	Mpa-d-A-1-cys	M374	Mpa-a-F-7-cys
M39	Mpa-5-J-i-cys	M87	Mpa-C-3-k-cys	M135	Mpa-a-1-C-cys	M183	Mpa-i-7-G-cys	M231	Mpa-5-d-E-cys	M279	Mpa-B-b-3-cys	M327	Mpa-k-H-1-cys	M375	Mpa-d-A-7-cys
M40	Mpa-5-E-h-cys	M88	Mpa-A-3-e-cys	M136	Mpa-j-1-J-cys	M184	Mpa-h-7-C-cys	M232	Mpa-5-j-G-cys	M280	Mpa-H-a-3-cys	M328	Mpa-g-D-1-cys	M376	Mpa-i-G-7-cys
M41	Mpa-6-C-d-cys	M89	Mpa-B-4-b-cys	M137	Mpa-j-2-I-cys	M185	Mpa-g-8-I-cys	M233	Mpa-6-d-E-cys	M281	Mpa-H-h-4-cys	M329	Mpa-a-G-2-cys	M377	Mpa-c-J-8-cys
M42	Mpa-6-I-c-cys	M90	Mpa-F-4-f-cys	M138	Mpa-a-2-H-cys	M186	Mpa-e-8-J-cys	M234	Mpa-6-k-H-cys	M282	Mpa-E-f-4-cys	M330	Mpa-g-H-2-cys	M378	Mpa-j-G-8-cys
M43	Mpa-6-J-a-cys	M91	Mpa-J-4-g-cys	M139	Mpa-e-2-C-cys	M187	Mpa-c-8-E-cys	M235	Mpa-6-f-D-cys	M283	Mpa-F-c-4-cys	M331	Mpa-h-E-2-cys	M379	Mpa-e-A-8-cys
M44	Mpa-6-A-k-cys	M92	Mpa-H-4-h-cys	M140	Mpa-b-2-J-cys	M188	Mpa-d-8-G-cys	M236	Mpa-6-g-F-cys	M284	Mpa-D-g-4-cys	M332	Mpa-k-D-2-cys	M380	Mpa-d-C-8-cys
M45	Mpa-6-H-h-cys	M93	Mpa-I-4-a-cys	M141	Mpa-f-2-D-cys	M189	Mpa-b-8-H-cys	M237	Mpa-6-a-G-cys	M285	Mpa-J-j-4-cys	M333	Mpa-e-J-2-cys	M381	Mpa-k-F-8-cys
M46	Mpa-6-G-j-cys	M94	Mpa-D-4-j-cys	M142	Mpa-h-2-B-cys	M190	Mpa-i-8-D-cys	M238	Mpa-6-e-B-cys	M286	Mpa-I-k-4-cys	M334	Mpa-c-F-2-cys	M382	Mpa-h-D-8-cys
M47	Mpa-6-E-f-cys	M95	Mpa-A-4-i-cys	M143	Mpa-c-2-E-cys	M191	Mpa-j-8-C-cys	M239	Mpa-6-h-I-cys	M287	Mpa-C-i-4-cys	M335	Mpa-f-I-2-cys	M383	Mpa-i-E-8-cys
M48	Mpa-6-F-g-cys	M96	Mpa-C-4-e-cys	M144	Mpa-g-2-G-cys	M192	Mpa-h-8-W-cys	M240	Mpa-6-b-J-cys	M288	Mpa-G-e-4-cys	M336	Mpa-i-C-2-cys	M384	Mpa-b-H-8-cys

b) Library 2: Polar

M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence
M385	Mpa-1-D-W-cys	M433	Mpa-7-E-u-cys	M481	Mpa-J-5-t-cys	M529	Mpa-v-3-F-cys	M577	Mpa-1-q-H-cys	M625	Mpa-7-m-G-cys	M673	Mpa-J-o-5-cys	M721	Mpa-u-D-3-cys
M386	Mpa-1-E-r-cys	M434	Mpa-7-B-I-cys	M482	Mpa-E-5-p-cys	M530	Mpa-W-3-H-cys	M578	Mpa-1-r-G-cys	M626	Mpa-7-n-A-cys	M674	Mpa-B-m-5-cys	M722	Mpa-q-C-3-cys
M387	Mpa-1-J-v-cys	M435	Mpa-7-I-t-cys	M483	Mpa-C-5-r-cys	M531	Mpa-s-3-G-cys	M579	Mpa-1-s-D-cys	M627	Mpa-7-o-E-cys	M675	Mpa-E-u-5-cys	M723	Mpa-I-A-3-cys
M388	Mpa-1-I-t-cys	M436	Mpa-7-G-n-cys	M484	Mpa-A-5-o-cys	M532	Mpa-I-3-C-cys	M580	Mpa-1-n-E-cys	M628	Mpa-7-r-I-cys	M676	Mpa-A-v-5-cys	M724	Mpa-v-F-3-cys
M389	Mpa-1-G-n-cys	M437	Mpa-7-J-r-cys	M485	Mpa-D-5-l-cys	M533	Mpa-q-3-A-cys	M581	Mpa-1-p-B-cys	M629	Mpa-7-q-B-cys	M677	Mpa-G-n-5-cys	M725	Mpa-o-H-3-cys
M390	Mpa-1-A-m-cys	M438	Mpa-7-D-s-cys	M486	Mpa-H-5-v-cys	M534	Mpa-o-3-B-cys	M582	Mpa-1-o-F-cys	M630	Mpa-7-u-F-cys	M678	Mpa-F-r-5-cys	M726	Mpa-s-J-3-cys
M391	Mpa-1-B-o-cys	M439	Mpa-7-C-p-cys	M487	Mpa-I-5-s-cys	M535	Mpa-u-3-D-cys	M583	Mpa-1-u-J-cys	M631	Mpa-7-t-J-cys	M679	Mpa-C-t-5-cys	M727	Mpa-m-I-3-cys
M392	Mpa-1-H-u-cys	M440	Mpa-7-F-v-cys	M488	Mpa-G-5-m-cys	M536	Mpa-r-3-J-cys	M584	Mpa-1-v-I-cys	M632	Mpa-7-v-C-cys	M680	Mpa-W-s-5-cys	M728	Mpa-t-G-3-cys
M393	Mpa-2-G-u-cys	M441	Mpa-8-G-r-cys	M489	Mpa-G-6-t-cys	M537	Mpa-r-4-F-cys	M585	Mpa-2-n-D-cys	M633	Mpa-8-t-F-cys	M681	Mpa-B-s-6-cys	M729	Mpa-m-E-4-cys
M394	Mpa-2-C-n-cys	M442	Mpa-8-A-p-cys	M490	Mpa-F-6-q-cys	M538	Mpa-o-4-G-cys	M586	Mpa-2-t-F-cys	M634	Mpa-8-v-E-cys	M682	Mpa-G-t-6-cys	M730	Mpa-o-C-4-cys
M395	Mpa-2-E-I-cys	M443	Mpa-8-D-s-cys	M491	Mpa-C-6-m-cys	M539	Mpa-v-4-C-cys	M587	Mpa-2-p-A-cys	M635	Mpa-8-m-H-cys	M683	Mpa-C-p-6-cys	M731	Mpa-v-H-4-cys
M396	Mpa-2-A-p-cys	M444	Mpa-8-F-n-cys	M492	Mpa-J-6-v-cys	M540	Mpa-n-4-E-cys	M588	Mpa-2-I-E-cys	M636	Mpa-8-o-J-cys	M684	Mpa-H-n-6-cys	M732	Mpa-p-D-4-cys
M397	Mpa-2-D-s-cys	M445	Mpa-8-B-I-cys	M493	Mpa-H-6-u-cys	M541	Mpa-t-4-J-cys	M589	Mpa-2-q-G-cys	M637	Mpa-8-n-I-cys	M685	Mpa-A-u-6-cys	M733	Mpa-n-F-4-cys
M398	Mpa-2-B-t-cys	M446	Mpa-8-J-v-cys	M494	Mpa-B-6-o-cys	M542	Mpa-m-4-A-cys	M590	Mpa-2-r-I-cys	M638	Mpa-8-u-C-cys	M686	Mpa-J-v-6-cys	M734	Mpa-r-A-4-cys
M399	Mpa-2-H-m-cys	M447	Mpa-8-E-u-cys	M495	Mpa-E-6-r-cys	M543	Mpa-p-4-I-cys	M591	Mpa-2-u-C-cys	M639	Mpa-8-r-G-cys	M687	Mpa-D-m-6-cys	M735	Mpa-q-I-4-cys
M400	Mpa-2-F-o-cys	M448	Mpa-8-I-m-cys	M496	Mpa-D-6-p-cys	M544	Mpa-I-4-B-cys	M592	Mpa-2-o-B-cys	M640	Mpa-8-s-B-cys	M688	Mpa-E-o-6-cys	M736	Mpa-u-G-4-cys
M401	Mpa-3-J-I-cys	M449	Mpa-C-1-s-cys	M497	Mpa-B-7-p-cys	M545	Mpa-r-5-C-cys	M593	Mpa-3-q-F-cys	M641	Mpa-C-s-1-cys	M689	Mpa-D-I-7-cys	M737	Mpa-s-H-5-cys
M402	Mpa-3-A-t-cys	M450	Mpa-J-1-m-cys	M498	Mpa-I-7-v-cys	M546	Mpa-m-5-J-cys	M594	Mpa-3-u-H-cys	M642	Mpa-J-o-1-cys	M690	Mpa-F-s-7-cys	M738	Mpa-v-C-5-cys
M403	Mpa-3-F-r-cys	M451	Mpa-A-1-v-cys	M499	Mpa-A-7-t-cys	M547	Mpa-n-5-I-cys	M595	Mpa-3-r-B-cys	M643	Mpa-D-p-1-cys	M691	Mpa-E-r-7-cys	M739	Mpa-t-A-5-cys
M404	Mpa-3-G-s-cys	M452	Mpa-D-1-u-cys	M500	Mpa-H-7-r-cys	M548	Mpa-I-5-E-cys	M596	Mpa-3-p-J-cys	M644	Mpa-E-r-1-cys	M692	Mpa-A-o-7-cys	M740	Mpa-o-J-5-cys
M405	Mpa-3-B-u-cys	M453	Mpa-E-1-t-cys	M501	Mpa-F-7-m-cys	M549	Mpa-t-5-A-cys	M597	Mpa-3-m-D-cys	M645	Mpa-A-u-1-cys	M693	Mpa-H-q-7-cys	M741	Mpa-q-G-5-cys
M406	Mpa-3-I-o-cys	M454	Mpa-G-1-p-cys	M502	Mpa-C-7-I-cys	M550	Mpa-v-5-F-cys	M598	Mpa-3-n-I-cys	M646	Mpa-G-v-1-cys	M694	Mpa-J-t-7-cys	M742	Mpa-m-B-5-cys
M407	Mpa-3-C-n-cys	M455	Mpa-I-1-q-cys	M503	Mpa-G-7-u-cys	M551	Mpa-s-5-B-cys	M599	Mpa-3-o-E-cys	M647	Mpa-H-n-1-cys	M695	Mpa-B-u-7-cys	M743	Mpa-p-E-5-cys
M408	Mpa-3-H-q-cys	M456	Mpa-B-1-I-cys	M504	Mpa-D-7-o-cys	M552	Mpa-p-5-H-cys	M600	Mpa-3-t-A-cys	M648	Mpa-I-q-1-cys	M696	Mpa-C-m-7-cys	M744	Mpa-r-D-5-cys
M409	Mpa-4-H-p-cys	M457	Mpa-A-2-p-cys	M505	Mpa-F-8-o-cys	M553	Mpa-u-6-C-cys	M601	Mpa-4-s-H-cys	M649	Mpa-A-n-2-cys	M697	Mpa-D-m-8-cys	M745	Mpa-o-W-6-cys
M410	Mpa-4-C-m-cys	M458	Mpa-J-2-s-cys	M506	Mpa-J-8-I-cys	M554	Mpa-t-6-G-cys	M602	Mpa-4-t-J-cys	M650	Mpa-C-r-2-cys	M698	Mpa-G-o-8-cys	M746	Mpa-t-F-6-cys
M411	Mpa-4-G-r-cys	M459	Mpa-F-2-v-cys	M507	Mpa-G-8-n-cys	M555	Mpa-v-6-F-cys	M603	Mpa-4-u-G-cys	M651	Mpa-D-t-2-cys	M699	Mpa-B-u-8-cys	M747	Mpa-u-J-6-cys
M412	Mpa-4-J-I-cys	M460	Mpa-H-2-o-cys	M508	Mpa-I-8-m-cys	M556	Mpa-s-6-H-cys	M604	Mpa-4-v-C-cys	M652	Mpa-E-o-2-cys	M700	Mpa-C-v-8-cys	M748	Mpa-v-E-6-cys
M413	Mpa-4-F-q-cys	M461	Mpa-E-2-q-cys	M509	Mpa-E-8-t-cys	M557	Mpa-r-6-I-cys	M605	Mpa-4-I-A-cys	M653	Mpa-B-u-2-cys	M701	Mpa-F-p-8-cys	M749	Mpa-s-C-6-cys
M414	Mpa-4-A-s-cys	M462	Mpa-G-2-n-cys	M510	Mpa-D-8-p-cys	M558	Mpa-p-6-E-cys	M606	Mpa-4-p-W-cys	M654	Mpa-J-I-2-cys	M702	Mpa-H-I-8-cys	M750	Mpa-n-G-6-cys
M415	Mpa-4-E-o-cys	M463	Mpa-I-2-u-cys	M511	Mpa-A-8-q-cys	M559	Mpa-o-6-D-cys	M607	Mpa-4-m-D-cys	M655	Mpa-I-m-2-cys	M703	Mpa-A-q-8-cys	M751	Mpa-I-A-6-cys
M416	Mpa-4-I-v-cys	M464	Mpa-B-2-W-cys	M512	Mpa-B-8-u-cys	M560	Mpa-I-6-A-cys	M608	Mpa-4-r-E-cys	M656	Mpa-F-p-2-cys	M704	Mpa-W-r-8-cys	M752	Mpa-r-B-6-cys
M417	Mpa-5-B-q-cys	M465	Mpa-I-3-m-cys	M513	Mpa-q-1-E-cys	M561	Mpa-q-7-F-cys	M609	Mpa-5-u-I-cys	M657	Mpa-C-n-3-cys	M705	Mpa-o-G-1-cys	M753	Mpa-o-E-7-cys
M418	Mpa-5-E-I-cys	M466	Mpa-E-3-t-cys	M514	Mpa-n-1-H-cys	M562	Mpa-u-7-I-cys	M610	Mpa-5-m-A-cys	M658	Mpa-H-p-3-cys	M706	Mpa-s-J-1-cys	M754	Mpa-t-I-7-cys
M419	Mpa-5-D-n-cys	M467	Mpa-J-3-r-cys	M515	Mpa-t-1-D-cys	M563	Mpa-p-7-J-cys	M611	Mpa-5-s-F-cys	M659	Mpa-D-I-3-cys	M707	Mpa-v-I-1-cys	M755	Mpa-p-H-7-cys
M420	Mpa-5-H-u-cys	M468	Mpa-F-3-p-cys	M516	Mpa-v-1-B-cys	M564	Mpa-n-7-E-cys	M612	Mpa-5-v-D-cys	M660	Mpa-G-q-3-cys	M708	Mpa-r-A-1-cys	M756	Mpa-s-J-7-cys
M421	Mpa-5-I-p-cys	M469	Mpa-D-3-n-cys	M517	Mpa-r-1-G-cys	M565	Mpa-v-7-C-cys	M613	Mpa-5-I-B-cys	M661	Mpa-B-t-3-cys	M709	Mpa-m-H-1-cys	M757	Mpa-n-G-7-cys
M422	Mpa-5-C-v-cys	M470	Mpa-A-3-I-cys	M518	Mpa-u-1-F-cys	M566	Mpa-s-7-G-cys	M614	Mpa-5-t-H-cys	M662	Mpa-J-v-3-cys	M710	Mpa-u-E-1-cys	M758	Mpa-r-W-7-cys
M423	Mpa-5-J-s-cys	M471	Mpa-H-3-v-cys	M519	Mpa-o-1-C-cys	M567	Mpa-t-7-D-cys	M615	Mpa-5-o-G-cys	M663	Mpa-A-s-3-cys	M711	Mpa-p-C-1-cys	M759	Mpa-u-F-7-cys
M424	Mpa-5-F-m-cys	M472	Mpa-G-3-u-cys	M520	Mpa-p-1-A-cys	M568	Mpa-m-7-H-cys	M616	Mpa-5-n-C-cys	M664	Mpa-I-m-3-cys	M712	Mpa-I-D-1-cys	M760	Mpa-I-C-7-cys
M425	Mpa-6-J-v-cys	M473	Mpa-A-4-o-cys	M521	Mpa-o-2-B-cys	M569	Mpa-n-8-B-cys	M617	Mpa-6-n-G-cys	M665	Mpa-A-s-4-cys	M713	Mpa-v-D-2-cys	M761	Mpa-v-F-8-cys
M426	Mpa-6-D-n-cys	M474	Mpa-C-4-q-cys	M522	Mpa-u-2-G-cys	M570	Mpa-t-8-C-cys	M618	Mpa-6-m-I-cys	M666	Mpa-B-u-4-cys	M714	Mpa-I-H-2-cys	M762	Mpa-s-D-8-cys
M427	Mpa-6-G-m-cys	M475	Mpa-H-4-t-cys	M523	Mpa-m-2-F-cys	M571	Mpa-s-8-E-cys	M619	Mpa-6-q-F-cys	M667	Mpa-I-n-4-cys	M715	Mpa-t-G-2-cys	M763	Mpa-t-C-8-cys
M428	Mpa-6-F-t-cys	M476	Mpa-I-4-m-cys	M524	Mpa-t-2-H-cys	M572	Mpa-r-8-J-cys	M620	Mpa-6-u-A-cys	M668	Mpa-D-p-4-cys	M716	Mpa-u-A-2-cys	M764	Mpa-u-J-8-cys
M429	Mpa-6-H-p-cys	M477	Mpa-D-4-s-cys	M525	Mpa-q-2-I-cys	M573	Mpa-v-8-I-cys	M621	Mpa-6-p-H-cys	M669	Mpa-G-I-4-cys	M717	Mpa-m-F-2-cys	M765	Mpa-q-H-8-cys
M430	Mpa-6-A-o-cys	M478	Mpa-J-4-I-cys	M526	Mpa-p-2-A-cys	M574	Mpa-q-8-H-cys	M622	Mpa-6-s-E-cys	M670	Mpa-E-q-4-cys	M718	Mpa-r-C-2-cys	M766	Mpa-r-B-8-cys
M431	Mpa-6-I-I-cys	M479	Mpa-B-4-v-cys	M527	Mpa-I-2-E-cys	M575	Mpa-I-8-A-cys	M623	Mpa-6-r-D-cys	M671	Mpa-H-r-4-cys	M719	Mpa-n-J-2-cys	M767	Mpa-n-E-8-cys
M432	Mpa-6-B-s-cys	M480	Mpa-E-4-r-cys	M528	Mpa-r-2-J-cys	M576	Mpa-p-8-D-cys	M624	Mpa-6-t-B-cys	M672	Mpa-J-t-4-cys	M720	Mpa-o-E-2-cys	M768	Mpa-p-A-8-cys

Figure S4. All macrocycles sequences. All codes are discussed in Figure 2. M# represent the macrocycle number. (a) Library 1 (M1 to M384). (b) Library 2 (M385 to M768). Figure S4 continues next page.

c) Library 3: Aromatic

M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence
M769	Mpa-1-C-w-cys	M817	Mpa-7-B-ag-cys	M865	Mpa-F-5-y-cys	M913	Mpa-ag-3-C-cys	M961	Mpa-1-ac-D-cys	M1009	Mpa-7-ae-A-cys	M1057	Mpa-E-x-5-cys	M1105	Mpa-w-H-3-cys
M770	Mpa-1-H-y-cys	M818	Mpa-7-H-w-cys	M866	Mpa-E-5-ad-cys	M914	Mpa-x-3-A-cys	M962	Mpa-1-ab-A-cys	M1010	Mpa-7-y-I-cys	M1058	Mpa-D-ae-5-cys	M1106	Mpa-ac-D-3-cys
M771	Mpa-1-A-ag-cys	M819	Mpa-7-A-ae-cys	M867	Mpa-I-5-ac-cys	M915	Mpa-ad-3-G-cys	M963	Mpa-1-w-I-cys	M1011	Mpa-7-ac-G-cys	M1059	Mpa-H-ac-5-cys	M1107	Mpa-x-I-3-cys
M772	Mpa-1-F-ah-cys	M820	Mpa-7-G-z-cys	M868	Mpa-A-5-x-cys	M916	Mpa-y-3-F-cys	M964	Mpa-1-af-B-cys	M1012	Mpa-7-af-F-cys	M1060	Mpa-C-ag-5-cys	M1108	Mpa-ab-A-3-cys
M773	Mpa-1-I-x-cys	M821	Mpa-7-C-ah-cys	M869	Mpa-G-5-ae-cys	M917	Mpa-z-3-H-cys	M965	Mpa-1-x-H-cys	M1013	Mpa-7-ab-H-cys	M1061	Mpa-B-y-5-cys	M1109	Mpa-ag-C-3-cys
M774	Mpa-1-D-af-cys	M822	Mpa-7-D-x-cys	M870	Mpa-H-5-ab-cys	M918	Mpa-ac-3-B-cys	M966	Mpa-1-ae-C-cys	M1014	Mpa-7-aa-J-cys	M1062	Mpa-F-ad-5-cys	M1110	Mpa-aa-E-3-cys
M775	Mpa-1-J-ad-cys	M823	Mpa-7-I-y-cys	M871	Mpa-B-5-w-cys	M919	Mpa-ae-3-E-cys	M967	Mpa-1-ah-F-cys	M1015	Mpa-7-w-B-cys	M1063	Mpa-A-ab-5-cys	M1111	Mpa-ae-G-3-cys
M776	Mpa-1-E-aa-cys	M824	Mpa-7-J-af-cys	M872	Mpa-C-5-ag-cys	M920	Mpa-ah-3-J-cys	M968	Mpa-1-ag-G-cys	M1016	Mpa-7-x-E-cys	M1064	Mpa-I-ah-5-cys	M1112	Mpa-ad-J-3-cys
M777	Mpa-2-H-ab-cys	M825	Mpa-8-G-aa-cys	M873	Mpa-H-6-ac-cys	M921	Mpa-ah-4-D-cys	M969	Mpa-2-af-A-cys	M1017	Mpa-8-aa-H-cys	M1065	Mpa-E-ab-6-cys	M1113	Mpa-ag-E-4-cys
M778	Mpa-2-C-aa-cys	M826	Mpa-8-B-ac-cys	M874	Mpa-B-6-ad-cys	M922	Mpa-y-4-I-cys	M970	Mpa-2-x-G-cys	M1018	Mpa-8-ae-E-cys	M1066	Mpa-G-ah-6-cys	M1114	Mpa-aa-J-4-cys
M779	Mpa-2-Iw-cys	M827	Mpa-8-E-ab-cys	M875	Mpa-I-6-ae-cys	M923	Mpa-x-4-B-cys	M971	Mpa-2-aa-H-cys	M1019	Mpa-8-ag-D-cys	M1067	Mpa-B-x-6-cys	M1115	Mpa-y-I-4-cys
M780	Mpa-2-E-z-cys	M828	Mpa-8-I-z-cys	M876	Mpa-C-6-ag-cys	M924	Mpa-ab-4-F-cys	M972	Mpa-2-ad-B-cys	M1020	Mpa-8-ah-F-cys	M1068	Mpa-J-af-6-cys	M1116	Mpa-ah-F-4-cys
M781	Mpa-2-F-ae-cys	M829	Mpa-8-A-ag-cys	M877	Mpa-D-6-af-cys	M925	Mpa-ag-4-J-cys	M973	Mpa-2-w-J-cys	M1021	Mpa-8-ad-B-cys	M1069	Mpa-H-aa-6-cys	M1117	Mpa-z-C-4-cys
M782	Mpa-2-A-x-cys	M830	Mpa-8-F-x-cys	M878	Mpa-J-6-ah-cys	M926	Mpa-aa-4-A-cys	M974	Mpa-2-ag-C-cys	M1022	Mpa-8-y-J-cys	M1070	Mpa-F-ac-6-cys	M1118	Mpa-ac-A-4-cys
M783	Mpa-2-G-ah-cys	M831	Mpa-8-D-ah-cys	M879	Mpa-A-6-x-cys	M927	Mpa-w-4-E-cys	M975	Mpa-2-ah-I-cys	M1023	Mpa-8-ac-C-cys	M1071	Mpa-I-z-6-cys	M1119	Mpa-x-B-4-cys
M784	Mpa-2-D-ad-cys	M832	Mpa-8-H-af-cys	M880	Mpa-E-6-z-cys	M928	Mpa-z-4-H-cys	M976	Mpa-2-ac-D-cys	M1024	Mpa-8-af-I-cys	M1072	Mpa-C-ad-6-cys	M1120	Mpa-af-H-4-cys
M785	Mpa-3-H-ac-cys	M833	Mpa-H-1-x-cys	M881	Mpa-B-7-ag-cys	M929	Mpa-x-5-F-cys	M977	Mpa-3-ac-J-cys	M1025	Mpa-G-ah-1-cys	M1073	Mpa-B-aa-7-cys	M1121	Mpa-ae-E-5-cys
M786	Mpa-3-D-af-cys	M834	Mpa-A-1-ag-cys	M882	Mpa-A-7-ab-cys	M930	Mpa-w-5-H-cys	M978	Mpa-3-w-A-cys	M1026	Mpa-D-ad-1-cys	M1074	Mpa-F-ad-7-cys	M1122	Mpa-z-C-5-cys
M787	Mpa-3-G-ag-cys	M835	Mpa-J-1-z-cys	M883	Mpa-H-7-w-cys	M931	Mpa-aa-5-C-cys	M979	Mpa-3-z-I-cys	M1027	Mpa-I-af-1-cys	M1075	Mpa-C-z-7-cys	M1123	Mpa-af-J-5-cys
M788	Mpa-3-E-ah-cys	M836	Mpa-B-1-ae-cys	M884	Mpa-I-7-y-cys	M932	Mpa-ah-5-E-cys	M980	Mpa-3-aa-F-cys	M1028	Mpa-E-ac-1-cys	M1076	Mpa-D-ac-7-cys	M1124	Mpa-ad-H-5-cys
M789	Mpa-3-C-y-cys	M837	Mpa-F-1-aa-cys	M885	Mpa-D-7-x-cys	M933	Mpa-ad-5-A-cys	M981	Mpa-3-ae-G-cys	M1029	Mpa-B-x-1-cys	M1077	Mpa-J-y-7-cys	M1125	Mpa-ah-B-5-cys
M790	Mpa-3-F-aa-cys	M838	Mpa-G-1-ah-cys	M886	Mpa-C-7-ad-cys	M934	Mpa-z-5-B-cys	M982	Mpa-3-af-H-cys	M1030	Mpa-J-ae-1-cys	M1078	Mpa-A-ag-7-cys	M1126	Mpa-x-D-5-cys
M791	Mpa-3-J-ae-cys	M839	Mpa-C-1-y-cys	M887	Mpa-J-7-aa-cys	M935	Mpa-ae-5-J-cys	M983	Mpa-3-y-C-cys	M1031	Mpa-H-y-1-cys	M1079	Mpa-I-ae-7-cys	M1127	Mpa-w-G-5-cys
M792	Mpa-3-A-ad-cys	M840	Mpa-D-1-ab-cys	M888	Mpa-F-7-ac-cys	M936	Mpa-af-5-I-cys	M984	Mpa-3-x-D-cys	M1032	Mpa-F-z-1-cys	M1080	Mpa-G-ab-7-cys	M1128	Mpa-ag-I-5-cys
M793	Mpa-4-B-ah-cys	M841	Mpa-A-2-ac-cys	M889	Mpa-I-8-z-cys	M937	Mpa-ag-6-A-cys	M985	Mpa-4-ab-B-cys	M1033	Mpa-A-ad-2-cys	M1081	Mpa-J-w-8-cys	M1129	Mpa-ac-C-6-cys
M794	Mpa-4-H-w-cys	M842	Mpa-H-2-w-cys	M890	Mpa-E-8-ae-cys	M938	Mpa-aa-6-G-cys	M986	Mpa-4-aa-I-cys	M1034	Mpa-B-ag-2-cys	M1082	Mpa-G-ae-8-cys	M1130	Mpa-ah-G-6-cys
M795	Mpa-4-C-ad-cys	M843	Mpa-B-2-ah-cys	M891	Mpa-G-8-ag-cys	M939	Mpa-ah-6-E-cys	M987	Mpa-4-ae-J-cys	M1035	Mpa-G-ab-2-cys	M1083	Mpa-B-y-8-cys	M1131	Mpa-aa-F-6-cys
M796	Mpa-4-G-ag-cys	M844	Mpa-I-2-y-cys	M892	Mpa-F-8-ab-cys	M940	Mpa-ad-6-D-cys	M988	Mpa-4-x-G-cys	M1036	Mpa-H-ae-2-cys	M1084	Mpa-I-x-8-cys	M1132	Mpa-w-E-6-cys
M797	Mpa-4-A-ab-cys	M845	Mpa-G-2-ag-cys	M893	Mpa-B-8-ah-cys	M941	Mpa-ae-6-J-cys	M989	Mpa-4-ag-D-cys	M1037	Mpa-F-y-2-cys	M1085	Mpa-E-ah-8-cys	M1133	Mpa-x-J-6-cys
M798	Mpa-4-F-af-cys	M846	Mpa-D-2-ab-cys	M894	Mpa-C-8-ac-cys	M942	Mpa-z-6-C-cys	M990	Mpa-4-ac-F-cys	M1038	Mpa-I-ah-2-cys	M1086	Mpa-H-ag-8-cys	M1134	Mpa-z-B-6-cys
M799	Mpa-4-I-ac-cys	M847	Mpa-J-2-z-cys	M895	Mpa-A-8-ad-cys	M943	Mpa-w-6-I-cys	M991	Mpa-4-ah-C-cys	M1039	Mpa-D-z-2-cys	M1087	Mpa-D-ab-8-cys	M1135	Mpa-y-I-6-cys
M800	Mpa-4-J-ae-cys	M848	Mpa-F-2-ad-cys	M896	Mpa-H-8-y-cys	M944	Mpa-ab-6-H-cys	M992	Mpa-4-w-H-cys	M1040	Mpa-J-w-2-cys	M1088	Mpa-A-aa-8-cys	M1136	Mpa-ae-A-6-cys
M801	Mpa-5-J-ac-cys	M849	Mpa-F-3-aa-cys	M897	Mpa-aa-1-F-cys	M945	Mpa-ac-7-J-cys	M993	Mpa-5-af-A-cys	M1041	Mpa-E-x-3-cys	M1089	Mpa-aa-D-1-cys	M1137	Mpa-ag-E-7-cys
M802	Mpa-5-D-y-cys	M850	Mpa-B-3-ah-cys	M898	Mpa-y-1-I-cys	M946	Mpa-ad-7-H-cys	M994	Mpa-5-y-I-cys	M1042	Mpa-A-w-3-cys	M1090	Mpa-y-H-1-cys	M1138	Mpa-y-G-7-cys
M803	Mpa-5-F-ah-cys	M851	Mpa-E-3-ad-cys	M899	Mpa-ab-1-C-cys	M947	Mpa-ab-7-E-cys	M995	Mpa-5-aa-C-cys	M1043	Mpa-H-af-3-cys	M1091	Mpa-x-E-1-cys	M1139	Mpa-ad-J-7-cys
M804	Mpa-5-H-ae-cys	M852	Mpa-I-3-x-cys	M900	Mpa-ad-1-B-cys	M948	Mpa-ag-7-D-cys	M996	Mpa-5-x-H-cys	M1044	Mpa-I-ae-3-cys	M1092	Mpa-ag-B-1-cys	M1140	Mpa-ae-A-7-cys
M805	Mpa-5-G-w-cys	M853	Mpa-G-3-ae-cys	M901	Mpa-x-1-D-cys	M949	Mpa-z-7-G-cys	M997	Mpa-5-ac-D-cys	M1045	Mpa-C-aa-3-cys	M1093	Mpa-af-F-1-cys	M1141	Mpa-aa-C-7-cys
M806	Mpa-5-I-x-cys	M854	Mpa-C-3-ag-cys	M902	Mpa-w-1-J-cys	M950	Mpa-ah-7-B-cys	M998	Mpa-5-ae-B-cys	M1046	Mpa-D-y-3-cys	M1094	Mpa-w-G-1-cys	M1142	Mpa-w-B-7-cys
M807	Mpa-5-B-af-cys	M855	Mpa-H-3-af-cys	M903	Mpa-ag-1-A-cys	M951	Mpa-y-7-C-cys	M999	Mpa-5-w-F-cys	M1047	Mpa-J-ab-3-cys	M1095	Mpa-ah-I-1-cys	M1143	Mpa-ab-F-7-cys
M808	Mpa-5-C-aa-cys	M856	Mpa-D-3-y-cys	M904	Mpa-ac-1-H-cys	M952	Mpa-aa-7-I-cys	M1000	Mpa-5-z-E-cys	M1048	Mpa-F-z-3-cys	M1096	Mpa-ae-C-1-cys	M1144	Mpa-z-D-7-cys
M809	Mpa-6-A-aa-cys	M857	Mpa-C-4-ag-cys	M905	Mpa-ad-2-C-cys	M953	Mpa-ag-8-G-cys	M1001	Mpa-6-ag-J-cys	M1049	Mpa-J-y-4-cys	M1097	Mpa-w-I-2-cys	M1145	Mpa-x-B-8-cys
M810	Mpa-6-J-af-cys	M858	Mpa-G-4-ah-cys	M906	Mpa-ae-2-F-cys	M954	Mpa-w-8-D-cys	M1002	Mpa-6-ah-I-cys	M1050	Mpa-D-ag-4-cys	M1098	Mpa-ag-H-2-cys	M1146	Mpa-af-C-8-cys
M811	Mpa-6-I-ad-cys	M859	Mpa-D-4-w-cys	M907	Mpa-aa-2-D-cys	M955	Mpa-ac-8-B-cys	M1003	Mpa-6-y-G-cys	M1051	Mpa-B-ae-4-cys	M1099	Mpa-x-G-2-cys	M1147	Mpa-ah-G-8-cys
M812	Mpa-6-B-x-cys	M860	Mpa-B-4-y-cys	M908	Mpa-ag-2-G-cys	M956	Mpa-ah-8-A-cys	M1004	Mpa-6-z-D-cys	M1052	Mpa-A-af-4-cys	M1100	Mpa-ad-F-2-cys	M1148	Mpa-z-I-8-cys
M813	Mpa-6-C-ae-cys	M861	Mpa-H-4-x-cys	M909	Mpa-z-2-B-cys	M957	Mpa-af-8-E-cys	M1005	Mpa-6-x-F-cys	M1053	Mpa-F-ad-4-cys	M1101	Mpa-y-J-2-cys	M1149	Mpa-ab-E-8-cys
M814	Mpa-6-F-w-cys	M862	Mpa-F-4-ac-cys	M910	Mpa-w-2-J-cys	M958	Mpa-aa-8-F-cys	M1006	Mpa-6-ae-E-cys	M1054	Mpa-C-z-4-cys	M1102	Mpa-ae-B-2-cys	M1150	Mpa-ad-F-8-cys
M815	Mpa-6-D-ah-cys	M863	Mpa-J-4-ab-cys	M911	Mpa-ab-2-A-cys	M959	Mpa-ab-8-J-cys	M1007	Mpa-6-aa-H-cys	M1055	Mpa-E-aa-4-cys	M1103	Mpa-ab-A-2-cys	M1151	Mpa-ag-A-8-cys
M816	Mpa-6-G-ag-cys	M864	Mpa-A-4-af-cys	M912	Mpa-ac-2-H-cys	M960	Mpa-x-8-I-cys	M1008	Mpa-6-af-A-cys	M1056	Mpa-H-ah-4-cys	M1104	Mpa-aa-E-2-cys	M1152	Mpa-y-J-8-cys

d) Library 4: Aliphatic

M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence	M#	Sequence
M1153	Mpa-1-A-al-cys	M1201	Mpa-7-F-ak-cys	M1249	Mpa-B-5-ao-cys	M1297	Mpa-ao-3-D-cys	M1345	Mpa-1-ar-H-cys	M1393	Mpa-7-ap-D-cys	M1441	Mpa-B-W-5-cys	M1489	Mpa-aj-D-3-cys
M1154	Mpa-1-F-am-cys	M1202	Mpa-7-B-ar-cys	M1250	Mpa-C-5-as-cys	M1298	Mpa-W-3-I-cys	M1346	Mpa-1-an-J-cys	M1394	Mpa-7-ai-F-cys	M1442	Mpa-C-am-5-cys	M1490	Mpa-aq-H-3-cys
M1155	Mpa-1-D-as-cys	M1203	Mpa-7-G-al-cys	M1251	Mpa-H-5-ar-cys	M1299	Mpa-aj-3-E-cys	M1347	Mpa-1-as-E-cys	M1395	Mpa-7-aj-I-cys	M1443	Mpa-F-al-5-cys	M1491	Mpa-ai-C-3-cys
M1156	Mpa-1-Iap-cys	M1204	Mpa-7-A-aj-cys	M1252	Mpa-E-5-al-cys	M1300	Mpa-am-3-B-cys	M1348	Mpa-1-al-I-cys	M1396	Mpa-7-aq-C-cys	M1444	Mpa-I-ai-5-cys	M1492	Mpa-ao-B-3-cys
M1157	Mpa-1-J-ao-cys	M1205	Mpa-7-H-aq-cys	M1253	Mpa-J-5-ak-cys	M1301	Mpa-as-3-A-cys	M1349	Mpa-1-ak-C-cys	M1397	Mpa-7-an-E-cys	M1445	Mpa-G-ao-5-cys	M1493	Mpa-ap-J-3-cys
M1158	Mpa-1-G-W-cys	M1206	Mpa-7-J-ai-cys	M1254	Mpa-G-5-ap-cys	M1302	Mpa-al-3-G-cys	M1350	Mpa-1-ap-D-cys	M1398	Mpa-7-al-J-cys	M1446	Mpa-E-ar-5-cys	M1494	Mpa-as-I-3-cys
M1159	Mpa-1-H-ar-cys	M1207	Mpa-7-E-ap-cys	M1255	Mpa-F-5-am-cys	M1303	Mpa-aq-3-H-cys	M1351	Mpa-1-aj-F-cys	M1399	Mpa-7-ao-G-cys	M1447	Mpa-H-an-5-cys	M1495	Mpa-an-A-3-cys
M1160	Mpa-1-B-aq-cys	M1208	Mpa-7-I-an-cys	M1256	Mpa-A-5-an-cys	M1304	Mpa-ai-3-C-cys	M1352	Mpa-1-ai-B-cys	M1400	Mpa-7-as-H-cys	M1448	Mpa-J-aq-5-cys	M1496	Mpa-ar-E-3-cys
M1161	Mpa-2-D-ar-cys	M1209	Mpa-8-B-ao-cys	M1257	Mpa-D-6-ap-cys	M1305	Mpa-ak-4-C-cys	M1353	Mpa-2-am-H-cys	M1401	Mpa-8-ar-E-cys	M1449	Mpa-J-aj-6-cys	M1497	Mpa-ak-E-4-cys
M1162	Mpa-2-J-al-cys	M1210	Mpa-8-G-ak-cys	M1258	Mpa-F-6-ar-cys	M1306	Mpa-ap-4-J-cys	M1354	Mpa-2-aj-A-cys	M1402	Mpa-8-ap-G-cys	M1450	Mpa-H-am-6-cys	M1498	Mpa-ao-F-4-cys
M1163	Mpa-2-E-ak-cys	M1211	Mpa-8-J-ar-cys	M1259	Mpa-I-6-am-cys	M1307	Mpa-ai-4-D-cys	M1355	Mpa-2-an-E-cys	M1403	Mpa-8-aj-D-cys	M1451	Mpa-A-aq-6-cys	M1499	Mpa-ar-C-4-cys
M1164	Mpa-2-C-am-cys	M1212	Mpa-8-C-ap-cys	M1260	Mpa-B-6-al-cys	M1308	Mpa-ar-4-G-cys	M1356	Mpa-2-al-C-cys	M1404	Mpa-8-W-J-cys	M1452	Mpa-E-ai-6-cys	M1500	Mpa-al-A-4-cys
M1165	Mpa-2-F-ao-cys	M1213	Mpa-8-F-am-cys	M1261	Mpa-A-6-as-cys	M1309	Mpa-al-4-I-cys	M1357	Mpa-2-ao-F-cys	M1405	Mpa-8-ak-B-cys	M1453	Mpa-C-ak-6-cys	M1501	Mpa-ai-D-4-cys
M1166	Mpa-2-G-an-cys	M1214	Mpa-8-I-an-cys	M1262	Mpa-J-6-ai-cys	M1310	Mpa-am-4-F-cys	M1358	Mpa-2-aq-G-cys	M1406	Mpa-8-ao-C-cys	M1454	Mpa-I-ap-6-cys	M1502	Mpa-an-I-4-cys
M1167	Mpa-2-H-aq-cys	M1215	Mpa-8-A-al-cys	M1263	Mpa-G-6-aq-cys	M1311	Mpa-an-4-H-cys	M1359	Mpa-2-ar-D-cys	M1407	Mpa-8-aq-I-cys	M1455	Mpa-F-as-6-cys	M1503	Mpa-ap-J-4-cys
M1168	Mpa-2-I-ai-cys	M1216	Mpa-8-E-ai-cys	M1264	Mpa-E-6-an-cys	M1312	Mpa-ao-4-E-cys	M1360	Mpa-2-ai-I-cys	M1408	Mpa-8-as-A-cys	M1456	Mpa-D-al-6-cys	M1504	Mpa-am-G-4-cys
M1169	Mpa-3-H-ap-cys	M1217	Mpa-F-1-am-cys	M1265	Mpa-G-7-ar-cys	M1313	Mpa-al-5-A-cys	M1361	Mpa-3-ap-F-cys	M1409	Mpa-H-as-1-cys	M1457	Mpa-J-aj-7-cys	M1505	Mpa-ao-G-5-cys
M1170	Mpa-3-E-am-cys	M1218	Mpa-D-1-aq-cys	M1266	Mpa-C-7-ap-cys	M1314	Mpa-am-5-F-cys	M1362	Mpa-3-am-H-cys	M1410	Mpa-E-ao-1-cys	M1458	Mpa-G-ak-7-cys	M1506	Mpa-aj-I-5-cys
M1171	Mpa-3-B-ak-cys	M1219	Mpa-E-1-ai-cys	M1267	Mpa-I-7-ak-cys	M1315	Mpa-ai-5-D-cys	M1363	Mpa-3-ar-E-cys	M1411	Mpa-G-ar-1-cys	M1459	Mpa-B-ar-7-cys	M1507	Mpa-as-A-5-cys
M1172	Mpa-3-D-ar-cys	M1220	Mpa-I-1-ar-cys	M1268	Mpa-J-7-aj-cys	M1316	Mpa-ap-5-I-cys	M1364	Mpa-3-ak-I-cys	M1412	Mpa-F-aj-1-cys	M1460	Mpa-E-am-7-cys	M1508	Mpa-al-D-5-cys
M1173	Mpa-3-I-an-cys	M1221	Mpa-C-1-an-cys	M1269	Mpa-B-7-ai-cys	M1317	Mpa-an-5-H-cys	M1365	Mpa-3-ai-C-cys	M1413	Mpa-B-ap-1-cys	M1461	Mpa-C-aq-7-cys	M1509	Mpa-ai-E-5-cys
M1174	Mpa-3-J-ai-cys	M1222	Mpa-J-1-ap-cys	M1270	Mpa-D-7-an-cys	M1318	Mpa-ao-5-E-cys	M1366	Mpa-3-al-G-cys	M1414	Mpa-J-aq-1-cys	M1462	Mpa-A-as-7-cys	M1510	Mpa-ar-J-5-cys
M1175	Mpa-3-C-al-cys	M1223	Mpa-G-1-aj-cys	M1271	Mpa-E-7-as-cys	M1319	Mpa-ar-5-J-cys	M1367	Mpa-3-as-D-cys	M1415	Mpa-C-ak-1-cys	M1463	Mpa-W-an-7-cys	M1511	Mpa-ak-C-5-cys
M1176	Mpa-3-A-as-cys	M1224	Mpa-B-1-ao-cys	M1272	Mpa-A-7-ao-cys	M1320	Mpa-aj-5-B-cys	M1368	Mpa-3-ao-B-cys	M1416	Mpa-A-ai-1-cys	M1464	Mpa-H-ap-7-cys	M1512	Mpa-aq-F-5-cys
M1177	Mpa-4-G-ak-cys	M1225	Mpa-D-2-ap-cys	M1273	Mpa-H-8-aq-cys	M1321	Mpa-ar-6-I-cys	M1369	Mpa-4-ai-D-cys	M1417	Mpa-B-am-2-cys	M1465	Mpa-I-al-8-cys	M1513	Mpa-an-F-6-cys
M1178	Mpa-4-F-ai-cys	M1226	Mpa-W-2-ak-cys	M1274	Mpa-G-8-aj-cys	M1322	Mpa-ak-6-B-cys	M1370	Mpa-4-ap-B-cys	M1418	Mpa-A-ak-2-cys	M1466	Mpa-A-aq-8-cys	M1514	Mpa-aj-D-6-cys
M1179	Mpa-4-E-ar-cys	M1227	Mpa-C-2-aj-cys	M1275	Mpa-A-8-ak-cys	M1323	Mpa-aj-6-C-cys	M1371	Mpa-4-as-H-cys	M1419	Mpa-E-ap-2-cys	M1467	Mpa-C-ap-8-cys	M1515	Mpa-as-C-6-cys
M1180	Mpa-4-D-al-cys	M1228	Mpa-A-2-aq-cys	M1276	Mpa-B-8-ao-cys	M1324	Mpa-al-6-D-cys	M1372	Mpa-4-al-G-cys	M1420	Mpa-H-ar-2-cys	M1468	Mpa-F-aj-8-cys	M1516	Mpa-ao-I-6-cys
M1181	Mpa-4-J-aq-cys	M1229	Mpa-J-2-al-cys	M1277	Mpa-F-8-as-cys	M1325	Mpa-as-6-G-cys	M1373	Mpa-4-aq-F-cys	M1421	Mpa-F-ai-2-cys	M1469	Mpa-H-as-8-cys	M1517	Mpa-ap-W-6-cys
M1182	Mpa-4-I-an-cys	M1230	Mpa-H-2-ao-cys	M1278	Mpa-E-8-ap-cys	M1326	Mpa-ao-6-J-cys	M1374	Mpa-4-aj-W-cys	M1422	Mpa-G-an-2-cys	M1470	Mpa-B-ao-8-cys	M1518	Mpa-ar-E-6-cys
M1183	Mpa-4-C-am-cys	M1231	Mpa-I-2-ar-cys	M1279	Mpa-C-8-ai-cys	M1327	Mpa-aq-6-A-cys	M1375	Mpa-4-an-A-cys	M1423	Mpa-I-al-2-cys	M1471	Mpa-J-an-8-cys	M1519	Mpa-ak-J-6-cys
M1184	Mpa-4-B-aj-cys	M1232	Mpa-G-2-as-cys	M1280	Mpa-I-8-an-cys	M1328	Mpa-ap-6-E-cys	M1376	Mpa-4-ar-J-cys	M1424	Mpa-D-aj-2-cys	M1472	Mpa-E-ak-8-cys	M1520	Mpa-ai-G-6-cys
M1185	Mpa-5-H-an-cys	M1233	Mpa-C-3-ap-cys	M1281	Mpa-am-1-A-cys	M1329	Mpa-ak-7-A-cys	M1377	Mpa-5-al-E-cys	M1425	Mpa-G-ao-3-cys	M1473	Mpa-aj-J-1-cys	M1521	Mpa-ak-F-7-cys
M1186	Mpa-5-G-aj-cys	M1234	Mpa-E-3-ak-cys	M1282	Mpa-an-1-H-cys	M1330	Mpa-ap-7-E-cys	M1378	Mpa-5-an-A-cys	M1426	Mpa-I-am-3-cys	M1474	Mpa-ai-A-1-cys	M1522	Mpa-ai-B-7-cys
M1187	Mpa-5-A-ai-cys	M1235	Mpa-G-3-aj-cys	M1283	Mpa-ak-1-J-cys	M1331	Mpa-an-7-G-cys	M1379	Mpa-5-aq-J-cys	M1427	Mpa-C-ai-3-cys	M1475	Mpa-ar-I-1-cys	M1523	Mpa-ap-D-7-cys
M1188	Mpa-5-F-ar-cys	M1236	Mpa-D-3-ai-cys	M1284	Mpa-aq-1-B-cys	M1332	Mpa-ar-7-F-cys	M1380	Mpa-5-ao-B-cys	M1428	Mpa-J-aq-3-cys	M1476	Mpa-al-H-1-cys	M1524	Mpa-al-J-7-cys
M1189	Mpa-5-C-ak-cys	M1237	Mpa-A-3-an-cys	M1285	Mpa-aj-1-C-cys	M1333	Mpa-ao-7-D-cys	M1381	Mpa-5-am-H-cys	M1429	Mpa-F-an-3-cys	M1477	Mpa-ap-D-1-cys	M1525	Mpa-ao-E-7-cys
M1190	Mpa-5-D-ao-cys	M1238	Mpa-I-3-al-cys	M1286	Mpa-ap-1-I-cys	M1334	Mpa-al-7-I-cys	M1382	Mpa-5-ai-F-cys	M1430	Mpa-E-al-3-cys	M1478	Mpa-as-C-1-cys	M1526	Mpa-aq-C-7-cys
M1191	Mpa-5-B-aq-cys	M1239	Mpa-H-3-am-cys	M1287	Mpa-as-1-E-cys	M1335	Mpa-as-7-B-cys	M1383	Mpa-5-ak-C-cys	M1431	Mpa-D-ar-3-cys	M1479	Mpa-ak-E-1-cys	M1527	Mpa-as-I-7-cys
M1192	Mpa-5-E-ap-cys	M1240	Mpa-F-3-as-cys	M1288	Mpa-ar-1-F-cys	M1336	Mpa-ai-7-H-cys	M1384	Mpa-5-ar-I-cys	M1432	Mpa-A-ak-3-cys	M1480	Mpa-am-F-1-cys	M1528	Mpa-an-A-7-cys
M1193	Mpa-6-H-ak-cys	M1241	Mpa-H-4-am-cys	M1289	Mpa-ap-2-H-cys	M1337	Mpa-am-8-A-cys	M1385	Mpa-6-aq-C-cys	M1433	Mpa-B-ak-4-cys	M1481	Mpa-am-A-2-cys	M1529	Mpa-aq-G-8-cys
M1194	Mpa-6-B-an-cys	M1242	Mpa-B-4-ao-cys	M1290	Mpa-ao-2-B-cys	M1338	Mpa-ap-8-E-cys	M1386	Mpa-6-ao-F-cys	M1434	Mpa-F-ap-4-cys	M1482	Mpa-ao-B-2-cys	M1530	Mpa-ap-E-8-cys
M1195	Mpa-6-D-ao-cys	M1243	Mpa-F-4-ai-cys	M1291	Mpa-aq-2-I-cys	M1339	Mpa-ak-8-H-cys	M1387	Mpa-6-an-E-cys	M1435	Mpa-C-ar-4-cys	M1483	Mpa-aq-J-2-cys	M1531	Mpa-aj-J-8-cys
M1196	Mpa-6-I-al-cys	M1244	Mpa-E-4-ar-cys	M1292	Mpa-am-2-D-cys	M1340	Mpa-al-8-D-cys	M1388	Mpa-6-al-I-cys	M1436	Mpa-A-an-4-cys	M1484	Mpa-as-C-2-cys	M1532	Mpa-ao-F-8-cys
M1197	Mpa-6-J-ar-cys	M1245	Mpa-I-4-ak-cys	M1293	Mpa-al-2-E-cys	M1341	Mpa-ar-8-F-cys	M1389	Mpa-6-ar-J-cys	M1437	Mpa-E-ai-4-cys	M1485	Mpa-an-G-2-cys	M1533	Mpa-am-A-8-cys
M1198	Mpa-6-A-aj-cys	M1246	Mpa-C-4-an-cys	M1294	Mpa-ai-2-C-cys	M1342	Mpa-an-8-G-cys	M1390	Mpa-6-as-G-cys	M1438	Mpa-G-as-4-cys	M1486	Mpa-ak-F-2-cys	M1534	Mpa-al-H-8-cys
M1199	Mpa-6-G-ap-cys	M1247	Mpa-A-4-aq-cys	M1295	Mpa-as-2-F-cys	M1343	Mpa-as-8-C-cys	M1391	Mpa-6-ak-H-cys	M1439	Mpa-J-aq-4-cys	M1487	Mpa-al-I-2-cys	M1535	Mpa-ar-B-8-cys
M1200	Mpa-6-F-ai-cys	M1248	Mpa-J-4-as-cys	M1296	Mpa-aj-2-G-cys	M1344	Mpa-aj-8-J-cys	M1392	Mpa-6-ai-B-cys	M1440	Mpa-D-aj-4-cys	M1488	Mpa-ap-D-2-cys	M1536	Mpa-as-C-8-cys

Figure S4 continued. (c) Library 3 (M769 to M1152). (d) Library 4 (M1153 to M1536).

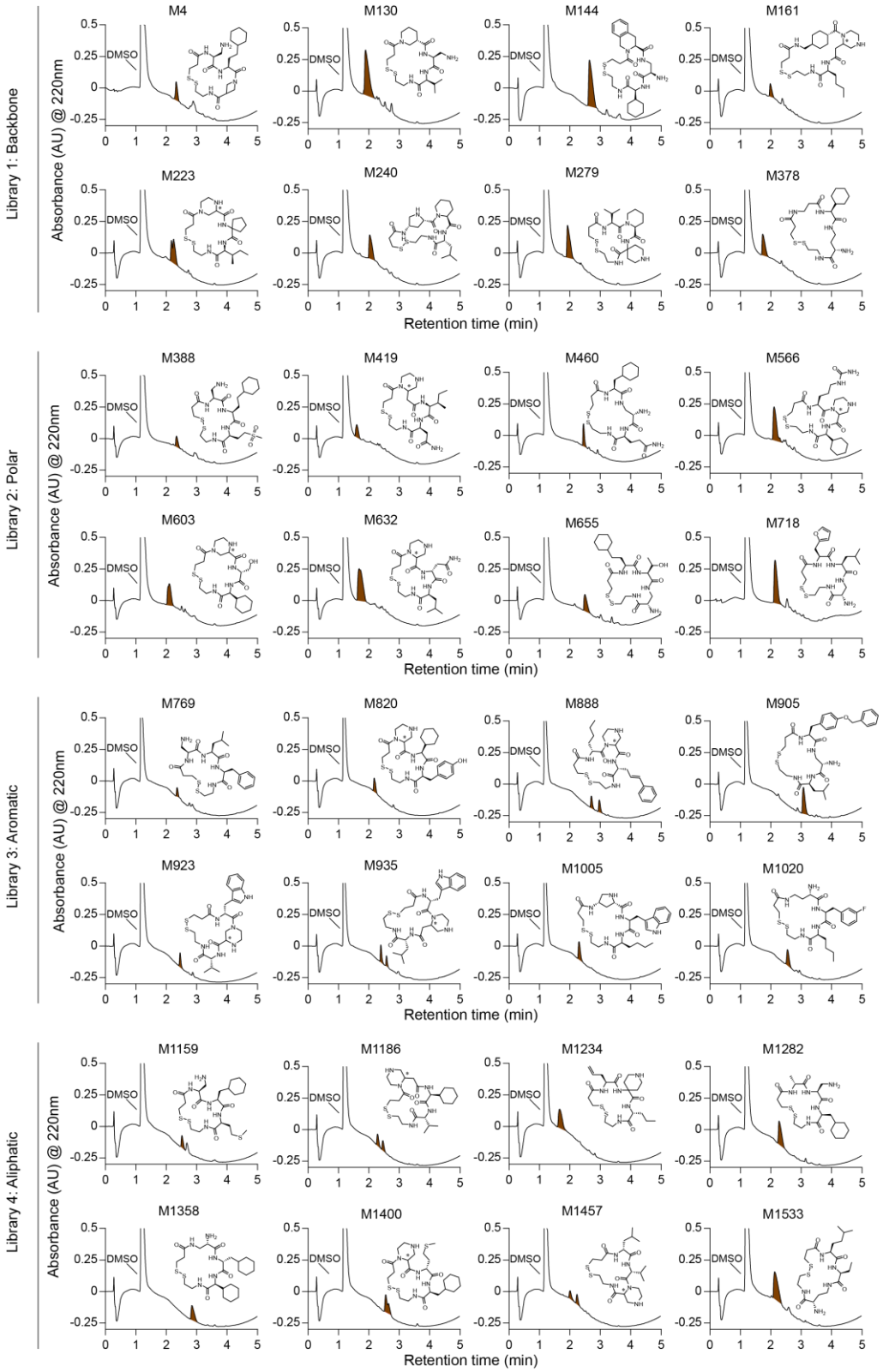


Figure S5. Quality control of the four library. LCMS of 8 randomly selected macrocycles per library. Highlighted in brown are the peaks corresponding to the desired macrocycles, which are all represented for each spectra. Above is written the macrocycle number (M#). Their corresponding sequences are written in Table S1 and in Figure S4.

Table S1. Quality control of the four library. Purity (%) was considered as the sum of desired product, isomer and linear. All UV spectrums used for the analysis of the macrocycle purities, as well as their structures, are presented in Figure S5.

	M#	Macrocycle sequence	Desired product (%)	Isomer (%)	Linear (%)	+ tBu (%)	Dimer (%)	Dimer + tBu (%)	Dimer + 2*tBu (%)	Other (%)	Purity (%)
Library 1: Backbone	M4	Mpa-1-I-d-cys	49.5	0	0	0	35.9	4.2	2.4	8	49.5
	M130	Mpa-f-1-B-cys	86.3	0	0	2.8	4.6	0	0	6.3	86.3
	M144	Mpa-g-2-G-cys	88.7	0	0	4.9	0	0.6	5.8	0	88.7
	M161	Mpa-i-5-F-cys	67.5	0	0	2.6	21.9	1.8	0	6.2	67.5
	M223	Mpa-4-k-D-cys	30.1	58	0	5.7	0	0	0	6.2	88.1
	M240	Mpa-6-b-J-cys	81.2	0	0	3.5	5.9	0	0	9.4	81.2
	M279	Mpa-B-b-3-cys	90.8	0	0	0.3	4.7	0	1.4	2.8	90.8
	M378	Mpa-j-G-8-cys	85.6	0	0	0	7.4	3	1.3	2.7	85.6
Library 2: Polar	M388	Mpa-1-I-t-cys	58.9	0	0	0	31.8	0	5.2	4.1	58.9
	M419	Mpa-5-D-n-cys	76.7	0	4	8.9	2.3	0	1.4	6.7	80.7
	M460	Mpa-H-2-o-cys	84	0	0	2.7	6	3	0	4.3	84.0
	M566	Mpa-s-7-G-cys	84.4	0	3	1.1	6.6	3.6	0	1.3	87.4
	M603	Mpa-4-u-G-cys	77.5	0	0	6.9	0	7.3	5.6	2.7	77.5
	M632	Mpa-7-v-C-cys	90.5	0	0	3.2	3.7	0.8	0.6	1.2	90.5
	M655	Mpa-I-m-2-cys	67.3	0	0	9.8	2.6	1.8	5.1	13.4	67.3
	M718	Mpa-r-C-2-cys	85.7	0	0	1.2	9.2	1.5	0	2.4	85.7
Library 3: Aromatic	M769	Mpa-1-C-w-cys	52.6	0	4.8	0	42.6	0	0	0	57.4
	M820	Mpa-7-G-z-cys	87.6	0	0	11.7	0	0	0	0.7	87.6
	M888	Mpa-F-7-ac-cys	45.2	49.9	0	0	2.5	0	0	2.4	95.1
	M905	Mpa-ad-2-C-cys	80.1	0	0	1.6	7.3	0	0	11	80.1
	M923	Mpa-x-4-B-cys	80.6	0	0	18	1.4	0	0	0	80.6
	M935	Mpa-ae-5-J-cys	52.7	30.6	0	10.4	0	0	4.3	2	83.3
	M1005	Mpa-6-x-F-cys	87.4	0	0	8	0	0	0	4.6	87.4
	M1020	Mpa-8-ah-F-cys	78.9	0	0	0	16.8	0	0	4.3	78.9
Library 4: Aliphatic	M1159	Mpa-1-H-ar-cys	53	0	33.6	0	0	0	0	13.4	86.6
	M1186	Mpa-5-G-aj-cys	51.9	42.3	0	0	0	0	0	5.8	94.2
	M1234	Mpa-E-3-ak-cys	91.2	0	0	0	2.1	0	3.1	3.6	91.2
	M1282	Mpa-an-1-H-cys	87.8	0	0	0.4	2.7	0	2.4	6.7	87.8
	M1358	Mpa-2-aq-G-cys	92.1	0	0	1.6	0	0	0	6.3	92.1
	M1400	Mpa-7-as-H-cys	55.8	38.8	0	0	0	0	0	5.4	94.6
	M1457	Mpa-J-aj-7-cys	44.7	42.9	0	0	5.5	0	1.4	5.5	87.6
	M1533	Mpa-am-A-8-cys	88.1	0	0	0	6	0	2.5	3.4	88.1

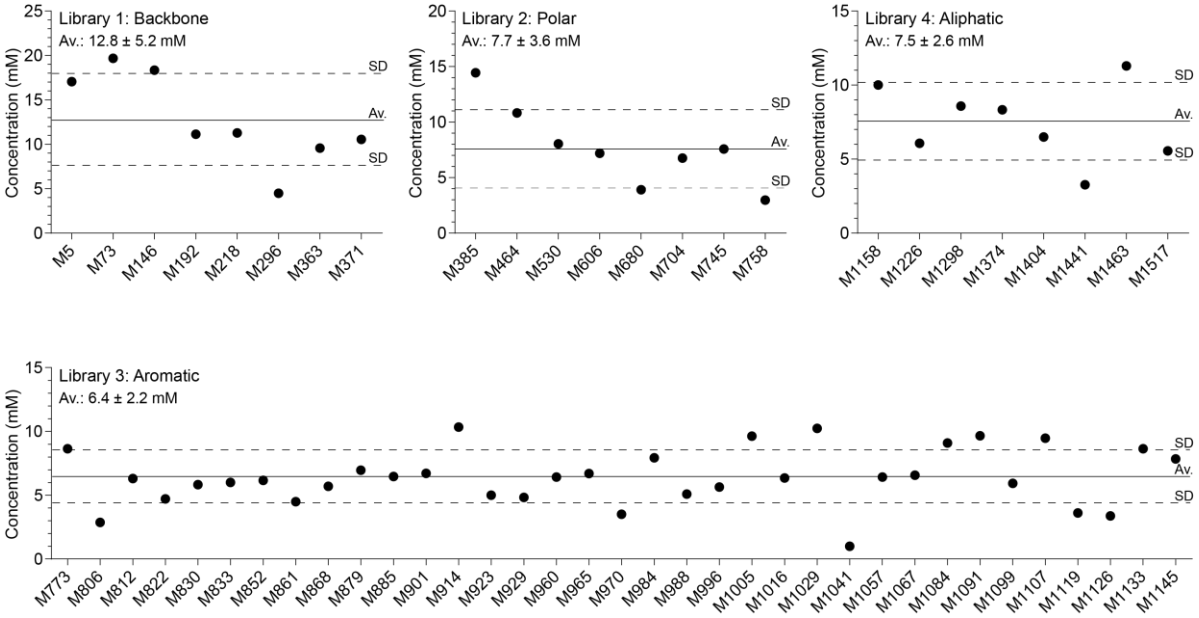


Figure S6. Library concentration determination. Quantification was performed by nanodrop at 280 nm. For libraries 1,2 and 4, one amino acid (other than the diamino acids) of 8 random macrocycles were replaced by tryptophan ("W"). For the library 3, tryptophan was one of the amino acids under the code "x" and the corresponding macrocycles containing it are presented. All corresponding macrocycles sequences are written in Figure S4.

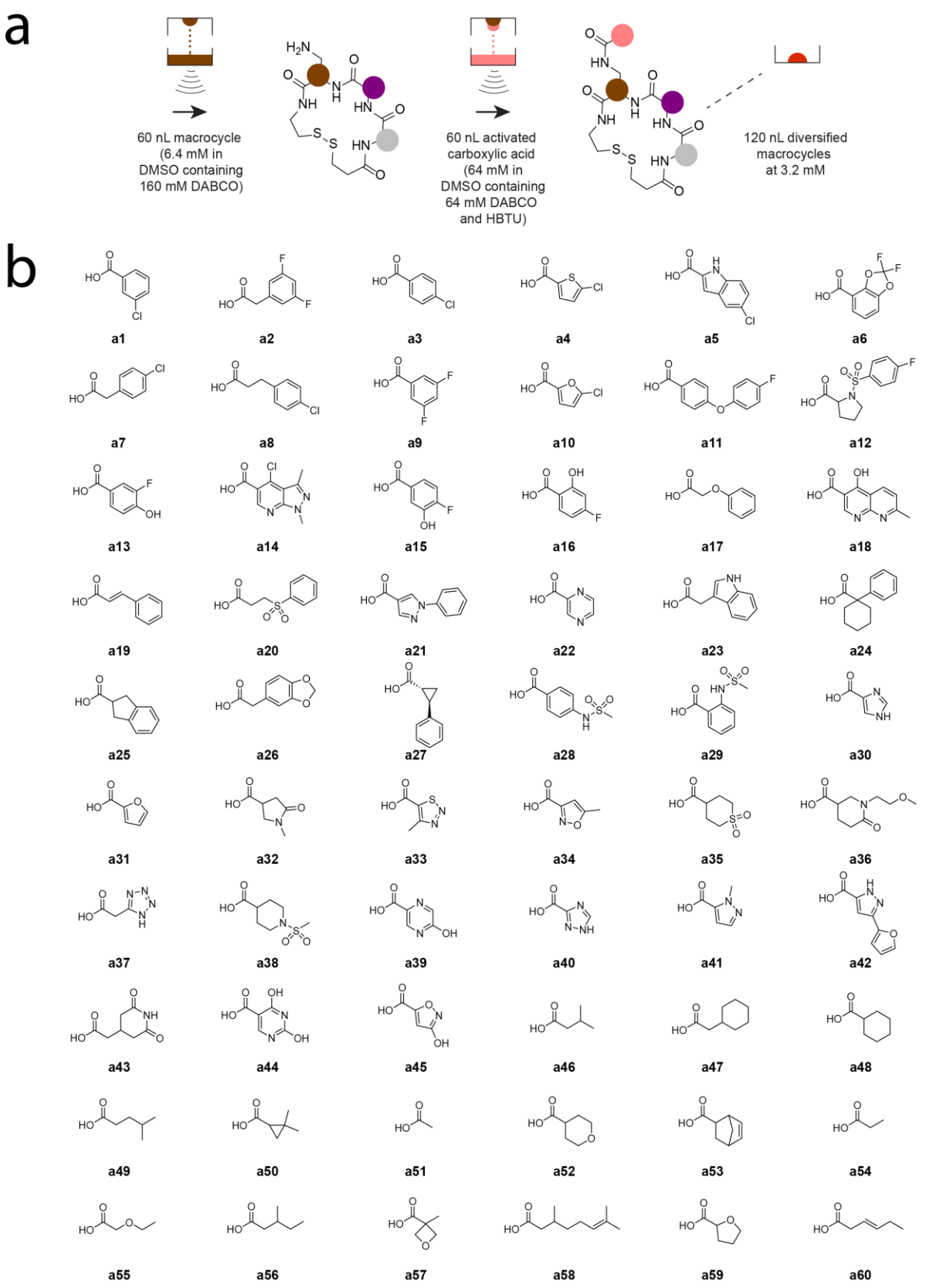


Figure S7. Peripheral diversification of macrocycles. (a) Strategy to diversify the macrocycle free amine by coupling activated carboxylic acids at picomole scale directly in assay plates. (b) All acids used to diversify the four libraries.

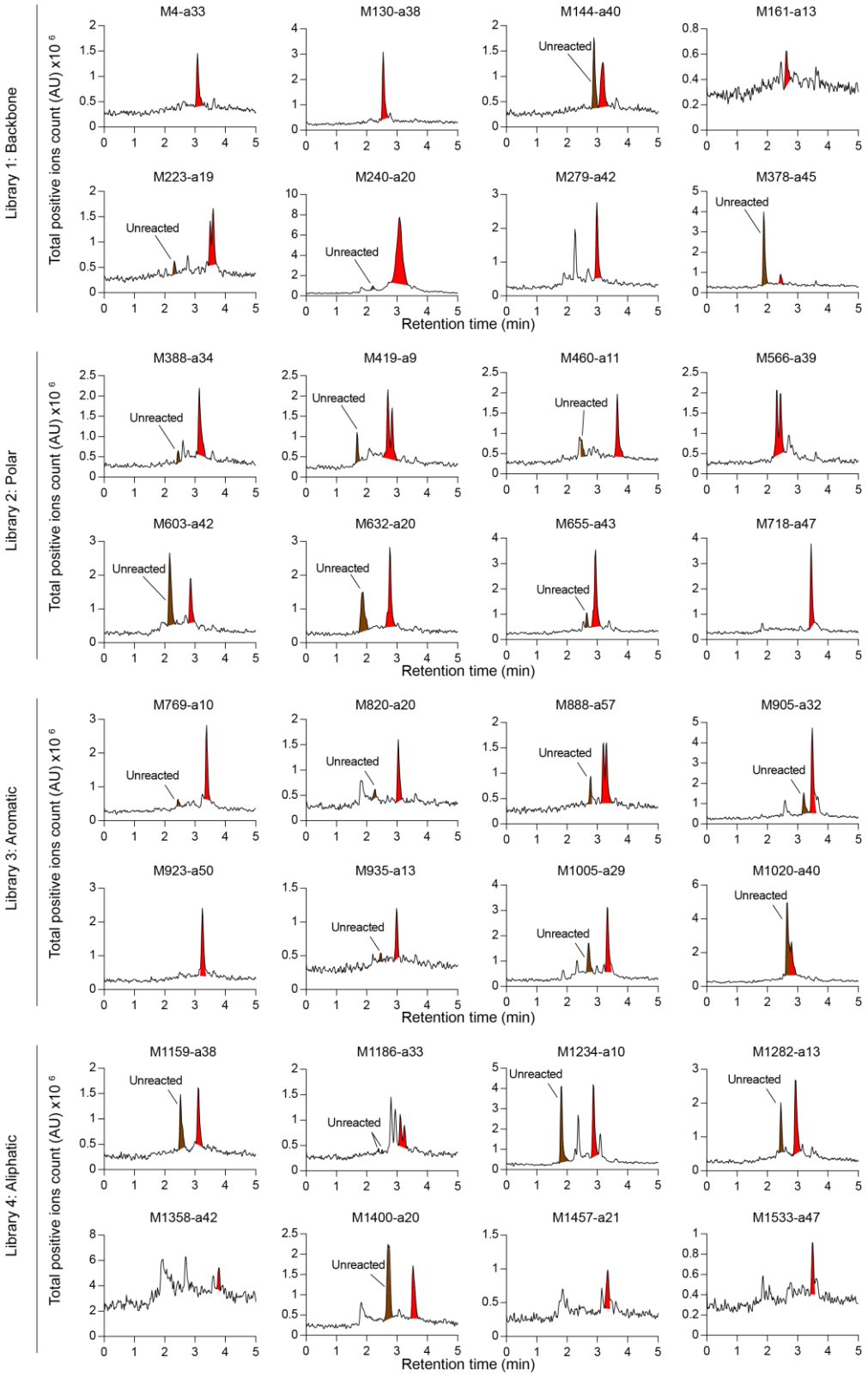


Figure S8. Total ions count of peripheral diversification of macrocycle in screening. The same macrocycles analysed in Figure S5/ Table S1 undergoing peripheral diversification through coupling of a random carboxylic acids to a free amine are presented. Desired acylated macrocycles peaks are highlighted in red. In brown are the unreacted species when it could be represented. Above is written the macrocycle number (M#) and the corresponding acid used for diversification (a#).

Table S2. Qualitative analysis of peripheral diversification of macrocycle in screening. Peaks of acylated macrocycles usually being too small at the concentration of analysis, only a qualitative analysis of the reaction can be performed by doing a ratio comprising the TIC (+) area of the desired specie over the sum of both unreacted and desired species.

[(a)] Small peaks were observable by TIC, which will affect the reliability of the result.

	M#	Macrocycle sequence	a#	TIC (+) Unreacted (AU)	TIC (+) Desired (AU)	Ratio desired/ (unreacted+desired)*100 (%)
MLibrary 1: Backbone	M4	Mpa-1-l-d-cys	a33	133223	2209482	94
	M130	Mpa-f-1-B-cys	a38	381715	5777368	94
	M144	Mpa-g-2-G-cys	a40	3390052	2304220	40
	M161	Mpa-i-5-F-cys	a13	36560	466526	93[(a)]
	M223	Mpa-4-k-D-cys	a19	60754	3883864	98
	M240	Mpa-6-b-J-cys	a20	435268	41572021	91
	M279	Mpa-B-b-3-cys	a42	206357	5726336	97
	M378	Mpa-j-G-8-cys	a45	10951633	332068	3
Library 2: Polar	M388	Mpa-1-l-t-cys	a34	748304	426625	86
	M419	Mpa-5-D-n-cys	a9	1704370	2988181	64
	M460	Mpa-H-2-o-cys	a11	625085	3899590	86
	M566	Mpa-s-7-G-cys	a39	309014	6180280	95
	M603	Mpa-4-u-G-cys	a42	6008851	3522189	37
	M632	Mpa-7-v-C-cys	a20	2117561	5915787	74
	M655	Mpa-l-m-2-cys	a43	1337024	6936828	84
	M718	Mpa-r-C-2-cys	a47	194899	7193262	97
Library 3: Aromatic	M769	Mpa-1-C-w-cys	a10	554907	5537353	91
	M820	Mpa-7-G-z-cys	a20	376649	2829952	88
	M888	Mpa-F-7-ac-cys	a57	1048123	4212220	80
	M905	Mpa-ad-2-C-cys	a32	1259079	9977405	89
	M923	Mpa-x-4-B-cys	a50	229635	4319444	95
	M935	Mpa-ae-5-J-cys	a13	50678	1042976	95
	M1005	Mpa-6-x-F-cys	a29	953814	9317007	91
	M1020	Mpa-8-ah-F-cys	a40	12877394	4404412	25
Library 4: Aliphatic	M1159	Mpa-1-H-ar-cys	a38	2909657	2751814	49
	M1186	Mpa-5-G-aj-cys	a33	177997	2149955	92
	M1234	Mpa-E-3-ak-cys	a10	11934767	7176334	38
	M1282	Mpa-an-1-H-cys	a13	3735829	4554541	55
	M1358	Mpa-2-aq-G-cys	a42	70596	325906	82[(a)]
	M1400	Mpa-7-as-H-cys	a20	5628841	3149578	36
	M1457	Mpa-J-aj-7-cys	a21	15109	1903939	99[(a)]
	M1533	Mpa-am-A-8-cys	a47	18616	1262412	99[(a)]

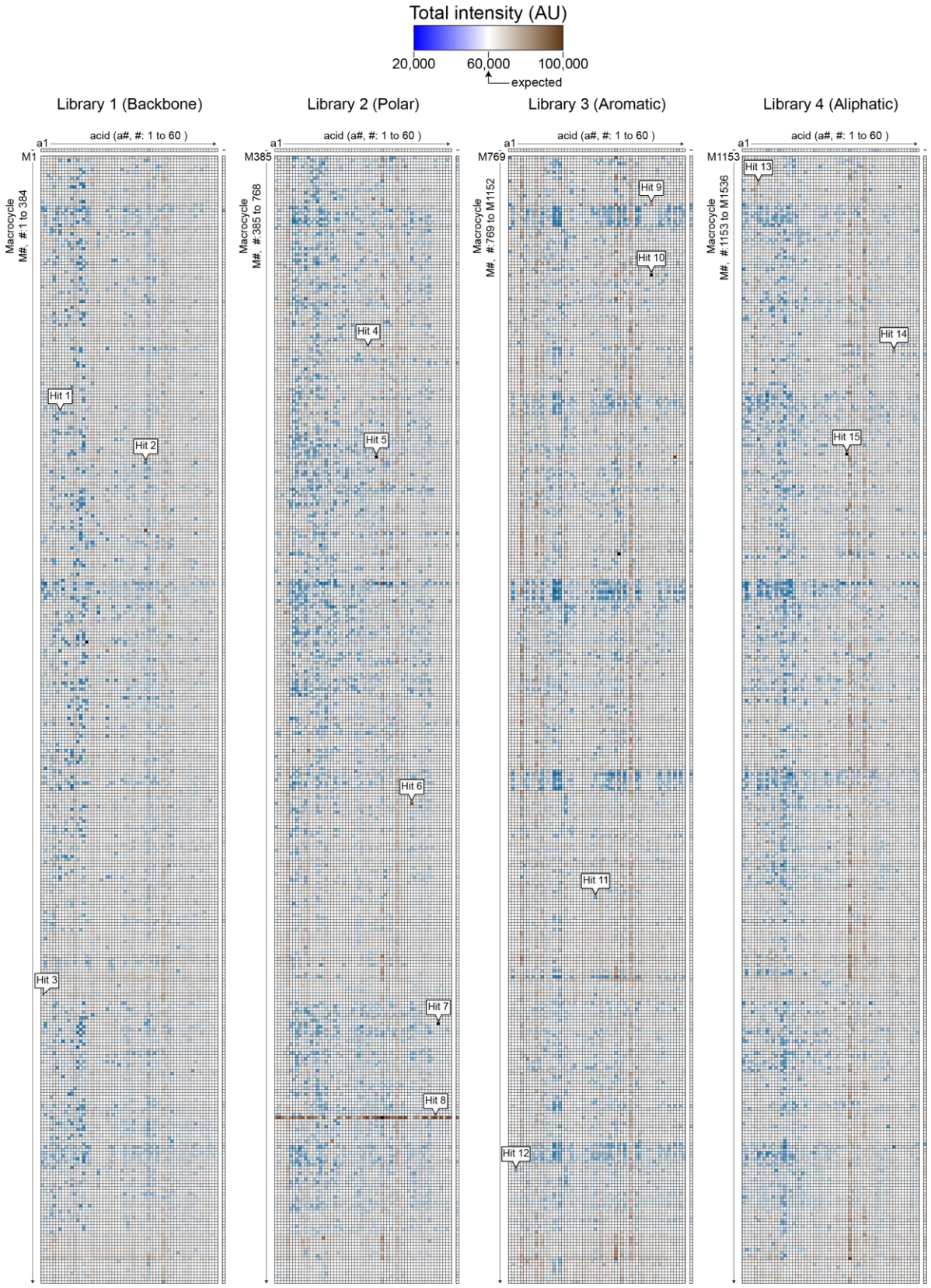


Figure S9. Total polarized fluorescent intensity of screen. Raw data of the total polarized fluorescent intensity observed in the screen for the test of all the 92,160 macrocycles. In black are values that were deleted for coloring as there were too high (around 5 x intensity above normal). Just the first macrocycle and acid are written for each library. Control without macrocycles or acids are represented with "-" horizontally and vertically, respectively.

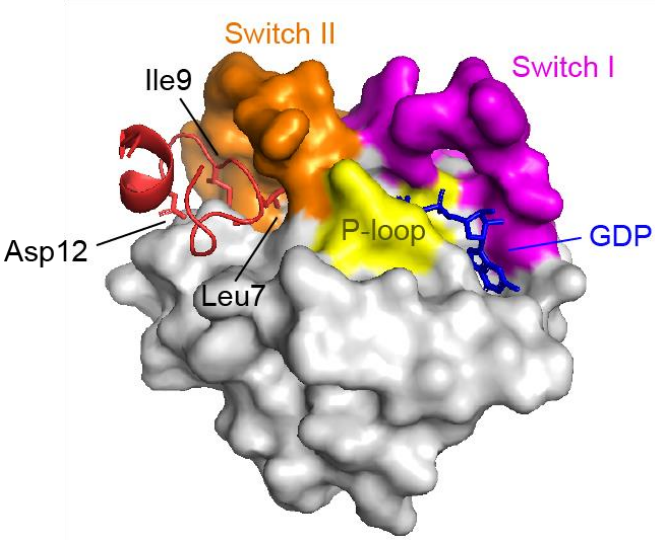


Figure S10. Structure of KRpep-2d bound to KRAS(G12D). X-ray structure (PDB entry 5XCO) of cyclic peptide (KRpep-2d in red, aa 1-19) bound to KRAS(G12D) (aa 1-169) loaded with GDP (in blue). Key amino acids involved in the binding of the peptide to KRAS(G12D) are highlighted. Key regions of KRAS(G12D) are highlighted; in yellow is the P-loop (aa 10-17), in purple the Switch I region (aa 25-40), in orange the Switch II region (aa 60-74).

5.8 References

(1) A.T., B.; D., X.; C.J., D. Inhibition of Ras for cancer treatment: The search continues. *Future Med. Chem.* **2011**.

(2) Santos, E.; Nebreda, A. R. Structural and functional properties of ras proteins . *FASEB J.* **1989**.

(3) Chang, E. H.; Gonda, M. A.; Ellis, R. W.; Scolnick, E. M.; Lowy, D. R. Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses. *Proc. Natl. Acad. Sci. U. S. A.* **1982**.

(4) Cooper, G. M. Cellular transforming genes. *Science.* **1982**.

(5) Malumbres, M.; Barbacid, M. RAS oncogenes: The first 30 years. *Nature Reviews Cancer*. **2003**.

(6) Thein, K. Z.; Biter, A. B.; Hong, D. S. Therapeutics Targeting Mutant KRAS. *Annual Review of Medicine*. **2021**.

(7) Eser, S.; Schnieke, A.; Schneider, G.; Saur, D. Oncogenic KRAS signalling in pancreatic cancer. *British Journal of Cancer*. **2014**.

(8) Prior, I. A.; Lewis, P. D.; Mattos, C. A comprehensive survey of ras mutations in cancer. *Cancer Research*. **2012**.

(9) Kessler, D.; Gmachl, M.; Mantoulidis, A.; Martin, L. J.; Zoephel, A.; Mayer, M.; Gollner, A.; Covini, D.; Fischer, S.; Gerstberger, T.; Gmaschitz, T.; Goodwin, C.; Greb, P.; Häring, D.; Hela, W.; Hoffmann, J.; Karolyi-Oezguer, J.; Knesl, P.; Kornigg, S.; Koegl, M.; Kousek, R.; Lamarre, L.; Moser, F.; Munico-Martinez, S.; Peinsipp, C.; Phan, J.; Rinnenthal, J.; Sai, J.; Salamon, C.; Scherbantin, Y.; Schipany, K.; Schnitzer, R.; Schrenk, A.; Sharps, B.; Siszler, G.; Sun, Q.; Waterson, A.; Wolkerstorfer, B.; Zeeb, M.; Pearson, M.; Fesik, S. W.; McConnell, D. B. Drugging an undruggable pocket on KRAS. *Proc. Natl. Acad. Sci. U. S. A.* **2019**.

(10) Wang, X.; Allen, S.; Blake, J. F.; Bowcut, V.; Briere, D. M.; Calinisan, A.; Dahlke, J. R.; Fell, J. B.; Fischer, J. P.; Gunn, R. J.; Hallin, J.; Laguer, J.; Lawson, J. D.; Medwid, J.; Newhouse, B.; Nguyen, P.; O'Leary, J. M.; Olson, P.; Pajk, S.; Rahbaek, L.; Rodriguez, M.; Smith, C. R.; Tang, T. P.; Thomas, N. C.; Vanderpool, D.; Vigers, G. P.; Christensen, J. G.; Marx, M. A. Identification of MRTX1133, a Noncovalent, Potent, and Selective KRASG12D Inhibitor. *J. Med. Chem.* **2021**.

(11) Patgiri, A.; Yadav, K. K.; Arora, P. S.; Bar-Sagi, D. An orthosteric inhibitor of the Ras-

Sos interaction. *Nat. Chem. Biol.* **2011**.

(12) Trinh, T. B.; Upadhyaya, P.; Qian, Z.; Pei, D. Discovery of a Direct Ras Inhibitor by Screening a Combinatorial Library of Cell-Permeable Bicyclic Peptides. *ACS Comb. Sci.* **2016**.

(13) Sakamoto, K.; Kamada, Y.; Sameshima, T.; Yaguchi, M.; Niida, A.; Sasaki, S.; Miwa, M.; Ohkubo, S.; Sakamoto, J. ichi; Kamaura, M.; Cho, N.; Tani, A. K-Ras(G12D)-selective inhibitory peptides generated by random peptide T7 phage display technology. *Biochem. Biophys. Res. Commun.* **2017**.

(14) Sakamoto, K.; Masutani, T.; Hirokawa, T. Generation of KS-58 as the first K-Ras(G12D)-inhibitory peptide presenting anti-cancer activity in vivo. *Sci. Rep.* **2020**.

(15) Kauke, M. J.; Traxlmayr, M. W.; Parker, J. A.; Kiefer, J. D.; Knihtila, R.; McGee, J.; Verdine, G.; Mattos, C.; Dane Wittrup, K. An engineered protein antagonist of K-Ras/B-Raf interaction. *Sci. Rep.* **2017**.

(16) Janes, M. R.; Zhang, J.; Li, L. S.; Hansen, R.; Peters, U.; Guo, X.; Chen, Y.; Babbar, A.; Firdaus, S. J.; Darjania, L.; Feng, J.; Chen, J. H.; Li, S.; Li, S.; Long, Y. O.; Thach, C.; Liu, Y.; Zarieh, A.; Ely, T.; Kucharski, J. M.; Kessler, L. V.; Wu, T.; Yu, K.; Wang, Y.; Yao, Y.; Deng, X.; Zarrinkar, P. P.; Brehmer, D.; Dhanak, D.; Lorenzi, M. V.; Hu-Lowe, D.; Patricelli, M. P.; Ren, P.; Liu, Y. Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **2018**.

(17) Iqbal, E. S.; Hartman, M. C. T. Shaping molecular diversity. *Nat. Chem.* **2018**.

(18) Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Lee, M. D.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W. D.; Duvall, J. R.; Kihlberg, J. Structural and conformational determinants of macrocycle cell permeability. *Nat. Chem. Biol.* **2016**.

(19) Sogabe, S.; Kamada, Y.; Miwa, M.; Niida, A.; Sameshima, T.; Kamaura, M.; Yonemori, K.; Sasaki, S.; Sakamoto, J.; Sakamoto, K. Crystal Structure of a Human K-Ras G12D Mutant in Complex with GDP and the Cyclic Inhibitory Peptide KRpep-2d. *ACS Med. Chem. Lett.* **2017**.

(20) Niida, A.; Sasaki, S.; Yonemori, K.; Sameshima, T.; Yaguchi, M.; Asami, T.; Sakamoto, K.; Kamaura, M. Investigation of the structural requirements of K-Ras(G12D) selective inhibitory peptide KRpep-2d using alanine scans and cysteine bridging. *Bioorganic Med. Chem. Lett.* **2017**.

(21) Kimple, A.; Yasgar, A.; Hughes, M.; Jadhav, A.; Willard, F.; Muller, R.; Austin, C.;

Inglese, J.; Ibeanu, G.; Siderovski, D.; Simeonov, A. A High Throughput Fluorescence Polarization Assay for Inhibitors of the GoLoco Motif/G-alpha Interaction. *Comb. Chem. High Throughput Screen.* **2008**.

(22) Gray, C. H.; Konczal, J.; Mezna, M.; Ismail, S.; Bower, J.; Drysdale, M. A fully automated procedure for the parallel, multidimensional purification and nucleotide loading of the human GTPases KRas, Rac1 and RalB. *Protein Expr. Purif.* **2017**.

6. General conclusions

PPIs are promising drug targets due to their key role in numerous pathologies. However, the discovery of PPI modulators has been challenging, especially when the interaction of interest is located intracellularly and drugs thus require being small for efficiently crossing cell membranes. Macrocycles are a promising drug format as they have shown to be efficient at modulating protein-protein interactions while having the potential to be cell-permeable and even orally available. However, the generation of macrocycle-based ligands to new PPI targets is overall hampered by the lack of large libraries of macrocyclic compounds needed for HTS. To fill this gap, I chose in my Ph.D. to develop and apply novel strategies for the ultra-parallel synthesis and screening of large macrocyclic compound libraries. I have built on a new type of library synthesis and screening principle recently introduced in our laboratory in which combinatorial macrocyclic compound libraries are synthesized in microwell plates through mixing and reacting building blocks, and where crude macrocyclic products are screened directly without prior purification.

In a first project, I aimed at increasing the throughput of the approach by applying acoustic dispensing, a technology that was used for the high-throughput screening of compound libraries but not for the combinatorial synthesis of macrocycle libraries. In the first project, I also evaluated a recently developed strategy for combinatorially synthesizing macrocyclic compound libraries, and I established a screening assay for the PPI target MDM2:p53. In a second project, I developed a TR-FRET-based assay for the high-throughput screening of macrocyclic compound libraries against the important PPI target IL-23:IL-23R. In a third project, I established a screening assay focusing on KRAS:effectors PPIs, and I performed a screen using a large macrocycle library that was synthesized using acoustic dispensing.

The key achievement of my Ph.D. project are i) the demonstration that acoustic dispensing is highly suited for the synthesis and screening of combinatorial macrocycle libraries, and ii) the establishment of a robust high-throughput screening assay for the important PPI target IL-23:IL-23R. In the following two sub-chapters, I am describing the two achievements in detail, as well as their potential impact for drug development.

Macrocycle library synthesis at a picomole scale using acoustic droplet ejection

Before starting my Ph.D. project, macrocycle libraries developed in our laboratory were generated by transferring microliter volumes of reagents to wells of 384-microwell plates, using disposable pipetting tips. Based on work by colleagues in the lab, who showed that macrocycle

synthesis reagents can efficiently be transferred by acoustic dispending in nanoliter droplets at high speed, I started to use the technology for synthesizing a combinatorial macrocycle library. An initial test strategy was based on the diversification of *m* linear peptides containing an N-terminal bromoacetamide and a C-terminal thiol protected by a disulfide bond with a peptidic tag, with *n* primary amines, followed by the reduction of the disulfide bond and cyclization with *o* linkers yielding *m* × *n* × *o* macrocycles. As proof of concept, I tested a target-focused library of 2,700 macrocycles against the MDM2:p53 PPI, which yielded low to sub-micromolar binders. While many of the macrocyclization reactions were not quantitative and side-products were formed, the approach allowed still to identify potent inhibitors of the MDM2:p53 PPI. The use of acoustic dispensing enabled reducing the synthesis scale by 250-fold, therefore drastically increasing the synthesis throughout as the combinatorial diversification could be performed directly in assay plates.

The successful generation and screening of macrocycle libraries at a picomole scale in my work encouraged our laboratory to apply acoustic dispensing to also other macrocycle synthesis strategies. Given the partially incomplete reactions and side-products observed in my project, new macrocycle synthesis strategies were followed that involved fewer reaction steps and less complex reactions. One such combinatorial strategy is based on the combinatorial cyclization of *m* di-thiol peptide with *n* bis-electrophile linkers, which I used in my second project, and for which I applied acoustic dispensing too. Another important new macrocycle synthesis strategy of the lab is based on the diversification of *m* disulfide-cyclized peptides with *n* carboxylic acids yielding *m* × *n* macrocycles, which I used in my third project to produce 92,160 macrocycles at a picomole scale.

Altogether, the use of acoustic dispensing for combinatorial synthesis of macrocycle libraries at picomole scale proved to be a robust approach that is now used routinely by several members of our laboratory. Macrocycle synthesis by acoustic dispensing and assaying directly in assay plates allows the rapid generation and screening of large libraries of individual untagged macrocycles, making the molecule class of macrocycles more accessible for ligands and, hopefully drug development.

Development of an assay for the IL-23:IL23R PPI

In several projects of my Ph.D., I aimed at generating bioassays that were suited for screening crude macrocycle libraries against PPIs. The first assay I developed was based on fluorescence polarization to detect inhibitors of the MDM2:p53 interaction. FP was chosen due to the simplicity of setting up the assay that was reported in literature and that required only a fluorescent probe and the target protein. Upon binding of the fluorescent peptide to the protein,

an increase in polarization can be monitored. The inhibition event can then be observed by a decrease in polarization. This assay proved robust to all reagents used for the macrocycles assembly and allowed clear identification of the most active macrocycles. Additionally, no false hits were observed as each hits identified in the screen could be reproduced.

Based on the success for MDM2:p53, I chose fluorescence polarization also for developing assays for the IL-23:IL-23R and KRAS(G12D):effectors PPIs. In both cases, after some optimization, the assays seemed robust. However, in the screening of synthesized macrocycles libraries, numerous artifacts leading to higher/lower polarization could be observed. These artifacts often correlated with important changes in total fluorescence intensity (increase or decrease), suggesting potential effects from compounds on the fluorescent peptide probe. These effects proved to be much more pronounced with the assay for IL-23:IL-23R, where identification of hits became impossible. Additionally, for KRAS(G12D), all hits observed in the primary screen did not reproduce. For all these reasons, it became clear that another assay format was needed for the high-throughput screening of PPI inhibitors.

I subsequently used TR-FRET as assay format as it is an established method in the pharma industry for screening PPI inhibitors, and because this assay type tolerates high concentrations of detergent and blocking agents such as BSA, which we thought would be important to prevent the artifacts in the FP-based assays. A fluorescent donor is used to transfer energy to an acceptor for a prolonged time, allowing measuring emission of the acceptor after a time delay, which can prevent numerous interferences. I used TR-FRET to monitor the protein-protein interaction between IL-23 and IL-23R. The signal proved much more robust than the one obtained with the fluorescence polarization assay. A higher signal to background was obtained as well as lower fluctuation of the controls. In addition, this assay format proved compatible with the combinatorial synthesis approach used as no effect on the signals was observed. Even when testing some compounds previously identified as problematic with the fluorescence polarization assay, no artifacts were observed. Screening a library of 3,072 cyclic peptides, the assay allowed simple identification of active hits. While the most potent compounds identified were side-products of macrocycle synthesis reactions, the screening assay is robust and is currently used for the screening of larger macrocycle libraries. Based on the successes with TR-FRET as assay format in my work, TR-FRET has recently been applied by other members of our laboratory for establishing screening assays of other PPIs.

7. CV

Gontran SANGOUARD

+41 79 888 03 12 | gosangouard@gmail.com | Linkedin
10 avenue de la gare, 1022 Chavannes-près-Renens, Switzerland

KEY STRENGTHS

- **Creative problem solving:** bringing innovative solutions to unforeseen problems while focusing on decreasing general cost and operation time
- **Qualitative/quantitative analytical skills:** Rational decision making based on thorough data treatment and analysis
- **Teamwork:** leading several projects with international teams from various backgrounds in scientific or business-oriented context

EDUCATION

2018-now
- **PhD in chemical biology** | EPFL, Switzerland
  -**Thesis title**: Generating macrocyclic inhibitors of protein-protein interactions

2015-2017
- **MSc in molecular and biological chemistry** | EPFL, Switzerland
  -**Minor** in management of technology and entrepreneurship (MTE)

2012-2015
- **BSc in chemistry and chemical engineering** | EPFL, Switzerland

CORE EXPERIENCE

2018-now (expected March 2022)
- **Doctoral assistant** | Laboratory of protein and peptide therapeutics (LPPT), École polytechnique fédérale de Lausanne (EPFL), Switzerland
  -Optimized and validated patentable methodology that decreased raw material consumption by 250 times in a compound library synthesis strategy (1 pending patent)
  -Discovered several compounds against cancer and autoimmune disease
  -Managed 3 long-term projects in parallel (1 published, 1 to be published)
  -Collaborated on 4 other projects (1 published, 2 to be published)
  -Supervised up to 3 students simultaneously for 6 months
  -Assisted in teaching (biochemistry courses, lab training for students)
  -Handled instruments purchase (contact several suppliers, organize comparative testing, negotiate prices)
  -Responsible for lab key instruments (involved training and maintenance)

ADDITIONAL EXPERIENCE

2020 (during PhD)
- **Inossuisse Business Concept** | EPFL, Switzerland
  -Start-up launcher workshop
  -Presented concept selected during the pitching contest
  -Involved interviews of field experts, pitching, IP, business, accounting and presentation training

2019 (during PhD)
- **Bench2Biz** | Zurich, Switzerland
  -Start-up launcher workshop
  -Submitted concept selected (involved a presentation and an interview)
  -Developed a business plan/roadmap to translate a start-up concept into a potential business with the help of a team of dedicated industrial experts
  -Involved pitching, IP training, interviews and business courses

2019 (during PhD)
- **Talk at Life Sciences Switzerland** | Zurich, Switzerland
  -Chemical biology short talk at scientific conference

2017

- **Internship** | Chemical glycobiology laboratory, Simon Fraser University (SFU), Canada
  -Improved the reliability of a test by 60%
  -Worked on 2 projects in parallel (1 published)

2016-2017

- **Minor in management of technology and entrepreneurship (MTE)** | EPFL, Switzerland
  -One entire semester in a non-scientific field
  -Training in business, entrepreneurship, design thinking, negotiation techniques, presentation skills, managerial and financial accounting

2015-2017

- **Teaching assistant** | ISIC, EPFL, Switzerland
  -Led general and organic chemistry teaching sessions in French and in English
  -Presented in front of 600 students
  -Taught private courses in general and organic chemistry

COMPETENCIES

Soft

Polyvalent, teamwork spirit, hard-worker, analytical reasoning, problem solving, project management, communication

Informatics

Microsoft Office (Word, Excel, PowerPoint), Adobe Illustrator, GraphPad, MestreNova, ChemDraw, Mendeley, analysis and development (basics in VBA)

Technical

- **Biological:** P1 lab working experience, PCR, plasmid engineering, protein expression (bacteria and mammalian cells), protein purification (SEC, affinity column, AKTA), SDS-PAGE, Electrophoresis
- **Chemical:** Peptide chemistry (SPPS), peptide purification, semi-prep HPLC, lyophilization, speedvac, rotavap
- **Analytical:** HPLC, LCMS, 1H-NMR, 13C-NMR
- **High-Throughput Screening:** Assay development (FP, TR-FRET), Screening platform (ACCESS), acoustic droplet ejection (ECHO), Bulk dispenser (Certus, Biotek dispenser), plate reader (Tecan, PheraStar, Perkin Elmer), 96-, 384- and 1536-well plates handling

Languages

French: Mother tongue **English:** Fluent (C2-level) **Spanish:** Basics (B1-level)

PATENT

2021 (pending)

- **Co-inventor:** Methods for synthesis and screening of large combinatorial macrocyclic compound libraries (ER patent application Nr 1174036.0)

LEISURE

Fitness (3 times a week), historical model making, 20 years winter sport practice (snowboarding and skiing), 15 years musical practice (music theory, drums and piano), Karaté (brown belt)

PERSONAL INFORMATION

- French
- 12th October 1994
- Single
- Driving license (Permis B)

PUBLICATIONS

2021 (first author)

- **Sangouard, G.** et al. Picomole-Scale Synthesis and Screening of Macrocyclic Compound Libraries by Acoustic Liquid Transfer. Angew. Chemie (2021).

2021 (co-author)

- Ceballos, J., Grinhagena, E., **Sangouard, G.**, Heinis, C. & Waser, J. Cys–Cys and Cys–Lys Stapling of Unprotected Peptides Enabled by Hypervalent Iodine Reagents. Angew. Chemie (2021).

2020 (co-author)

- Alteen, M. G. et al. A Direct Fluorescent Activity Assay for Glycosyltransferases Enables Convenient High-Throughput Screening: Application to O -GlcNAc Transferase . Angew. Chemie (2020).