New applications of dynamic combinatorial chemistry to medicinal chemistry

Alwin Mathijs Hartman

rijksuniversiteit groningen

UNIVERSITÄT DES SAARLANDES

HIPS HELMHOLTZ Institut für Pharmazeutische Forschung Saarland

New applications of dynamic combinatorial chemistry to medicinal chemistry

Alwin Mathijs Hartman
Ph.D. Thesis
University of Groningen, The Netherlands
Universität des Saarlandes, Saarbrücken, Germany
Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) — Helmholtz Centre for Infection Research (HZI), Department of Drug Design and Optimization, Saarbrücken, Germany

The research described in this thesis was carried out at the Stratingh Institute for Chemistry, University of Groningen, The Netherlands, at the Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Germany and at the Faculty of Natural Sciences and Technology of Saarland University, Germany.
In compliance with the requirements of the Graduate School of Science and Engineering of the Faculty of Science and Engineering, University of Groningen, The Netherlands, as well as with requirements of the Faculty of Natural Sciences and Technology of Saarland University.

This work was financially supported by the University of Groningen and the Netherlands Organization for Scientific Research (VIDI grant: 723.014.008).

Netherlands Organisation for Scientific Research

Cover was designed by Lysette Hartman.

Printing of this thesis was generously supported by the University of Groningen and the Graduate School of Science and Engineering.

Printed by Ipskamp Drukkers BV, Enschede, The Netherlands.

ISBN: 978-94-034-1970-1 (printed)
978-94-034-1969-5 (digital)

rijksuniversiteit
groningen

UNIVERSITÄT
DES
SAARLANDES

New applications of dynamic combinatorial chemistry to medicinal chemistry

PhD thesis

to obtain the degree of PhD of the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

and

to obtain the degree of Doctor of Natural Sciences at the
Faculty of Natural Sciences and Technology of
Saarland University

by

Alwin Mathijs Hartman

Groningen/ Saarbrücken
2019

Supervisors
Prof. Dr. A. J. Minnaard
University of Groningen (UG)

Prof. Dr. A. K. H. Hirsch
Saarland University (UdS)

Assessment committee (UG)
Prof. F. J. Dekker
Prof. R. Müller
Prof. S. Otto
Prof. M. D. Witte

Assessment committee (UdS)
Prof. A. K. H. Hirsch
Prof. A. J. Minnaard
Prof. M. D. Witte

Tag des Kolloquiums:	28.11.2019
Dekan:	Prof. Dr. G. Kickelbick
Berichterstatter:	Prof. Dr. A.K.H. Hirsch
	Prof. Dr. A.J. Minnaard
	Prof. Dr. M.D. Witte
Vorsitzender:	Prof. Dr. A. Luzhetskyy
Akad. Mitarbeiter:	Dr. S. Boettcher
Weitere Mitglieder:	Prof. Dr. W.R. Browne
	Prof. Dr. S. Otto
	Prof. Dr. F.J. Dekker

"Luck is what happens when preparation meets opportunity."

Seneca, roman philosopher 5 b.c.

" Man ist nie fertig mit der Wissenschaft, höchstens ist die Wissenschaft fertig mit einem."

Martin Empting

Aan mijn ouders

Summary

Applying dynamic combinatorial chemistry (DCC) to medicinal chemistry projects can be a helpful strategy for finding starting points in the drug-discovery process. As relevant drug target, 14-3-3 proteins play a role in several diseases and many biological processes. Proteins of this family engage in protein-protein interactions (PPIs), and can up-or down-regulate their binding partner's activity. Another family of relevant targets are glucansucrases, which are important enzymes in the initiation and development of cariogenic dental biofilms, commonly known as dental plaque. In the last two chapters, endothiapepsin was used for protein-templated DCC (ptDCC). Endothiapepsin belongs to the family of the aspartic proteases, which are involved in for example the maturation of the HIV virus particle.

Throughout this thesis, we focus on applying DCC to various projects. The main achievements are: 1) the description of the in-house protocol of DCC, in which aspects like solubility of building blocks and products, protein stability and more need to be taken in to account, 2) the application of acylhydrazone-based DCC to two targets, a (PPI)-target and a glucansucrase, 3) the identification of small-molecules, which stabilise PPIs of 14-3-3/ synaptopodin, 4) expanding the reaction toolbox of ptDCC by two additional reactions: nitrone and thiazolidine formation.

Zusammenfassung

Die Verwendung dynamisch kombinatorischer Chemie (DCC) in medizinisch-chemischen Projekten kann eine sehr hilfreiche Strategie sein, um Anknüpfungspunkte für die Wirkstoffentdeckung zu finden. 14-3-3 Proteine spielen eine Rolle in verschiedenen Krankheiten und vielen biologischen Prozessen. Proteine dieser Familie beteiligen sich an Protein-Protein-Interaktionen (PPIs) und können die Aktivität der Bindungspartner sowohl hoch- als auch herabregulieren. Eine andere Familie relevanter Targets sind die Glukansucrasen, welche wichtige Enzyme in der Initiierung und Entwicklung von kariogenen dentalen Biofilmen, allgemein bekannt als Plaque, sind. In den letzten beiden Kapiteln wurde Endothiapepsin für Protein-vermittelte DCC (ptDCC) verwendet. Endothiapepsin gehört zur Familie der Aspartylproteasen, welche zum Beispiel an der Reifung des HIV Viruspartikels beteiligt sind.

Im Verlauf dieser Arbeit fokussieren wir uns auf die Anwendung von DCC in verschiedenen Projekten. Die Hauptleistungen sind: 1) die Beschreibung des hausinternen DCC-Protokolls, in welchem Aspekte wie Löslichkeit von Bausteinen und Produkten, Proteinstabilität und weiteres wichtige zu beachten sind, 2) die Anwendung von Acylhydrazon-basierter DCC auf zwei Targets, eine Glukansucrase und ein PPI-Target, 3) die Identifikation kleiner Moleküle, die PPIs von 14-3-3/ Synaptopodin stabilisieren, 4) die Erweiterung des Reaktionsspielraums der ptDCC durch zwei zusätzliche Reaktionen: Nitron- und Thiazolidinbildung.

Table of Contents

Chapter 1

Introduction to dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCC) is a powerful tool to identify bioactive compounds. This efficient technique allows the target to select its own binders and circumvents the need for synthesis and biochemical evaluation of all individual derivatives. An ever-increasing number of publications report the use of DCC on biologically relevant targets. The work here complements reviews by focusing on the experimental protocol and giving detailed examples of essential steps and factors that need to be considered, such as protein stability, buffer composition and cosolvents.

This chapter has been published as a review article:

A. M. Hartman, R. M. Gierse, A. K. H. Hirsch, *Eur. J. Org. Chem.* **2019**, 3581–3590.

A. M. Hartman and R. M. Gierse contributed equally to the work in this chapter.

Section 1.4 was taken from the submitted review article:

P. Kirsch, A. M. Hartman, A. K. H. Hirsch, M. Empting,

A. M. Hartman was involved in writing the DCC example and editing the review.

1.1 Introduction

Since its dawn more than two decades ago, combinatorial chemistry approaches[1–5] have developed into target-directed dynamic combinatorial chemistry (tdDCC) and have matured as a hit- identification tool. There has been an increasing number of work published in this niche of supramolecular chemistry.[6–11] A growing number of groups have shown the general applicability and scope of tdDCC for the identification of modulators of targets.[6,12–20] tdDCC refers to general pharmacologically relevant targets which next to proteins also include DNA and RNA, whereas protein-templated DCC (ptDCC) only refers to proteins. Several reviews and book chapters on tdDCC have been published in recent years.[21–23] This chapter covers our work on ptDCC and provides the key features of our protocol, explaining the essential steps in designing a successful ptDCC experiment.

Carefully chosen building blocks are connected in a reversible manner *via* covalent or noncovalent bonds to form a dynamic combinatorial library (DCL) (Figure 1). Biocompatibility, pH dependence, temperature, solubility and stability of the components are important factors, which should be taken into account. The ideal DCLs do not require cosolvents, however, it can occur that the formed products have a lower solubility than the building blocks and in order to keep all compounds in solution, a cosolvent such as dimethyl sulfoxide (DMSO) is commonly used. Precipitation of DCL components could lead to an undesired shift in the equilibrium. By contrast, a desired shift of the equilibrium can be obtained by the addition of an external stimulus, such as a protein target. There are in general two different approaches that can be followed in ptDCC: 'adaptive DCC', in which the target is present during the formation of the DCL and 'pre-equilibrated DCC', in which the target is added after the DCL is established. An advantage of pre-equilibrated DCC is that the exchange chemistry can be applied in conditions which are not tolerated by the protein. A disadvantage is that the screening step is performed under static conditions and no amplification effects can be observed since the protein does not alter the equilibrium.

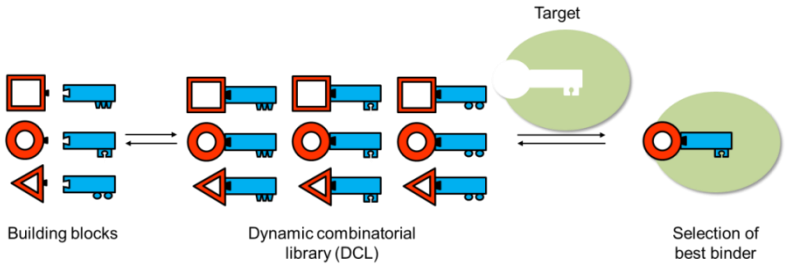


Figure 1. Schematic representation of target-directed dynamic combinatorial chemistry.

In ptDCC, the member(s) of DCLs, which bind best will be amplified, leading to an increase in their concentration compared to a control reaction without the external stimulus. These binders can then be further evaluated for their biochemical properties.

To enable a comparative analysis of DCLs, a blank reaction, without the target, should be run concurrent with a templated reaction. Another approach of DCC is non-comparative, in which the hits can be analysed in complex with the target or after being released from the target. There are different techniques that can be used to analyse the DCLs: liquid and size-exclusion chromatography coupled to mass spectrometry, NMR spectroscopy, fluorescence spectroscopy and X-ray crystallography. Figure 2 illustrates the comparative approach versus the non-comparative approach, which can be adopted in DCC. The reaction mixture can be 'frozen', in order to prevent the library from re-equilibrating during the analysis. In the case of acylhydrazone chemistry, this can be achieved by an increase in pH. Denaturation by heat, addition of a solvent or (ultra-fast) centrifugation ensures that all binders are released from the protein before analysis.

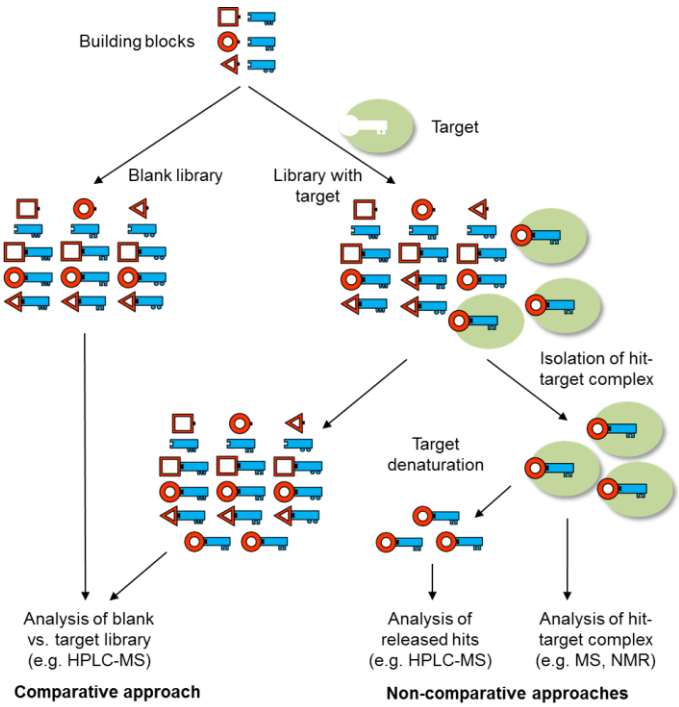


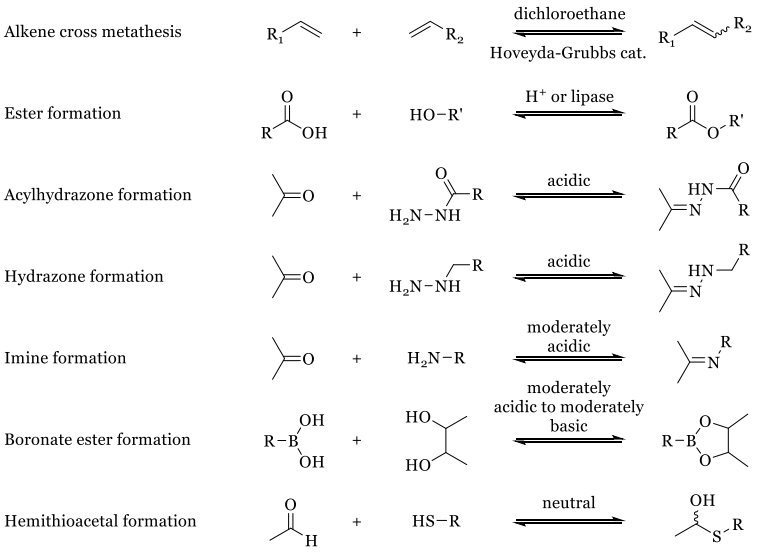
Figure 2. DCC approaches: comparative and non-comparative. In the comparative approach the library in presence of a target is compared to the library in absence of the target. In the non-comparative approaches, the hit–target complexes will be separated from the mixture and analysed as a complex or as released hits. The figure was adapted from Frei et al.[21]

1.1.1 Reversible reactions suitable for DCC

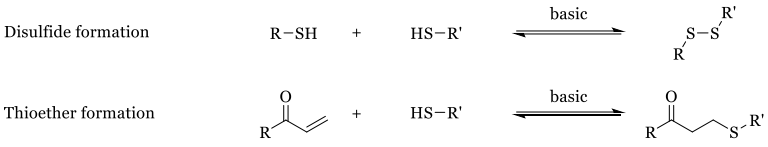
Only a limited number of reversible reactions have been used thus far, they are summarised in Scheme 1. One of the most frequently used reactions is the (acyl-)hydrazone formation, which combines ketone or aldehyde building blocks with (acyl-)hydrazides. This condensation reaction can take place in water, making it biocompatible.[24] The synthesis of the building blocks is generally straightforward or they may be commercially available.

At physiological conditions, neutral pH and room temperature, acylhydrazone formation and exchange are relatively slow. At acidic pH, the equilibrium is reached rapidly. However, Greaney and coworkers have shown that the pH dependence can be influenced by the addition of a nucleophilic catalyst. They were able to reach equilibrium reasonably fast at a comparatively high pH of 6.2 by using aniline, as a nucleophilic catalyst.[12] Previously Dawson and coworkers have shown that aniline could serve as a catalyst for acylhydrazone formation and oxime ligation.[25,26] Derivatives of aniline, which bear substituents at the aryl ring, are even more effective catalysts.[27]

The acylhydrazone linkage is reversible under acidic conditions and stable against hydrolysis at physiological pH values, allowing for the 'freezing' of the reversible reaction upon increasing the pH.[24]



Scheme 1. Reversible reactions used in target-directed DCC to identify bioactive compounds. Scheme was adapted from Van der Vlag and Hirsch.[23]



Scheme 1 continued. Reversible reactions used in target-directed DCC to identify bioactive compounds. Scheme was adapted from Van der Vlag and Hirsch.[23]

An overview of studies published over the past five years in the field of ptDCC is given in Table 1. It must be noted that much more work has been published applying DCC for the formation of diverse libraries in the drug-discovery process. For example the coupling of DCC to DNA-encoded libraries, creating so called DNA-encoded dynamic combinatorial chemical libraries (EDCCLs). Iminobiotin and homotetrameric streptavidin were used as a model system to identify a bidentate protein/ligand interaction. The addition of an external stimulus, for example a target protein, can shift the thermodynamic equilibrium and hence a DNA amplification can be observed after sequencing.[28]

Table 1. Protein-templated DCC studies reported over the past five years, in which a target was used as a template to influence the equilibrium. Therefore, only articles using an adaptive approach are listed, pre-equilibrated DCC examples are omitted.[29–31] The table is adapted from Frei et al and complemented.[21]

Target	Reversible reaction	Analysis	Library size	Equili-bration time	Method applied for affinity measurement	Best affinity	Ref.
Wt Tau RNA	Disulfide	HPLC-MS and NMR	21	2 days	Fluorescence titration	EC_{50} = 70 nM	[32]
HIV FSS RNA	Disulfide	MS	12	4 days	n.a.	n.a.	[33]
Vascular endothelial growth factor receptor (VEGFR) 2	Imine	HRMS	297	24 h	In vitro activity against cancer cell lines	IC_{50} = 2.4 µM	[34]
Endothiapepsin	Acylhydrazone	HPLC-MS	90	20 h	Inhibition assay	IC_{50} = 54.5 nM K_i = 25.4 nM	[35]
FimH	Acylhydrazone	HPLC	8	3 days	SPR	K_D = 273 nM	[36]
UDP-galacto-pyranose mutase	Acylhydrazone	HPLC	11	24 h	Fluorescence-based assay and MIC	K_D = 3 µM MIC = 26 µg mL^{-1}	[37]

Myeloperoxidase (MPO)	Hydrazone	Activity assay	6	n.a.	in vivo *a*ctivity assay	IC_{50} = 79 nM	[38]
*ec*FabH	Acylhydrazone	^{19}F-NMR	5	12 h	Enzymatic assay	IC_{50} = 3 mM	[39]
Multi-protein strategy on AlkB oxygenases: FTO, ALKBH3 and ALKBH5	Acylhydrazone	DSF and HPLC	10	5 h	HPLC-based demethylase and DSF assays	IC_{50} = 2.6 µM	[40]
*Trypanosoma cruzi bromodomain-containing (Tc*BDF3)	Acylhydrazone	HPLC-MS	30	n.a.	DSF	IC_{50} = 13–23 µM	[41]

DSF = differential scanning fluorimetry, HPLC = high-performance liquid chromatography, IC_{50} = half maximal inhibitory concentration, ITC = isothermal titration calorimetry, K_D = dissociation constant, K_i = inhibition constant, MIC = minimum inhibitory concentration, MS = mass spectrometry, n.a. = not available, NMR = nuclear magnetic resonance, SPR = surface plasmon resonance, T_m = thermal shift.

1.2 A closer look on the templating protein

To obtain meaningful results from DCC experiments, the quality of the input template is critical. As the equilibrium of the library shifts by the templating effect of the added protein sample, it should consist of the target protein as close to its native state as possible. The quantity of the used template depends on the protein target, there are reported successful DCC projects with 0.1 to 1.5 equivalents of protein. [29,42] DCC experiments are also possible with a mixture of proteins, but a well-defined sample eases up downstream data analysis and reduces the number of false positives for the desired target.[40] The condition of the protein sample depends on various variables. For DCC experiments the purity, concentration, tertiary and quaternary structure of the protein, additives and contaminations, as well as the pH-value are of particular importance. During the experiment, which can take up to several days, protein degradation and precipitation could occur. The tests described herein should give an overview and help to choose suitable experimental conditions to plan new DCC experiments. In the next paragraphs, we will briefly discuss the influence of those factors and suitable analytical methods to monitor them.

1.2.1 Purity

In the case of a mixed or impure protein sample, there might be several templated reactions proceeding in parallel. It is impossible to differentiate between a small fraction of the sample showing a strong template effect and a large fraction of the

protein pool showing only a weak amplification of a binder. This will result in overlapping data, which are difficult to analyse, and may result in false positives. We therefore recommend starting with the highest protein purity available.

1.2.2 Stability

Not only the initial state, but also the stability of the templating protein during the reaction should be checked by preliminary tests before conducting a DCC experiment. The time span over which a DCC experiment, pre-equilibrated or adaptive, is monitored can vary. It depends on the reaction rate and concentration and should ideally be monitored until the library reaches an equilibrium state. Usually, the DCL reaches a new equilibrium within the first few days, depending on the reversible reaction and conditions (Table 1). However, if the protein is stable for longer periods of time, longer equilibration times are possible, for example up to 20 days for the very stable protease endothiapepsin (see Section 1.2.4).[24]

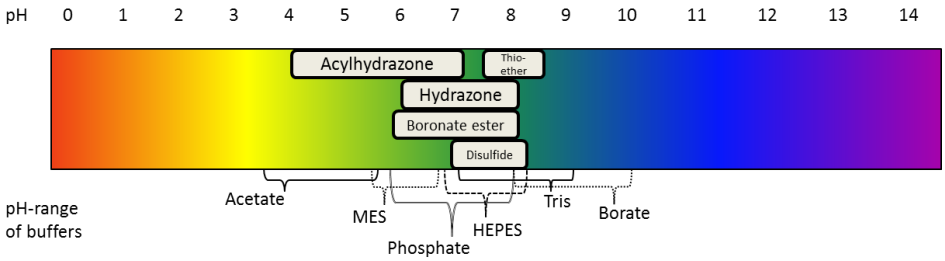
It is important that the protein is not precipitating or degrading during the experiment. Precipitation of the protein will remove the template from the solution. Denaturation of the template will lead to entirely new templates, which would affect the equilibrium state of the DCL. This can lead to random and irreproducible amplification of compounds by the unordered protein and a decrease of initially already amplified best binders of the native template. If the protein target is labile, it is therefore necessary to follow the reaction over time to identify the temporary, templated equilibrium of the DCC library. In this, compounds amplified by the native state of the template can be found.

Eventually, after prolonged incubation time, nearly every protein will degrade and, by this, change the equilibrium of the DCL again. Compounds amplified in this step can be ignored, as they were not templated by the native protein. Observation of the DCC experiment for longer timeframes than the template integrity can be guaranteed is therefore of no use.

1.2.3 Buffer and pH

When choosing a buffer for DCC experiments, several different requirements have to be met. Attention should be paid to possible side reactions with the DCL or chelation effects. For example, Tris-buffer could form imines with aldehyde building blocks, which might influence the formation of the DCL. Some stabilization of the protein is beneficial, but strong interactions of the buffer with the target protein should also be avoided, for instance, a phosphate buffer for a phosphate binding protein. The phosphate could compete with possible binders; possible effects of competition are discussed in more detail in Section 1.2.5. So far, in most cases common buffer systems have been used, which are shown in

Table 2 and Scheme 2. The choice of buffer is, however, not limited to the established systems.



Scheme 2. Example of possible buffers and the pH ranges of reactions used in DCC experiments.

Table 2. Buffers commonly used in different DCC reactions. *Tris buffer requires special attention.

Reaction	Buffer described in literature
Acylhydrazone formation [6,24,36]	Ammonium- and Sodium acetate, Phosphate, Tris*
Hydrazone formation [43,44]	Phosphate, Tris*
Disulfide [45,46]	Phosphate, Borate
Thioether [47]	Water/DMSO
Imine [13]	Water
Boronate ester [14,48]	Ammonium acetate, Water

For many protein targets, the stability at room temperature and the optimal buffer conditions are not known. We therefore recommend determining these conditions prior to performing DCC experiments. As several interdependent factors, like pH, buffer, ionic strength and ions influence the stability of a protein, it is difficult to suggest a stepwise flow scheme for the determination of the ideal buffer composition for a given protein.[28] Not only the protein but also the exchange chemistry might be affected significantly by varying these parameters. We propose to first measure the effect of pH, buffer and ionic strength over a wide

range in parallel. Afterwards, a small selection (2 to 5) of the most stabilizing combinations can be evaluated for their long-term effect on the protein. Subsequently, the best condition will then be used to determine the influence of DMSO (Section 2.6) and, if of interest, additives (Section 2.5). The selection of the initial buffers could be broadened, in case no suitable condition was found.

Two or more buffers should be screened per pH value to distinguish the influence of the buffer component and the pH value on the stability of the protein. It is also possible to use a so-called "superbuffer",[49] a mix of three or more buffer components, enabling the adjustment of a wide pH-range, without changing the buffer composition or concentration.

The effect of the buffer components on a protein can be measured in a straightforward way, by determining the melting point of the target protein *via* a thermal-shift assay / differential scanning fluorimetry (TSA /DSF).[28] In this method, the protein is incubated together with a lipophilic dye, for example sypro orange. The dye shows an increase in fluorescence after binding to the hydrophobic parts of a protein. These are often located at the inside of a protein and become exposed during temperature-induced unfolding/melting. The temperature–dependent increase in fluorescence can be measured in a RT-PCR apparatus and yields the T_m of the protein.

Other methods, like DSC, ITC and CD (differential scanning calorimetry, isothermal titration calorimetry and circular dichroism spectroscopy) and the determination of melting points by CD could also be used to gain information on the interaction and possible stabilization of the protein with its buffer, but require a high amount of protein and/or long measurement time. The TSA, however, offers high throughput and a short assay time, together with already several published or commercially available kits in 96-well format.[50,51] These kits were originally intended to screen for optimal crystallization conditions and cover several stability-influencing conditions. When performing DCC experiments, the design of an individual 96-well plate layout, tailored to the buffers and conditions compatible with the planned DCC reaction, might be useful. This is a short time investment, which might pay off quickly in the future, if ptDCC is used on several different targets.

After a DCC–compatible, stabilizing buffer condition has been identified, the protein should be checked for its long-term stability. To check for cleavage of the protein backbone an analysis by SDS-PAGE is of sufficient sensitivity (Figure 3). To determine if the protein folding is affected, TSA is again the method of choice, since the signal directly depends on the unfolding process of the protein. With prolonged degradation, the melting point decreases slightly. As a secondary effect of the degradation, the fluorescence curve can show bi- and multi-phasic melting

points and an overall decrease in signal intensity and resolution. A fully denatured enzyme will just show a decreasing fluorescence signal with no peak from protein unfolding. As controls, a fresh and a heat-treated sample of the target protein should be included in the experiment.

The tendency of a protein to precipitate is concentration-dependent. Because of this, the assays determining the protein stability should be performed with the same protein concentration that is intended to be used in the DCC experiment. If this is not possible, due to limited protein availability, the first experiments might be done with less protein. However, at least for the chosen final condition, the stability assessment should be repeated with the protein concentration that will be used in the DCC experiments.

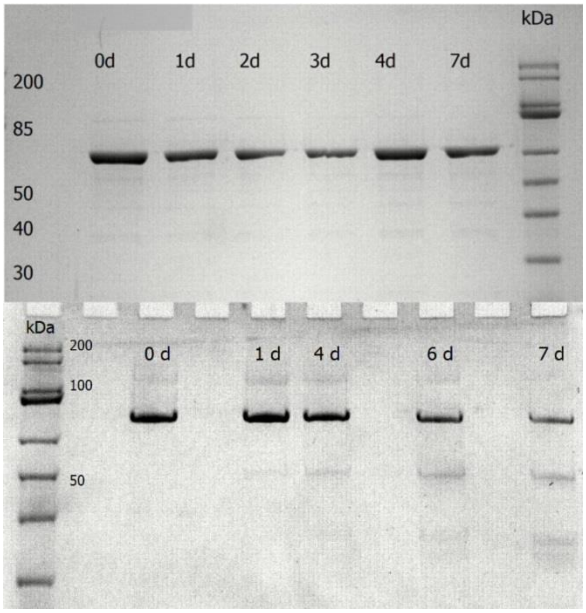


Figure 3. 12% SDS-PAGE of different homologues of the enzyme 5-Deoxyxylulose 5-Phosphate Synthase (DXS) after incubation at RT. The protein on the upper gel shows no sign of degradation. The second protein, shown on the lower gel, shows signs of degradation, starting already at day one with a very faint band around 50 kDa. From day 6 on a decrease of the main protein band also becomes clear. In the top left corner a gel-label was removed using image processing software.

1.2.4 Functional enzyme assay

For enzymatic protein targets, a functional assay can be used instead of TSA and PAGE measurements for the assessment of long-term stability. The analysis of activity data of a functional assay to determine the best experimental conditions of the DCC experiments leaves less room for interpretation than the analysis of the results of a melting-point analysis. Therefore, if a functional assay is available, and the enzyme is showing catalytic activity in the desired pH range of the DCC reaction, the activity assay should be the method of choice.

In a previous study from 2014, we could monitor the activity of the target protein endothiapepsin by performing a fluorescence-based assay (Figure 4). The pH-optimum of endothiapepsin is 4.5, and the enzymatic activity was not affected even after 20 days incubation at RT and a pH of 4.6. Considering this high stability, no buffer optimization was needed.[24]

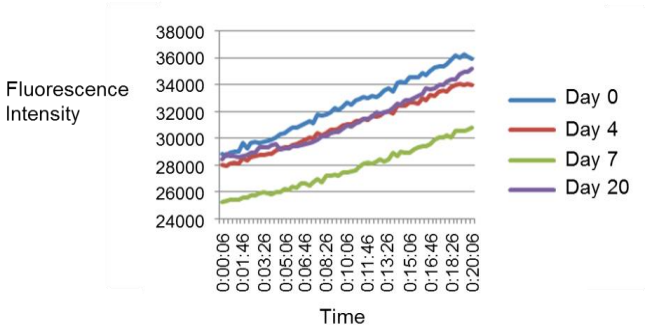


Figure 4. Activity of endothiapepsin, a pepsin-like aspartic protease, in a fluorescence-based assay at different time intervals of incubation at room temperature. Figure was adapted from Mondal et al.[24]

1.2.5 Additives and contaminations

During the purification, the protein might be in contact with different buffers and conditions. Some of the buffer components might remain bound to the protein, even after a buffer exchange. These contaminants might influence the experiment. It is therefore recommended to critically evaluate the composition of the protein sample. Not only should the protein storage buffer be evaluated, but also the origin of the sample.

Common substances that could be found in protein samples are for example imidazole as a leftover from an IMAC (immobilised metal affinity chromatography) purification step. Protein samples are often supplemented with reducing agents like 2-ME, DTT or TCEP (2-Mercaptoethanol, Dithiothreitol or Tris(2-carboxyethyl)phosphine) in concentrations up to 10 mM to keep the

protein in a reducing environment. If disulfide formation is the reversible reaction of choice, the final reducing agent concentration should be evaluated to make sure that the formation of disulfide bonds is not inhibited.

The effect of additives and contaminations is related to the volume of the protein sample used in the individual DCC experiments. This not only determines the final concentration of protein, but also the concentrations of the contaminants. If the batch-to-batch concentration of the protein varies and its volume is adjusted to reach the same final concentration in the DCC experiment, it should be noted that the concentrations and effects of the additives in the DCC experiment might vary.

Compounds that remain in the protein sample can have an influence on the DCC reaction or on the target protein. One should critically check every buffer component on possible interference with the planned exchange chemistry. Screening literature for known reactions between the DCL members and sample components can be considered. Performing a control experiment with all buffer components, in the absence of protein, can assure that no side reactions are taking place. If the exact composition of the protein sample is unknown, a small volume of the buffer might be gained by concentration of the protein using an ultrafiltration device and using the flow-through for the control experiment.

Some agents used during protein purification, such as cryoprotectants like glycerol or detergents like Tween, will interact in a non-specific way with the protein surface. From our experience, if there is no hint that they might affect the experiment, leftover cryoprotectants and detergents can be tolerated. Special care should be taken if cofactors, coenzymes or ions are supplemented during the purification process to stabilise the enzyme. The same holds true for buffer components structurally related to those supplements. Everything that binds to the targeted binding pocket is competing with the DCC library. If a natural, tight binding cofactor is present during the experiment, it could prevent the building blocks from binding and therefore also inhibit their amplification. However, the use of tight binders can be beneficial in control experiments. If a compound with a known binding site is inhibiting the formation of some previously observed binders this can be taken as a hint that the templated binders are targeting the same protein pocket.

1.2.6 DMSO

Addition of a small percentage of DMSO to the reaction solution is a common practice in the design of enzymatic assays to improve the solubility of hydrophobic compounds. For biochemical assays, DMSO concentrations up to 10% are regularly used.[52]

In DCC experiments, the building blocks of the library are typically dissolved in DMSO stock solutions to enable the easy assembly of a library. Depending on the library composition and number of compounds used, the final DMSO concentration would vary. To keep the reaction conditions comparable, we recommend adding DMSO up to a concentration that can be kept constant for all experiments of a project. This fixed concentration should be evaluated and chosen beforehand, to ensure the protein tolerates it.

DMSO has a very broad range of effects on proteins, it can even decrease the solubility and induce precipitation.[53] Both, rate acceleration, as well as inhibition of the enzyme-catalysed reaction by DMSO have been observed. An influence of already low percentages of DMSO on the enzymatic activity often hints to DMSO acting as an unspecific effector, interacting with the active site of the enzyme.[54] If the enzymatic activity is reduced by DMSO at higher concentrations (>10% DMSO), it is often by influencing the overall protein conformation by displacing water molecules bound to the surface and unfolding the protein.[55] On the other hand, there are DMSO-tolerant enzymes known which show activity up to 80% DMSO.[54] Enzyme activity assays are the method of choice to estimate the effect of DMSO on an enzyme. If no activity assay is available, the effect of DMSO could also be measured using TSA, however, interactions with the active site are difficult to detect with this method. We often observe a small effect on the T_m of a protein, but a strong effect on the enzymatic activity. Taken together, the DMSO concentration has several effects on the protein structure. The benefits of DMSO addition need to be weighed against the risk of creating an artificial enzymatic fold, which could amplify compounds that would not bind under native conditions. Therefore, the DMSO concentration should be as low as possible, in our lab up to 5% are regularly used.

1.2.7 Temperature

To speed up the rate at which the DCL reaches equilibrium, the experiments are normally performed at room temperature. For labile proteins, a lower reaction temperature may be necessary, which can improve the stability of the proteins. At the same time, the equilibration rate is decreased, leading to a prolonged incubation time. The optimal temperature for protein stability in DCC could vary from enzyme to enzyme and thus needs to be evaluated in each individual case but room temperature is used in most cases.

1.3 Setting up a ptDCC experiment

When crystal structures are available, or even cocrystal structures, a structure-based approach can be undertaken to design promising building blocks. In this case also non-binders could be designed as control elements, which are not supposed to emerge as hits. The type of reversible linkage should be carefully

selected because it influences the molecular recognition by the target. For example, the acylhydrazone linkage resembles the amide functionality; and features hydrogen-bond donors and acceptors. We showed that by combining DCC with de novo structure-based design, the risks associated with this attractive approach are reduced.[24]

1.3.1 Formation of the DCLs

The building blocks might have to be dissolved in DMSO, allowing them as well as the formed products to stay soluble in the final mixture. In principle, they could also be dissolved in the desired buffer, which would be most ideal. In 2014, we coupled DCC to saturation-transfer difference (STD)-NMR spectroscopy, which requires lower concentrations of protein than a general DCC experiment (Table 3). STD-NMR spectroscopy enables selection of the binders from the DCL, since the intensity of these signals is stronger due to a more efficient saturation transfer. As a result of only observing binders, STD-NMR spectra cannot be used to determine concentrations of DCL members and therefore amplification cannot be calculated. In follow-up experiments, it is possible to determine the K_D value of a ligand *via* STD-NMR or other biophysical assays.[56]

The ratio of hydrazides versus aldehydes should allow for the formation of all possible products, therefore at least one equivalent of each hydrazide per aldehyde should be used. For example, if three aldehydes are used then at least three equivalents of each hydrazide should be added, making sure that there is an excess of hydrazides. When required, a nucleophilic catalyst like aniline could be added. The most frequently used concentration of DMSO lies around 5–10%.

Table 3. General protocol for DCC and protocol for DCC coupled to ^{1}H-STD-NMR. * Aniline or another nucleophilic catalyst could be added when required. ** In a control experiment, no protein is added. *** Buffer conditions to guarantee protein stability should be determined a priori.

	Final concentration in general DCC	Final concentration used in DCC coupled to ^{1}H-STD-NMR[24]
Aldehyde	0.1 mM	0.4 mM
Hydrazide	0.1–0.3 mM	1 mM (for each of the five hydrazides)
DMSO	5–10^%	5–10^%
Aniline*	10 mM	–
Protein**	10–100 μM	4 μM
Buffer***	0.1 M	Ammonium acetate in D_2O (0.1 M, pH 4.6)
pH*	Acidic–neutral	pH 4.6

Control experiments should be considered, which should clarify where binding of molecules to the protein occurs and if it is specific or unspecific. This could for example be performed by the addition of a known inhibitor. If the previously observed amplification is not observed any longer, then the hit compounds are competitive binders. Based on the work of Danieli et al., B. Ernst and coworkers propose that the use of bovine serum albumin (BSA), as a negative control template for which no amplification is expected since the binding pocket is different, is not a good control since it could influence the library composition, whilst the use of a competitive inhibitor is better. BSA has been used in DCC to show that the applied library only gives hits with the real target and that BSA would yield the same result as the blank.[21,57] BSA is commonly known for its stability and was thought not to interfere with biological reactions, however recently DCC experiments have even been used to target BSA.[58]

1.3.2 Analysis of the DCLs

Different techniques such as fluorescence-polarization, SPR, ITC, MST, STD-NMR, crystallography and others can be used to evaluate and possibly optimise obtained hits. We and Rademann and coworkers have reviewed the analytical methods used in protein-templated dynamic combinatorial chemistry to detect hit compounds. [23][59]

A commonly applied method to analyse DCC experiments is the recording of HPLC-MS chromatograms of the libraries. As an illustrative example of the comparative approach, we drew HPLC chromatograms of a blank library and a target library (Figure 5). When we compare both chromatograms, we see that peak number five has increased in the library containing the target, whereas peaks three and six have decreased. The total amount of building blocks stays the same, only the equilibrium can be shifted towards one or more products.

In order to accurately determine the amplification or decrease of peaks, their relative peak areas (RPA) should be compared. The fictional RPAs of both chromatograms in Figure 5 are given in Table 4. The amplification factor in percentage can be calculated by equation 1, where the amplification factor in 'fold' is given by equation 2. Using these two equations, the product at peak five has increased by 100% or twofold. Frei et al. report on a particularly thorough analysis of a DCL using the lectin FimH as a target, using HPLC analysis with an optimised DCC protocol.[36]

Equation 1: amplification factor (%) = $\frac{RPA_{target} - RPA_{blank}}{RPA_{blank}} * 100\%$

Equation 2: amplification fold = $\frac{New}{Old}$

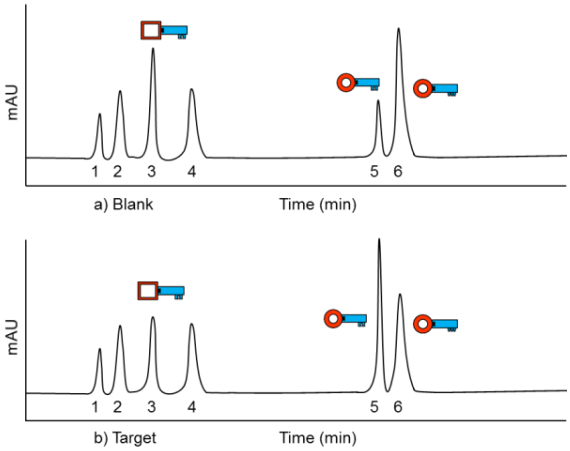


Figure 5. Schematic example of HPLC chromatograms: a) blank library chromatogram, b) target library chromatogram.

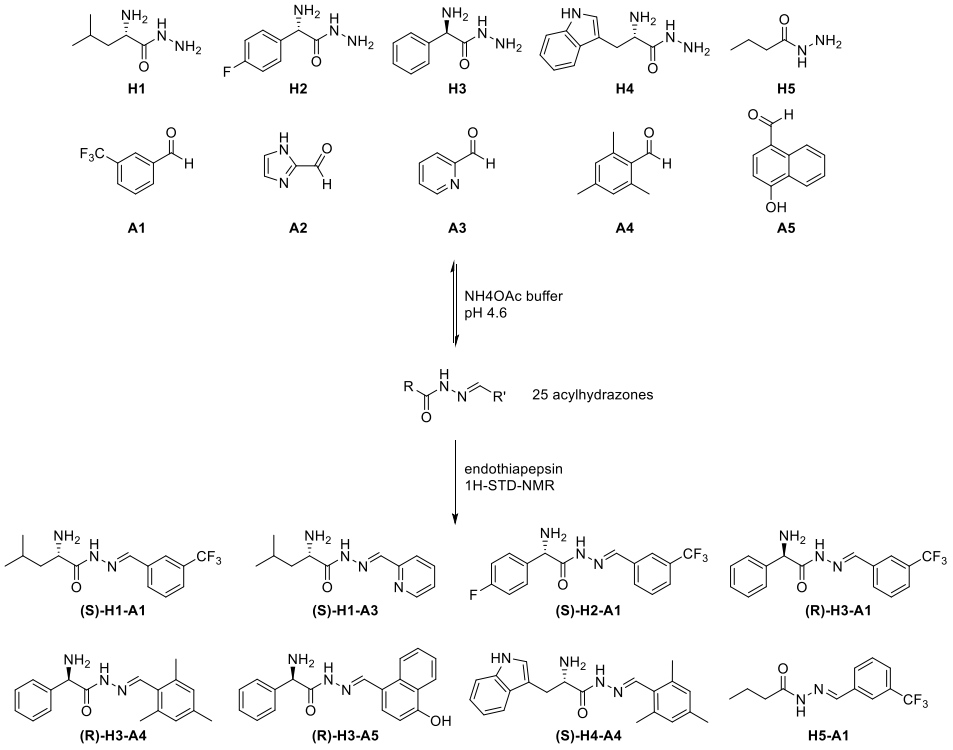
Table 4. Example of relative peak areas (RPA) obtained from HPLC chromatograms from Figure 4.

Peak number	Relative peak area in blank (%)	Relative peak area in target (%)	Amplification in %	Amplification in 'fold'
1	10	10	-	1
2	15	15	-	1
3	20	16	–20%	0.8
4	16	16	-	1
5	12	24	100%	2
6	27	19	–30%	0.7
Total	100%	100%		

1.3.3 DCL analysed with STD-NMR spectroscopy

Inspired by the work of Ramström and coworkers[20], we analysed the formed DCLs by STD-NMR spectroscopy (Scheme 3). We used the model enzyme endothiapepsin as target. As a control with a known binder we used saquinavir (K_i = 48 nM), a potent peptidic inhibitor, to differentiate specific from nonspecific binding. Each sub-library contained all five hydrazides and one of the aldehyde building blocks and was allowed to equilibrate for 24 hours before adding the target. By analysing the imine-type proton signals of the acylhydrazone products in the ^{1}H-STD-NMR spectra (Figure 6) we identified in total eight binders. To confirm the results from STD-NMR, we performed an enzyme-inhibition assay and showed that the hits were inhibitors with IC_{50} values ranging from 12.8 µM to 365 µM. The high hit rate in this publication is due to use of five sublibraries detecting the best binder of each library, whereas in a regular ptDCC setup only the overall best binders will be discovered. In addition, the high hit rate is also a

result of the synergistic combination of de novo structure based drug design (SBDD) and DCC. In STD-NMR the protein is used as a tool to analyse the library, whereas in a ptDCC experiment the protein influences the equilibrium and hence the concentrations.



Scheme 3. Formation of dynamic combinatorial library and enzymatic selection of the best binders by 1H-STD-NMR analysis; Scheme adapted from Mondal et al.[24]

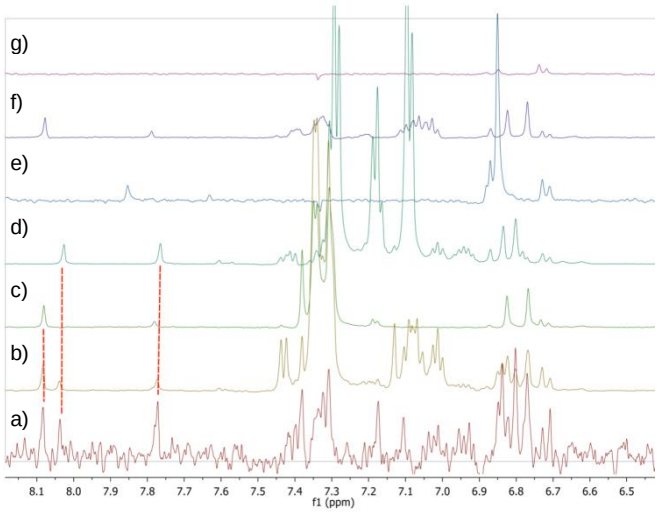


Figure 6. DCL generated from **H1–5** + **A4**: (aromatic region) a) ^{1}H-STD-NMR spectrum of **H1–5** + **A4**, b) ^{1}H-NMR spectrum of **H1–5** + **A4**, c) ^{1}H-NMR spectrum of **H3**+**A4**, d) ^{1}H-NMR spectrum of **H4**+**A4** (2 singlets correspond to the *E*/*Z* isomers), e) **H1**+**A4**, f) **H2**+**H4** and g) **H5**+**A4**. Figure was adapted from Mondal et al.[24]

1.3.4 How to proceed after obtaining hits

Having obtained a validated hit, identified by de novo structure-based drug design in combination with DCC and STD-NMR, we have used a structure-based design approach to improve the molecular recognition by the target.[60] In this specific case, we were fortunate to have an x-ray crystal structure of the target endothiapepsin in complex with the hit. If this is not the case, optimization is still possible, relying on structure–activity relationships.

1.4 DCC in a synergistic combination with fragment linking

For fragment linking, two or more fragments have to bind to different but adjacent sites of the enzyme active site.[61] This approach introduces one additional component into the ligand system: a linker moiety. Finding the right linker motif, whilst maintaining the binding poses of both fragments, which orients the individual fragment units in the favourable geometry in relation to each other without introducing too much flexibility, can be challenging. The combination of two fragments with rather low affinity could result in significantly higher affinity and has the potential to result in 'superadditive' contributions of both binding motifs. The challenge in fragment linking is the exploration of the binding mode of both fragments and the identification of an optimal linker fitting in between. Only in this case, the overall reduced so-called 'rigid body entropy'

translates into synergistically improved affinity. By binding of a fragment to a target protein, rotational and translational entropy is lost. This entropy penalty has to be overcompensated by attractive interactions formed between the ligand and the target. When two fragments bind in parallel to adjacent sites, each has to pay this entropy penalty. When these two fragments are linked together in an ideal way, the resulting singular compound only encounters the loss of rigid body entropy once. Hence, the observed affinity will be much greater than only the sum of the individual affinities.[62] The additional binding energy gained, is often also referred to as linker energy. To overcome the challenges associated with fragment linking, we pioneered a synergistic combination with DCC. For this proof-of-concept study, we again used the model enzyme endothiapepsin.[35] X-ray crystal structures of endothiapepsin in complex with fragment inhibitors **1** and **2** (PDB IDs: 4KUP and 3T7P) identified by DCC were used as a starting point for fragment linking studies facilitated by DCC. Hits **1** and **2** display IC_{50} values of 12.8 μM and 14.5 μM and LEs of 0.27 and 0.29, respectively. The linking of **1** and **2** should generate an inhibitor that occupies two binding pockets of endothiapepsin (Figure 7).

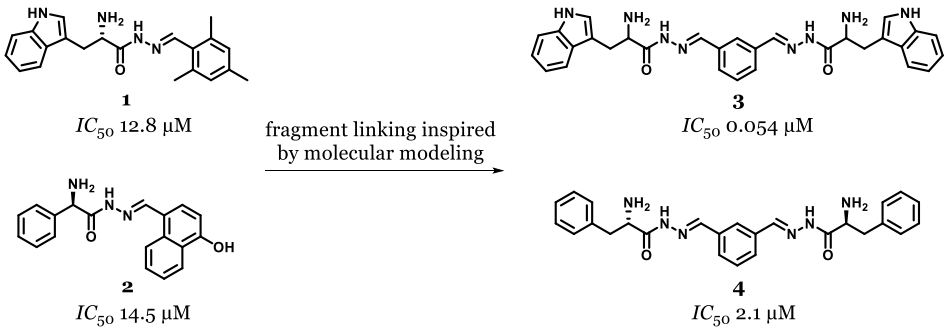


Figure 7. Structures of hits **1** and **2** and linked bisacylhydrazone linked inhibitors **3** and **4**.[35]

The homo-bis-acylhydrazones **3** and **4** were hits from the DCC experiments and were synthesised and evaluated accordingly. Compared to compound **2**, the potency of inhibitor **3** was increased 240-fold, yielding an IC_{50} value of 0.054 μM and a LE value of 0.29. For inhibitor **4** an IC_{50} of 2.1 μM and a LE value of 0.25 was determined (Figure 7).[35] Obviously, only the symmetric linking modality resulted in efficient cooperative binding.

1.5 Conclusions

There are a number of steps which should be carefully taken into account, in order to obtain active hits by DCC. If information on the target is available, e.g. a crystal-structure, one could consider a structure-based design when choosing the building blocks. The type of reversible linkage to be used can be chosen at this stage. Conditions necessary for the equilibration to take place should be compatible with the target. After establishing conditions, which will ensure the target remains folded, the actual DCC experiment can be started. To do so, stock solutions of building blocks, catalyst and protein should be prepared. The formed DCLs can be analysed by different techniques such as STD-NMR or HPLC-MS. Compounds that have been selected by the target, and their biochemical properties should be evaluated and possibly optimised in further studies.

1.6 Outline of this thesis

Dynamic combinatorial chemistry (DCC) has evolved over the past decades from a tool to easily generate a pool of derivatives to an efficient technique to find hit compounds when applied on a target. To be able to use target-directed DCC (tdDCC), a number of criteria must be met ranging from biocompatibility, solubility, stability, pH dependence to temperature and type of reversible reaction.

The protein family of 14-3-3 has been selected as a target, since it allows for all of these criteria to be met. 14-3-3 proteins are involved in protein-protein interactions (PPIs) in many different biological processes, ranging from diseases to cell-cycle control and signal transduction. Modulating 14-3-3 proteins for binding is therefore an important class of research.

The first aim of this thesis is applying DCC on biological relevant targets; such as 14-3-3 proteins and glucansucrases. The second aim of this thesis is on extending the list of reversible reactions, which can be applied in tdDCC. This allows medicinal chemists more freedom, by being able to use different scaffolds. Therefore, the overall objective is to find new applications of DCC to medicinal chemistry.

In chapter 3, we apply tdDCC on the 14-3-3ζ isoform. We use the PPI complex of 14-3-3ζ/synaptopodin in acylhydrazone-based DCC, aiming to find small drug-like molecules which can stabilise this PPI.

Chapter 4 describes the application of tdDCC on a glucantransferase, which is found to be causative for adhesion of bacteria to the tooth enamel, which can lead to dental caries.

Finally, in chapters 5 and 6 the design and first applications of new scaffolds for tdDCC is described. In chapter 5, the application of nitrone-based DCC with endothiapepsin is evaluated. And in chapter 6, the thiazolidine scaffold is investigated and DCC conditions optimised for endothiapepsin.

1.7 References

[1] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84.

[2] R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* **1991**, *354*, 84–86.

[3] R. A. Houghten, *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 5131–5135.

[4] H. M. Geysen, R. H. Meloen, S. J. Barteling, *Proc. Natl. Acad. Sci.* **1984**, *81*, 3998 LP – 4002.

[5] R. Frank, W. Heikens, G. Heisterberg-Moutsis, H. Blöcker, *Nucl. Acid. Res.* **1983**, *11*, 4365–4377.

[6] I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci.* **1997**, *94*, 2106–2110.

[7] M. H. Ohlmeyer, R. N. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci.* **1993**, *90*, 10922 LP – 10926.

[8] S. J. Rowan, P. S. Lukeman, D. J. Reynolds, J. K. M. Sanders, *New J. Chem.* **1998**, *22*, 1015–1018.

[9] V. A. Polyakov, M. I. Nelen, N. Nazarpack-Kandlousy, A. D. Ryabov, A. V Eliseev, *J. Phys. Org. Chem.* **1999**, *12*, 357–363.

[10] A. Ganesan, *Angew. Chemie Int. Ed.* **1998**, *37*, 2828–2831.

[11] C. Karan, B. L. Miller, *Drug Discov. Today* **2000**, *5*, 67–75.

[12] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* **2010**, *2*, 490–497.

[13] Z. Fang, W. He, X. Li, Z. Li, B. Chen, P. Ouyang, K. Guo, *Bioorganic Med. Chem. Lett.* **2013**, *23*, 5174–5177.

[14] M. Demetriades, I. K. H. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y.-M. Tian, T. D. W. Claridge, P. J. Ratcliffe, E. C. Y. Woon, et al., *Angew. Chemie Int. Ed.* **2012**, *51*, 6672–6675.

[15] E. C. Y. Woon, M. Demetriades, E. A. L. Bagg, W. Aik, S. M. Krylova, J. H. Y. Ma, M. Chan, L. J. Walport, D. W. Wegman, K. N. Dack, et al., *J. Med. Chem.* **2012**, *55*, 2173–2184.

[16] S. Sakai, Y. Shigemasa, T. Sasaki, *Tetrahedron Lett.* **1997**, *38*, 8145–8148.

[17] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Angew. Chemie Int. Ed.* **2002**, *41*, 3405–3407.

[18] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Tetrahedron* **2004**, *60*, 771–780.

[19] B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* **2006**, *128*, 8459–8467.

[20] R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero, O. Ramström, *Angew. Chemie Int. Ed.* **2010**, *49*, 589–593.

[21] P. Frei, R. Hevey, B. Ernst, *Chem. - A Eur. J.* **2019**, *25*, 60–73.

[22] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455–2488.

[23] R. Van der Vlag, A. K. H. Hirsch, in *Compr. Supramol. Chem. 2*, Elsevier, **2017**, pp. 487–509.

[24] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259–3263.

[25] A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* **2006**, *128*, 15602–15603.

[26] A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem. Int. Ed* **2006**, *45*, 7581–7584.

[27] P. Crisalli, E. T. Kool, *J. Org. Chem.* **2013**, *78*, 1184–1189.

[28] F. V Reddavide, W. Lin, S. Lehnert, Y. Zhang, *Angew. Chemie Int. Ed.* **2015**, *54*, 7924–7928.

[29] L. Monjas, L. J. Y. M. Swier, I. Setyawati, D. J. Slotboom, A. K. H. Hirsch, *ChemMedChem* **2017**, *12*, 1693–1696.

[30] F. T. Kern, K. T. Wanner, *ChemMedChem* **2015**, *10*, 396–410.

[31] F. Kern, K. T. Wanner, *Bioorg. Med. Chem.* **2019**, *27*, 1232–1245.

[32] G. Artigas, P. López-Senín, C. González, N. Escaja, V. Marchán, *Org. Biomol. Chem.* **2015**, *13*, 452–464.

[33] J. D. McAnany, J. P. Reichert, B. L. Miller, *Bioorganic Med. Chem.* **2016**, *24*, 3940–3946.

[34] Z. Yang, Z. Fang, W. He, Z. Wang, H. Gan, Q. Tian, K. Guo, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1671–1674.

[35] M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch, *Angew. Chemie Int. Ed.* **2016**, *55*, 9422–9426.

[36] P. Frei, L. Pang, M. Silbermann, D. Eris, T. Mühlethaler, O. Schwardt, B. Ernst, *Chem. Eur. J.* **2017**, *23*, 11570–11577.

[37] J. Fu, H. Fu, M. Dieu, I. Halloum, L. Kremer, Y. Xia, W. Pan, S. P. Vincent, *Chem. Commun.* **2017**, *53*, 10632–10635.

[38] J. Soubhye, M. Gelbcke, P. Van Antwerpen, F. Dufrasne, M. Y. Boufadi, J. Nève, P. G. Furtmüller, C. Obinger, K. Zouaoui Boudjeltia, F. Meyer, *ACS Med. Chem. Lett.* **2017**, *8*, 206–210.

[39] A. G. Ekström, J. T. Wang, J. Bella, D. J. Campopiano, *Org. Biomol. Chem.* **2018**, *16*, 8144–8149.

[40] M. Das, T. Yang, J. Dong, F. Prasetya, Y. Xie, K. H. Q. Wong, A. Cheong, E. C. Y. Woon, *Chem. – An Asian J.* **2018**, *13*, 2854–2867.

[41] P. García, V. L. Alonso, E. Serra, A. M. Escalante, R. L. E. Furlan, *ACS Med. Chem. Lett.* **2018**, *9*, 1002–1006.

[42] A. J. Clipson, V. T. Bhat, I. McNae, A. M. Caniard, D. J. Campopiano, M. F. Greaney, *Chem. - A Eur. J.* **2012**, *18*, 10562–10570.

[43] M. Sindelar, K. T. Wanner, *ChemMedChem* **2012**, *7*, 1678–1690.

[44] R. Nguyen, I. Huc, *Chem. Commun.* **2003**, 942–943.

[45] D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon, J. A. Wells, *Proc. Natl. Acad. Sci.* **2000**, *97*, 9367–9372.

[46] R. F. Ludlow, S. Otto, *J. Am. Chem. Soc.* **2008**, *130*, 12218–12219.

[47] B. Shi, M. F. Greaney, *Chem. Commun.* **2005**, 886–888.

[48] I. K. H. Leung, T. Brown, C. J. Schofield, T. D. W. Claridge, *Medchemcomm* **2011**, *2*, 390–395.

[49] J. Newman, *Acta Crystallogr. Sect. D* **2004**, *60*, 610–612.

[50] L. Reinhard, H. Mayerhofer, A. Geerlof, J. Mueller-Dieckmann, M. S. Weiss, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2013**, *69*, 209–214.

[51] F. H. Niesen, H. Berglund, M. Vedadi, *Nat. Protoc.* **2007**, *2*, 2212–2221.

[52] G. S. Sittampalam, N. P. Coussens, K. Brimacomber, A. Grossman, M. Arkin, D. Auld, C. Austin, J. Baell, B. Bejcek, J. M. M. Caaveiro, et al., **2004**.

[53] T. Arakawa, Y. Kita, S. N. Timasheff, *Biophys. Chem.* **2007**, *131*, 62–70.

[54] D. H. Rammler, *Ann. N. Y. Acad. Sci.* **1967**, *141*, 291–299.

[55] M. Jackson, H. H. Mantsch, *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1991**, *1078*, 231–235.

[56] A. Viegas, J. Manso, F. L. Nobrega, E. J. Cabrita, *J. Chem. Educ.* **2011**, *88*, 990–994.

[57] B. Danieli, A. Giardini, G. Lesma, D. Passarella, B. Peretto, A. Sacchetti, A. Silvani, G. Pratesi, F. Zunino, *J. Org. Chem.* **2006**, *71*, 2848–2853.

[58] C. Qiu, Z. Fang, L. Zhao, W. He, Z. Yang, C. Liu, K. Guo, *React. Chem. Eng.* **2019**, *4*, 658–662.

[59] M. Jaegle, E. L. Wong, C. Tauber, E. Nawrotzky, C. Arkona, J. Rademann, *Angew. Chem. internat. Ed.* **2017**, *56*, 7358–7378.

[60] A. M. Hartman, M. Mondal, N. Radeva, G. Klebe, A. K. Hirsch, *Int. J. Mol. Sci.* **2015**, *16*, DOI 10.3390/ijms160819184.

[61] D. A. Erlanson, W. Jahnke, *Fragment-Based Drug Discovery: Lessons and Outlook*, Wiley-VCH Verlag GmbH, **2016**.

[62] C. W. Murray, M. L. Verdonk, *J. Comput. Aided. Mol. Des.* **2002**, *16*, 741–753.

Chapter 2

Molecular Insight into Specific 14-3-3 Modulators: Inhibitors and Stabilisers of Protein-Protein Interactions of 14-3-3

The 14-3-3 protein family is implicated in several diseases and biological processes. Several recent reviews have summarised knowledge on certain aspects of 14-3-3 proteins, ranging from a historic overview to the structure, function and regulation. This chapter focuses on the structures and molecular recognition of the modulators by the 14-3-3 proteins, and small modifications of certain modulators are proposed where cocrystal structures have been reported. Our analysis opens up possibilities for the optimisation of the reported compounds. It is very timely to analyse the current status of recently developed modulators given that the field has seen a lot of activity in recent years. This chapter provides an overview combined with a critical analysis of each class of modulators, keeping their suitability for future development in mind.

This chapter has been published as a review article:

A. M. Hartman, A. K. H. Hirsch, *Eur. J. Med. Chem.* **2017**, 136, 573–584.

2.1 Introduction

Protein-protein interactions (PPIs) play a significant role in many diseases. While a single polypeptide chain can have enzymatic or structural activity, interactions with other proteins allow an infinite variety of ways to modulate the activity. It is assumed that between 130,000 and 650,000 PPIs exist, making PPIs important therapeutic targets. This especially holds for diseases, which are difficult to treat, for example when targets like ion channels, enzymes and GPCRs have not yet been identified or are not present at all.[1]

One class of proteins in which PPIs play an important role is the 14-3-3 protein family. This protein family consists of seven isoforms in humans: beta (β), epsilon (ε), eta (η), gamma (γ), tau (τ), sigma (σ) and zeta (ζ). Each of these human isoforms is encoded by a distinct genetic sequence. While plants have been shown to have more than ten isoforms, eukaryotic micro-organisms have fewer isoforms, for instance, two isoforms have been revealed in yeast.[2] All isoforms have a high conformational and functional conservation. The two isoforms found in yeast can be interchanged with plant or mammalian isoforms and still be active.[2] It has been shown that the binding groove of all isoforms has three conserved binding motifs: RSXpSXP (mode 1), RXXXpSXP (mode 2) and pS/TX-COOH (mode 3), in which pS stands for a phosphoserine residue.[3,4] 14-3-3 proteins primarily form dimers, which both feature a U-shaped binding groove.[5,6] A superimposition of the seven isoforms is shown in Figure 1, illustrating the conserved conformation of the monomers.[7] Wang and Shakes have made multiple alignments of forty-six sequences of plant, animal and fungal species to determine the molecular evolution of the 14-3-3 protein family.[8] Wilkert *et al.* also show a sequence conformation among the 14-3-3 proteins.[9]

Figure 1. Superimposition of monomers of the seven human isoforms.

The 14-3-3 monomers are acidic (pI ≈ 4.6)[10], have a molecular weight of 25–30 kDa and can combine to form homo- and heterodimers.[11] The non-trivial name 14-3-3 was assigned to this protein family by the researchers who first identified it from a classification study on brain proteins, due to position of the bands on 2D diethylaminoethyl (DEAE)-cellulose chromatography and starch gel electrophoresis.

The 14-3-3 family is implicated in several diseases and biological processes. Studies on binding partners of the 14-3-3 proteins have revealed over 500 proteins, and this number will continue to grow with ongoing research.[12] Among these binding partners are many biologically relevant targets. In some diseases, the 14-3-3 proteins have inhibitory effects, while in other cases they lead to stabilisation. It is even observed that binding of 14-3-3 enhances the activity of the enzyme *N*-acetyltransferase (AANAT).[3] Examples of such processes are: regulation of metabolism, signal transduction, cell-cycle control, apoptosis, protein trafficking, transcription, stress response and malignant transformation.[5] A few examples of the many diseases in which 14-3-3 proteins are involved are presented in Table 1.

Table 1. Few examples of many diseases in which 14-3-3 proteins are involved.

14-3-3 isoform	Disease/ function	Binding partner
(σ) and (ζ)	Noonan Syndrome[13]	C-Raf
(ζ)	Cell death[14,15]	Pathogenic protein Exoenzyme S
(σ) specifically	Oncogenesis and cellular response in DNA damage[9,16]	p53
all	Mutant (breast) cancer[17,18]	Raf
(σ) and (ζ)	Alzheimer's[19]	Tau

In order to develop treatments for these diseases, it will be necessary to modulate PPIs of the relevant proteins. An approach to modulate the PPIs is by the use of small-molecule modulators. Modulators of the PPIs of 14-3-3 with partner proteins can come in two types: inhibitors or stabilisers. Small-molecules modulators are usually derived from natural products such as proteins, peptides or secondary metabolites. These modulators have complex structures, making them specific and selective for 14-3-3 proteins.[20] As the isoforms are encoded by distinct genes, they have a unique but similar amino acid sequence. Due to this

uniqueness it might be possible to selectively modulate a single isoform, leaving the other isoforms untouched. Notably, this selectivity is quite essential from a therapeutic point of view, as the drug should not disturb any other biological processes. Research into the discovery and optimisation of small-molecule modulators should therefore take the selectivity of the compounds into consideration, wherever applicable.

Here, we systematically explore the binding of small-molecule modulators of 14-3-3 proteins at the molecular level, that is, the exact binding of the modulators to the amino acid residues lining the binding groove, aiming to provide an insight into the way these modulators work. For almost all modulators in this chapter of which a crystal structure is known we propose modifications based on a structure-based analysis. Therefore, this chapter gathers knowledge, which can be used for structure-based design of (new) modulators. Several recent reviews have summarised knowledge on other aspects of 14-3-3 proteins: the structural basis,[6] insights from genome-wide studies,[21] a historic overview,[22] the structure, function and regulation,[7] an update on 14-3-3 sigma related to human cancer[11] and small-molecule modulators with a focus on the discovery and the biochemical relevance of the compounds.[12]

2.2 Structure-based optimisation

There are cocrystal structures available of almost all 14-3-3 isoforms (some of which are in complex with interaction partners). 14-3-3 can therefore be conveniently targeted using a structure-based approach. The analysis of the binding modes of the modulators described in this chapter gives insight into where the compounds might be extended or modified in order to achieve higher affinity.

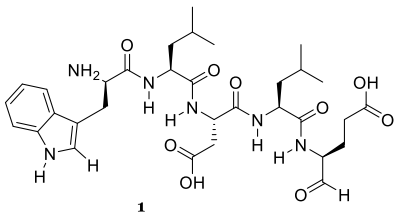
For the known cocrystallised modulators, we suggest small modifications, based on a first inspection of the experimentally determined binding mode analysed *in silico*. More advanced docking studies and biochemical evaluations should be performed to confirm the improved affinities. The computer programme SeeSAR was used for the evaluation, using the in-built scoring function HYDE.[23–25] As explained by Reulecke *et al.*, HYDE focusses on hydration and desolvation, taking into consideration the local hydrophobicity, solvent accessible surface and contact surface area. It calculates this affinity for every atom present in the interface.[25]

2.3 Inhibitors

The first compounds reported to inhibit the PPIs of 14-3-3 with partner proteins were **1** and its dimer difopein (dimeric fourteen-three-three peptide inhibitor) (**2**) (Figure 2).[26] The former, a 20-amino acid peptide of Mw 2309.6, was identified by Wang *et al.* from a phage display to 14-3-3(τ).[27] Petosa *et al.* showed *via* a crystal structure (Figure 3) that R18 binds in the amphipathic groove and that the central WLDLE motif is hosted in the part of the pocket that binds phosphonic acids, also called the phospho-accepting pocket.[28] Moreover, the peptide does not require phosphorylation to bind to 14-3-3, as the negatively charged Asp and Glu mimic the contact points of phosphoserine.[7,26] As seen from HYDE scoring in the programme SeeSAR, there are multiple favourable and some unfavourable interactions between **1** and 14-3-3(ζ) (Figure S1 in the supporting information).[23,24] The C-terminal Leucine of **1** is involved in lipophilic interactions with Gly169 and Ile217. The carboxylic acid terminal group of **1** shows electrostatic interactions with Arg56 and Arg127. Also the oxygen atom of the most left amide bond shows to lose energy due to desolvation.

Compound **1** does not have selectivity between isoforms of 14-3-3. However, it is selective for 14-3-3 proteins.[26] Binding by **1** inhibits binding of partners like Raf-1, ASK1 and exoenzyme S from binding to 14-3-3 *in* vitro. Fu and Masters reported that **2**, enhances the ability of the antineoplastic drug cisplatin to kill cells. Compound **2** can bind competitively to one or two monomers of 14-3-3 given that the two **1** repeats are separated by a flexible linker. Cao *et al.* showed that **2**, as a general antagonist of 14-3-3, effectively hinders proliferation and triggers apoptosis of tumour cells, which were implanted in nude mice. *Via* flow cytometric analysis it was shown that apoptosis is time-dependent.[29]

The first small-molecule inhibitor of 14-3-3 proteins was compound **3** (Figure 4), which specifically and competitively inhibits 14-3-3/ligand binding *in vitro*. Wu *et al.* demonstrated that **3** also causes apoptosis and cell-cycle arrest in mammals. The central motif of **3** is a phosphorylated peptide-like small molecule; its non-phosphorylated analogue shows no activity in competing with 14-3-3(σ)/ligands. Molecular docking was used by the authors to confirm the binding mode of **3** with 14-3-3(σ); no cocrystal structure is available.[30] Optimisation of **3** led to compound **4**. This prodrug is a phosphoserine mimetic and potent 14-3-3 inhibitor. When compound **4** is used as a prodrug, the active compound, the phosphonic acid is delivered more effectively into the cells.[31]



1

Pro-His-Cys-Val-Pro-Arg-Asp-Leu-Ser-**Trp-Leu-Asp-Leu-Glu**-Ala-Asn-Met-Cys-Leu-Pro

Figure 2. Structure of the WLDLE motif which is found in the phospho-accepting pocket of 14-3-3 proteins.[27] And the sequence of **1**, highlighted is the WLDLE motif.

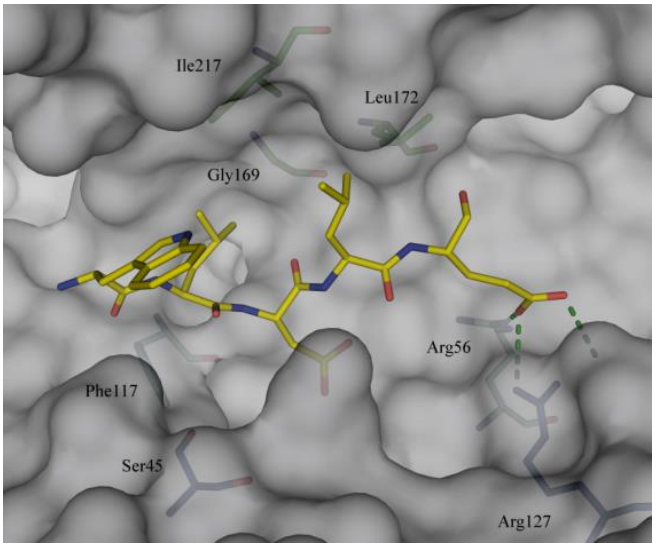
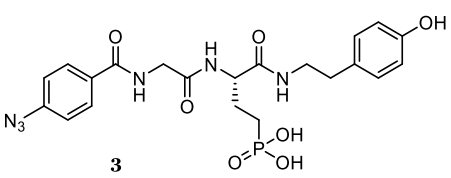
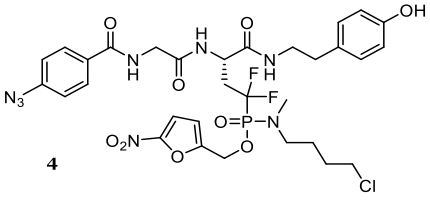


Figure 3. Cocrystal structure of 14-3-3(ζ) with **1** (PDB code: 1A38). Colour code: protein surface: grey; interacting residues: sticks; inhibitor skeleton: C: yellow; N: blue; O: red, Cl: green; H-bonds, below 3.2 Å: green, dashed lines. This colour code is maintained throughout this chapter.[28]



3



4

Figure 4. Inhibitors **3** and prodrug **4**, identified *via* an small-molecule microarray (SMM) technique consisting of fragment-based high-throughput identification combined with a reconstitution step and biological validation.[30]

Almost concurrently with the discovery of **4**, Corradi and *et al.* disclosed the non-peptidic small-molecule **5**, an inhibitor of the 14-3-3/c-Abl complex (Figure 5). The authors used an *in silico* virtual screening campaign and docking simulations, based on a structure-based pharmacophore model, against 200.000 compounds. Compound **5** can be used to treat Bcr-Abl-associated diseases, which are resistant to traditional drugs.[32] Unfortunately, it was found that **5** is chemically unstable, as it cyclises at room temperature.[33] The authors carried out a virtual-screening campaign with the aim of identifying compounds that score better than **5**. Compounds **6** and **7** emerged as hits that are cytotoxic against Bcr-Abl-expressing Ba/F3 cells, and according to molecular modelling bind to 14-3-3(σ) in a manner similar to that of **5** and **1**.[33]

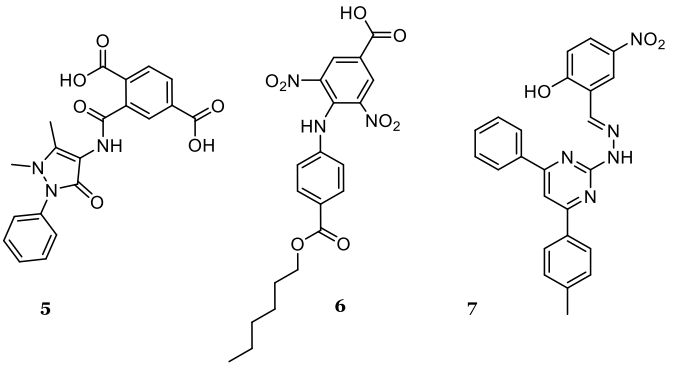


Figure 5. Inhibitors **5**, **6** and **7**, identified *via* a virtual screening campaign.[32]

Another small-molecule inhibitor of all 14-3-3 isoforms in 14-3-3 PPIs, FOBISIN101 (**8**), was reported in 2011 (Figure 6).[34] FOBISIN (FOurtien-three-three BInding Small molecule INhibitor) 101 is a phosphoSer/Thr-mimetic agent, which upon X-ray irradiation covalently binds to 14-3-3(ζ) at Lys120. Furthermore, it has lipophilic interactions with Pro167, Gly171 and Ile219 and polar interactions to the flexible Lys49. Zhao *et al.* observed a colour change from bright orange before X-ray irradiation to brownish then yellowish and finally colourless after irradiation. In addition to this covalent linkage, which plays a role in its inhibitory activity, the authors suggest that the phosphate moiety might play a critical role. This conclusion was drawn from the observation that a non-phosphorylated derivative of **8** is significantly less effective in blocking 14-3-3/ligand binding. Mutation of Lys120 to Glu was shown to inactivate 14-3-3(ζ), this was observed in an assay to determine the activation of ExoS ADP-ribosyltransferase. Therefore, this indicates that Lys120 plays an important role in the binding to ligands/modulators.[34] The authors noted that another covalent linkage might also occur through imine formation, which was proposed later by Röglin *et al.* and is shown in Figure 7. Röglin *et al.* propose a likely mechanism

involving imine hydrolysis to a pyridoxal-phosphate (PLP) intermediate followed by imine formation with Lys122 of 14-3-3(ζ).[35] The differences in experimental conditions appear to determine whether the covalent linkage takes place. With the Cu-Kα edge radiation they did not observe a colour change/covalent linkage as previously described. However, synchrotron irradiation did result in the covalent modification at Lys120 as described before. They agree that the risk of nonspecific imine formation might be prevented by eliminating the aldehyde functionality, thereby making it a more potent class of compounds.[36]

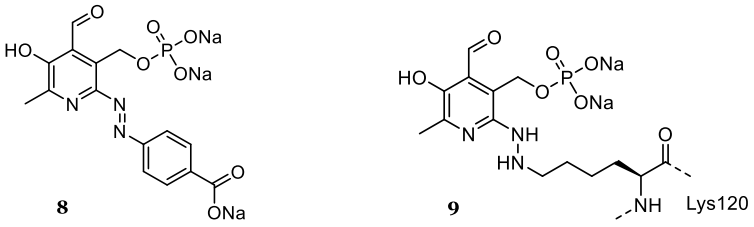


Figure 6. Inhibitors FOBISIN101 (**8**) and the covalent adduct **9** with 14-3-3(ζ). Compound **8** was identified from a fluorescence polarisation-based 14-3-3 binding assay, screening compounds that disrupt the 14-3-3(γ)-pS259-Raf-1 peptide complex.[34]

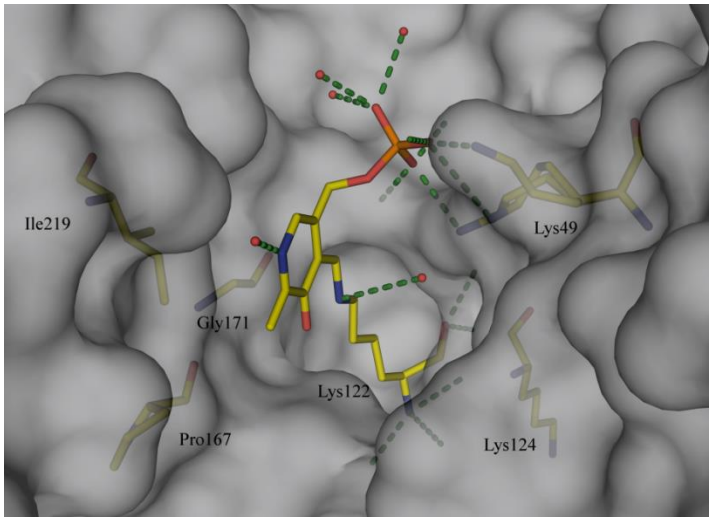


Figure 7. Cocrystal structure of **9**: compound **8** covalently linked to Lys122 of 14-3-3(σ) *via* an imine bond (PDB code: 3U9X).[35]

Shortly after the publication on **8**, Park *et al.* reported on xanthine (**10**) and lacosamide (**11**) derivatives as specific binders of 14-3-3(ζ) (Figure 8). Using *in vitro* labelling, they showed that the enantiomer *R*-**11** binds specifically to 14-3-3(ζ). The stereoisomer *S*-**11** was found to be less active, and *R*-**12** outcompetes *R*-**11** for binding to 14-3-3(ζ). The authors were unable to explain the isoform selectivity, cocrystallographic data could help to gain a better understanding.[37]

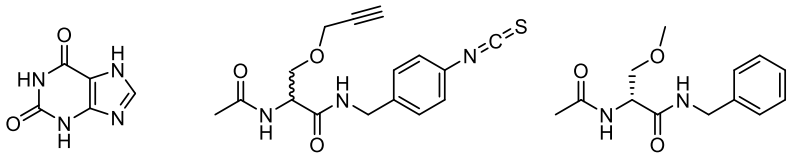
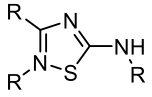


Figure 8. Modulator Xanthine (**10**) and inhibitors lacosamide (*R*) and (*S*)-**11** and derivative (*R*)-**12**. 14-3-3 was identified as a binding partner to these compounds from a ligand affinity and chemical reporter strategy.[37]

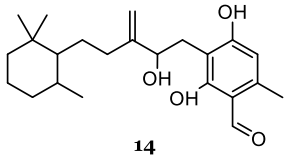
A high-throughput screen using an *in vitro* fluorescence-polarisation assay afforded another class of scaffolds that modulate 14-3-3/ligand interactions, namely compounds with the core structure **13** (Figure 9).[38] The main part of the PRLX-scaffold is a 1,2,4-thiadiazole, decorated with substituents (R), which were not specified as they were patented compounds. There is no x-ray crystal structure published with these modulators. The authors used multiple biophysical assays to assess the activity of the compounds, starting with isolated cells and followed by intact tissue. Five different 14-3-3 isoforms from *E. coli* ((β, ζ η, ε and γ) were tested and it was found that the compounds had highest binding affinity for the (γ) isoform.[38] The origin of this specificity was left uninvestigated. The thiadiazole scaffold of compound **13** (Figure 4) inhibits the interaction of 14-3-3(γ) and a phosphorylated HSP20 peptide (6-FAM-WLRRApSAP)**.**



13

Figure 9. Inhibitor scaffold **13** identified by high-throughput screening.[38]

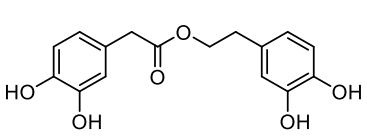
In 2011, the inhibitor **14** was identified in a study aimed at making derivatives of the natural product class of the moverastins, which target cell migration of tumour cells for a possible treatment of tumour metastasis. Compound **14** (Figure 10) was found to interact differently than moverastins, therefore suggesting a different target than the original target of the study, farnesyltransferase (FTase).[39] A follow-up study confirmed 14-3-3(ζ) as the target, and from the seven human isoforms it was found that (ζ) had the strongest binding affinity. The authors suggest that this selectivity stems from the fact that the C-terminal region of 14-3-3 is involved in the binding, which is the most variable region of the isoforms. Compound **14** presumably inhibits cell migration by competing with the 14-3-3(ζ)/Tiam1 complex.[40]



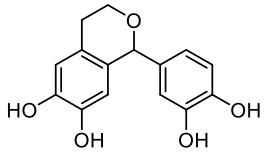
14

Figure 10. Inhibitor **14**, a moverastin derivative that targets 14-3-3(ζ).[39]

In 2012, Yan *et al.* identified Blapsins A (**15**) and B (**16**) as potent small-molecule inhibitors. These compounds (Figure 11) were isolated from *Blaps japanesis*, characterised and evaluated for their inhibitory potencies on 14-3-3(γ).[41] They were found to inhibit 14-3-3 binding to the fluorescence-labelled TMR-Raf pSer259 peptide, used for monitoring protein interactions.[42]



15



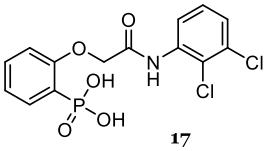
16

Figure 11. Inhibitors Blapsin A (**15**) and Blapsin B (**16**), isolated from *Blaps japanesis*.[41]

A virtual-screening campaign and subsequent experimental validation revealed compound **17** (Figure 12) as a potent inhibitor of 14-3-3(ζ).[43] Compound **17** has a structure similar to that of **15**, except that **17** has a phosphonic acid, which serves as an anchor in the binding pocket of 14-3-3. The authors screened the ZINC *all now* subset in order to find a non-peptidic and non-covalently binding small-molecule inhibitor of the extracellular PPI of 14-3-3 with the membrane receptor aminopeptidase N (APN). In addition to disruption of the 14-3-3(σ)/ APN complex, compound **17** was also found to inhibit the overexpression of

matrix metalloproteinase (MMP)-1 in human fibroblasts, which was induced by signalling of 14-3-3 proteins. The binding mode of **17** to 14-3-3 proteins was validated by solving a cocrystal structure in complex with 14-3-3(σ) (Figure 13).[43]

Compound **17** binds strongly due to the phosphonic acid. Analysis with the programme SeeSAR revealed that both chloride atoms are involved in lipophilic interactions with Arg60, Arg56 and Ala57, the oxygen atoms at the phosphonic acid anchor are involved in strong H-bonding interactions with Arg129 and Arg56. The carbon atom *para* to the phosphonic acid does not contribute to affinity. This problem might be overcome by installing a hydroxyl moiety on that carbon, as was modelled in the SeeSAR programme (see the supporting information Figure S2 and S3).



17

Figure 12. Structure of virtual-screening hit **17**, a potent 14-3-3 inhibitor.[43]

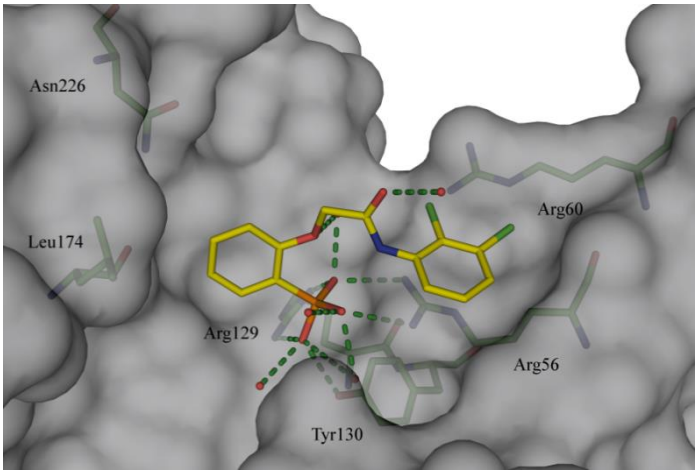


Figure 13. Crystal structure of virtual-screening hit **17** in complex with 14-3-3(σ) (PDB code: 4DHU).[43]

In 2014, Glas and coworkers reported that constrained peptides can serve as inhibitors of pathogenic PPIs. They used target-adapted cross-links to stabilise secondary structures in order to inhibit the interaction of 14-3-3 with virulence vector exoenzyme S (ExoS).[15] The ESp sequence was used as a starting point for the design of macrocyclic structures. This ESp sequence, ^{420}QGLLDALDLAS^{430}, had previously been identified as a binding motif.[14] The authors compared the binding mode of their cross-linked peptide, **18** (Figure 14), to that of ESp by

solving the crystal structure of both ligands with 14-3-3(ζ) (Figure 15 and supporting information Figure S4). Both ligands were shown to lie in the same binding site, **18**, however, has a 20-fold higher binding affinity. The authors suggest, *via* NMR, CD and ITC experiments, that the improved binding entropy for the cross-linked peptide is due to the reduction of conformational flexibility.[15]

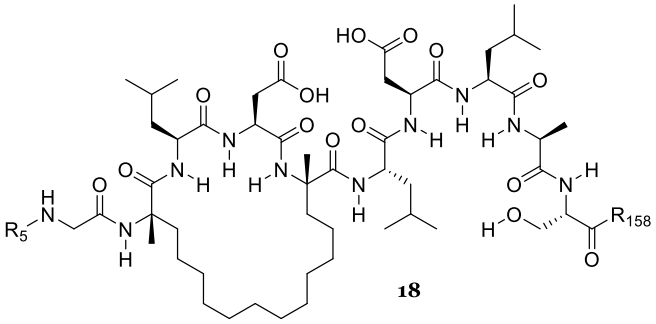


Figure 14. Cocrystal structure of cross-linked peptide **18** with 14-3-3(ζ) (PDB code: 4N84).[15]

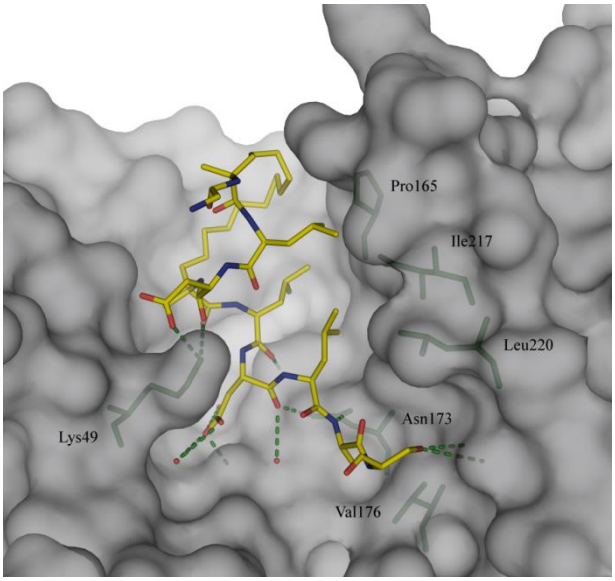


Figure 15. Structure of the constrained peptide **18**, an inhibitor of 14-3-3(ζ).[15]

A novel modulator of 14-3-3/ligand complexes was introduced by Milroy *et al.*. They started with the stabiliser fusicoccin A (FC) (**19**) (Figure 17) in order to

design inhibitors, since they found that the region on 14-3-3 where **19** binds partially overlaps with the region where known inhibitors bind. The part of the active site from which stabilisation is observed lies next to the part where inhibition is observed, that is, there is a small interface between the modulating effects of these parts. The Tau pSer214 epitope was modified with different amines to obtain higher affinity. Milroy *et al.* hypothesised that using the OMe substituent of **19** would give high affinities due to interactions with the FC pocket, consisting of Phe, Met and Lys residues.[19]

The Tau epitope, AcNH-RTPpSLPTP-OH, decorated with compound **20** (Figure 16) showed to have good affinities for 14-3-3 due to the OMe group. Yet, tau epitope expanded with **21** emerged as the most active inhibitor, the bulky aromatic groups increased the affinity. This compound inhibits the binding of phosphorylated full-length Tau (protein) to 14-3-3. It has a 225-fold higher affinity than the Tau epitope. Inhibitors of 14-3-3/Tau might be useful in the treatment of Alzheimer's disease. However, the cocrystal structure they reported is in complex with amine **20,** which has a fifteen-fold worse IC_{50} value than **21** (see the supporting information Figure S5).[19]

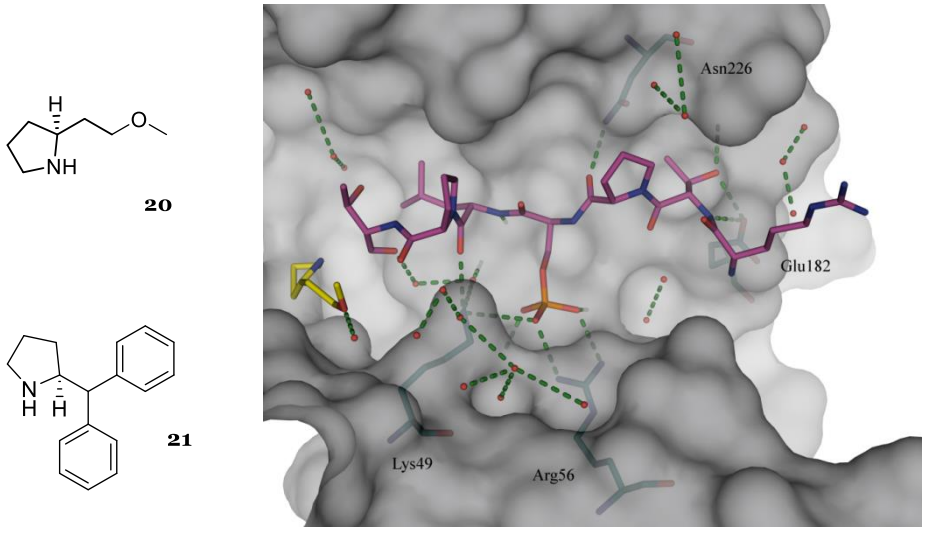


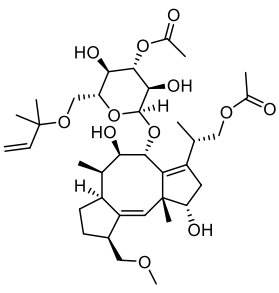
Figure 16. Cocrystal structure of 14-3-3(σ) and phosphorylated Tau epitope with amine **20**, identified using a structure based drug design (SBDD) approach (PDB code: 4Y3B).[19]

2.4 Stabilisers

The natural product fusicoccin A (**19**) was reported as the first stabiliser of PPIs of 14-3-3/ligands. This diterpene glycoside, member of a class of diterpenes bearing a 5-8-5 ring structure called fusicoccanes, was found to bind to 14-3-3 receptors.[44] **19** molecularly glues together the 14-3-3/ plasma membrane ATPase complex (PMA), while it does not bind to the individual partners.[45] The PMA is bound at the C-terminus *via* a phosphorylated Thr. The scaffold of **19** provides a starting point for the design of further stabilisers of 14-3-3/ligand complexes, as it interacts on multiple sites with both protein and ligands.

19 was found to not only stabilise the plant plasma membrane pump but also the binding of 14-3-3 to estrogen receptor alpha (ERα). The ERα is an anticancer target given that stabilising the 14-3-3/ERα complex has a downstream effect and leads to decreased cell proliferation.[18]

Binding of **19** to PMA2 can probably be enhanced by changing the oxygen on the oxane ring for a carbon atom (Figure 17 and supporting information Figures S6–S8). In this case, the programme SeeSAR predicts a better affinity, and Hyde evaluation only reveals favourable interactions. A similar improvement can be achieved in **19** with the 14-3-3/ cystic fibrosis transmembrane conductance regulator (CFTR) complex: the oxygen atom bridging the two rings is found to partially disrupt binding. The binding is predicted to improve when the oxygen is replaced by a carbon atom (Figure 18 and supporting information Figures S9 – S11).



19

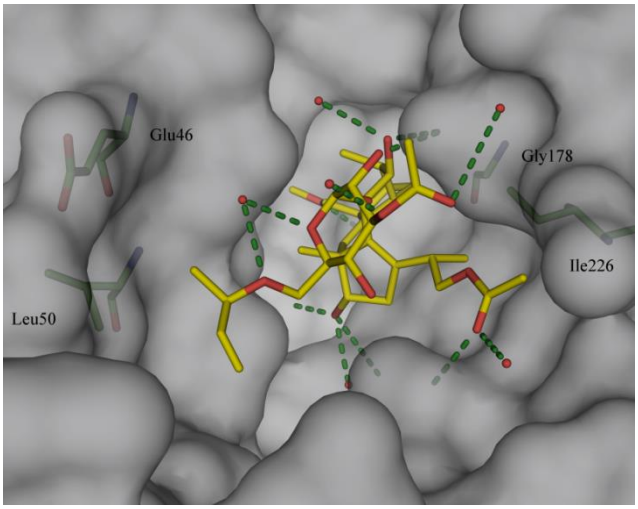


Figure 17. Crystal structure of the stabiliser fusicoccin (**19**)-plasma membrane ATPase complex 2-C52 in the tobacco 14-3-3c isoform (PDB code: 2O98).[46]

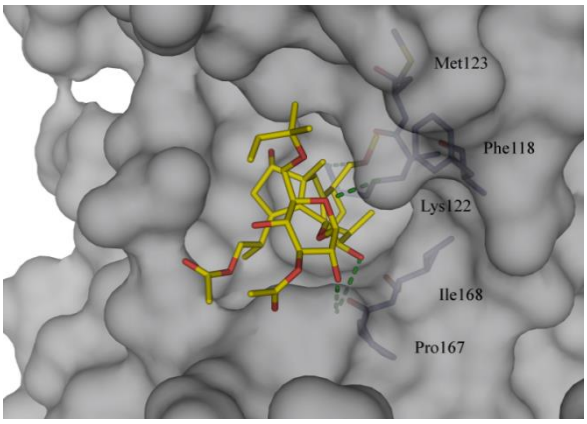
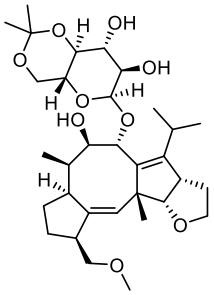


Figure 18. Fusicoccin A (**19**) and cystic fibrosis transmembrane conductance regulator R6-domain with 14-3-3(ζ) (PDB code: 5D3F).[47]

Using **19** as a starting point, Anders *et al.* have made a semisynthetic derivative (**22**) (Figure 19). This derivative targets 14-3-3 in the human potassium channel TASK-3 and interacts in the mode 3 motif of the conserved binding groove. **22** leads to a 19-fold increased binding of 14-3-3 to TASK-3.[48]



22

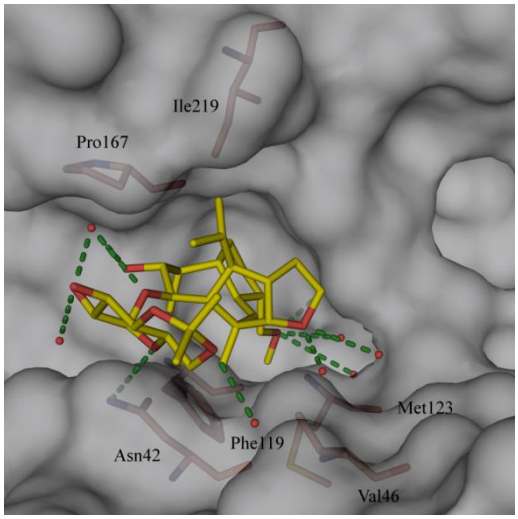
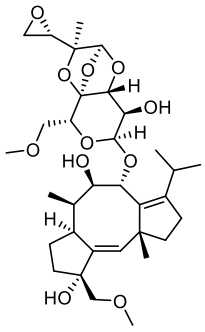


Figure 19. Cocrystal structure of fusiccocin-derivative (**22**) bound to 14-3-3(σ) (PDB code: 3SPR).[48]

This potent stabiliser has some clashes of the oxygen atom of the oxane ring with the carbonyl of Asn42, as observed from analysing the cocrystal structure. This unfavourable interaction can be avoided by replacing the oxygen for a carbon atom, which according to the Hyde-scoring of SeeSAR, would give better

interactions and a better affinity than **22** (see supporting information Figures S12 – S14).

Cotylenin A (CN-A) (**23**) is a natural product closely related to **19**. Compound **23** has anti-tumour activity by stabilising certain PPIs involved in human cancers. It is assumed that the hydroxyl moiety on the cyclopentenyl ring in **19** might determine the structural specificity compared to **23** and related fusicoccanes.[49] Four years after the work of Ottmann, Molzan *et al.* reported that the 14-3-3/C-Raf complex is stabilised by **23** (Figure 20). Cotylenin A interacts with pSer233 and pSer259 of C-Raf. At these positions, 14-3-3 binding inhibits C-Raf, whereas at pSer621, 14-3-3 binding activates C-Raf.[50,51] The *in vivo* activity of **23** on 14-3-3/C-Raf is increased compared to the *in vitro* activity. This difference indicates that other processes are also influenced by **23**.[17]



23

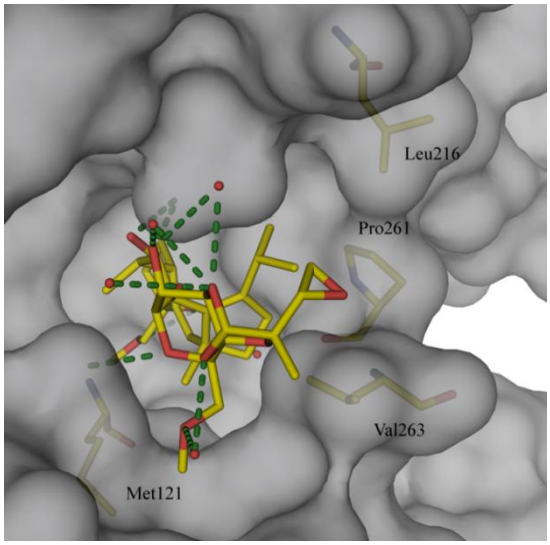
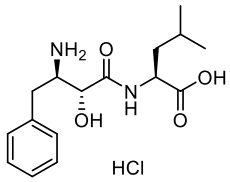


Figure 20. Cocrystal structure of stabiliser cotylenin A (**23**) and C-Raf in 14-3-3(ζ) (PDB code: 4IHL).[17]

Compound **23** displays an unfavourable interaction at the oxygen atom bridging the two ring structures. Substituting this oxygen for a nitrogen atom could increase the affinity by having an additional H-bond towards Asp-213 (Figures S15 – S17 in the supporting information). The two six-membered rings are now less involved in binding, however changing the bridging oxygen atom for a nitrogen atom should lead to an increased affinity.

In 2010, Rose *et al.* identified small-molecule stabilisers of 14-3-3 PPIs, namely a tri-substituted pyrrolidone (**24**), and a dipeptide epibestatin (**25**). The authors screened a 37000-member compound library using green fluorescent protein

(GFP) fused with 14-3-3 against binding to surface-immobilised gluthatione S-transferase fused with PMA2-CT52. The selective stabilisation of 14-3-3/PMA2 by **25** displays a slower association than **19** and **24**, but a similar K_d to **19**. In other words, stabilising the PPI with **25** is relatively slow but once it has been established it will bind strongly (Figure 21).[52] It was found that compound **26** has a stabilising effect on 14-3-3/PMA2 that is three times stronger than that of **24** (Figures 21 and 22), presumably because more contact points with the PMA2 residues are established.[53]



25

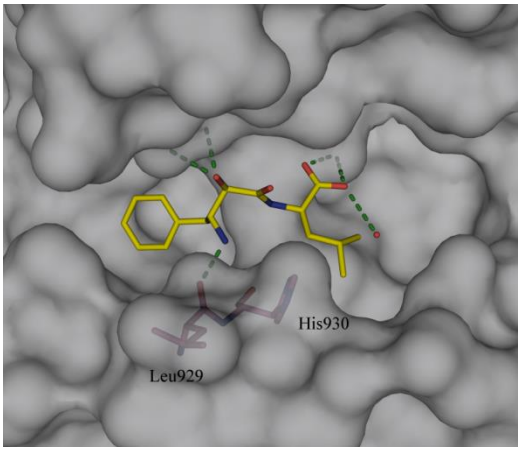
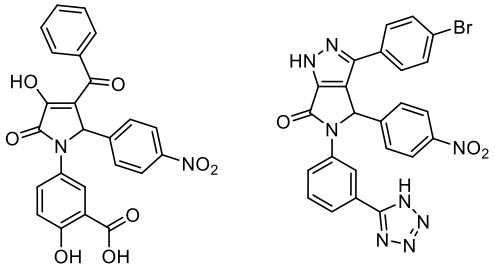


Figure 21. Crystal structure of stabiliser epibestatin (**25**)-plasma membrane ATPase complex 2-CT30 in 14-3-3 (PDB code: 3M50). 25 was identified by screening a compound library against binding of 14-3-3 to plasma membrane ATPase complex 2-C52.[52]

The programme SeeSAR was used to suggest some modifications.[23] Substituting the free amine in epibestatin (**25**) by a carbon atom, for instance, enhances the binding to the PMA moiety. However, this still gives clashes; changing it into a hydroxyl group makes it a bit smaller and again an H-bond acceptor/donor. With an alcohol moiety the affinity most probably increases again, but clashes are still seen with Leu929 and His930 (see Figures S18 – S20 in the supporting information).

The hydroxyl group of the central five-membered ring in **24** appears to have some unfavourable interactions with Val45 and Asn49. Replacing the hydroxyl group for an amine improves the affinity a bit, going to fluoride improves the binding even more (Figures S21 – S23 in the supporting information).



24 **26**

Figure 22. Stabilisers **24** and **26**, identified by screening a compound library against binding of 14-3-3 to plasma membrane ATPase complex 2-CT52.[52]

Mizoribine (**27**), or bredinin (Figure 24), is a compound isolated from *Eupenicillium brefeldianum* and found to have activities in multiple animal studies as well as an immunosuppressive activity.[54] This imidazole nucleoside enhances the interaction of glucocorticoid receptor (GR) with 14-3-3(η) *in vitro*.[55] Mizoribine has been an approved drug in Japan for many years for the treatment of lupus nephritis, rheumatoid arthritis and for immunosuppressive therapy after renal transplantation.[55]

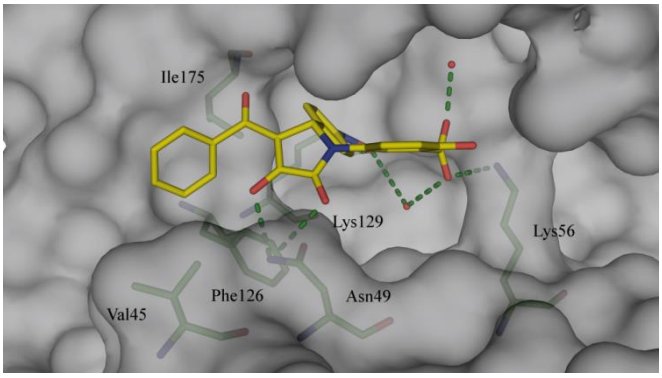
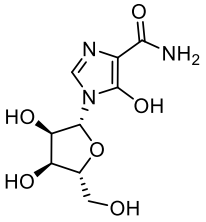


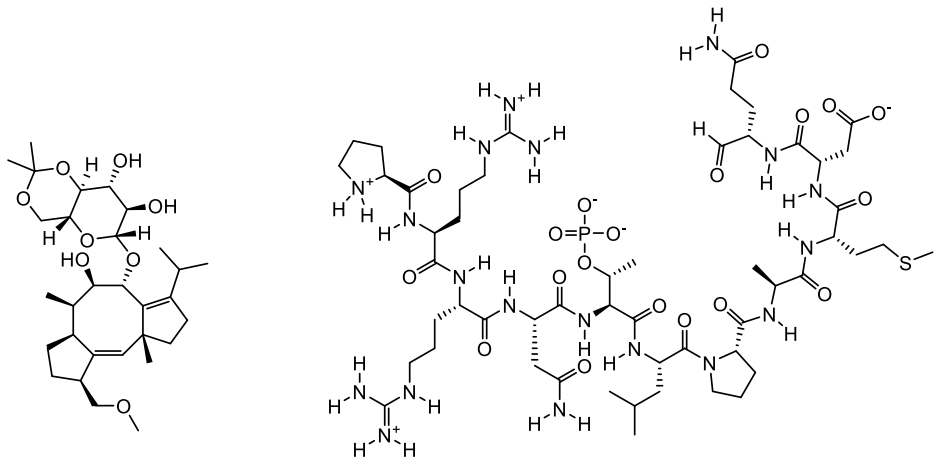
Figure 23. Cocrystal structure of stabiliser **24** and plasma membrane ATPase complex 2-CT30 in 14-3-3 (PDB code: 3M51).[52]



27

Figure 24. Stabiliser mizoribine (**27**), isolated from *Eupenicillium brefeldianum*, stabilises the glucocorticoid receptor (GR) with 14-3-3(η).[54]

The most recently published small-molecule stabiliser **28** holds promise to be used in new therapeutic approaches to target Gab2-related cancers. Gab2 is a signalling protein,[56] which is correlated to the mitogen-activated protein kinase (MAPK) pathways involved in certain cancer types (*e.g.*, breast cancer).[57] Bier *et al.* showed in an elaborate study that the FC-derivative **28** stabilises the binding of pT391 peptides, which are derived from the binding moiety of Gab2 (Figure 25). The stabilisation is specific for the pT391 motif, as no enhanced binding was observed for the other phosphorylation site of Gab2, the Gab2pS210 binding motif. The authors examined the binding interactions of Gab2pT391 and Gab2pS210 *via* cocrystal structures.[58] It was shown that, in the pT391 motif, Leu392 of Gab2 can be best accommodated in the hydrophobic contact surface of Leu172, Ile217 and Leu220 of 14-3-3(ζ). Gab2 motifs are anchored through the phosphonic acid in the binding triad of Arg56, Arg127 and Tyr128, as is frequently observed and described in this chapter. The cocrystal structure of 14-3-3(ζ) with Gab2 and **28** is given in Figure 26. The hydrophobic pocket in 14-3-3(ζ), consisting of the following amino acids Val46, Phe117, Met121, Pro165, Ile166, Gly169 and Ile217, is accommodated by the terpenoid core of **28** and by Leu392 of Gab2. The Gab2pT391 peptide has a sharp turn, most probably induced by Pro393, which therefore gives space for **28** to bind and stabilise the complex. This turn is not observed in Gab2pS210, which leads to the postulation by the authors that the difference of one or two amino acids in the 14-3-3 recognition motifs of Gab2 could be used to design PPI-target-specific stabilisers.[58]



28 **Gab2pT391 peptide**

Figure 25. Stabiliser **28**, which stabilises the complex of Gab2pT391 peptide and 14-3-3(ζ).[58]

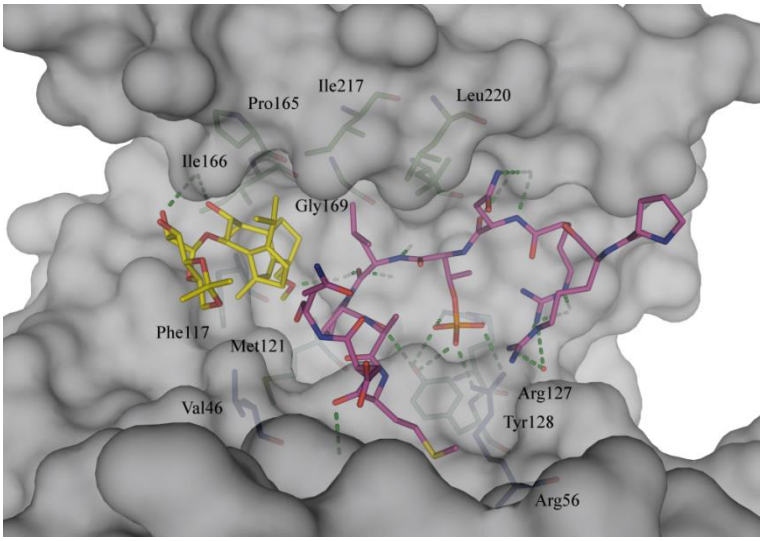


Figure 26. Cocrystal structure of stabiliser **28** and Gab2pT391 peptide in 14-3-3(ζ) (PDB code: 5EXA).[58]

The programme SeeSar was used to analyse the binding affinity of **28** to Gab2pT391 peptide and 14-3-3(ζ). The methoxy carbon atom at the terpenoid core has strong interactions with Met121 and Phe117. The remaining carbon atoms are also involved in hydrophobic interactions with the protein. For example, the isopropyl moiety has strong interactions with Asp213, Ile217 and Leu392. However, the bridging oxygen atom towards the two six-membered rings shows to be non-favourable. It forms an intramolecular H-bond, but it loses affinity due to desolvation. Substituting this oxygen atom for a carbon atom might improve the binding affinity, as this carbon atom could then have lipophilic interactions with Asp213. The terpenoid core should remain in the hydrophobic part, and the two six-membered rings which are solvent exposed move slightly. (Figures S24 – S26 in the supporting information).[59]

2.5 Development in the discovery of modulators of 14-3-3 proteins since 2016

Eighteen years ago Camoni *et al*. have shown that Adenosine 5'-Monophosphate (AMP) inhibits 14-3-3 proteins binding to the plant plasma membrane H^+-ATPase (PMA). At a concentration of 1 mM, 5´-AMP inhibits binding of 14-3-3 to a peptidic analogue of the binding site of the H^+-ATPase.[60] However, it has been shown in 2016 that AMP can also act as a stabiliser. This small molecule was found to stabilise the complex of 14-3-3 with carbohydrate-response element-binding protein (ChREBP). 5´-AMP binds the complex with a K_d of 7.6 ± 0.7 μM,

where under the same conditions no binding was observed to the complex partners alone.[61]

Cyclotides, plant-derived circular peptides, are a promising class of molecules in drug discovery and development. Hellinger *et al.* have used photo-affinity crosslinking of the circular peptide [T20K]kalata B1 (Table 2) and found that it targets 14-3-3 proteins. It was found that this cyclotide exerts a stabilizing effect on the 14-3-3/Foxo3a protein complex. By stabilizing the complex, [T20K]kB1 inhibited the function of Foxo3a with an IC_{50} of 5.3 ± 0.7 μM. A similar way of stabilizing 14-3-3 complexes as is known for fusicoccin-A, was observed for the cyclotide.[62]

Table 2. Cyclotide probe, found to modulate 14-3-3 proteins.

Name	Sequence	Molecular weight (Da) (mono.)
Kalata B1 (kB1)	CGETCVGGTCNTPGCKCSWPVCTRNGLPV	2917.19

In a joint effort, teams of the University of Eindhoven, University of Duisburg-Essen and AstraZeneca reported on the first small-molecule stabiliser of the 14-3-3—p53 interaction in 2017. They showed that fusiccocin-A **19** stabilises the PPI of 14-3-3 with p53, rather than the previously hypothesised opposite. Biophysical experiments show a clear stabilising effect, however the cocrystal structure showed more disorder. The authors hypothesise that this could be due to artefacts of soaking the crystal or that **19** acts allosterically instead of the expected mode-of-action as a `molecular glue´. The binding of p53 C-terminal domain 15mer peptide to 14-3-3 was found to be stabilized by **19** by a factor of 4.5 — 8 (K_d 23.6 ± 2.2 μM in absence of **19**, and 5.40 ± 0.84 μM in presence).[63]

In 2018, Andrei *et al.* have reported on the rational design of new semisynthetic analogues of fusiccocin-A to stabilise PPIs (Figure 27). They included a hydrogen-bond donor by exchanging the 19-acetoxy group for an acetamide, moreover they removed the 3´-acetyl group which did not significantly contributed to the potency. The stabilisation factor S was increased by **29** and **30** up to 16-fold for different 14-3-3 complexes, compared to **19**. In a cell growth assay, using a human ovarian carcinoma cell line, EC_{50} values of 4.3 ± 0.3 μM for **29** and 2.5 ± 0.1 μM for **30** were found.[64]

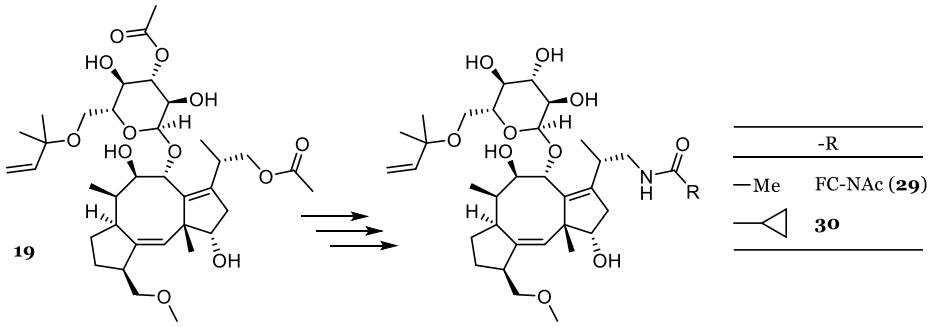


Figure 27. Semisynthetic analogues of fusicoccin-A (19) *via* Structure-based design and molecular dynamics.[64]

In a different study by this author, it was demonstrated that full length Tau protein could be inhibited from binding to 14-3-3 by synthetic peptide analogues of Tau. These compounds bind 14-3-3σ with K_d values in the range of 2.7—20.6 μM.[65]

The groups of Schmuck and Ottmann have recently put forth a new class of supramolecular peptide-based ligands which stabilise 14-3-3ζ and two effectors, Tau and C-Raf. The binding by positively charged compounds is designed around the anionic region in the central binding channel of 14-3-3 (Figure 28). They implemented a guanidiniocarbonyl pyrrole (GCP) moiety, an arginine mimic, to form stable ion pairs with carboxylates and phosphates. Compound **31** was the most potent derivative, enhancing the interaction of 14-3-3ζ to C-Raf and Tau with S factors of 84 and 26. Binding partner C-Raf´s K_d value improved from 16.00 ± 0.60 μM to 0.19 ± 0.01 μM and binding partner Tau´s K_d value improved from 7.00 ± 0.20 μM to 0.27 ± 0.01 μM.[66]

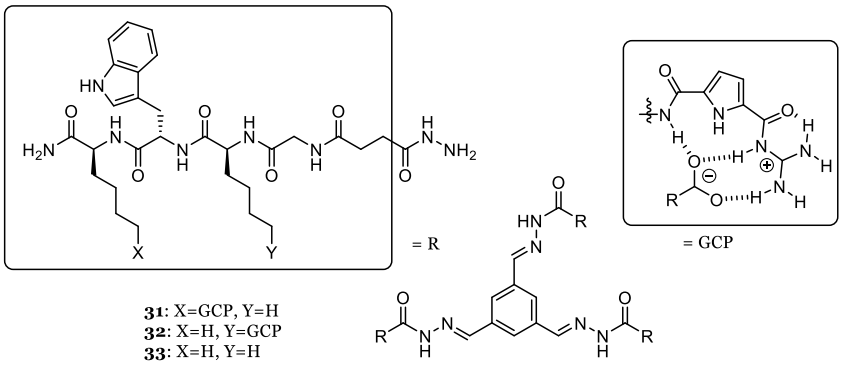


Figure 28. Supramolecular peptide-based ligands, **31**, **32** and **33**, stabilising 14-3-3ζ and two effectors, Tau and C-Raf. Binding was mainly due to the formation of ion pairs through a guanidiniocarbonyl pyrrole moiety with carboxylates and phosphates.[66]

Another recent publication was inspired by **5** (BV02) in order to find chemically more stable compounds. It was found that the bioactive form of **5** is the phtalimide **34**, which is formed *via* a hydration/dehydration pathway (Figure 29). Botta and coworkers have removed a carbonyl group of **34**, which was found to take place in the hydrolysis and made the chemically more stable compound **35**. It was found, in an *in vitro* antiproliferative activity study, that **35** has an IC_{50} value of 7.7 ± 2.0 μM, similar to the IC_{50} value of 5.2 ± 0.7 μM for **34**.[67]

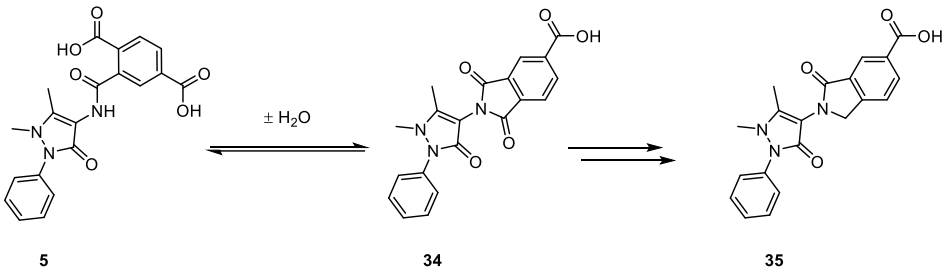


Figure 29. Bioactive form of **5** was shown to be the phtalimide **34**. Botta and coworkers synthesised **35** in order to increase the chemical stability of the small-molecule inhibitor of 14-3-3.[67]

The interaction of 14-3-3 with a phosphorylated motif derived from Estrogen Receptor α (ERα-pp) was used as a PPI complex in a site-directed fragment-based screening campaign. Based on disulfide trapping of fragments, Ottmann and coworkers have found orthosteric stabilisers, increasing the 14-3-3/ERα affinity up to 40-fold (Figure 30). The apparent K_d value of 14-3-3σ/ERα-pp improved from 1.3 μM to 32 nM in presence of **36**.[68]

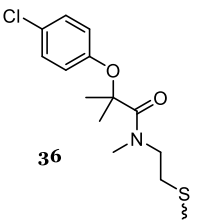


Figure 30. Small-molecule **36**, found *via* disulfide trapping with 14-3-3, stabilising the complex of 14-3-3/ERα.[68]

2.6 Conclusions

Research has focussed, for almost 20 years now, on the development of 14-3-3 inhibitors and stabilisers. We have analysed the structures of the reported inhibitors and stabilisers. Visual inspection of the cocrystal structures of the published small-molecule modulators at the molecular level aided by the programme SeeSAR has led to the proposal of small modifications, which might improve the binding affinity. The knowledge gained from exploring the cocrystal structures sets the stage for structure-based design of (new) modulators. Designing molecules to stabilise or even activate 14-3-3/partner protein complexes has the potential to become an important part of modulating 14-3-3 PPIs, since research has focussed so far more on inhibiting PPIs.

Since there are seven isoforms, with a high degree of conservation, it might be wiser to target the interaction partner rather than 14-3-3. This approach might lead to more specific small-molecule modulators that target a specific PPI instead of all PPIs from the 14-3-3 protein. This would require a paradigm shift, as research has focussed on binding to 14-3-3. Focussing on the interaction partner might give more freedom and creativity when aiming to influence PPIs.

2.7 References

[1] M. Skwarczynska, C. Ottmann, *Future Med. Chem.* **2015**, *7*, 2195–2219.

[2] G. Tzivion, J. Avruch, *J. Biol. Chem.* **2002**, *277*, 3061–3064.

[3] S. Ganguly, J. L. Weller, A. Ho, P. Chemineau, B. Malpaux, D. C. Klein, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 1222–1227.

[4] B. Coblitz, M. Wu, S. Shikano, M. Li, *FEBS Lett.* **2006**, *580*, 1531–1535.

[5] X. Yang, W. H. Lee, F. Sobott, E. Papagrigoriou, C. V Robinson, J. G. Grossmann, M. Sundström, D. A. Doyle, J. M. Elkins, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 17237–17242.

[6] T. Obsil, V. Obsilova, *Semin. Cell Dev. Biol.* **2011**, *22*, 663–672.

[7] H. Fu, R. R. Subramanian, S. C. Masters, *Annu. Rev. Pharmacol. Toxicol.* **2000**, *40*, 617–647.

[8] W. Wang, D. C. Shakes, *J. Mol. Evol.* **1996**, *43*, 384–398.

[9] E. W. Wilkert, R. A. Grant, S. C. Artim, M. B. Yaffe, *J. Biol. Chem.* **2005**, *280*, 18891–18898.

[10] M. B. Yaffe, K. Rittinger, S. Volinia, P. R. Caron, A. Aitken, H. Leffers, S. J. Gamblin, S. J. Smerdon, L. C. Cantley, *Cell* **1997**, *91*, 961–971.

[11] P. Mhawech, *Cell Res.* **2005**, *15*, 228–236.

[12] C. Ottmann, *Bioorganic Med. Chem.* **2013**, *21*, 4058–4062.

[13] M. Molzan, B. Schumacher, C. Ottmann, A. Baljuls, L. Polzien, M. Weyand, P. Thiel, R. Rose, M. Rose, P. Kuhenne, et al., *Mol. Cell. Biol.* **2010**, *30*, 4698–711.

[14] C. Ottmann, L. Yasmin, M. Weyand, J. L. Veesenmeyer, M. H. Diaz, R. H. Palmer, M. S. Francis, A. R. Hauser, A. Wittinghofer, B. Hallberg, *EMBO J.* **2007**, *26*, 902–913.

[15] A. Glas, D. Bier, G. Hahne, C. Rademacher, C. Ottmann, T. N. Grossmann, *Angew. Chemie - Int. Ed.* **2014**, *53*, 2489–2493.

[16] H. Hermeking, C. Lengauer, K. Polyak, T.-C. He, L. Zhang, S. Thiagalingam, K. W. Kinzler, B. Vogelstein, *Mol. Cell* **1997**, *1*, 3–11.

[17] M. Molzan, S. Kasper, L. Röglin, M. Skwarczynska, T. Sassa, T. Inoue, F. Breitenbuecher, J. Ohkanda, N. Kato, M. Schuler, et al., *ACS Chem. Biol.* **2013**, *8*, 1869–1875.

[18] I. J. De Vries-van Leeuwen, D. da Costa Pereira, K. D. Flach, S. R. Piersma, C. Haase, D. Bier, Z. Yalcin, R. Michalides, K. A. Feenstra, C. R. Jiménez, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8894–8899.

[19] L. G. Milroy, M. Bartel, M. A. Henen, S. Leysen, J. M. C. Adriaans, L. Brunsveld, I. Landrieu, C. Ottmann, *Angew. Chemie - Int. Ed.* **2015**, *54*, 15720–15724.

[20] L. G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld, C. Ottmann, *Chem. Rev.* **2014**, *114*, 4695–4748.

[21] G. P. H. van Heusden, *Genomics* **2009**, *94*, 287–293.

[22] A. Aitken, *Semin. Cancer Biol.* **2006**, *16*, 162–172.

[23] BioSolveIT GmbH, SeeSAR 4.1**.**

[24] BioSolveIT GmbH, "HYDE binding assessment," can be found under https://www.biosolveit.de/HYDE/.

[25] I. Reulecke, G. Lange, J. Albrecht, R. Klein, M. Rarey, *ChemMedChem* **2008**, *3*, 885–897.

[26] S. C. Masters, H. Fu, *J. Biol. Chem.* **2001**, *276*, 45193–45200.

[27] B. Wang Yang, H., Liu, Y.S., Jelinek, T., Zhang, L., Ruoslahti, E., *Biochemistry* **1999**, *12499–504*, 12499–12504.

[28] C. Petosa, S. C. Masters, L. A. Bankston, J. Pohl, B. Wang, H. Fu, R. C. Liddington, *J. Biol. Chem.* **1998**, *273*, 16305–16310.

[29] W. Cao, X. Yang, J. Zhou, Z. Teng, L. Cao, X. Zhang, Z. Fei, *Apoptosis* **2010**, *15*, 230–241.

[30] H. Wu, J. Ge, S. Q. Yao, *Angew. Chemie* **2010**, *122*, 6678–6682.

[31] A. Arrendale, K. Kim, J. Y. Choi, W. Li, R. L. Geahlen, R. F. Borch, *Chem. Biol.* **2012**, *19*, 764–771.

[32] V. Corradi, M. Mancini, F. Manetti, S. Petta, M. A. Santucci, M. Botta, *Bioorganic Med. Chem. Lett.* **2010**, *20*, 6133–6137.

[33] V. Corradi, M. Mancini, M. A. Santucci, T. Carlomagno, D. Sanfelice, M. Mori, G. Vignaroli, F. Falchi, F. Manetti, M. Radi, et al., *Bioorganic Med. Chem. Lett.* **2011**, *21*, 6867–6871.

[34] J. Zhao, Y. Du, J. R. Horton, a. K. Upadhyay, B. Lou, Y. Bai, X. Zhang, L. Du, M. Li, B. Wang, et al., *Proc. Natl. Acad. Sci.* **2011**, *108*, 16212–16216.

[35] L. Roglin, P. Thiel, O. Kohlbacher, C. Ottmann, *Proc. Natl. Acad. Sci.* **2012**, *109*, E1051–E1053.

[36] A. K. Upadhyay, J. R. Horton, Y. Du, Y. Bai, X. Cheng, H. Fu, *Proc. Natl. Acad. Sci.* **2012**, *109*, E1054–E1054.

[37] K. D. Park, D. Kim, O. Reamtong, C. Eyers, S. J. Gaskell, R. Liu, H. Kohn, *J. Am. Chem. Soc.* **2011**, *133*, 11320–11330.

[38] S. S. An, P. S. Askovich, T. I. Zarembinski, K. Ahn, J. M. Peltier, M. von Rechenberg, S. Sahasrabudhe, J. J. Fredberg, *Respir. Res.* **2011**, *12*, 8.

[39] M. Sawada, S. I. Kubo, K. Matsumura, Y. Takemoto, H. Kobayashi, E. Tashiro, T. Kitahara, H. Watanabe, M. Imoto, *Bioorganic Med. Chem. Lett.* **2011**, *21*, 1385–1389.

[40] H. Kobayashi, Y. Ogura, M. Sawada, R. Nakayama, K. Takano, Y. Minato, Y. Takemoto, E. Tashiro, H. Watanabe, M. Imoto, *J. Biol. Chem.* **2011**, *286*, 39259–39268.

[41] Y. M. Yan, H. Q. Dai, Y. Du, B. Schneider, H. Guo, D. P. Li, L. X. Zhang, H. Fu, X. P. Dong, Y. X. Cheng, *Bioorganic Med. Chem. Lett.* **2012**, *22*, 4179–4181.

[42] Y. Du, S. C. Masters, F. R. Khuri, H. Fu, *J. Biomol. Screen.* **2006**, *11*, 269–276.

[43] P. Thiel, L. Röglin, N. Meissner, S. Hennig, O. Kohlbacher, C. Ottmann, *Chem. Commun. (Camb).* **2013**, *49*, 8468–8470.

[44] C. Oecking, C. Eckerskorn, E. W. Weiler, *FEBS Lett.* **1994**, *352*, 163–166.

[45] C. Ottmann, S. Marco, N. Jaspert, C. Marcon, N. Schauer, M. Weyand, C. Vandermeeren, G. Duby, M. Boutry, A. Wittinghofer, et al., *Mol. Cell* **2007**, *25*, 427–440.

[46] B. Kostelecky, A. T. Saurin, A. Purkiss, P. J. Parker, N. Q. McDonald, *EMBO Rep* **2009**, *10*, 983–989.

[47] L. M. Stevers, C. V. Lam, S. F. R. Leysen, F. A. Meijer, D. S. van Scheppingen, R. M. J. M. de Vries, G. W. Carlile, L. G. Milroy, D. Y. Thomas, L. Brunsveld, et al., *Proc. Natl. Acad. Sci.* **2016**, E1152–E1161.

[48] C. Anders, Y. Higuchi, K. Koschinsky, M. Bartel, B. Schumacher, P. Thiel, H. Nitta, R. Preisig-Müller, G. Schlichthörl, V. Renigunta, et al., *Chem. Biol.* **2013**, *20*, 583–593.

[49] C. Ottmann, M. Weyand, T. Sassa, T. Inoue, N. Kato, A. Wittinghofer, C. Oecking, *J. Mol. Biol.* **2009**, *386*, 913–919.

[50] N. Dumaz, R. Marais, *J. Biol. Chem.* **2003**, *278*, 29819–29823.

[51] G. Tzivion, Z. Luo, J. Avruch, *Nature* **1998**, *394*, 88–92.

[52] R. Rose, S. Erdmann, S. Bovens, A. Wolf, M. Rose, S. Hennig, H. Waldmann, C. Ottmann, *Angew. Chemie - Int. Ed.* **2010**, *49*, 4129–4132.

[53] A. Richter, R. Rose, C. Hedberg, H. Waldmann, C. Ottmann, *Chem. - A Eur. J.* **2012**, *18*, 6520–6527.

[54] K. Kamata, M. Okubo, E. Ishigamori, Y. Masaki, H. Uchida, K. Watanabe, N. Kashiwagi, *Transplantation* **1983**, *35*, 144–149.

[55] S. Takahashi, H. Wakui, J. a Gustafsson, J. Zilliacus, H. Itoh, *Biochem. Biophys. Res. Commun.* **2000**, *274*, 87–92.

[56] K. Nishida, Y. Yoshida, M. Itoh, T. Fukada, T. Ohtani, T. Shirogane, T. Atsumi, M. Takahashi-Tezuka, K. Ishihara, M. Hibi, et al., *Blood* **1999**, *93*, 1809–16.

[57] T. Brummer, D. Schramek, V. M. Hayes, H. L. Bennett, C. E. Caldon, E. A. Musgrove, R. J. Daly, *J. Biol. Chem.* **2006**, *281*, 626–637.

[58] D. Bier, M. Bartel, K. Sies, S. Halbach, Y. Higuchi, Y. Haranosono, T. Brummer, N. Kato, C. Ottmann, *ChemMedChem* **2016**, *11*, 911.

[59] BioSolveIT GmbH, SeeSAR 5.5.

[60] L. Camoni, S. Visconti, M. Marra, P. Aducci, *J. Biol. Chem.* **2001**, *276*, 31709–31712.

[61] S. Sato, H. Jung, T. Nakagawa, R. Pawlosky, T. Takeshima, W. R. Lee, H. Sakiyama, S. Laxman, R. M. Wynn, B. P. Tu, et al., *J. Biol. Chem.* **2016**, *291*, 10515–10527.

[62] R. Hellinger, K. Thell, M. Vasileva, T. Muhammad, S. Gunasekera, D. Kümmel, U. Göransson, C. W. Becker, C. W. Gruber, *Front. Chem.* **2017**, *5*, DOI 10.3389/fchem.2017.00073.

[63] R. G. Doveston, A. Kuusk, S. A. Andrei, S. Leysen, Q. Cao, M. P. Castaldi, A. Hendricks, L. Brunsveld, H. Chen, H. Boyd, et al., *FEBS Lett.* **2017**, *591*, 2449–2457.

[64] S. A. Andrei, P. de Vink, E. Sijbesma, L. Han, L. Brunsveld, N. Kato, C. Ottmann, Y. Higuchi, *Angew. Chemie Int. Ed.* **2018**, *57*, 13470–13474.

[65] S. A. Andrei, F. A. Meijer, J. F. Neves, L. Brunsveld, I. Landrieu, C. Ottmann, L.-G. Milroy, *ACS Chem. Neurosci.* **2018**, *9*, 2639–2654.

[66] A. Gigante, J.-N. Grad, J. Briels, M. Bartel, D. Hoffmann, C. Ottmann, C. Schmuck, *Chem. Commun.* **2019**, *55*, 111–114.

[67] L. Iralde-Lorente, Y. Cau, L. Clementi, L. Franci, G. Tassone, D. Valensin, M. Mori, A. Angelucci, M. Chiariello, M. Botta, *J. Enzyme Inhib. Med. Chem.* **2019**, *34*, 657–664.

[68] E. Sijbesma, K. K. Hallenbeck, S. Leysen, P. J. de Vink, L. Skóra, W. Jahnke, L. Brunsveld, M. R. Arkin, C. Ottmann, *J. Am. Chem. Soc.* **2019**, *141*, 3524–3531.

Chapter 3

Discovery of small-molecule modulators of 14-3-3 PPIs *via* dynamic combinatorial chemistry

Protein-protein interactions (PPIs) play an important role in many biological processes such as cell-cycle regulation and multiple diseases. The family of 14-3-3 proteins is an attractive target as they serve as binding partner to various proteins, and are therefore capable of regulating their biological activities. Discovering small-molecule modulators of such complexes via traditional screening approaches is a challenging task. Herein, we pioneered the first application of dynamic combinatorial chemistry (DCC) to a PPI, to find modulators of 14-3-3 proteins. The amplified hits from the DCC experiments were evaluated for their binding affinity via surface plasmon resonance (SPR) technique, indicating that they are 14-3-3/synaptopodin PPI stabilisers. Ongoing crystallization studies will hopefully provide with structural knowledge.

A. M. Hartman, W. A. M. Elgaher, N. Hertrich, S. A. Andrei, C. Ottmann and A. K. H. Hirsch, *manuscript submitted*.

A. M. Hartman was involved in the design of the project, performing the DCC experiments, synthesis of compounds, screening via SPR and writing of the manuscript. W. A. M. Elgaher was involved in the screening via SPR and editing of the manuscript, N. Hertrich was involved in the DCC experiments and the synthesis of compounds, S. A. Andrei was involved with the expression and purification of the protein and the fluorescence polarisation assay, C. Ottmann and A. K. H. Hirsch were involved in editing the manuscript and supervision of the project.

3.1 Introduction

The family of 14-3-3 proteins is present in all eukaryotic cell types, and its members are involved in several processes in the human body.[1] They play significant roles ranging from signal transduction, regulation of metabolism to cell-death, and they are correlated to diseases such as Alzheimer's and Noonan Syndrome.[2,3] This is attributed to the ability of 14-3-3 to establish protein–protein interactions (PPIs) with more than 500 protein partners.[4] There are seven human isoforms of 14-3-3 known: beta (β), epsilon (ε), eta (η), gamma (γ), tau (τ), sigma (σ) and zeta (ζ). The binding partners feature three conserved binding motifs for the binding groove in 14-3-3: RSXpSXP (mode 1), RXXXpSXP (mode 2) and pS/TX-COOH (mode 3), where pS denotes a phosphoserine residue.[5–7] 14-3-3 proteins are potential drug targets and an increasing number of chemical classes that modulate 14-3-3 PPIs have been reported, as was listed in Chapter 2. Modulators of 14-3-3 PPIs can be inhibitors, mostly small synthetic molecules, and stabilisers, which include bigger scaffolds (e.g., pyrrolidone1) and (semi-) natural products, e.g., fusicoccin-A.[8–12] Dynamic combinatorial chemistry (DCC) has become an established technique for hit identification. Briefly, it allows a target-based amplification of the best binder(s) from a pool of reacting building blocks and the corresponding products existing under thermodynamic equilibrium. The types of reversible linkages that can be applied in DCC, reaction conditions, and analysis of the dynamic combinatorial library (DCL) have been comprehensively reviewed before.[13–15] In the presented work, we exploited the power of DCC to identify new PPI modulators targeting the 14-3-3(ζ) isoform. For the DCC experiments, we used acylhydrazone formation from the corresponding hydrazide and aldehyde as a reversible reaction. The acylhydrazone linkage can take part in binding with the desired target, as it offers H-bond donor and -acceptor sites. The acylhydrazone formation is sufficiently reversible in acidic media, but also stable against hydrolysis.[16] In basic media the reaction is too slow, therefore high pH values are used to freeze the equilibrium prior to DCL analysis. Bhat *et al.* showed that the use of aniline can accelerate the formation of the equilibrium to only six hours at the near physiological pH value of 6.2.[17] As a prerequisite for the acylhydrazone-based DCC, we studied the stability of our target protein 14-3-3(ζ) under acidic conditions using different buffers and pH values. We found that in MES buffer at pH 6.5, 14-3-3(ζ) is stable at room temperature up to seven days. Therefore, we selected these conditions in the presence of 10 mM aniline as a nucleophilic catalyst for the DCC experiments. In order to confirm that the protein is folded correctly, we measured the circular dichroism (CD) of 14-3-3. The obtained CD-spectrum matches that in literature (supporting information, Figure S4).[18]

3.2 Results and Discussion

The initial design of the DCL and choice of building blocks were inspired by compound **1**, a small-molecule inhibitor of 14-3-3 that was discovered by virtual screening.[19] We envisioned the acylhydrazone linkage between the two aromatic rings, resulting in compound **2** (Figure 1). This modification should maintain the length of the linker between the two aromatic moieties of **1** owing to the restricted flexibility of the acylhydrazone group. Accordingly, compound **2** was synthesised in three consecutive steps starting with a condensation reaction between the commercially available 2-iodobenzhydrazide and 2,3-dichlorobenzaldehyde to form the acylhydrazone **3** (Scheme 1). A palladium-catalysed coupling between the aryl iodide **3** and triethyl phosphite followed, to afford the phosphate ester **4**. This was achieved by first formation of a palladium phosphonate complex at 150 °C, followed by the addition of the acylhydrazone **3** at 100 °C. It was important to lower the temperature before adding the acylhydrazone **3** in order to prevent its decomposition. Deprotection of **4** was achieved by TMSBr, resulting in the target compound **2**. Noteworthy, trials to prepare the hydrazide building block with *o*-phosphonic acid moiety, required for DCC, using the above-mentioned coupling conditions were unsuccessful, and only an intramolecular *N*-arylation product was obtained. We therefore had to use compound **2** in our DCC experiments for *in situ* generation of the corresponding hydrazide and aldehyde.

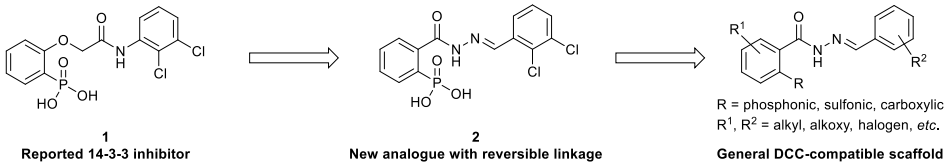
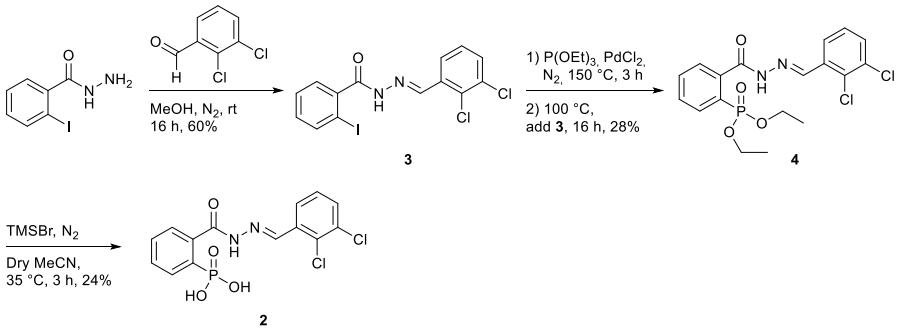


Figure 1. Design of DCL for 14-3-3 PPIs modulation based on the known small-molecule inhibitor **1**.



Scheme 1. Synthetic route towards compound **2**.

We evaluated compound **2** for its biochemical properties *via* fluorescence polarisation (FP) assay and surface plasmon resonance (SPR) (Figure 2 and supporting information Figure S5). The FP assay is based on the decrease of fluorescence polarisation of liberated fluorophore in this case a fluorescently labelled 21 amino acid long peptide (synaptopodin). A known stabilising molecule, fusicoccin, was used as a positive control, it results an increase of FP signal upon increased concentration (Figure 2). In agreement with our design approach, titrating compound **2** to the protein–peptide complex, lowers the signal, indicating an inhibitory effect with an EC_{50} value of 120 μM (Figure 2). Evaluation of the binding affinity using SPR revealed that compound **2** shows low millimolar affinity (K_D value 1.01 mM) towards 14-3-3(ζ) (Figure S5).

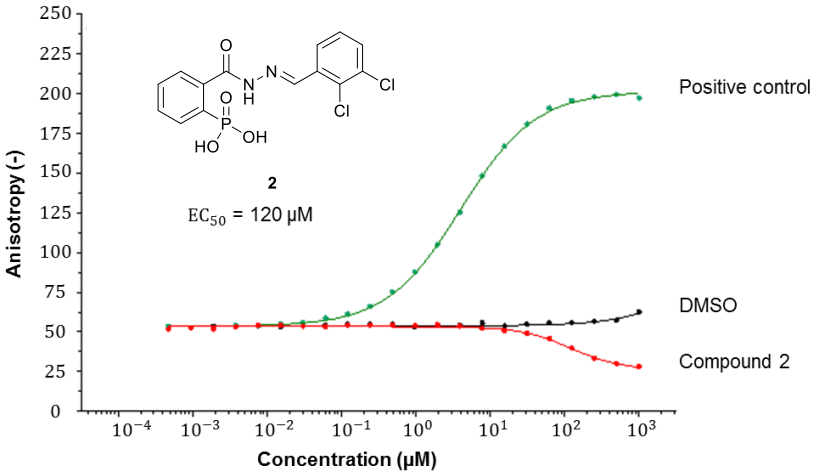
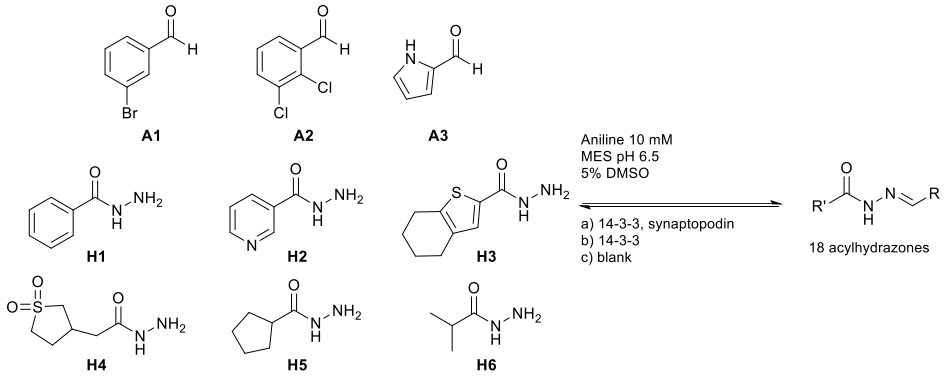


Figure 2. Fluorescent polarisation assay of compound **2**: Titration of **2** to 400 nM 14-3-3ζ with 10 nM fluorescently labelled peptide synaptopodin, resulting in the displacement of the peptide (EC_{50} 120 μM). Data obtained from single measurement.

Encouraged by these results we designed a DCL based on compound **2** (Figure 1 and Scheme S1). However, when we ran the DCC experiment, not all of the possible acylhydrazones bearing *o*-phosphonic or *o*-sulfonic acid moieties could be observed by LC-MS analysis (Supporting information Figure S1). We therefore modified the DCL by omitting the acidic motifs as shown in Scheme 2. The DCL consisted of 3 aldehydes **A1**–**A3** and 6 hydrazides **H1**–**H6**. Consequently, we ran three DCC experiments: (a) a library in which building blocks were present in combination with the 14-3-3(ζ)/unlabelled synaptopodin complex as a PPI model; (b) a library containing the building blocks and 14-3-3 protein; and (c) a 'blank' in which only building blocks were present (Scheme 2).

The DCLs were allowed to equilibrate for 6 h and were analysed *via* HPLC-MS, resulting in the chromatograms shown in Figure 3. The most obvious differences

are the two peaks at retention times of 13.5 and 13.9 minutes corresponding to the compounds **A1H3** and **A2H3**, respectively. In the presence of 14-3-3(ζ) only and in the PPI-complex (14-3-3(ζ)/synaptopodin), these two acylhydrazones show a significant amplification (> 2-fold, 148%) compared to the DCL in the absence of protein. Table 1 shows the ratios of the relative areas of each peak compared to the blank DCL. These two hits were then synthesised to confirm the identity of peaks (Figure S3) and for biochemical characterisation. Synthesis was accomplished through the reaction of the hydrazide **H3** with a stoichiometric amount of the appropriate aldehyde at room temperature overnight to afford the corresponding acylhydrazones in a good to quantitative yield (Scheme 3).



Scheme 2. Dynamic combinatorial library (DCL) of acylhydrazones with aldehydes (100 μM each), hydrazides (300 μM each), DMSO (5%), aniline (10 mM) and: a) 14-3-3(ζ) (10 μM) and synaptopodin (10 μM); b) control with 14-3-3(ζ) (10 μM); and c) control without protein or synaptopodin.

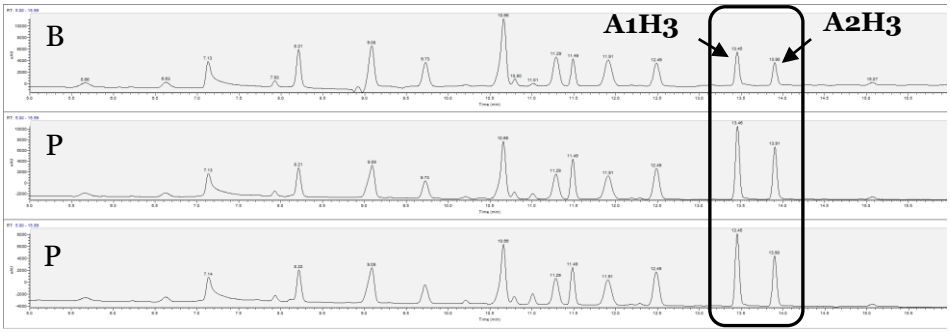
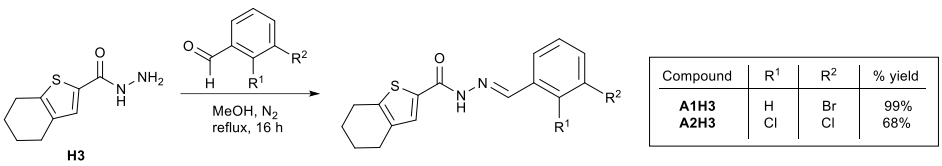


Figure 3. UV-chromatograms at 290 nm of DCLs in duplicates, blank (B), in the presence of protein (P), and in the presence of protein plus synaptopodin (PS) at 6 h. Compounds **A1H3** and **A2H3** are amplified in P as well as PS compared to B.

Table 1. Amplification factors of the products formed in the DCC experiments analysed *via* the relative surface areas of peaks in the UV-chromatograms. The experiments were performed in duplicate and the average values were taken.

Compound	Retention time (min)	Amplification factor ((RPA$_{protein}$– RPA$_{blank}$)/RPA$_{blank}$)*100 %	Amplification factor ((RPA$_{complex}$– RPA$_{blank}$)/RPA$_{blank}$)*100%
A3H2	4.76	–59.0	–44.0
A3H4	5.66	–33.9	–35.9
A3H6	6.62	–25.7	–16.4
A3H1	7.93	–33.1	–18.1
A1H2 & A3H5	8.21	–32.3	–35.1
A2H2 & A1H4	9.09	–28.0	–26.0
A2H4	9.73	–36.7	–32.9
A1H1 & A1H6	10.66	–20.4	–23.1
A3H3	10.79	–1.3	–15.0
A2H6	11.29	–22.5	–23.6
A2H1	11.49	32.2	29.5
A1H5	11.91	–12.2	–16.0
A2H5	12.49	12.8	14.6
A1H3	13.46	148.0	148.4
A2H3	13.91	148.5	153.6



Scheme 3. Synthetic route towards compounds **A1H3** and **A2H3**.

We used SPR binding assays to follow binding events of the synthesised DCC hits to the 14-3-3(ζ) protein. We checked first whether the immobilised protein is still in the native folded state and can engage in PPIs by determining the binding affinity (K_D) of synaptopodin to 14-3-3(ζ). The small peptide showed a clear binding response with the same K_D value of 1.38 μM obtained from either the Langmuir isotherm or the kinetic curves (Supporting information, Figure S6). Encouraged by this result, we determined the affinity and binding kinetics of the hit compounds in a similar manner. The acylhydrazones **A1H3** and **A2H3** showed low micromolar affinity to 14-3-3(ζ) (K_D values 16 and 15 μM, respectively). Interestingly, compound **A2H3** exhibited about 4.5-fold slower on-

rate and longer residence time compared to **A1H3** although they have similar binding affinities (Table 2 and supporting information, Figure S7 and S8).

Table 2. Kinetic parameters of hit compounds **A1H3** and **A2H3** to 14-3-3(ζ). Data obtained from single experiments.

Compound	R_{max} (RU)	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_D (μM)	Res sd
A1H3	4.6 ± 0.2	2.6 ± 0.2 ×10^3	0.041 ± 0.001	16 ± 1	0.38
A2H3	4.5 ± 0.2	5.8 ± 0.5 ×10^2	0.0087 ± 0.0004	15 ± 1	0.50

R_{max}: maximum analyte binding capacity; k_{on}: association rate constant; k_{off}: dissociation rate constant; K_D: equilibrium dissociation constant; Res sd: residual standard deviation.

Intrigued by these findings, we next investigated the mode of binding of the new 14-3-3 PPIs modulators (compounds **2**, **A1H3**, and **A2H3**) by SPR competition assays using synaptopodin as a reference, which occupies the 14-3-3 main binding pocket. Modulators that inhibit 14-3-3 PPIs bind in the phosphorylation binding pocket, whereas stabilisers bind allosterically to the binding pocket or at the interface between 14-3-3 and its protein partners.[20] To check whether the compounds bind to the active site or elsewhere, compound **2** (1000 μM), synaptopodin (1 μM), and a mixture of both at the same concentration were injected in sequence over immobilised 14-3-3(ζ). The obtained response unit (RU) value of the mixture was compared to the theoretical sum of the RU values for the individual compounds. If **2** bound to the main binding pocket of 14-3-3, it would compete with synaptopodin and the response of the mixture should be less than the sum of RU values determined for the single compounds. On the other hand, if **2** bound allosterically to the active site, no competition would occur and the response of the mixture should be equal to the sum of RU values of the individual compounds. We found that the RU value of **2** in combination with synaptopodin was less than the sum of the individual responses (Figure 4 and S9). This suggests that compound **2** competes with synaptopodin for the same binding pocket, however no compound could completely displace the other at the tested concentrations. To verify this, the sum of responses was recalculated considering a fractional occupancy (FO) of **2** and synaptopodin according to Perspicace *et al.*[21] Indeed, the experimental RU value of the mixture was equal to the new estimated one (Figure 4), indicating that the compounds compete for the same binding site. The same results were obtained using different concentrations of **2** and synaptopodin (1000 μM vs 25 μM; and 200 μM vs 1 μM), respectively (Figures S10 and S11). In accordance with the result of the FP-assay,

these results clearly indicate that compound **2** binds to the active site of 14-3-3, leading to disruption of PPIs.

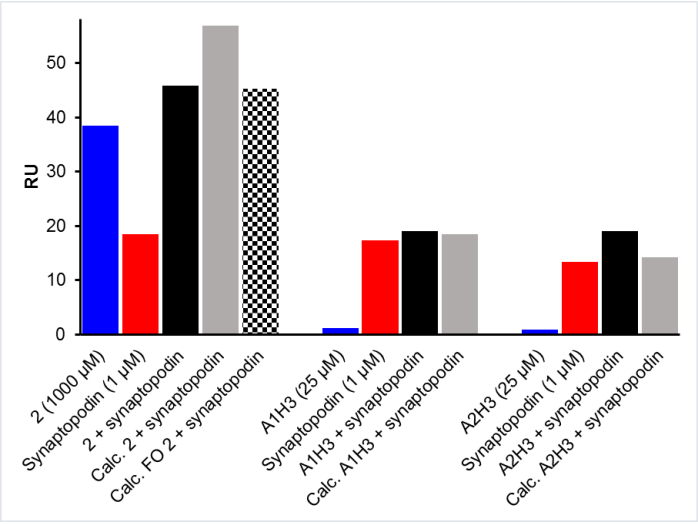


Figure 4. SPR responses of compounds **2**, **A1H3**, and **A2H3** in the competition assays, using synaptopodin as a reference compound binding to the active site of 14-3-3.

Using the same approach, the DCC hit compounds **A1H3** (25 µM) and **A2H3** (25 µM) were injected alone and as a mixture with synaptopodin (1 µM) over immobilised 14-3-3, and their binding responces were analysed (Figure 5).

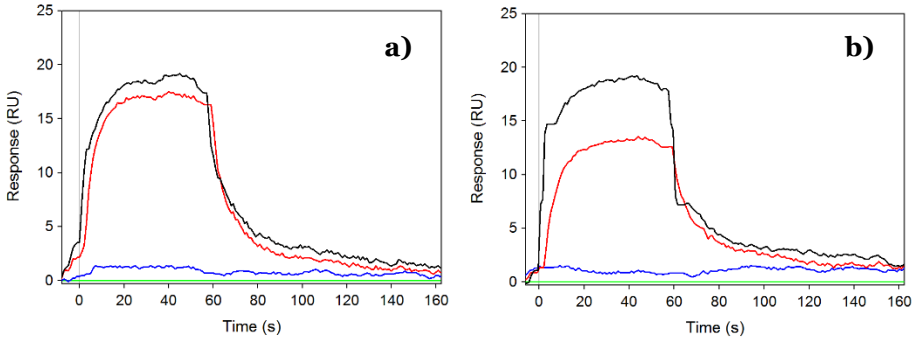


Figure 5. SPR competition assay: (a) Sensorgram overlay of **A1H3** (25 µM, blue), synaptopodin (1 µM, red) and a **A1H3**–synaptopodin mixture (black), showing an additive effect, indicating a non-competitive binding; (b) Overlay of sensorgrams of **A2H3** (25 µM, blue), synaptopodin (1 µM, red) and a **A2H3**–synaptopodin mixture (black), showing a synergistic effect, indicating a non-competitive binding and a stabilising effect. Curves obtained from single experiments.

Interestingly, the RU values of the mixtures containing **A1H3** or **A2H3** with synaptopodin were equal to or more than the sum of the individual responses (Figure 4 and 5). This indicates that these two compounds bind to 14-3-3 in a different pocket than that of synaptopodin. Moreover, the increased binding response of the mixture compared to the calculated sum indicates a stabilising effect of the acylhydrazones to the complex of synaptopodin with 14-3-3(ζ). This confirms our findings from the DCC experiments, where the same amplification factors for these hits were obtained in the presence of 14-3-3 alone as well as the 14-3-3(ζ)/synaptopodin complex (Table 1). Therefore, their binding site could be allosteric to the main binding pocket or at the interface of the 14-3-3(ζ)/synaptopodin complex. Co-crystallisation studies are ongoing and would help to confirm the exact binding pocket of the 14-3-3 protein.

3.3 Conclusions

We set out to develop novel PPI-modulators targeting the versatile 14-3-3 protein family. Firstly, we pursued a ligand-based design of the acylhydrazone **2**, which turned out to be an inhibitor of 14-3-3 PPIs. Next, we applied a DCC approach, using 14-3-3(ζ) in complex with synaptopodin as a PPI model, resulting in the discovery of two modulators **A1H3** and **A2H3**. No significant change was observed in the DCL composition in the presence of only 14-3-3 compared to the 14-3-3(ζ)/synaptopodin complex, indicating that the hit compounds bind independently to a different site than the main binding-groove of 14-3-3 involved in PPIs. Finally, we determined the binding affinities and kinetics of the hits *via* SPR and investigated their binding site on 14-3-3. Results of the SPR competition assays support our initial findings from the DCC experiments, suggesting that these small molecules stabilise the 14-3-3(ζ)/synaptopodin complex either directly or allosterically. Cocrystallisation studies are ongoing in order to determine their exact binding site.

3.4 Experimental

3.4.1 Materials and methods

Chemicals were purchased from commercial suppliers and used without pretreatment. Solvents used for the experiments were reagent-grade and dried, if necessary, according to standard procedures. The reactions were performed under nitrogen atmosphere, unless otherwise stated. The yields were calculated for the analytically pure compounds and were not optimised. The purifications were performed using column chromatography with Macherey-Nagel Silica 60 M 0.04–0.063 mm. Preparative high-performance liquid chromatography (HPLC, Ultimate 3000 UHPLC+ focused, Thermo Scientific) purification was performed on a reversed-phase column (C18 column, 5 µm, Macherey-Nagel, Germany). The solvents used for the chromatography were water (0.1 % formic acid) and MeCN

(0.1% formic acid), or EtOAc and DCM. ^{1}H-, ^{13}C and ^{31}P-NMR spectra were measured on a Bruker Fourier 500 spectrometer (500, 126 or 202 MHz), respectively. The chemical shifts were reported in parts per million (ppm) relative to the corresponding solvent peak. The coupling constants of the splitting patterns were reported in Hz and were indicated as singlet (s), doublet (d), triplet (t) and multiplet (m). Due to the presence of isomers for acylhydrazones, some of the signals are doubled. UPLC-MS and HRMS measurements were performed using Thermo Scientific systems, see supporting information.

3.4.2 DCC conditions

The corresponding acylhydrazides (each 3 µL, stock solutions 100 mM in DMSO) and the aldehydes (each 1 µL, stock solutions 100 mM in DMSO) were added to a MES buffer (0.1 M, 2 mM $MgCl_2$, pH 6.5). Aniline (10 µL, stock solution 1 mM), synaptopodin (10 µL, stock solution 1 mM) and 14-3-3(ζ) (5.35 µL, stock solution 1.87 mM) were added accordingly. DMSO was added to reach a final concentration of DMSO in the DCL of 5%. The end-volume was 1 mL. Final concentrations in the DCLs are shown in Table 3. The DCL was left shaking at room temperature and was frequently monitored *via* UPLC-MS. The reactions were performed in 1.5 mL Eppendorf cups and stirred on a Rotating Mixer (Benchmark Scientific).

Table 5. Final concentrations in the DCLs.

	14-3-3/synaptopodin	14-3-3	Blank
DMSO	5%	5%	5%
Aniline	10 mM	10 mM	10 mM
Aldehyde	100 µM (each)	100 µM (each)	100 µM (each)
Hydrazide	300 µM (each)	300 µM (each)	300 µM (each)
Protein	10 µM	10 µM	-
Synaptopodin	10 µM	-	-

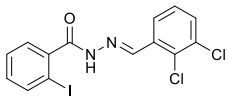
For monitoring *via* UPLC-MS, 10 µL of the corresponding library was withdrawn and diluted in 90 µL acetonitrile, the pH was raised to pH > 8 by adding 2 µL NaOH (1.0 M) to freeze the reaction. The mixture was centrifuged at 10.000 rpm for 2 minutes, and the supernatant was analysed *via* UPLC-MS.

3.4.3 Synthesis

General procedure for acylhydrazone formation:[16]

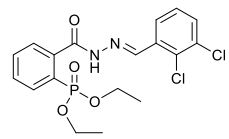
To the hydrazide (1 eq.) dissolved in MeOH, the corresponding aldehyde (1.2 eq.) was added. The reaction mixture was stirred at room temperature or refluxed until completion. After cooling to room temperature, the reaction mixture was concentrated *in vacuo*. Purification was performed by column chromatography, affording the corresponding acylhydrazone in 60 – 99% yield.

(*E*/*Z*)-*N*'-(2,3-dichlorobenzylidene)-2-iodobenzohydrazide (3)



The acylhydrazone was synthesised at room temperature following the general procedure, using 2-iodobenzohydrazide (1.00 g, 3.82 mmol) in MeOH (7.6 mL) and 2,3-dichlorobenzaldehyde (0.607 g, 3.47 mmol). After purification by column chromatography (DCM/EtOAc 1:0 → 0:1), the acylhydrazone **3** was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (866 mg, 60%). ^{1}H-NMR (500 MHz, DMSO-d$_6$) δ= 12.22 (s, 1H, *Z*), 12.15 (s, 1H, *E*), 8.73 (s, 1H, *E*), 8.49 (s, 1H, *Z*), 8.07 – 7.14 (m, 7H, mixture *E*/*Z*); ^{13}C-NMR (126 MHz, DMSO-d$_6$) (combined peaks of *E*/*Z*) δ= 171.03, 165.11, 143.64, 141.92, 140.86, 139.82, 139.27, 138.35, 133.77, 133.73, 132.41, 131.79, 131.59, 131.38, 131.08, 130.81, 130.57, 128.63, 128.57, 128.46, 128.28, 128.19, 127.80, 125.56, 124.85, 94.02, 93.59, 40.11, 40.02, 39.94, 39.85, 39.78, 39.69, 39.61, 39.52, 39.35, 39.19, 39.02; HRMS (ESI) calcd for $C_{14}H_9Cl_2IN_2O$ [*M*+H]$^+$: 418.9209, found 418.9190.

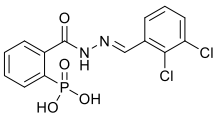
Diethyl (*E*/*Z*)-(2-(2-(2,3-dichlorobenzylidene)hydrazine-1-carbonyl)phenyl)phosphonate (4)



The phosphonic ester was synthesised according to an adapted protocol of Benin *et al.*[22] Triethyl phosphite (1.3 mL, 7.6 mmol) was added to [$PdCl_2$] (16.9 mg, 0.10 mmol) under N_2 atmosphere and heated at 150 °C for 3 h. The mixture was cooled to 100 °C, and compound **3** (400 mg, 0.95 mmol) was added. The reaction mixture was stirred for two days until complete conversion was observed. Water and DCM were added after the mixture had cooled down to room temperature. The organic layer was washed with water (3 x 5 mL), a saturated aqueous solution of $NaHCO_3$ (1x), dried over $MgSO_4$, filtered, and the solvent was evaporated *in vacuo*. The crude material was purified by column chromatography (SiO_2; DCM/EtOAc 3:1 → 0:1). The product **4** was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a colourless oil (113 mg, 28%). ^{1}H-NMR (500 MHz, DMSO-d$_6$) δ= 12.11 (s, 1H, *Z*), 12.01 (s, 1H, *E*), 8.69 (s, 1H, *E*), 8.41 (s, 1H, *Z*), 8.03 – 7.18 (m, 7H, mixture *E*/*Z*), 4.12 – 3.77 (m, 4H,

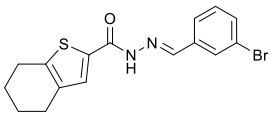
mixture *E*/*Z*), 1.16 (dt, *J* = 53.0, 7.0 Hz, 6H, mixture *E*/*Z*); ^{13}C-NMR (126 MHz, DMSO-d_6) (combined peaks of *E*/*Z, and coupling with* ^{31}*P*) δ= 170.9, 170.9, 164.6, 164.6, 142.9, 140.2, 140.1, 139.3, 139.2, 138.7, 134.0, 133.8, 133.0, 133.0, 132.5, 132.4, 132.4, 132.4, 132.3, 132.0, 132.0, 131.7, 131.1, 131.0, 130.6, 129.8, 129.7, 128.8, 128.7, 128.6, 128.3, 128.2, 127.6, 127.5, 127.4, 126.1, 125.9, 125.5, 124.7, 124.6, 62.0, 61.9, 61.7, 61.6, 40.1, 40.0, 40.0, 39.9, 39.8, 39.7, 39.5, 39.4, 39.2, 39.0, 16.1, 16.0, 15.9, 15.9; ^{31}P-NMR (202 MHz, DMSO-d_6) δ= 15.55 (dd, *J* = 45.1, 4.5 Hz); HRMS (ESI) calcd for $C_{18}H_{19}Cl_2N_2O_4P$ [*M*+H]$^+$: 429.0517, found 429.0516.

(*E*/*Z*)-(2-(2-(2,3-Dichlorobenzylidene)hydrazine-1-carbonyl)phenyl)phosphonic acid (2)



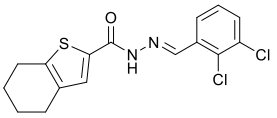
The phosphonic acid was obtained by deprotection of the ester by TMSBr. Compound **4** (93 mg, 0.2 mmol) was dissolved in 2.2 mL dry MeCN under N_2 atmosphere. To this mixture, was added TMSBr (143 µL, 1.08 mmol) and stirred at 35 °C for 3 h. The solvent was evaporated *in vacuo*, water/THF (1:1) was added, and the mixture was stirred at room temperature for 1 h. The crude mixture was lyophilised and purified by reversed-phase prep-HPLC (water (0.1% formic acid)/MeCN (0.1% formic acid), 9:1 → 1:9). Product **2** was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 7:3) as a white solid (21.2 mg, 24%). ^{1}H-NMR (500 MHz, DMSO-d_6) δ= 12.33 (s, *E*, 1H), 11.94 (s, *Z*, 1H), 8.68 (s, *E*, 1H), 8.37 (s, *Z*, 1H), 8.05 – 7.18 (m, 7H); ^{13}C-NMR (126 MHz, DMSO-d_6) δ= 171.6, 168.0, 142.9, 138.2, 137.6, 134.0 (d, *J* = 11.0 Hz), 132.4, 132.2, 131.6, 130. 9 (d, *J* = 16.6 Hz), 129.7 (d, *J* = 12.6 Hz), 128.5, 128.2, 125.4; ^{31}P-NMR (202 MHz, DMSO-d_6) δ= 10.6 (s); HRMS (ESI) calcd for $C_{14}H_{11}Cl_2N_2O_4P$ [*M*+H]$^+$: 372.9906, found 372.9902.

(*E*/*Z*)-*N*'-(3-Bromobenzylidene)-4,5,6,7-tetrahydrobenzo[b]thiophene-2-carbohydrazide (A1H3)



The acylhydrazone was synthesised by following the general procedure by using 4,5,6,7-tetrahydro-1-benzothiophene-2-carbohydrazide (43.5 mg, 0.22 mmol) in MeOH (1.6 mL) and 3-bromobenzaldehyde (40 µL, 0.34 mmol) and heating to reflux. After purification *via* column chromatography (DCM/EtOAc, 9:1 → 1:9), the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (81 mg, quantitative). ^{1}H-NMR (500 MHz, $CDCl_3$) δ= 9.35 (s, *E*, 1H), 7.92 (s, 1H), 7.89 – 7.74 (m, 1H), 7.70 (d, *J* = 7.7 Hz, 1H), 7.53 (d, *J* = 8.0 Hz, 1H), 7.31 (t, *J* = 7.8 Hz, 1H), 2.86 (t, *J* = 5.8 Hz, 2H), 2.67 (t, *J* = 5.8 Hz, 2H), 1.92 – 1.79 (m, 4H); ^{13}C-NMR (126 MHz, $CDCl_3$) δ=136.0, 133.2, 130.6, 130.5, 126.3, 123.1, 25.4, 25.4, 23.4, 22.8; HRMS (ESI) calcd for $C_{16}H_{15}BrN_2OS$ [*M*+H]$^+$: 363.0161, found 363.0158.

(*E*/*Z*)-*N*'-(2,3-Dichlorobenzylidene)-4,5,6,7-tetrahydrobenzo[b]thiophene-2-carbohydrazide (A2H3)



The acylhydrazone was synthesised by following the general procedure by using 4,5,6,7-tetrahydro-1-benzothiophene-2-carbohydrazide (34.4 mg, 0.18 mmol) in MeOH (1.6 mL) and 2,3-dichlorobenzaldehyde (27.8 mg, 0.16 mmol) and heating to reflux. After purification *via* column chromatography (DCM/EtOAc, 9:1 → 1:9), the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (38.2 mg, 68%). ^{1}H-NMR (500 MHz, $CDCl_3$) δ= 8.97 (s, *E*, 1H), 8.27 (s, *Z*, 1H), 8.10 (d, *J* = 7.9 Hz, 1H), 7.85 (s, 1H), 7.51 (dd, *J* = 7.9, 1.4 Hz, 1H), 7.31 (t, *J* = 7.9 Hz, 1H), 2.85 (t, *J* = 6.0 Hz, 2H), 2.66 (t, *J* = 6.0 Hz, 2H), 1.92 – 1.77 (m, 4H);^{13}C-NMR (126 MHz, $CDCl_3$) δ=133.5, 132.5, 131.7, 127.8, 126.2, 25.5, 25.4, 23.4, 22.8; HRMS (ESI) calcd for $C_{16}H_{14}Cl_2N_2OS$ $[M+H]^+$: 353.0277, found 353.0274.

3.4.4 Protein expression and purification

14-3-3ζ was expressed according to the general protocol as was published by Andrei *et al.*.[23]

3.4.5 Fluorescence polarisation assay (FP)

The performed FP assays were based on the protocol that was published by Andrei *et al.*.[23]

FP assays were performed in 10 mM HEPES, pH 7.4, 150 mM NaCl, 1.0 mg/mL BSA and 0.01% v/v Tween-20 buffer. Fluorescently labelled peptides were dissolved to 10 nM in FP buffer as a mastermix. The desired compounds were then added from DMSO stock as required, to a final DMSO content of 1%. This solution was used to fill Corning Low-binding Black Round Bottom 384-well plates (Corning #4514) with a final volume of 10 µL per well. A two-fold dilution series was then performed with 14-3-3(ζ), starting from 243 µM. FP was then measured in a Tecan Infinite F500 platereader, using 485 (20) nm excitation and 535 (20) nm emission filters. The obtained anisotropy values were then plotted and fitted against a 4-parameter one-site binding model in GraphPad Prism.

3.4.6 Binding studies by surface plasmon resonance (SPR)

The SPR experiments were based on the procedure that was published by Henn et al..[24]

The SPR experiments were performed using a Reichert SR7500DC surface plasmon resonance spectrometer (Reichert Technologies, Depew, NY, USA), and medium density carboxymethyl dextran hydrogel CMD500M sensor chips

(XanTec Bioanalytics, Düsseldorf, Germany). Double distilled (dd) water was used as the running buffer for immobilisation. HBS-EP buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.005% v/v tween 20, pH 7.4) containing 5% v/v DMSO was used as the running buffer for binding study. All running buffers were filtered and degassed prior to use. The 14-3-3 (30.3 kDa) was immobilised in one of the two flow cells by standard amine-coupling procedure. The other flow cell was left blank to serve as a reference. The system was initially primed with borate buffer 100 mM (pH 9.0), then the carboxymethyldextran matrix was activated by a 1:1 mixture of *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) 100 mM and *N*-hydroxysuccinimide (NHS) 100 mM at a flow rate of 10 µL/min for 7 min. The 14-3-3 was diluted to a final concentration of 57 µg/mL in 10 mM sodium acetate buffer (pH 4.5) and was injected at a flow rate of 10 µL/min for 7 min. Non-reacted surface was quenched by 1 M ethanolamine hydrochloride (pH 8.5) at a flow rate of 25 µL/min for 3 min. A series of 10 buffer injections was run initially on both reference and active surfaces to equilibrate the system resulting in a stable immobilisation level of approximately 2500 µ refractive index unit (µRIU). Binding experiments were performed at 20 °C. Compounds dissolved in DMSO were diluted with HBS-EP buffer (final DMSO concentration of 5% v/v) and were injected at a flow rate of 30 µL/min. Single-cycle kinetics were applied for K_D determination. The association time was set to 60 s, and the dissociation phase was recorded for 120 s. Ethylene glycol 80% in the running buffer was used for regeneration of the surface. Differences in the bulk refractive index due to DMSO were corrected by a calibration curve (nine concentrations: 3–7% v/v DMSO in HBS-EP buffer). Data processing and analysis were performed by Scrubber software (Version 2.0c, 2008, BioLogic Software). Sensorgrams were calculated by sequential subtractions of the corresponding curves obtained from the reference flow cell and the running buffer (blank). SPR responses are expressed in resonance unit (RU). The K_D values were calculated by global fitting of the kinetic curves.

3.4.7 SPR competition assays

The compounds were prepared as 100 µL samples in the same manner as mentioned above in the SPR binding assays. For each competition experiment, a series of 12 injections was applied including two blanks, the acylhydrazone, a blank for surface regeneration, synaptopodin, ethylene glycol 80% in buffer for surface regeneration, two blanks, acylhydrazone/synaptopodin mixture with the same concentrations of the single compounds, ethylene glycol 80% in buffer, and two blanks, respectively. Samples were injected at a flow rate of 30 µL/min for 60 s and the dissociation phase was recorded for 120 s. Data processing and analysis were performed as mentioned above.

3.5 References

[1] G. Paul, H. van Heusden, *IUBMB Life* **2005**, *57*, 623–629.

[2] M. Molzan, B. Schumacher, C. Ottmann, A. Baljuls, L. Polzien, M. Weyand, P. Thiel, R. Rose, M. Rose, P. Kuhenne, et al., *Mol. Cell. Biol.* **2010**, *30*, 4698–711.

[3] L. G. Milroy, M. Bartel, M. A. Henen, S. Leysen, J. M. C. Adriaans, L. Brunsveld, I. Landrieu, C. Ottmann, *Angew. Chem. Int. Ed.* **2015**, *54*, 15720–15724.

[4] C. Ottmann, *Bioorganic Med. Chem.* **2013**, *21*, 4058–4062.

[5] G. Tzivion, J. Avruch, *J. Biol. Chem.* **2002**, *277*, 3061–3064.

[6] B. Coblitz, M. Wu, S. Shikano, M. Li, *FEBS Lett.* **2006**, *580*, 1531–1535.

[7] S. Ganguly, J. L. Weller, A. Ho, P. Chemineau, B. Malpaux, D. C. Klein, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 1222–1227.

[8] V. Corradi, M. Mancini, F. Manetti, S. Petta, M. A. Santucci, M. Botta, *Bioorganic Med. Chem. Lett.* **2010**, *20*, 6133–6137.

[9] J. Zhao, Y. Du, J. R. Horton, a. K. Upadhyay, B. Lou, Y. Bai, X. Zhang, L. Du, M. Li, B. Wang, et al., *Proc. Natl. Acad. Sci.* **2011**, *108*, 16212–16216.

[10] L. M. Stevers, C. V. Lam, S. F. R. Leysen, F. A. Meijer, D. S. van Scheppingen, R. M. J. M. de Vries, G. W. Carlile, L. G. Milroy, D. Y. Thomas, L. Brunsveld, et al., *Proc. Natl. Acad. Sci.* **2016**, E1152–E1161.

[11] C. Anders, Y. Higuchi, K. Koschinsky, M. Bartel, B. Schumacher, P. Thiel, H. Nitta, R. Preisig-Müller, G. Schlichthörl, V. Renigunta, et al., *Chem. Biol.* **2013**, *20*, 583–593.

[12] A. M. Hartman, A. K. H. Hirsch, *Eur. J. Med. Chem.* **2017**, *136*, 573–584.

[13] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455–2488.

[14] R. Van der Vlag, A. K. H. Hirsch, in *Compr. Supramol. Chem. 2*, Elsevier, **2017**, pp. 487–509.

[15] P. Frei, R. Hevey, B. Ernst, *Chem. Eur. J.* **2019**, *25*, 60–73

[16] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259–3263.

[17] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* **2010**, *2*, 490–497.

[18] A. Ghosh, B. N. Ratha, N. Gayen, K. H. Mroue, R. K. Kar, A. K. Mandal, A. Bhunia, *PLoS One* **2015**, 1–21.

[19] P. Thiel, L. Röglin, N. Meissner, S. Hennig, O. Kohlbacher, C. Ottmann, *Chem. Commun. (Camb).* **2013**, *49*, 8468–8470.

[20] R. Rose, S. Erdmann, S. Bovens, A. Wolf, M. Rose, S. Hennig, H. Waldmann, C. Ottmann, *Angew. Chemie - Int. Ed.* **2010**, *49*, 4129–4132.

[21] S. Perspicace, D. Banner, J. Benz, F. Müller, D. Schlatter, W. Huber, *J. Biomol. Screen.* **2009**, *14*, 337–349.

[22] V. Benin, S. Durganala, A. B. Morgan, *J. Mater. Chem.* **2012**, *22*, 1180.

[23] S. A. Andrei, P. de Vink, E. Sijbesma, L. Han, L. Brunsveld, N. Kato, C. Ottmann, Y. Higuchi, *Angew. Chem. Int. Ed.* **2018**, *57*, 13470–13474.

[24] C. Henn, S. Boettcher, A. Steinbach, R. W. Hartmann, *Anal. Biochem.* **2012**, *428*, 28–30.

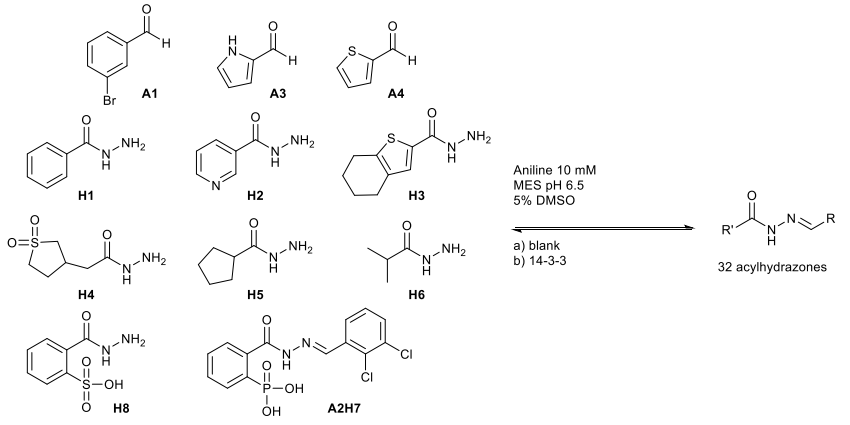
3.6 Supporting information

3.6.1 UPLC-MS analysis of DCC

UPLC-MS was carried out on a ThermoScientific Dionex Ultimate 3000 UHPLC System coupled to a ThermoScientific Q Exactive Focus with an electrospray ion source. An Acquity Waters Column (BEH, C8 1.7 µm, 2.1 x 150 mm, Waters, Germany) equipped with a VanGuard Pre-Column (BEH C8, 5 x 2.1 mm, 1,7 µm, Waters, Germany) was used for separation. At a flow rate of 0.250 mL/min, the gradient of H_2O (0.1% formic acid) and acetonitrile (0.1% formic acid) was held at 5% acetonitrile for 1 min and then increased to 95% over 16 min. It was held there for 1.5 min before the gradient was decreased to 5% over 0.1 min where it was held for 1.9 min. The mass spectrum was measured in positive mode in a range from 100 – 700 m/z.

3.6.2 HRMS analysis

High-resolution mass spectra were recorded with a ThermoScientific system where a Dionex Ultimate 3000 RSLC was coupled to a Q Exactive Focus mass spectrometer with an electrospray ion source. An Acquity UPLC® BEH C8, 150 x 2.1 mm, 1.7 µm column equipped with a VanGuard Pre-Column BEH C8, 5 x 2.1 mm, 1,7 µm (Waters, Germany) was used for separation. At a flow rate of 250 µL/min, the gradient of H2O (0.1% FA) and ACN (0.1% FA) was held at 10% B for 1 min and then increased to 95% B over 4 min. It was held there for 1.2 min before the gradient was decreased to 10% B over 0.3 min where it was held for 1 min. The mass spectrum was measured in positive mode in a range from 120 – 1000 m/z. UV spectrum was recorded at 254 nm.



Scheme S1. Original dynamic combinatorial library (DCL), from which not all products could be observed. Aldehydes (100 µM each), hydrazides (100 µM each) and **A2H7** (compound **2**) (100 µM), DMSO (5%), aniline (10 mM), and a) blank without protein; b) library with 14-3-3(ζ) (10 µM). See Figure S1 for the chromatograms.

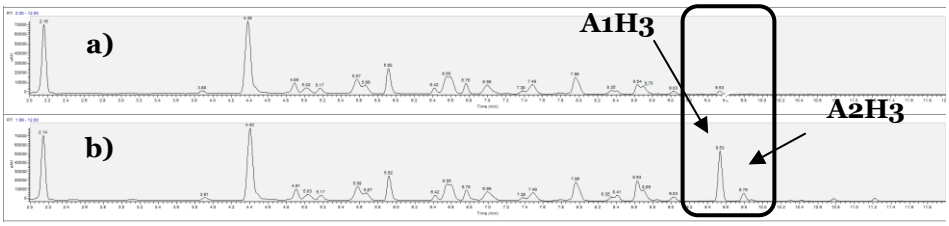
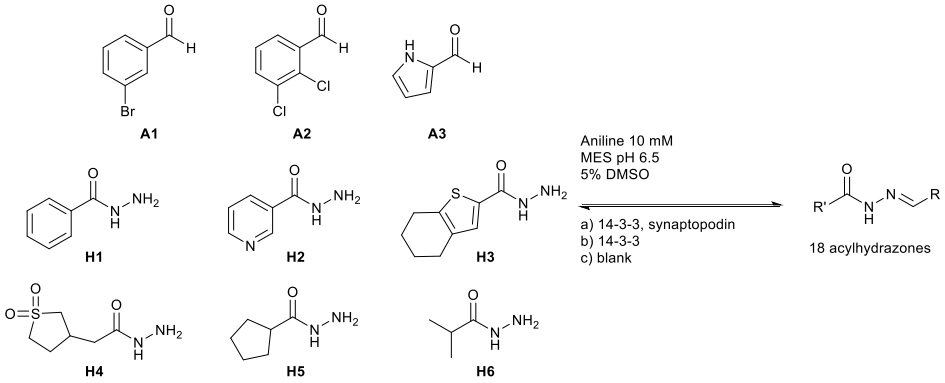


Figure S1. UV-chromatograms at 290 nm of DCLs of a) blank without protein; and b) library with 14-3-3(ζ) (10 μM). Data obtained from single experiment.

Products with **H8** and **H7** could not all be observed, but **A1H3** and **A2H3** show amplification in protein compared to blank. Therefore, we modified the DCL and omitted these structures, resulting in the DCL shown in Scheme S2.



Scheme S2. Adapted dynamic combinatorial library (DCL) with aldehydes (100 μM each), hydrazides (300 μM each), DMSO (5%), aniline (10 mM) and: a) 14-3-3(ζ) (10 μM) and synaptopodin (10 μM); b) control with 14-3-3(ζ) (10 μM); and c) control without protein or synaptopodin. See figure S2 for the chromatograms in duplicate.

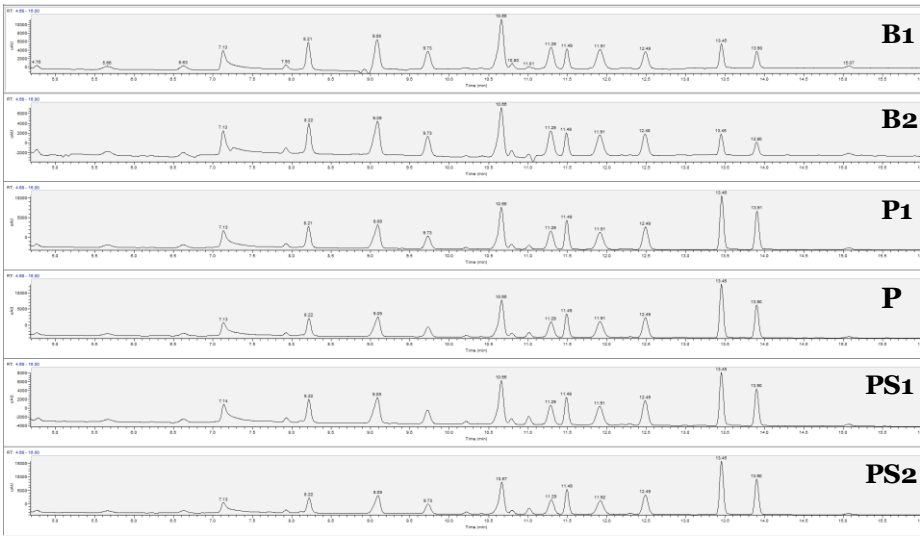


Figure S2. UV-chromatograms at 290 nm of DCLs in duplicate, blank (B), protein (P) and protein plus synaptopodin (PS) at 6 h.

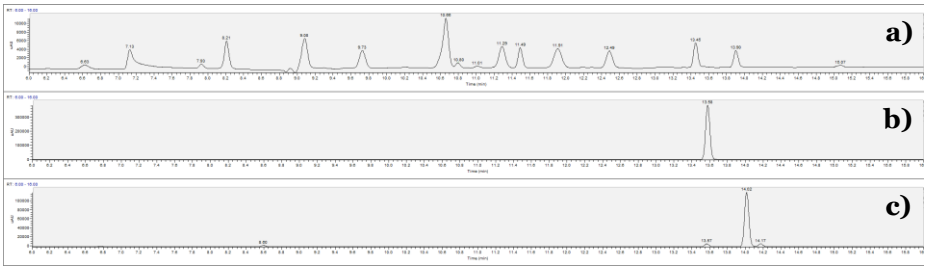


Figure S3. UV-chromatograms at 290 nm of a) the blank DCL at 6 h, b) synthesised **A1H3**, and c) synthesised **A2H3**. The synthesised compounds both have a bit longer retention time (0.12 min) due to being dissolved in pure MeCN, where the DCL also contains water and DMSO. Data obtained from single experiments.

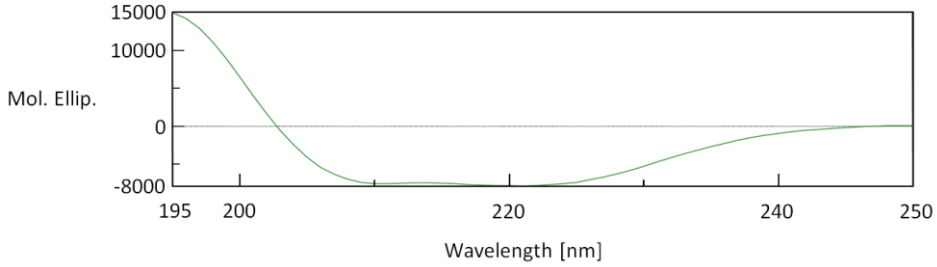


Figure S4. UV CD spectra of 14-3-3(ζ) 1μM protein in MiliQ. Measured on a J-1500 CD Spectrometer (Jasco). In agreement with Ghosh et al.[18] Data obtained from single experiment.

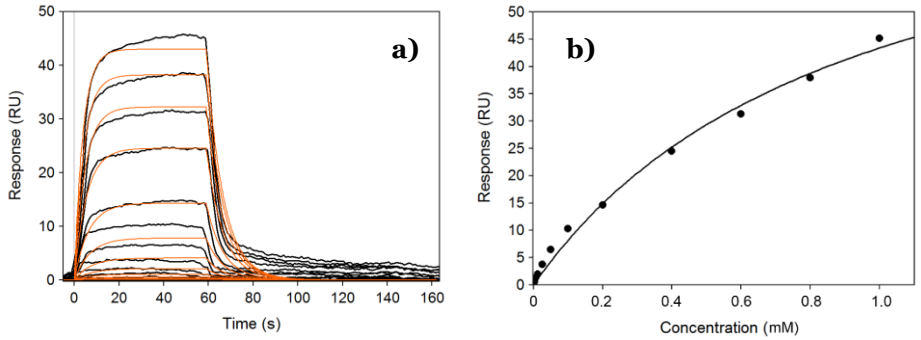


Figure S5. SPR binding assay of **2**: a) Overlay of sensorgrams (black) of **2** at concentrations of 1.6–1000.0 µM running over an immobilized 14-3-3(ζ). Global fitting of the association and dissociation curves (red), b) Langmuir binding isotherm (K_D = 1.01 ± 0.03 mM). Data obtained from single experiment.

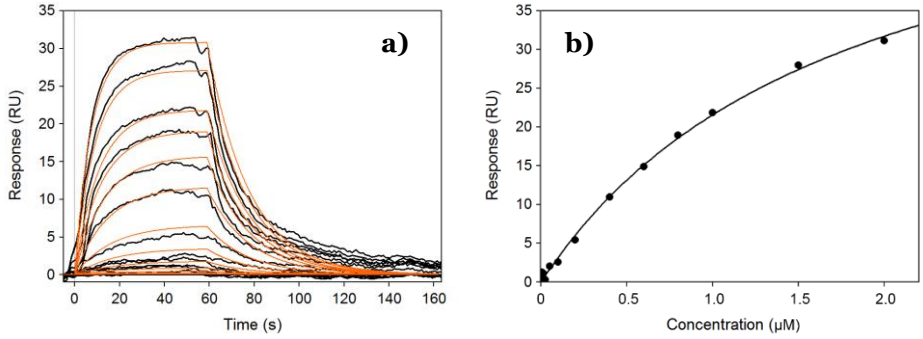


Figure S6. SPR binding assay of synaptopodin: (a) Overlay of sensorgrams (black) of synaptopodin at concentrations of 0.003–2.0 µM running over an immobilised 14-3-3(ζ). Global fitting of the association and dissociation curves (red). (b) Langmuir binding isotherm (K_D = 1.38 ± 0.02 µM). Data obtained from single experiment.

Table S1. The kinetic parameters of compound **2** and synaptopodin binding to 14-3-3. Data obtained from single measurement. R_{max}: maximum analyte binding capacity; k_{on}: association rate constant; k_{off}: dissociation rate constant; K_D: equilibrium dissociation constant; Res sd: residual standard deviation.

Compound	R_{max} (RU)	k_{on} ($M^{-1} s^{-1}$)	k_{off} (s^{-1})	K_D (µM)	Res sd
2	86.0 ± 1.0	1.32 ± 0.04 ×10²	0.133 ± 0.002	1010 ± 30	1.61
Synaptopodin	52.0 ± 5.0	3.92 ± 0.06 ×10⁴	0.0542 ± 0.0004	1.38 ± 0.02	0.90

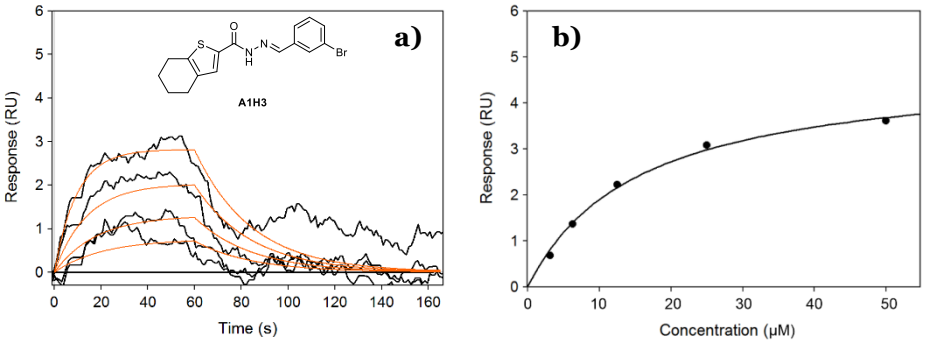


Figure S7. SPR binding assay of **A1H3**: a) Overlay of sensorgrams (black) of **A1H3** at concentrations of 3.1–50 μM running over an immobilized 14-3-3(ζ). Global fitting of the association and dissociation curves (red), b) Langmuir binding isotherm of **A1H3** (K_D =16 ± 1 μM). Data obtained from single experiment.

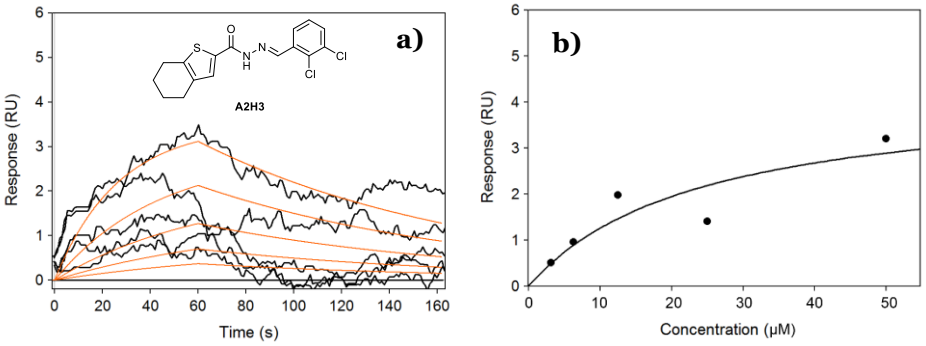
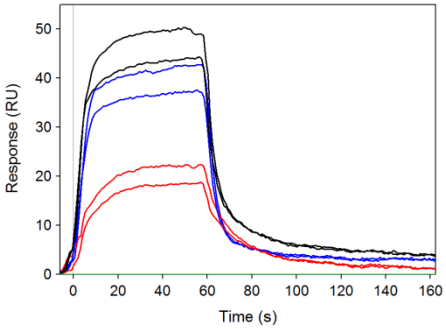


Figure S8. SPR binding assay of **A2H3**: a) Overlay of sensorgrams (black) of **A2H3** at concentrations of 3.1–50 μM running over an immobilized 14-3-3(ζ). Global fitting of the association and dissociation curves (red), b) Langmuir binding isotherm of **A2H3** (K_D =15 ± 1 μM). Data obtained from single experiment.



	Compound **2** (1000 μM)	Synaptopodin (1 μM)	**2** (1000 μM) + Synaptpodin (1 μM)
$RU_{(experiment\ 1)}$	42.08	22.21	49.67
$RU_{(experiment\ 2)}$	36.93	18.39	43.55
$RU_{(average)}$	39.51 (± 3.6)	20.30 (± 2.7)	46.61 (± 4.3)

Figure S9 and **Table S2.** SPR competition assay: Overlay of sensorgrams of **2** (1000 μM, blue), synaptopodin (1 μM, red) and **2**–synaptopodin mixture (black) in duplicate. RU values of the mixture are less than the sum of the individual RU responses (29–33% decrease of **2** response and 60–65% decrease of synaptopodin response), indicating a competitive effect.

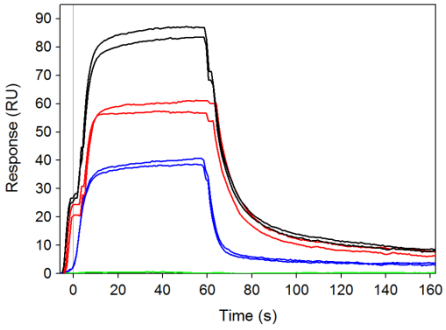
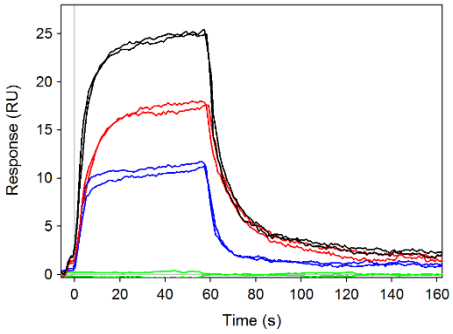


Figure S10. SPR competition assay: Overlay of sensorgrams of **2** (1000 μM, blue), synaptopodin (25 μM, red) and **2**–synaptopodin mixture (black) in duplicate.

Table S3. SPR competition assay in duplicate. RU values of the mixture are less than the sum of the individual RU responses (30–33% decrease of **2** response and 20–22% decrease of synaptopodin response), indicating a competitive effect.

	Compound **2** (1000 μM)	Synaptopodin (25 μM)	**2** (1000 μM) + Synaptpodin (25 μM)
$RU_{(experiment\ 1)}$	39.99	60.35	87.03
$RU_{(experiment\ 2)}$	37.18	55.73	81.94
$RU_{(average)}$	38.59 (± 1.9)	58.04 (± 3.3)	84.49 (± 3.6)



	Compound **2** (200 μM)	Synaptopodin (1 μM)	**2** (200 μM) + Synaptpodin (1 μM)
$RU_{(experiment\ 1)}$	10.55	16.90	24.85
$RU_{(experiment\ 2)}$	11.20	17.69	24.69
$RU_{(average)}$	10.88 (± 0.5)	17.30 (± 0.6)	24.77 (± 0.1)

Figure S11 and **Table S4.** SPR competition assay: Overlay of sensorgrams of **2** (200 μM, blue), synaptopodin (1 μM, red) and **2**–synaptopodin mixture (black) in duplicate. RU values of the mixture are less than the sum of the individual RU responses (25–37% decrease of **2** response and 15–24% decrease of synaptopodin response), indicating a competitive effect.

Chapter 4

Design and synthesis of glucosyltransferase inhibitors: dynamic combinatorial chemistry approach

Dynamic combinatorial chemistry was applied to find inhibitors of glucosyltransferase (GTF) 180. GTFs are the major producers of extracellular polysaccharides, which are important factors in the initiation and development of cariogenic dental biofilms. These biofilms are also known as dental plaque and are causative for caries. The work described herein focusses on the design and synthesis of building blocks for DCC, the analysis of the dynamic combinatorial libraries (DCLs) via UPLC-MS and evaluation of the synthesised hits by activity assays and SPR measurements.

A. M. Hartman, V. R. Jumde, W. A. M. Elgaher, E. M. te Poele, L. Dijkhuizen, A. K. H. Hirsch, *manuscript submitted*.

A. M. Hartman was involved in the analysis of the DCC, synthesis and analysis of hit compounds and preparation of the manuscript. V. R. Jumde designed and performed the DCC, and was involved in synthesis. W. A. M. Elgaher assisted in the SPR measurements and E. M. te Poele performed the activity assay. L. Dijkhuizen and A. K. H. Hirsch were involved in the preparation of the manuscript and supervision of the project.

4.1 Introduction

Cariogenic dental biofilms, also known as dental plaque, are causative for dental caries. Important factors for the initiation and development of this oral disease are the fermentation of dietary carbohydrates, of which sucrose is considered most cariogenic. It is a substrate for the synthesis of extracellular (EPS) and intracellular (IPS) polysaccharides, which form the biofilm, having glucan as one of the main components. The biofilm hosts bacteria and can promote their adhesion to the tooth enamel. Glucosyltransferases (GTFs) are the major producers of EPS, and are secreted by different strains of bacteria. These GTFs, also known as glucansucrases, are therefore potential targets in order to inhibit biofilm formation and therefore prevent dental caries.[1–3]

Glucansucrases are enzymes which are part of the glycoside hydrolase family GH70, consisting of four catalytically important conserved sequence motifs. To the superfamily of GH-H also belong the glycoside hydrolase families 13 and 77.[4] Cocrystal structures, containing the catalytic and C-terminal domains, of glucansucrase of *Lactobacillus reuteri 180* were previously reported and provided evidence for an α-retaining double displacement mechanism using one nucleophilic residue.[5] After the glycosidic bond is cleaved, the glucosyl moiety can elongate a glucan chain, or it can be transferred to a water molecule or an acceptor substrate such as maltose (Figure 1).

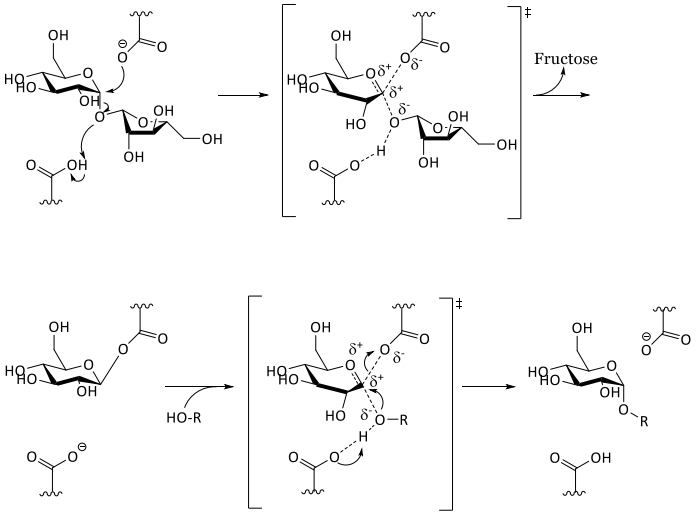


Figure 1. Reaction scheme of the proposed catalytic mechanism of GTF-180 *via* an a-retaining double displacement, leading to retention of the stereochemistry.[5,6]

Glucansucrases can be inhibited by small-molecules, as well as natural compounds. Natural inhibitors can for example be found in culture broths of

bacteria, as was the case for acarbose (Figure 2). In 1977, researchers of Bayer published the finding of α-amylase inhibitors from broths of Actinoplanes strains SE 50, SE 82 and SB18, of which BAY g 5421 (acarbose) was the most potent. Schmidt *et al.* postulated that acarbose could be a transition-state anologue.[7] Since then, acarbose has been used as an antidiabetic drug throughout the world, and was found to have cardiovascular benefits.[8–10] Newbrun *et al.* showed that acarbose also inhibits glucansucrases.[11]

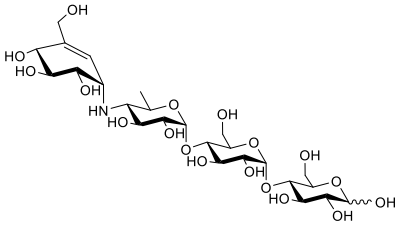
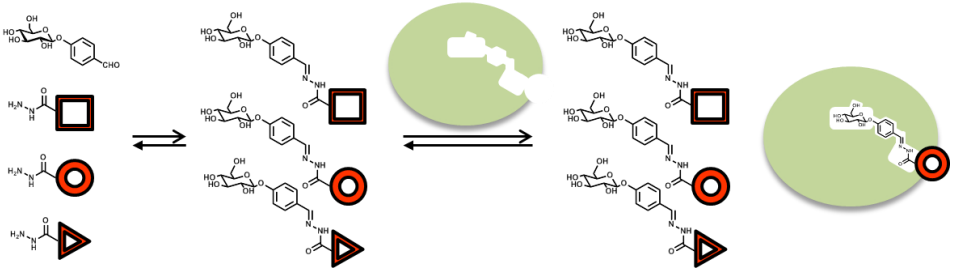


Figure 2. Acarbose, an α-glucosidase inhibitor acting as transition-state analogue. Identified from culture broths of Actinoplanes.[7]

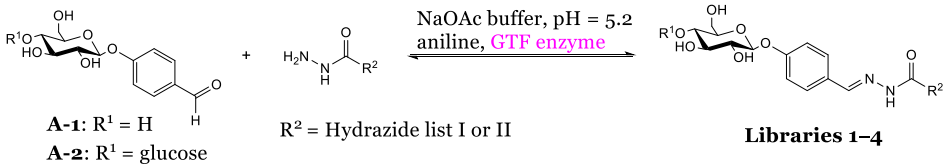
To identify hit compounds, dynamic combinatorial chemistry (DCC) has become an attractive strategy.[12–14] One of the most frequently used reversible reactions in DCC, in order to find bioactive molecules, is the acylhydrazone formation. DCC allows a target protein to alter the equillibrium of a product mixture, also known as the dynamic combinatorial library (DCL). Due to the change in equilibrium by the protein, good binders get amplified and will therefore be found as hits (Scheme 1).



Scheme 1. Schematic illustration of target-directed DCC using the acylhydrazone linkage, formed by the reaction between a glucose derived aromatic aldehyde and representative hydrazides. The protein target causes a change in equilibrium, which leads to the amplification of the hit compounds.

4.1.1 Dynamic combinatorial library design

We chose the acylhydrazone formation reaction for DCC. It is a reversible reaction between an aldehyde and a hydrazide. Inspired by the structure of acarbose, we designed the basic scaffold of our molecules. An aryl aldehyde linked to glucose and maltose (**A-1** and **A-2**, Scheme 2). We then selected eighteen chemically different and commercially available hydrazides (**H-1**—**H-18**) and divided them into two separate groups (Figure 3). Each aldehyde (**A-1** and **A-2**) was reacted separately with the two sets of hydrazides (**Hydrazides-I** and **Hydrazides-II**), resulting in four different DCLs (**DCL-1**–**4**, Scheme 2) which were analysed *via* UPLC-MS. **DCL-1** is the library formed by aldehyde **A-1** with the first set of hydrazides (**A1HI**). **DCL-2** is the library formed by **A-2** with first set of hydrazides (**A2HI**). **DCL-3** is the library formed by **A-1** with the second set of hydrazides (**A1HII**), and **DCL-4** is the library formed by **A-2** with the second set of hydrazides (**A2HII**).



Scheme 2. Each aldehyde was reacted separately with the two hydrazide libraries, resulting in the formation of four dynamic combinatorial libraries (**A1H1–8** (**DCL1**), **A2H1–8** (**DCL2**), **A1H9–18** (**DCL3**) and **A2H9–18** (**DCL4**)).

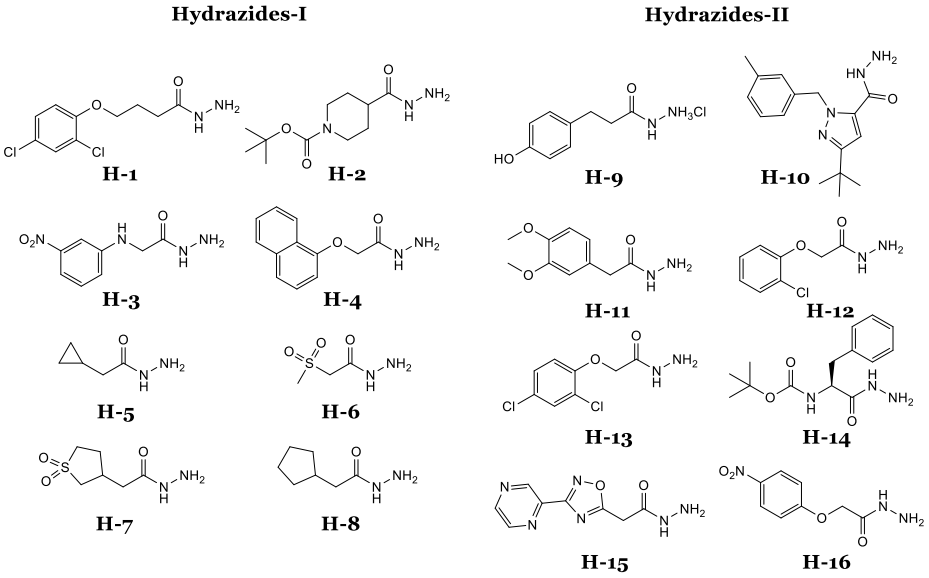


Figure 3. Hydrazide groups used in the DCC experiments.

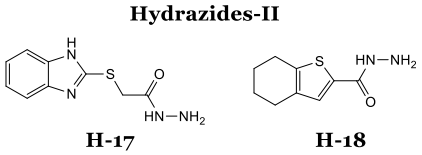
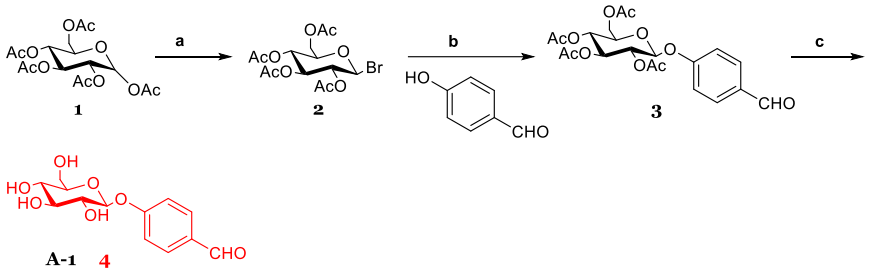


Figure 3 continued. Hydrazide groups used in the DCC experiments.

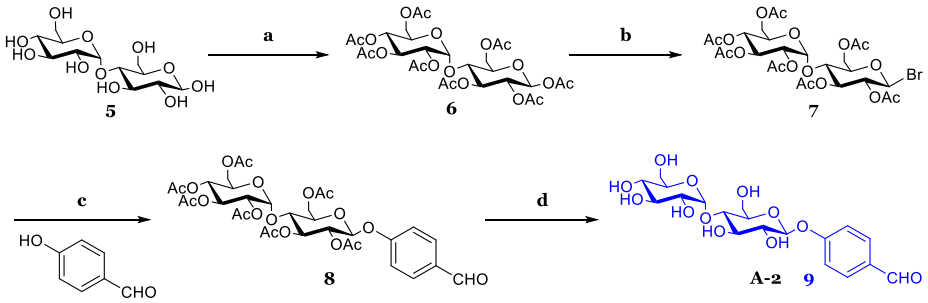
4.2 Results and Discussion

4.2.1 Synthesis of the building blocks

The aldehyde building blocks were synthesised according to the routes shown in Schemes 3 and 4. α-D-Glucose penta-acetate (**1**) was reacted with a solution of HBr in acetic acid, resulting in brominated glucose tetra-acetate (**2**). Without purification, **2** was coupled to 4-hydroxybenzaldehyde, using silver(I) oxide and yielding product **3**.[15] Deprotection of the acetate groups by sodium methoxide using the classical Zemplen deacetylation gave aldehyde **A-1** in quantitative yield (Scheme 3).[16] The same route was used for the synthesis of **A-2**, however β-maltose was acetylated first (Scheme 4).



Scheme 3. Synthetic route towards aldehyde **A-1**. *Reagents and conditions*: (a) Ac_2O, HBr in AcOH 33%, 0 °C – rt, 15 h, 70%; (b) Ag_2O, MeCN, rt, overnight 40%; (c) NaOMe, MeOH, Amberlite H^+ resin, quantitative yield.[15]



Scheme 4. Synthethic route towards aldehyde **A-2**. *Reagents and conditions*: (a) Ac_2O, $HClO_4$, AcOH, rt, 1 h; (b) 33% HBr in AcOH, 0 °C – rt, 15 h, 52% yield over two steps; (c) MeCN, Ag_2O, rt, overnight, 62%; (d) NaOMe, MeOH, Amberlite H^+ resin, 70%.[15,17]

4.2.2 Forming the DCLs

Each library consisted of one aldehyde, one group of hydrazides, aniline and DMSO in sodium acetate buffer. Aniline was added to enhance the rate at which the acylhydrazone formation reaches equilibrium, as it serves as nucleophilic catalyst to form Schiff bases with the corresponding aldehydes.[18] The use of DMSO as a cosolvent allows the building blocks and products to be soluble, preventing an undesired shift in equilibrium due to precipitation. A desired shift in the equilibrium, also known as the template effect, was achieved by the addition of the target protein GTF-180. The protein was added after pre-equilibrium was reached in the blank library.

4.2.3 Monitoring the DCLs

The DCL was left shaking at room temperature and was frequently monitored *via* UPLC-MS. Samples for UPLC-MS were prepared by taking 100 µL of the corresponding library and diluting it with 100 µL acetonitrile. The pH was raised to > 8 by the addition of 8 µL NaOH (2 M), to freeze the equilibrium. The mixture was centrifuged at 10,000 rpm for 2 minutes, and the supernatant was analysed *via* UPLC-MS.

The formation of the acylhydrazones reached equilibrium within three hours. It was at this time point that we added the GTF enzyme and continued the analysis *via* UPLC-MS. The distribution of the products in the DCLs of the blank library versus the protein library can be compared by the relative peak areas from the UV-chromatograms (Figure 4 and Table 1, and Figures S1–S3 and Tables S1–3). We selected the most amplified structure of each library and synthesised these compounds (Figure 5). For two product members of **DCL-2** (**A2H2** and **A2H8**), the glycosidic bond between the two sugar rings was cleaved after 6 hours (Figure S1 and Table S1) resulting in **A1H2** and **A1H8**. This indicates that these compounds enter the active site and get cleaved by the enzyme. Along with the hits, the cleaved products (**A1H2** and **A1H8**) will also be tested for their activities.

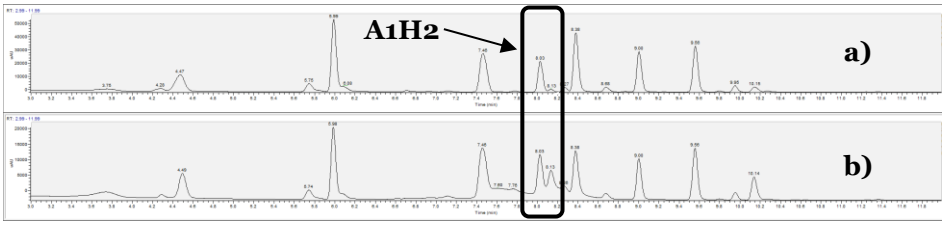


Figure 4. DCL-1. Dynamic combinatorial libraries of aldehyde **A-1** with hydrazide library 1: a) UV-chromatogram at 290 nm of the blank reaction at 6 h; b) UV-chromatogram at 290 nm of the protein-templated reaction at 7 h. Data obtained from single experiment.

Table 1. Amplification folds of the formed products in **DCL-1**; analysed *via* the relative surface areas of peaks in the UV-chromatograms of the protein-templated reaction (P) and blank reaction (B). Data obtained from single experiment.

Compound	UV-retention time (min)	Amplification fold (%P/%B)
A1H1	9.6	0.9
A1H2	**8.1**	**1.8**
A1H3	7.5	1.1
A1H4	9.0	0.8
A1H5	5.8	0.8
A1H6	3.8	1.2
A1H7	4.5	0.9
A1H8	7.5	1.1

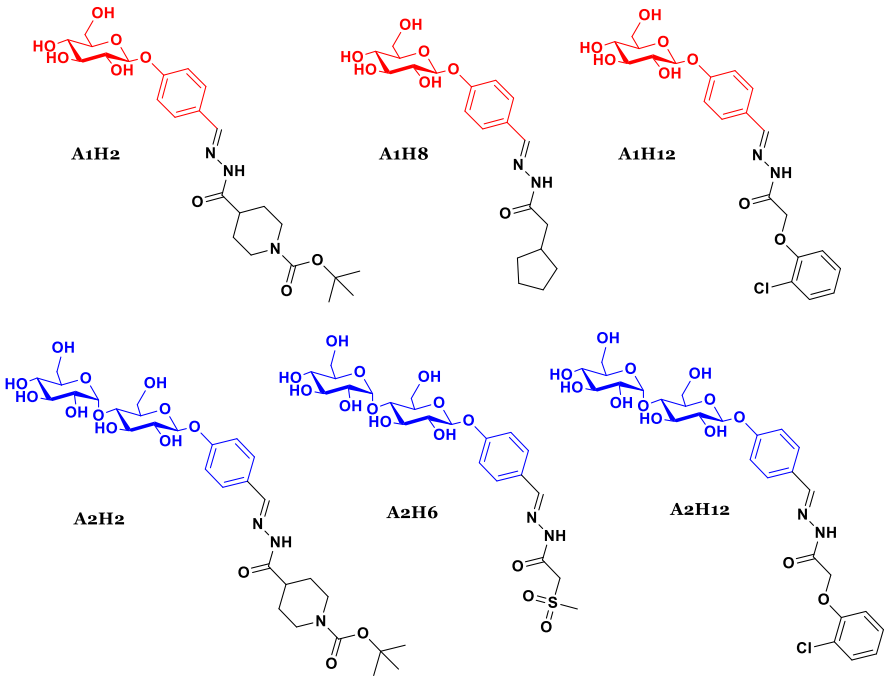


Figure 5. Hit compounds from **DCL-1 — DCL-4**.

4.2.4 Binding studies by surface plasmon resonance (SPR)

We immobilised GTF-180 to a sensor chip *via* amine coupling and followed the change in surface plasmon resonance upon injecting our hit compounds. Firstly, we ran a measurement with acarbose as a positive control, which showed that the setup was working. We then injected our hit compounds and the data are given in Table 2. Most hit compounds showed solubility problems above 1.5 mM, when 5% DMSO was used as cosolvent. We therefore could not reach the plateaus and could only fit the available data. In addition, the responses for most of the compounds were relatively low, as can be seen in the sensorgrams at the end of the experimental section (S4 – S10). Due to a very low response, the sensorgram obtained for compound **A2H6** could not be evaluated, and therefore its binding affinity could not be determined. As can be seen from Table 2, almost all compounds show weak binding, except for acarbose (10.7 µM). Compound **A1H12** showed to be the most promising binder with micromolar affinity. Since the active site of GTF-180 is polar, it is a very challenging task to address the binding pockets. We were therefore pleased to find **A1H12** as moderate binder.

Table 2. Binding study results obtained from surface plasmon resonance (SPR). Data obtained from single experiments.

Compound	R_{max} (RU)	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_D (mM)	Res sd
Acarbose	1.36 ± 0.03	1.8 ± 0.1 ×10³	0.0191 ± 0.0008	0.0107 ± 0.0008	0.41
A1H2	4.7 ± 0.2	1.7 ± 0.2 ×10²	0.25 ± 0.02	1.4 ± 0.1	0.41
A1H8	32 ± 4	18 ± 3	0.142 ± 0.004	8 ± 1	0.53
A1H12	8.7 ± 0.3	3.4 ± 0.2 ×10²	0.112 ± 0.004	0.33 ± 0.02	0.54
A2H2	40 ± 20	8 ± 5	0.099 ± 0.006	12 ± 8	0.74
A2H12	1000.0 ± 3000	1 ± 3	0.126 ± 0.003	100 ± 300	0.7

R_{max}: maximum analyte binding capacity; k_{on}: association rate constant; k_{off}: dissociation rate constant; K_D: equilibrium dissociation constant; Res sd: residual standard deviation.

4.2.5 GTF-180 activity assay

The activity of the protein can be followed *via* glucose oxidase/peroxidase (GOPOD) analysis. We followed the hydrolysis of sucrose, which forms glucose, and also monitored the transferase product fructose. The enzyme was incubated for half an hour at 37 °C with the compounds before sucrose was added. Figure 6 shows the results of the analysis of the hit compounds in the GTF activity assay. As can be seen from this figure, hardly any inhibition was observed. We used acarbose as a positive control, however it should be mentioned that acarbose is a moderate to weak inhibitor of GTF-180. The assay condition contained 20% of DMSO, which could have a bad influence on protein–ligand binding.[19] We have screened the binding affinities *via* SPR at 5% DMSO and the compounds were binding with low mM affinities. These results are in line with each other and therefore, repeating the activity assay at lower DMSO percentages will most probably not change the outcome

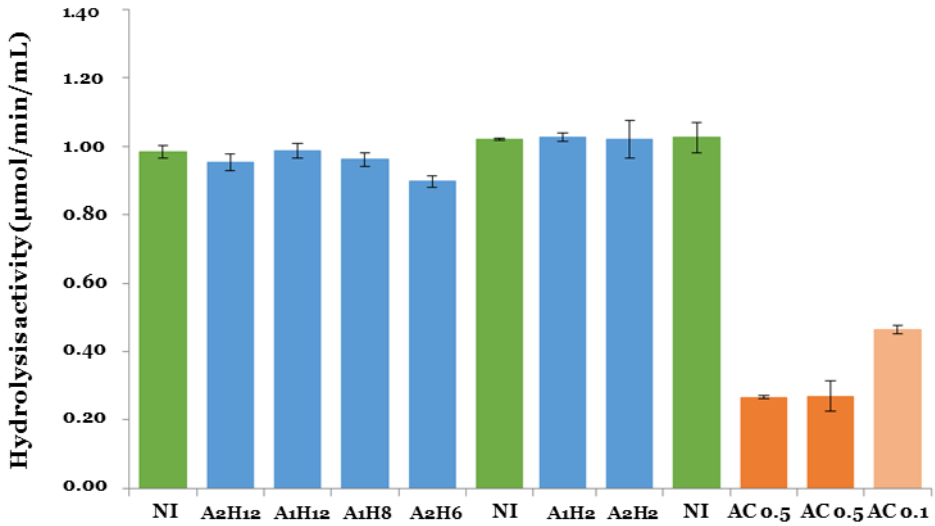


Figure 6. GTF-180 activity assay based on the hydrolysis of sucrose. No significant inhibition was observed for the hit compounds at 500 µM, and moderate inhibition by acarbose at 500 µM.

4.3 Conclusions

In this work, we showed the first application of DCC to glucansucrases. We have designed our compounds to bear a glucose or maltose head group, which resembles natural substrates, as a proper starting point. These molecules were then expected to grow into the pocket by eighteen different aromatic or aliphatic tails *via* DCC. By separating the complex libraries into four individual sublibraries, we were able to analyse the DCC experiments over time. Interestingly, we observed in the presence of enzyme that the glycosidic bond between the two glucose molecules in maltose was cleaved, indicating that it could also serve as a substrate for glucansucrases. Maltose is formerly known to be an acceptor for the glucosyl moiety. No cleavage was observed for the glycosyl moiety in the glucose derivative.

LC-MS analysis of the sublibraries resulted in six hit compounds, which were synthesised and analysed for their biophysical and biochemical properties *via* SPR and a GTF-activity assay. In comparison to acarbose, the hit compounds showed only moderate to low affinities or activities as inhibitors of GTF-180. The results of the SPR measurements are in line with the activity assay, even though they were performed at different concentrations of DMSO. In both SPR as the activity assay, acarbose was used as a positive control. It showed moderate binding, as well as moderate activity.

The first DCC approach for the quest of hits for GTFs leads to some interesting findings concerning solubility issues, active site of the enzyme and design of the scaffold. We may have failed in reaching strong binders since the active site is rather polar. This makes it difficult for compounds to enter and inhibit the active site. The cleavage of the glycosydic bond is a good indication that our compounds could enter the active site, but most likely they are not binding strong enough to inhibit the enzyme. This could be due to the fact that acylhydrazones are rather rigid and flat, possibly hindering the cavity to be filled properly.

Future work should focus on making these products more soluble to be able to test their activity with the least amount of DMSO possible. Moreover, mimicking the transition state, like acarbose does, might be the most rewarding strategy to achieve stable inhibitors. Furthermore, compounds should be flexible enough to be able to cover most of the cavity of the binding pocket.

4.4 Experimental section

4.4.1 Materials and methods

See section 3.4.1.

4.4.2 General procedure for DCC experiments

The reaction mixture composition for each sublibrary was obtained by adding the hydrazides (each 3 μL, stock solutions 100 mM in DMSO) and the aldehyde (3 μL, stock solutions 100 mM in DMSO) to a sodium acetate buffer (590.5 μL, 0.1 M, pH 5.2). Aniline (5.55 μL, stock solution 1.8 M) was added as well as DMSO, to reach a final concentration of DMSO in the DCL of 10%. Protein (309.5 μL, stock solution 96.93 μM) was added accordingly after 3 h of equilibration. Final concentrations in the DCL were: aniline (10 mM), aldehyde (300 μM), hydrazides (300 μM each), protein (30 μM) and DMSO (10%). The DCL was left shaking at room temperature and was frequently monitored *via* UPLC-MS. After 6–7 h of shaking with protein, the mixture was analysed *via* UPLC-MS.

For monitoring *via* UPLC-MS, 100 μL of the corresponding library was diluted in 100 μL acetonitrile, the pH was raised to pH > 8 by adding 8 μL NaOH (2.0 M) to freeze the reaction equilibrium. The mixture was centrifuged at 10,000 rpm for 2 min, and the supernatant was analysed *via* UPLC-MS.

4.4.3 Binding studies by surface plasmon resonance (SPR)

See section 3.4.6. Instead of 14-3-3, GTF-180 (5 μM, 117 kDa) was immobilised on the surface.

4.4.4 GTF-180 activity assay

In an 8-well PCR strip, 50 μL sodium acetate buffer (100 mM, pH 4.7, 8 mM $CaCl_2$), 40 μL DMSO, inhibitor 20 μL (5 mM), 20 μL GTF-180ΔN (4.5 μM) and 70 μL Milli-Q H_2O were added and incubated for 30 min at 37 °C. To the wells of a 96-wells PCR plate was added 12.5 μL NaOH (0.4 M). The wells of a different 8-well PCR strip were filled with 200 μL sucrose solution (100 mM) and incubated for 5 min before the start of the assay.

To start the assay, 20 μL of the 100 mM sucrose stock was added to the wells containing inhibitor and enzyme. Every 30 seconds, 25 μL sample was taken and mixed with 12.5 μL NaOH (0.4 M). After the last time point at 3.5 min, 12.5 μL HCl (0.4 M) was added to neutralise the samples.

For the glucose oxidase/peroxidase (GOPOD) analysis, 12.5 μL sample was mixed with 187.5 μL GOPOD. The glucose standard ranged from 25 to 0.195 mM.

4.4.5 Synthesis

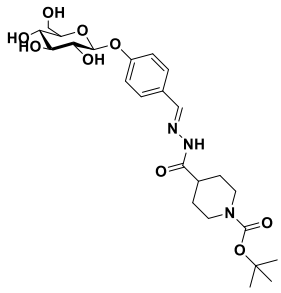
General procedure for acylhydrazone formation (GP1):[20]

To the hydrazide (1 eq.) dissolved in MeOH, the corresponding aldehyde (1.2 eq.) was added. The reaction mixture was stirred at room temperature or refluxed until completion. After cooling to room temperature, the reaction mixture was concentrated *in vacuo*. Purification of acetylated products was performed by column chromatography and deprotected sugars were purified by preparative high-performance liquid chromatography, affording the corresponding acylhydrazone in 60% to quantitative yields. For acylhydrazones, some ^{13}-NMR signals doubled due to presence of two isomers (*E* & *Z*).

General procedure for the deprotection of the acetyl groups (GP2):[16]

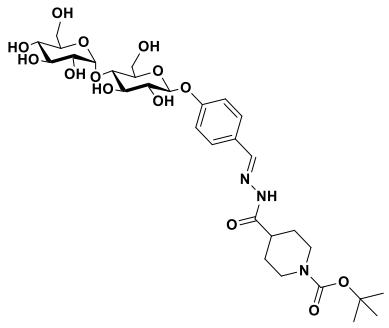
The classical Zemplén deacetylation method of the *O*-acetyl protecting groups with sodium methoxide in methanol at room temperature was used. The *O*-acetyl protected sugar was dissolved in methanol (0.01 M), and a catalytic amount of sodium methoxide (0.15 eq) was added. The reaction mixture was stirred at room temperature until complete deprotection was achieved.

tert-Butyl 4-(2-(4-(((2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)benzylidene)hydrazine-1-carbonyl)piperidine-1-carboxylate (A1H2)



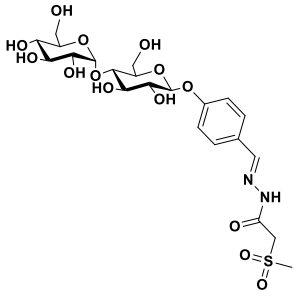
The acylhydrazone was synthesised according to **GP1** by using 1-Boc-isonipecotic acid hydrazide (26 mg, 0.1 mmol) in methanol (1.0 mL) and para-glucose-benzaldehyde (14.5 mg, 0.05 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 7:3) as a white solid (12 mg, 46%). ^{1}H-NMR (500 MHz, MeOD) δ = 8.06 (s, 1H, *E*), 7.90 (s, 1H, *Z*), 7.72 (d, *J* = 8.8 Hz, 1H), 7.62 (d, *J* = 8.8 Hz, 1H), 7.13 (dd, *J* = 8.8, 4.0 Hz, 2H), 5.00 – 4.95 (m, 1H), 4.14 (d, *J* = 13.3 Hz, 2H), 3.93 – 3.86 (m, 1H), 3.71 (dd, *J* = 12.1, 5.6 Hz, 1H), 3.52 – 3.44 (m, 3H), 3.44 – 3.37 (m, 1H), 2.47 (tt, *J* = 11.4, 3.7 Hz, 1H), 1.89 – 1.76 (m, 2H), 1.75 – 1.56 (m, 2H), 1.47 (s, 9H), 1.36 – 1.26 (m, 1H), 1.15 (d, *J* = 6.2 Hz, 1H), 0.94 – 0.83 (m, 1H); ^{13}C-NMR (126 MHz, MeOD) δ = 178.6, 173.9, 161.0, 160.6, 156.4 (d, *J* = 10.1 Hz), 149.3, 145.5, 130.2, 129.8, 129.5, 117.9 (d, *J* = 10.7 Hz), 101.9 (d, *J* = 6.1 Hz), 81.1 (d, *J* = 12.8 Hz), 78.2 (d, *J* = 3.5 Hz), 77.9 (d, *J* = 4.2 Hz), 74.8, 71.3 (d, *J* = 3.1 Hz), 62.5, 42.7, 39.5, 29.5, 28.9, 28.7; HRMS (ESI) calcd for $C_{24}H_{35}N_3O_9$ $[M+H]^+$: 510.2446, found 510.2432

***tert*-Butyl-4-(2-((*Z*)-4-(((2*S*,3*R*,4*R*,5*S*,6*R*)-3,4-dihydroxy-6-(hydroxymethyl)-5-(((2*R*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)tetrahydro-2*H*-pyran-2-yl)oxy)benzylidene)hydrazine-1-carbonyl)piperidine-1-carboxylate (A2H2)**



The acylhydrazone was synthesised according to **GP1** by using 1-boc-isonipecotic acid hydrazide (38 mg, 0.15 mmol) in methanol (1.0 mL) and *para*-maltose-benzaldehyde (22 mg, 0.05 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (9.1 mg, 28%); ^{1}H-NMR (500 MHz, MeOD) δ = 8.07 (s, 1H, *E*), 7.90 (s, 1H, *Z*), 7.73 (d, *J* = 8.8 Hz, 1H), 7.63 (d, *J* = 8.8 Hz, 1H), 7.13 (dd, *J* = 8.8, 3.5 Hz, 2H), 5.21 (d, *J* = 3.8 Hz, 1H), 5.01 (dd, *J* = 7.8, 1.9 Hz, 1H), 4.14 (d, *J* = 13.4 Hz, 2H), 3.95 – 3.42 (m, 13H), 2.47 (tt, *J* = 11.5, 3.8 Hz, 1H), 1.90 – 1.77 (m, 2H), 1.77 – 1.57 (m, 3H), 1.47 (s, *J* = 6.5 Hz, 9H); ^{13}C NMR (126 MHz, MeOD) δ 178.57 (s), 173.93 (s), 160.73 (d, *J* = 40.2 Hz), 156.45 (d, *J* = 10.2 Hz), 149.25 (s), 145.48 (s), 130.25 (s), 129.54 (d, *J* = 8.3 Hz), 117.84 (d, *J* = 11.2 Hz), 102.94 (s), 101.71 (d, *J* = 6.4 Hz), 81.14 (d, *J* = 12.5 Hz), 80.83 (d, *J* = 5.5 Hz), 77.66 (d, *J* = 3.9 Hz), 76.80 (d, *J* = 2.8 Hz), 75.08 (s), 74.85 (s), 74.44 (s), 74.18 (s), 71.50 (s), 62.76 (s), 61.90 (s), 42.69 (s), 39.54 (s), 29.51 (s), 28.67 (s); HRMS (ESI) calcd for $C_{30}H_{45}N_3O_{14}$ [*M*+H]$^{+}$: 672.2974, found 672.2958

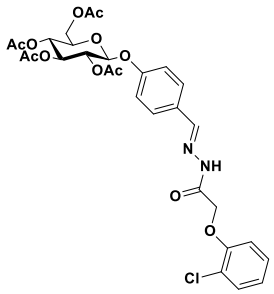
***N*'-((*Z*)-4-(((2*S*,3*R*,4*R*,5*S*,6*R*)-3,4-Dihydroxy-6-(hydroxymethyl)-5-(((2*R*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)tetrahydro-2*H*-pyran-2-yl)oxy)benzylidene)-2-(methylsulfonyl)acetohydrazide (A2H6)**



The acylhydrazone was synthesised according to **GP1** by using 2-(methylsulfonyl)acetic acid hydrazide (12.8 mg, 0.08 mmol) in methanol (0.8 mL) and *para*-maltose-benzaldehyde (30.2 mg, 0.07 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (4.1 mg, 10%). ^{1}H-NMR (500 MHz, MeOD) δ = 8.10 (s, 1H, *E*), 7.95 (s, 1H, *Z*), 7.75 (d, *J* = 8.8 Hz, 1H), 7.66 (d, *J* = 8.8 Hz, 1H), 7.14 (d, *J* = 8.8 Hz, 2H), 5.21 (d, *J* = 3.8 Hz, 1H), 5.01 (dd, *J* = 7.7, 4.5 Hz, 1H), 4.15 (s, 1H), 3.94 – 3.43 (m, 13H), 3.20 (d, *J* = 7.9 Hz, 3H); ^{13}C-NMR (126 MHz, MeOD) δ = 165.9, 161.1 (d, *J*

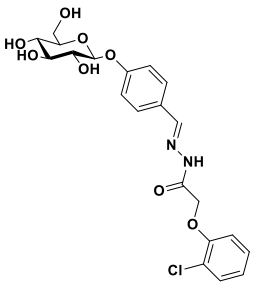
= 15.2 Hz), 160.8, 150.8, 146.8, 130.5, 129.8, 129.4, 129.1, 117.9 (d, *J* = 6.4 Hz), 102.9, 101.7 (d, *J* = 5.4 Hz), 80.8 (d, *J* = 5.4 Hz), 77.7 (d, *J* = 3.2 Hz), 76.8, 75.1, 74.8, 74.4, 74.2, 71.5, 62.8, 61.9 (d, *J* = 2.2 Hz), 60.0, 57.4, 49.6, 49.5, 49.3, 49.3, 49.2, 49.0, 48.8, 48.7, 48.5, 42.6, 42.0, 40.4; HRMS (ESI) calcd for $C_{22}H_{32}N_2O_{14}S$ [*M*+H]$^+$: 581.1647, found 581.1620.

(2*R*,3*R*,4*S*,5*R*,6*S*)-2-(Acetoxymethyl)-6-(4-((*Z*)-(2-(2-(2-chlorophenoxy)acetyl)hydrazineylidene)methyl)phenoxy)tetrahydro-2*H*-pyran-3,4,5-triyl triacetate (Acetylated A1H12)



The acylhydrazone was synthesised according to **GP1** by using 2-chlorophenoxyacetic acid hydrazide (26.6 mg, 0.13 mmol) in methanol (1.8 mL) and para-(*O*-acetylatedglucose)-benzaldehyde (50.0 mg, 0.11 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (43 mg, 61%). ^{1}H-NMR (500 MHz, MeOD) δ = 8.19 (s, 1H, *E*), 7.93 (s, 1H, *Z*), 7.72 (dd, *J* = 54.3, 8.8 Hz, 2H), 7.40 (ddd, *J* = 24.3, 7.9, 1.6 Hz, 1H), 7.31 – 7.20 (m, 1H), 7.14 – 6.92 (m, 4H), 5.44 – 5.35 (m, 2H), 5.25 (s, 1H), 5.20 – 5.08 (m, 2H), 4.75 (d, *J* = 6.6 Hz, 1H), 4.31 (dd, *J* = 12.3, 5.3 Hz, 1H), 4.20 – 4.06 (m, 2H), 2.10 – 1.95 (m, 12H); ^{13}C-NMR (126 MHz, MeOD) δ = 171.9 (d, *J* = 82.7 Hz), 171.4 (s), 171.3 (s), 171.1 (s), 167.1 (s), 159.9 (d, *J* = 44.7 Hz), 155.3 (d, *J* = 74.1 Hz), 150.6 (s), 146.1 (s), 131.4 (d, *J* = 19.4 Hz), 130.6 (s), 130.3 (s), 130.1 (s), 129.8 (s), 129.3 (s), 128.9 (s), 124.4 (s), 124.1 (s), 124.0 (s), 123.1 (s), 117.9 (d, *J* = 10.9 Hz), 116.1 (s), 115.3 (s), 99.1 (d, *J* = 6.4 Hz), 74.1 (s), 73.0 (s), 72.6 (s), 69.7 (d, *J* = 3.6 Hz), 69.2 (s), 67.3 (s), 63.1 (d, *J* = 4.4 Hz), 20.9 – 20.1 (m); HRMS (ESI) calcd for $C_{29}H_{31}ClN_2O_{12}$ [*M*+H]$^+$: 635.1638, found 635.1638.

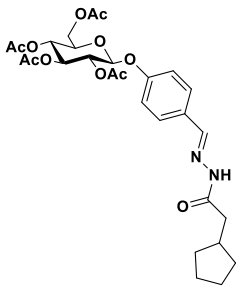
2-(2-Chlorophenoxy)-*N*'-((*Z*)-4-(((2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)benzylidene)acetohydrazide (A1H12)



The acylhydrazone was synthesised according to **GP2** by using **acetylated A1H12** (26.1 mg, 0.04 mmol) in methanol (4 mL) and sodium methoxide (0.33 mg, 0.006 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:7) as a white solid in quantitative yield (20 mg). ^{1}H-NMR (500 MHz, DMSO) δ = 11.53 (bs, 1H), 8.22 (s, 1H, *E*), 7.96 (s, 1H, *Z*), 7.64 (t, *J* = 8.4 Hz, 2H), 7.44 (ddd, *J* = 12.5, 7.9, 1.5 Hz, 1H), 7.33 – 7.22 (m, 1H), 7.10 – 7.04 (m, 2H), 7.04 – 6.92 (m, 2H), 5.37 (bs, 1H, OH), 5.25 (s, 2H), 5.16 (bs, 1H, OH), 5.08 (bs, 1H,

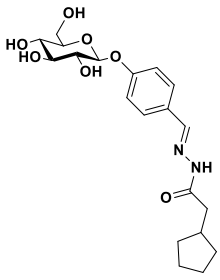
OH), 4.96 – 4.91 (m, 1H), 4.74 (s, 1H), 4.58 (s, 1H), 3.69 (d, *J* = 11.3 Hz, 1H), 3.47 (s, 1H), 3.29 – 3.21 (m, 2H), 3.17 (t, *J* = 9.0 Hz, 1H); ^{13}C-NMR (126 MHz, DMSO) δ =168.4, 158.7, 153.7, 143.6, 130.1, 130.0, 128.6, 128.4, 128.3, 128.1, 127.7, 122.1, 121.5, 121.1, 116.4, 116.4, 114.1, 113.8, 100.0, 77.1, 76.6, 73.2, 69.7, 65.3, 60.6; HRMS (ESI) calcd for $C_{21}H_{23}ClN_2O_8$ [*M*+H]$^+$: 467.1216, found 467.1205.

(2*R*,3*R*,4*S*,5*R*,6*S*)-2-(Acetoxymethyl)-6-(4-((*Z*)-(2-(2-cyclopentylacetyl)hydrazineylidene)methyl)phenoxy)tetrahydro-2*H*-pyran-3,4,5-triyl triacetate (Acetylated A1H8)



The acylhydrazone was synthesised according to **GP1** by using 2-Chlorophenoxyacetic acid hydrazide (18.9 mg, 0.13 mmol) in methanol (1.8 mL) and para-(*O*-acetylatedglucose)-benzaldehyde (50.0 mg, 0.11 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 7:3) as a white solid (59 mg, 93%). ^{1}H-NMR (500 MHz, MeOD) δ = 8.05 (s, 1H, *E*), 7.89 (s, 1H, *Z*), 7.69 (dd, *J* = 60.2, 8.8 Hz, 2H), 7.07 (d, *J* = 8.8 Hz, 2H), 5.45 – 5.34 (m, 2H), 5.22 – 5.06 (m, 2H), 4.37 – 4.25 (m, 1H), 4.21 – 4.08 (m, 2H), 2.74 (d, *J* = 7.4 Hz, 1H), 2.30 (d, *J* = 1.9 Hz, 2H), 2.08 – 1.96 (m, 12H), 1.85 (m, 2H), 1.74 – 1.64 (m, 2H), 1.63 – 1.54 (m, 2H), 1.34 – 1.18 (m, 2H); ^{13}C-NMR (126 MHz, MeOD) δ = 172.3 (d, *J* = 12.7 Hz), 171.4 (d, *J* = 42.3 Hz), 171.1, 159.9, 148.6, 130.3, 129.5, 117.9, 117.8, 99.1, 74.1, 73.1, 72.7, 69.7, 63.1, 49.6, 49.5, 49.5, 49.3, 49.2, 49.0, 48.8, 48.7, 48.5, 41.6, 39.4, 38.6, 38.0, 33.5, 33.4, 25.9, 20.6, 20.5; HRMS (ESI) calcd for $C_{28}H_{36}N_2O_{11}$ [*M*+H]$^+$: 577.2392, found 577.2360.

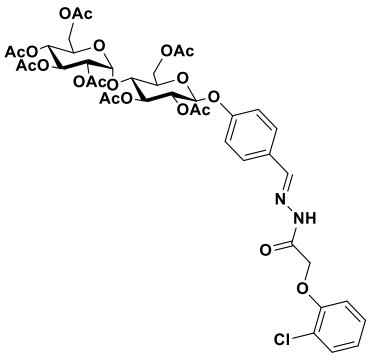
2-Cyclopentyl-*N*'-((*Z*)-4-(((2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)benzylidene)acetohydrazide (A1H8)



The acylhydrazone was synthesised according to **GP2** by using compound **acetylated A1H8** (23.0 mg, 0.04 mmol) in methanol (4 mL) and sodium methoxide (0.32 mg, 0.006 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 7:3) as a white solid (13 mg, 82%). ^{1}H-NMR (500 MHz, MeOD) δ = 8.05 (s, 1H, *E*), 7.88 (s, 1H, *Z*), 7.73 (d, *J* = 8.8 Hz, 1H), 7.61 (d, *J* = 8.8 Hz, 1H), 7.13 (dd, *J* = 8.8, 1.9 Hz, 2H), 4.98 – 4.95 (m, 1H), 3.94 – 3.86 (m, 1H), 3.71 (dd, *J* = 12.1, 5.6 Hz, 1H), 3.52 – 3.44 (m, 3H), 3.44 – 3.36 (m, 1H), 2.74 (d, *J* = 7.5 Hz, 1H), 2.34 – 2.24 (m, 2H), 1.92 – 1.78 (m, 2H), 1.74 – 1.64 (m, 2H), 1.64 – 1.49 (m, 2H), 1.34 – 1.18 (m, 2H); ^{13}C-NMR (126 MHz, MeOD) δ = 177.5, 172.3, 160.9, 160.6, 149.0, 145.1, 130.2,

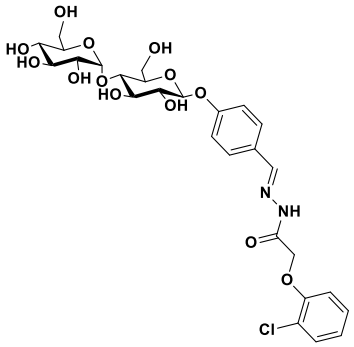
129.9, 129.6, 129.4, 117.9, 117.8, 101.9, 101.9, 78.2, 78.0, 77.9, 74.9, 71.3, 71.3, 62.5, 62.5, 41.6, 39.4, 38.5, 38.0, 33.5, 33.4, 25.9, 25.9;HRMS (ESI) calcd for $C_{20}H_{28}N_2O_7$ [*M*+H]$^+$: 409.1969, found 409.1964.

(2*R*,3*R*,4*S*,5*R*,6*R*)-2-(Acetoxymethyl)-6-(((2*R*,3*R*,4*S*,5*R*,6*S*)-4,5-diacetoxy-2-(acetoxymethyl)-6-(4-((*Z*)-(2-(2-(2-chlorophenoxy)acetyl)hydrazineylidene)methyl)phenoxy)tetrahydro-2*H*-pyran-3-yl)oxy)tetrahydro-2*H*-pyran-3,4,5-triyl triacetate (Acetylated A2H12)



The acylhydrazone was synthesised according to **GP1** by using 2-Chlorophenoxyacetic acid hydrazide (8.9 mg, 0.04 mmol) in methanol (0.5 mL) and para-(*O*-acetylatedmaltose)-benzaldehyde (25.0 mg, 0.03 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (30 mg, 48%). ^{1}H NMR (500 MHz, MeOD) δ = 8.19 (s, 1H, *E*), 7.92 (s, 1H, *Z*), 7.71 (dd, *J* = 53.8, 8.8 Hz, 2H), 7.45 – 7.36 (m, 1H), 7.32 – 7.20 (m, 1H), 7.14 – 6.91 (m, 4H), 5.47 – 5.34 (m, 4H), 5.24 (s, 1H), 5.07 (dt, *J* = 13.6, 9.4 Hz, 2H), 4.88 (dd, *J* = 9.9, 3.3 Hz, 1H), 4.76 (s, 1H), 4.58 – 4.51 (m, 1H), 4.30 (dd, *J* = 10.8, 8.3 Hz, 1H), 4.27 – 4.11 (m, 1H), 4.09 (d, *J* = 6.3 Hz, 2H), 2.23 – 1.89 (m, 21H); ^{13}C NMR (126 MHz, MeOD) δ = 172.3 (d, *J* = 6.0 Hz), 171.9, 171.8, 171.6, 171.3, 171.2, 167.1, 160.1, 159.7, 155.5, 154.9, 154.8, 150.6, 146.1, 131.5, 131.4, 131.4, 130.6, 130.3, 130.0, 129.8, 129.3 (d, *J* = 6.5 Hz), 128.9, 124.4, 124.1, 124.0, 123.9, 123.1, 117.9, 117.8, 116.1, 115.7, 115.3, 98.7 (d, *J* = 4.7 Hz), 97.3, 76.4, 74.94 (d, *J* = 7.6 Hz), 73.6, 73.3, 71.7, 70.7, 69.9, 69.7, 69.2, 68.8, 67.3, 64.3, 63.1, 21.2, 20.8 (d, *J* = 3.5 Hz), 20.6 (d, *J* = 6.1 Hz), 20.6 (d, *J* = 1.4 Hz).; HRMS (ESI) calcd for $C_{41}H_{47}ClN_2O20$ [*M*+H]$^+$: 923.2484, found 923.2446.

2-(2-Chlorophenoxy)-*N*'-((*Z*)-4-(((2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)benzylidene)acetohydrazide (A2H12)



The acylhydrazone was synthesised according to **GP2** by using compound **acetylated A2H12** (27.0 mg, 0.03 mmol) in methanol (3 mL) and sodium methoxide (0.36 mg, 0.007 mmol). After purification, the acylhydrazone was obtained as a mixture of *E* and *Z* isomers (*E*:*Z* = 3:2) as a white solid (17.2 mg, 93%). ^{1}H NMR (500 MHz, MeOD) δ = 8.19 (s, 1H, *E*), 7.93 (s, 1H, *Z*), 7.70 (dd, *J* = 56.2, 8.8 Hz, 2H), 7.40 (ddd, *J* = 25.3, 7.9, 1.6 Hz, 1H), 7.32 – 7.20 (m, 1H), 7.17 – 6.92 (m, 4H), 5.26 (s, 1H), 5.21 (d, *J* = 3.7 Hz, 1H), 5.01 (dd, *J* = 7.7, 5.7 Hz, 1H), 4.76 (s, 1H), 3.98 – 3.42 (m, 12H); ^{13}C NMR (126 MHz, MeOD) δ = 171.36, 167.10, 161.10, 160.73, 155.57, 155.00, 150.95, 146.43, 131.51, 131.34, 130.48, 129.71, 129.30 (d, *J* = 5.5 Hz), 128.90, 124.43, 124.13, 124.02, 123.11, 117.85 (d, *J* = 7.0 Hz), 116.09, 115.24, 102.94, 101.70 (d, *J* = 5.6 Hz), 80.83 (d, *J* = 4.9 Hz), 77.67 (d, *J* = 3.1 Hz), 76.82, 75.09, 74.85, 74.45, 74.19, 71.51, 69.18, 67.30, 62.78, 61.92; HRMS (ESI) calcd for $C_{21}H_{23}ClN_2O_8$ [*M*+H]$^{+}$: 629.1744, found 629.1744.

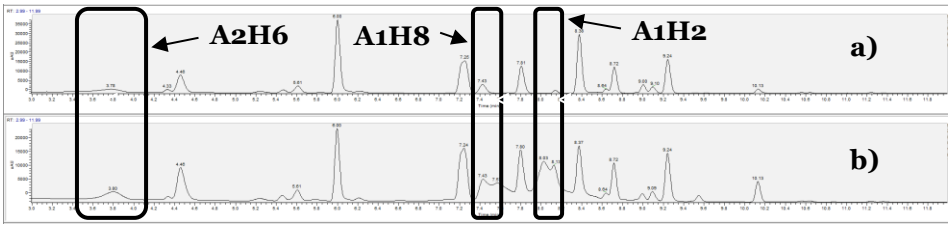


Figure S1. DCL-2. Dynamic combinatorial libraries of aldehyde **A-2** with hydrazide library 1: a) UV-chromatogram at 290 nm of the blank reaction at 3 h; b) UV-chromatogram at 290 nm of the protein-templated reaction at 6 h. Data obtained from single experiment.

Table S1. Amplification folds of the formed products in **DCL-2**; analysed *via* the relative surface areas of peaks in the UV-chromatograms of the protein-templated reaction (P) and blank reaction (B). Data obtained from single experiment.

Compound	UV-retention time (min)	Amplification folds (%P/%B)
A2H1	9.2	0.8
A2H2	7.8	0.9
A2H3	7.3	0.8
A2H4	8.7	0.8
A2H5	5.6	0.7
A2H6	***3.8***	***1.5***
A2H7	4.5	1.1
A2H8	7.3	0.8
A1H2	***8.1***	***16.4***
A1H8	***7.4***	***4.4***

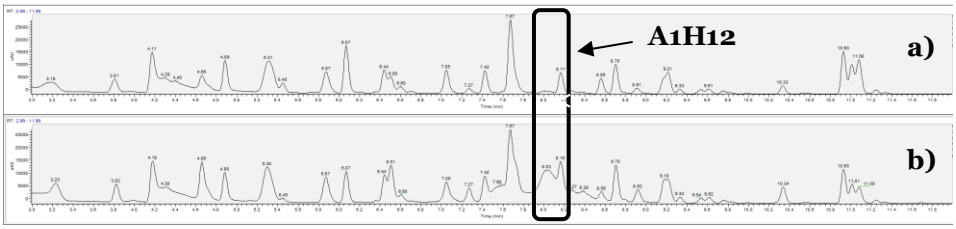


Figure S2. DCL-3. Dynamic combinatorial libraries of aldehyde **A-1** with hydrazide library 2: a) UV-chromatogram at 290 nm of the blank reaction at 3 h; b) UV-chromatogram at 290 nm of the protein-templated reaction at 6 h. Data obtained from single experiment.

Table S2. Amplification folds of the formed products in **DCL-3**; analysed *via* the relative surface areas of peaks in the UV-chromatograms of the protein-templated reaction (P) and blank reaction (B). Data obtained from single experiment.

Compound	UV-retention time (min)	Amplification fold (%P/%B)
A1H9	5.9	0.8
A1H10	10.9	0.7
A1H11	6.5	1.2
A1H12	***8.2***	***2.1***
A1H13	9.2	1.1
A1H14	9.2	1.1
A1H15	5.4	0.8
A1H16	7.4	1.0
A1H17	5.4	0.8
A1H18	8.7	1.0

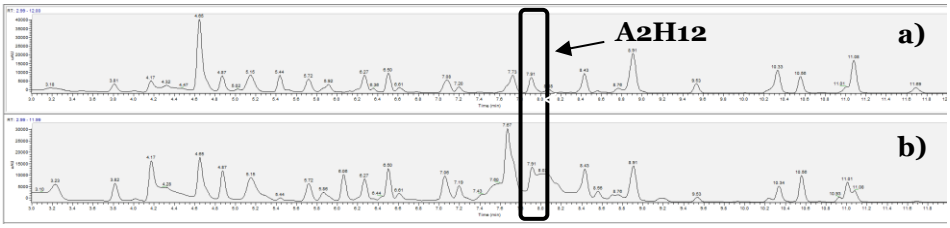


Figure S3. DCL-4. Dynamic combinatorial libraries of aldehyde **A-2** with hydrazide library 2:(a) UV-chromatogram of the blank reaction at 3 h; b) UV-chromatogram of the protein-templated reaction at 6 h. Data obtained from single experiment.

Table S3. Amplification folds of the formed products in **DCL-4**; analysed *via* the relative surface areas of peaks in the UV-chromatograms of the protein-templated reaction (P) and blank reaction (B). Data obtained from single experiment.

Compound	UV-retention time (min)	Amplification fold (%P/%B)
A2H9	5.7	1.0
A2H10	10.6	1.0
A2H11	6.2	1.0
A2H12	***7.9***	***3.2***
A2H13	8.9	0.6
A2H14	8.9	0.6
A2H15	-	-
A2H16	7.2	2.2
A2H17	5.2	0.1
A2H18	8.4	1.2

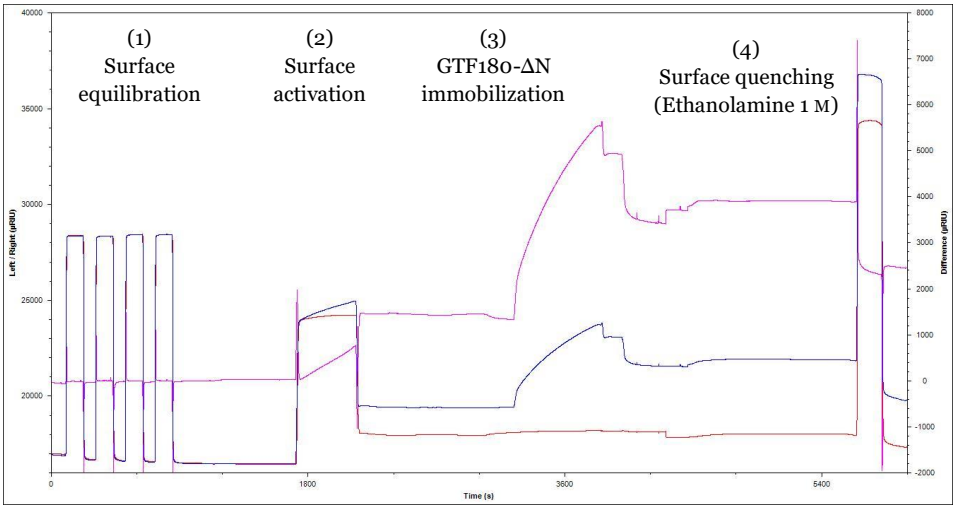


Figure S4. Sensorgram of the immobilisation procedure for GTF180-ΔN on CMD500M sensor chip: (1) Four injections of cleaning solution, (2) activation solution, (3) GTF180-ΔN, and (4) quenching solution. The blue, red, and magenta curves represent the left (active) channel, right (reference) channel, and the difference, respectively.

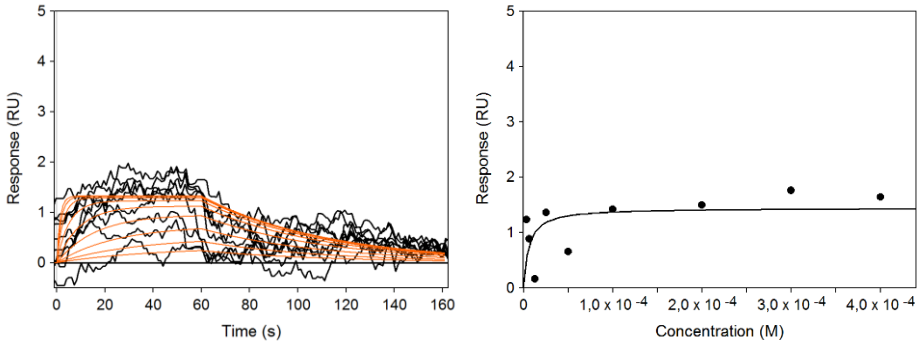


Figure S5. Overlay of sensorgrams (black) of **acarbose** at concentrations of 0.8–500 µM running over an immobilised GTF-180. Global fitting of the association and dissociation curves (red). Data obtained from single experiment.

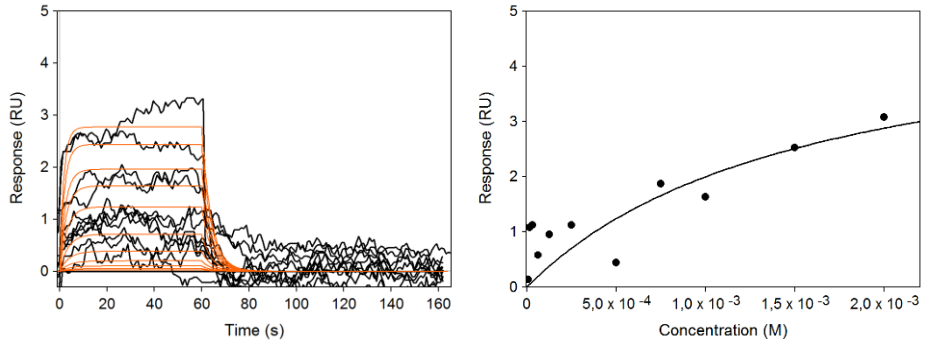


Figure S6. Overlay of sensorgrams (black) of **A1H2** at concentrations of 7.6–2000 μM running over an immobilised GTF-180. Global fitting of the association and dissociation curves (red). Data obtained from single experiment.

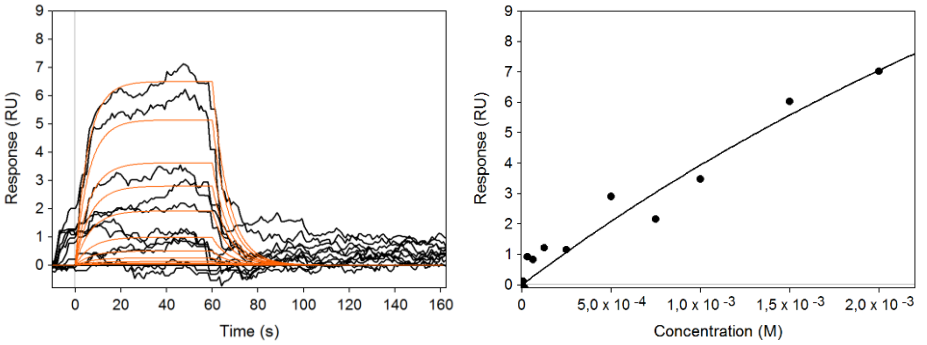


Figure S7. Overlay of sensorgrams (black) of **A1H8** at concentrations of 7.6–2000 μM running over an immobilised GTF-180. Global fitting of the association and dissociation curves (red). Data obtained from single experiment.

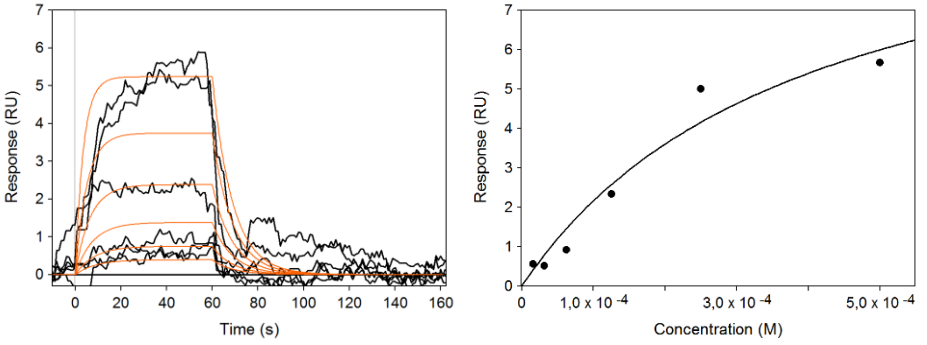


Figure S8. Overlay of sensorgrams (black) of **A1H12** at concentrations of 7.6–500 μM running over an immobilised GTF-180. Global fitting of the association and dissociation curves (red). Data obtained from single experiment.

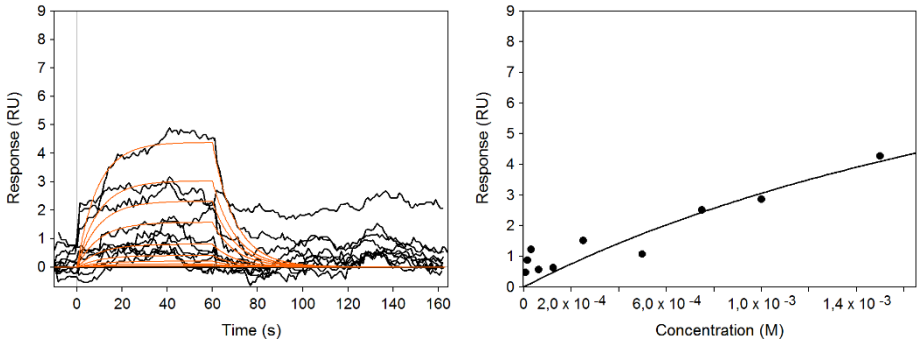


Figure S9. Overlay of sensorgrams (black) of **A2H2** at concentrations of 7.6–1500 µM running over an immobilised GTF-180. Global fitting of the association and dissociation curves (red). Data obtained from single experiment.

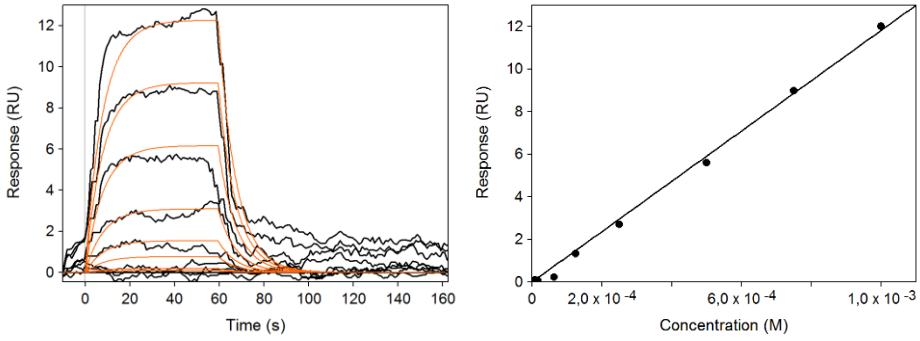


Figure S10. Overlay of sensorgrams (black) of **A2H12** at concentrations of 7.6–1000 µM running over an immobilised GTF-180. Global fitting of the association and dissociation curves (red). Data obtained from single experiment.

4.5 References

[1] W. H. Bowen, *Crit. Rev. Oral Biol. Med.* **2002**, *13*, 126–131.

[2] G. Rolla, *Scand. J. Dent. Res.* **1989**, *97*, 115.

[3] A. F. P. Leme, H. Koo, C. M. Bellato, G. Bedi, J. A. Cury, *J. Dent. Res.* **2006**, *85*, 878–887.

[4] "Carbohydrate-Active Enzymes database (CAZy)," can be found under http://www.cazy.org/Glycoside-Hydrolases.html, **n.d.**

[5] A. Vujicic-Zagar, T. Pijning, S. Kralj, C. A. López, W. Eeuwema, L. Dijkhuizen, B. W. Dijkstra, *PNAS* **2010**, *107*, 50.

[6] C. Moulis, G. Joucla, D. Harrison, E. Fabre, G. Potocki-Veronese, P. Monsan, M. Remaud-Simeon, *J. Biol. Chem.* **2006**, *281*, 31254–31267.

[7] D. D. Schmidt, W. Frommer, B. Junge, L. Müller, W. Wingender, E. Truscheit, *Naturwissenschaften* **1977**, *64*, 535.

[8] J. Weng, S. Soegondo, O. Schnell, W. H.-H. Sheu, W. Grzeszczak, H. Watada, N. Yamamoto, S. Kalra, *Diabetes Metab. Res. Rev.* **2015**, *31*, 155–167.

[9] M. Hanefeld, F. Schaper, *Expert Rev. Cardiovasc. Ther.* **2008**, *6*, 153–163.

[10] J. Chiasson, J. RG, R. Gomis, et al, *JAMA* **2003**, *290*, 486–494.

[11] E. Newbrun, C. I. Hoover, G. J. Walker, *Arch. Oral Biol.* **1983**, *28*, 531–536.

[12] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455–2488.

[13] R. Van der Vlag, A. K. H. Hirsch, in *Compr. Supramol. Chem. 2*, Elsevier, **2017**, pp. 487–509.

[14] P. Frei, R. Hevey, B. Ernst, *Chem. Eur. J.* **2019**, *25*, 60–73.

[15] N. Kuźnik, A. Chrobaczyński, M. Mika, P. Miler, R. Komor, M. Kubicki, *Eur. J. Med. Chem.* **2012**, *52*, 184–192.

[16] G. Zemplén, A. Kunz, *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.* **1924**, *57*, 1194.

[17] D. Gueyrard, R. Iori, A. Tatibouët, P. Rollin, *European J. Org. Chem.* **2010**, 3657–3664.

[18] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* **2010**, *2*, 490–497.

[19] D. Cubrilovic, R. Zenobi, *Anal. Chem.* **2013**, *85*, 2724–2730.

[20] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259–3263.

Chapter 5

Nitrone-based DCC

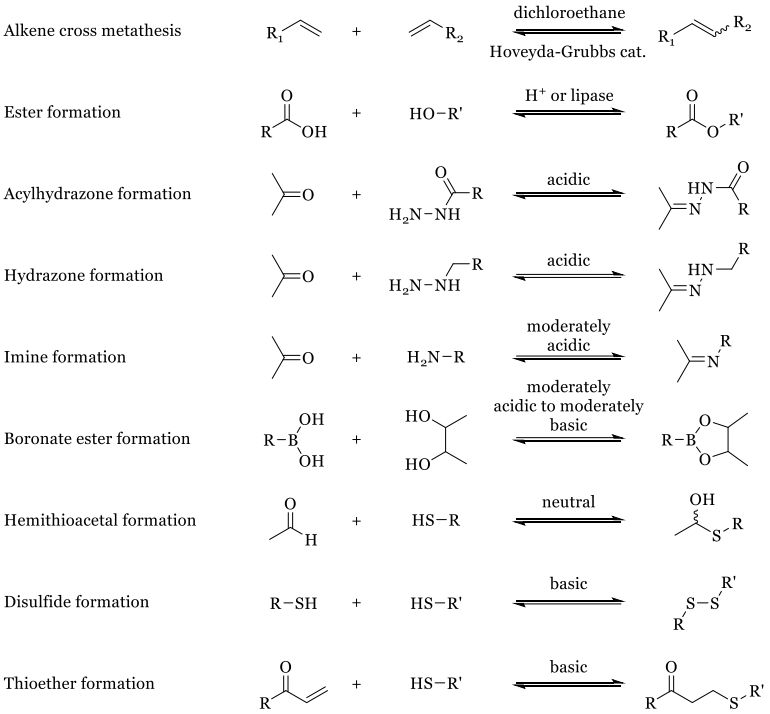
The reaction scope for dynamic combinatorial chemistry (DCC) is relatively limited. Since special requirements for proteins have to be met in protein-templated (pt) DCC, the reversible linker has to be biocompatible. Requirements like buffered aqueous media, relatively low DMSO concentrations and neutral pH values have therefore limited an extension of the portfolio of chemical reactions. In this chapter, we show the first proof of principle of ptDCC based on the reversible formation and exchange of nitrones.

A. M. Hartman, A. K. H. Hirsch, *manuscript in preparation*.

A. M. Hartman was involved in the design of the project, performing the DCC experiments, synthesis of compounds, and writing of the manuscript. A. K. H. Hirsch was involved in editing the manuscript and supervision of the project.

5.1 Introduction

As discussed in Chapter 1, the reaction scope for protein-templated dynamic combinatorial chemistry (ptDCC) is rather small (Scheme 1).[1–3] This is mostly due to the fact that the reaction needs to be biocompatible, performed at physiological pH values and most often cosolvents have to be added to guarantee solubility of building blocks and products. In this chapter, we describe our work on the first application of nitrone-based protein-templated DCC. We chose the model enzyme endothiapepsin as a protein.



Scheme 1. Reversible reactions used in target-directed DCC to identify bioactive compounds. Scheme was adapted from Van der Vlag and Hirsch.[2]

5.1.1 Biochemical relevance of nitrones

The nitrone functionality can be found in drug candidates, ranging from having anticancer activities (NXY-059), as antioxidant for neurodegenerative disorders (glycolipidic nitrone), as free radical-scavenger in the treatment of acute ischemic strokes (TBN), and as acidic and microbial corrosion inhibitors (Figure 1).[4–9]

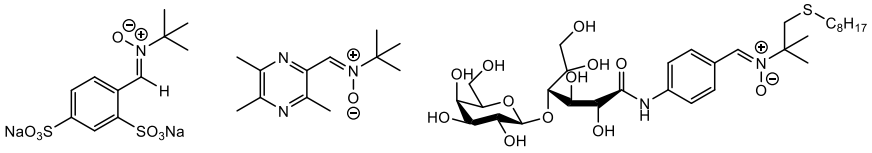


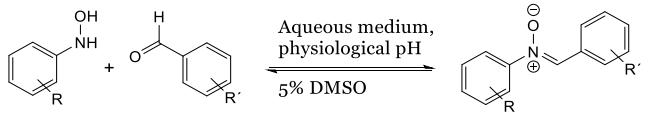
Figure 1. Examples of nitrone-based clinical candidates.[4,6,9]

5.1.2 Nitrone-based DCC

In 2005, M. Hochgürtel reported on the generation of a dynamic combinatorial library (DCL) using nitrone formation.[10] The patent reports on the reversible nitrone formation from aldehydes and hydroxylamines, reversible nitrone exchange between nitrone and aldehydes, reversible nitrone exchange between nitrone and hydroxylamines and the scrambling of nitrones. The general formation of DCLs was described using aliphatic hydroxylamines and aliphatic aldehydes, both possibly bearing chemically different substituents. Then in 2008, a PhD thesis by J. Sadownik illustrated the use of nitrone-based DCC for evolving complex systems. Nitrones were also used in reversible exchange with imines and the DCLs were afterwards coupled to the irreversible formation of cycloadducts.[11] In 2008, the Philp group published that nitrones can be formed reversibly in chloroform at room temperature.[12]

Sadownik concluded in his thesis that the choice of solvent proved to be a great challenge. The libraries were either formed in chloroform or DCM which were saturated with *p*-toluenesulfonic acid monohydrate. An important remark was made that both water and acid promote the exchange, and removing them slows down the exchange to a great extent. Having an aqueous system would therefore be a benefit for the nitrone formation. Aromatic building blocks were used for the applied exchange reaction between nitrones and imines, in contrast to the aliphatic building blocks in the patent.

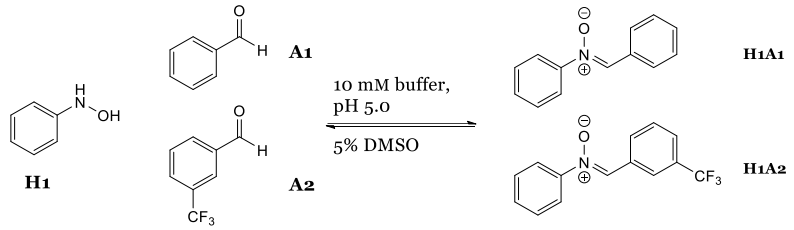
We were interested if the aromatic nitrone exchange could take place in aqueous media at physiological pH and in a reasonable time frame (Scheme 2). We therefore set out to investigate the formation and reversibility of aromatic nitrones in a pH window of 5–8. Subsequently, we wanted to apply it for the first time in a protein-templated DCC experiment.



Scheme 2. Reversible formation of nitrones based on the condensation reaction of hydroxylamines and aldehydes.

5.2 Results and Discussion

In order to prove the reversibility of nitrone formation under physiological conditions, we performed three experiments. The solvent of choice was a 10 mM ammonium acetate buffer at pH 5.0 using DMSO (v/v 5%) as a colsovent. In the first experiment, we showed that nitrones can be formed under the specified conditions using hydroxylamine (20 μM) and two aldehydes (each 200 μM, Scheme 3).



Scheme 3. Dynamic combinatorial library-N1: reversible formation of aromatic nitrones.

For the analysis of the library, a sample was diluted with acetonitrile, and the pH was increased > 7 by the addition of NaOH. This mixture was then analysed via UPLC-MS/MS. The relative peak area of the ion counts of each peak was determined and monitored over time (Figure 1). As can be seen from Figure 1, the equilibrium was reached after 6.5 hours.

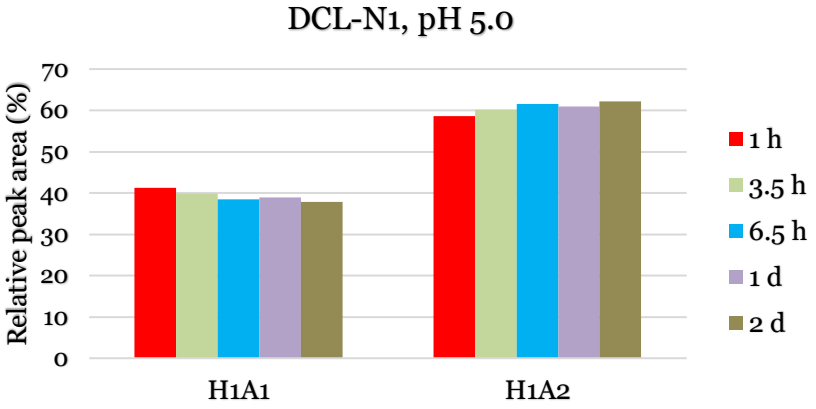
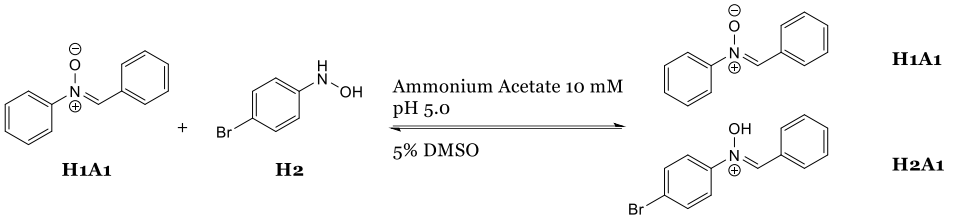


Figure 2. Analysis of dynamic combinatorial library-N1 at pH 5.0, relative peak area of each product over time. Data obtained from single experiment.

To prove that the nitrones can be formed reversibly, we used the pre-synthesised nitrone **H1A1** (20 μM) and added a hydroxylamine **H2** (200 μM, Scheme 2). We then monitored, *via* LC-MS, for the decrease in concentration of **H1A1** and the increase in concentration of **H2A1** (Figure 3).



Scheme 4. Dynamic combinatorial library-N2: exchange of nitrone **H1A1** by hydroxylamine **H2**.

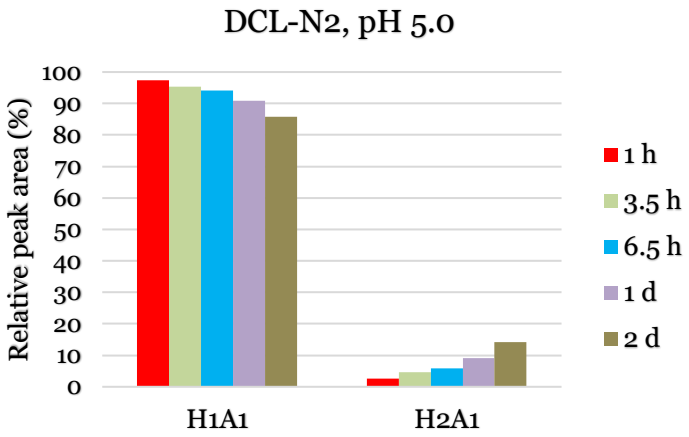
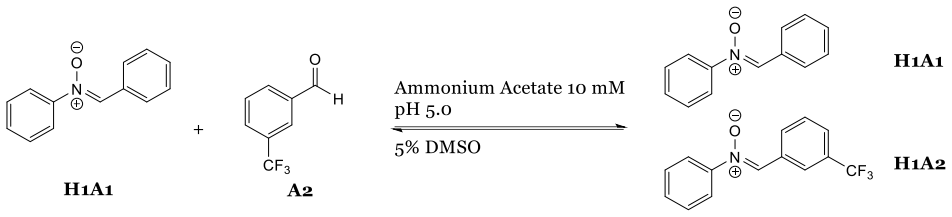


Figure 3. Analysis of dynamic combinatorial library-N2 at pH 5.0, relative peak area of each product over time. Data obtained from single experiment.

Over time, the concentration of **H2A1** is lowered, as **H1A1** is being formed. The concentrations of both products have not reached equilibrium yet, as can be seen from Figure 3. However, the formation of **H2A1** indicates that the nitrone linkage is indeed able to be broken and formed again. This shows that the system is reversible, when hydroxylamines like **H2** are added to pre-formed nitrones.

In the next experiment, DCL-N3, we showed the reversibility of the nitrone **H1A1** (20 μM) in presence of aldehyde **A2** (200 μM, Scheme 5). For the nitrone **H1A2** being formed, the pre-synthesised nitrone **H1A1** must first undergo hydrolysis to free hydroxylamine **H1**. Figure 4 shows the analysis of DCL-N3 *via* LC-MS over time.



Scheme 5. Dynamic combinatorial library-N3: exchange of nitrone **H1A1** by aldehyde **A2**.

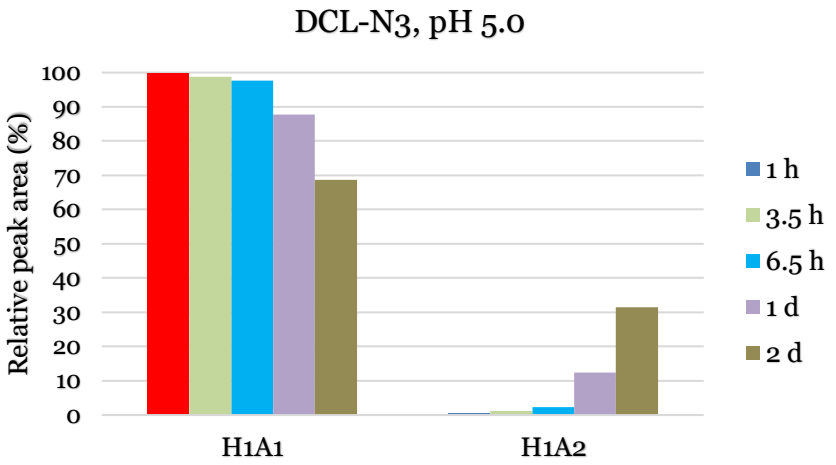


Figure 4. Analysis of dynamic combinatorial library-N3 at pH 5.0, relative peak area of each product over time. Data obtained from single experiment.

Just like in DCL-N2, equilibrium was not reached after 2 days. However, the formation of **H1A2** clearly shows that reversibility of the nitrone linkage is possible through a hydrolysis mechanism.

5.2.1 pH window

Most proteins have a pH optimum around 7.4, for this reason we set out to determine the pH window of the nitrone-exchange reaction. To investigate if the nitrone-exchange reaction could be applied on a wide range of protein targets. Therefore, we made buffer solutions of pH 5.5 (10 mM sodium acetate), 6.0, 6.5, 7.0, 7.5 and 8.0 (10 mM potassium acetate). We repeated the libraries N1, N2 and N3 with the different pH values, and took samples up to three days (see supporting information).

In DCL-N1, we could observe product formation from the individual building blocks up to pH 6.5. However, no product was observed at pH 7.0, 7.5 and 8.0. The equilibria seem to be reached after 1 day, after which the sample

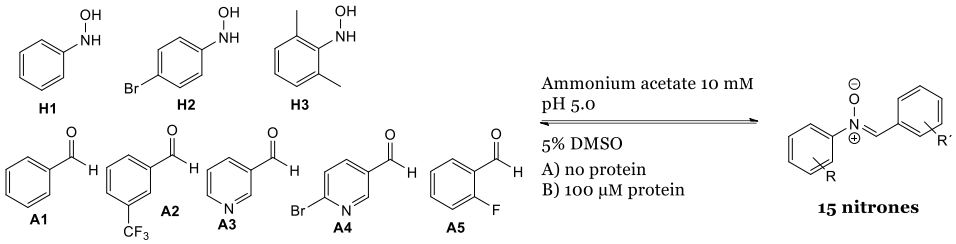
concentrations did not change significantly. To test if the library mixture would be stable under the analytical conditions, acetonitrile and pH > 7, we re-analysed the sample at pH 5.5 after 24 hours in the LC-MS vial. Previously observed strong signals of nitrones **H1A1** and **H1A2** in the fresh samples were not observed after 24 h. This indicates that samples need to be measured directly after being taken, and cannot be left for longer time periods in acetonitrile and pH > 7.

In DCL-N2, we observed the formation of nitrone **H2A1** at all pH values. Indicating that the reaction is reversible when sufficient amount of nucleophile is present.

In DCl-N3, the formation of nitrone **H1A2** was observed for pH values up to 5.5. At pH 6.0, hardly any **H1A2** could be detected. At pH 6.5 and higher, only nitrone **H1A1** could be detected, as this was added as a pre-synthesised product. The reversibility of the nitrone by an aldehyde was therefore limited to a maximum pH 5.5 – 6.0. However, this opens up the possibility of `freezing` the reaction when the aldehydes are used in excess, by adjusting the pH to > 7. An important observation from DCL-N3 is that nitrone **H1A1** seems to be stable against hydrolysis from pH 6.5 – 8.0. Therefore, 'freezing' the reaction by an increase in pH will most likely not change the library composition. The stability problem in the LC-MS vial, observed in DCL-N1, must therefore be related to acetonitrile and not to the pH.

5.2.2 protein-templated DCC

We then performed a protein-templated DCC, using the aspartic protease endothiapepsin as a model enzyme. The library consisted of three hydroxylamines (**H1**–**H3**, 20 μM each) and five aldehydes (**A1**–**A5**, 200 μM each). In the blank reaction, only DMSO was added to reach v/v 5%. In the templated reaction, also 5% DMSO was used, and endothiapepsin (100 μM) was added (Scheme 6). Both reactions were monitored *via* UPLC-MS over time. Again, the mass of each product was selected, and the corresponding peak area of the ion counts was obtained. Relative peak areas where calculated by dividing the ion counts of each product by the sum of the ion counts of all products, and multiplied by 100%. The relative peak areas where used to compare products in the protein sample to the blank sample, yielding the corresponding amplification of each product.



Scheme 6. Dynamic combinatorial library-N4: building blocks for the libraries in presence and absence of the protein endothiapepsin.

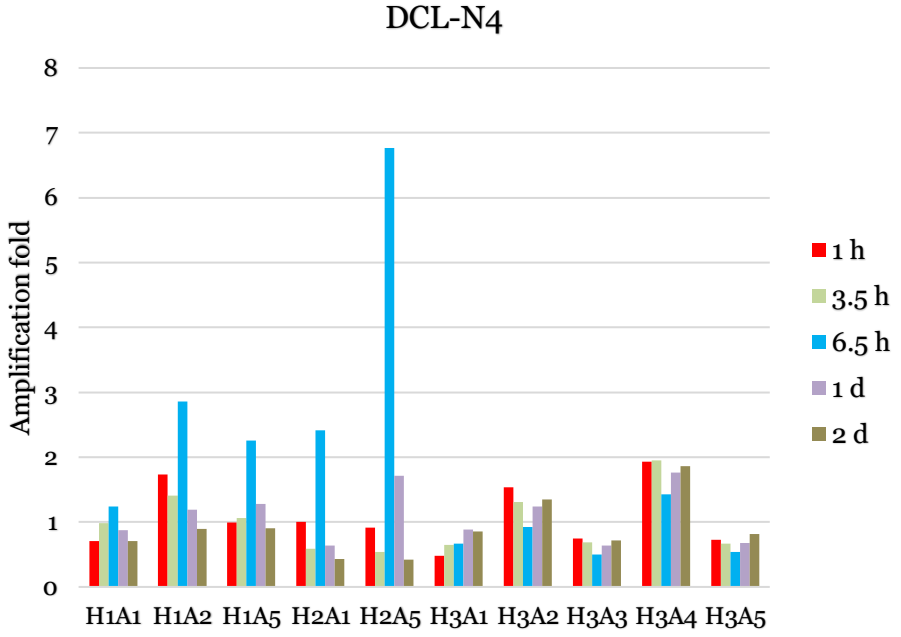


Figure 5. Amplification folds of products over time of DCL-N4: product's relative peak area of the sample with protein divided by the relative peak area in the blank reaction. Data obtained from single experiment.

Almost all compounds in DCL-4 showed a similar fold of amplification at day 1 as at day 2. We have therefore not taken samples after 2 days. In Figure 5, we have removed outliers for clarity reasons. The amplification fold of **H1A3** fluctuates too much, which might be due to the instability of the compound. **H1A4** and **H2A3** have exactly the same mass and the same retention time. Their amplification folds fluctuate too much, which might also be due the instability of the compounds. In general, it is important to monitor the stability of the compounds during the experiment. One of the most interesting observations is

that products **H2A2** and **H2A4** are hardly or not formed in the blank, but they are formed in the protein sample. For these products, no amplification can be calculated. Determining hit compounds only from amplification folds would have therefore overlooked these two compounds. They should, however, be considered as hits, since the only difference between both libraries is the absence/ presence of protein.

5.2.3 Cytotoxicity assay

We were interested to see if the hit compounds (H1A1, H2A2, H2A4) as well as the hydroxylamines (H1—H3) would show toxicity against a human cell line. A cytotoxicity assay, in which the viable cell mass is determined *via* OD measurements, would be sufficient for a first screening. We chose the HepG2 cell line, and determined the growth inhibition by the compounds at 100 μM each. None of these compounds showed inhibition of the growth. Therefore, none are cytotoxic. This preliminary result is in line with the fact that nitrones were already used in clinical studies.

5.2.4 Endothiapepsin activity assay

The same compounds as in the cytotoxicity were evaluated for their activity on the protein target, endothiapepsin. As a positive control we used saquinavir (supporting information, Figure S15). We considered **H1A1** as a negative control, since the amplification fold in DCL-4 is 1. The two hits of DCL-4, **H2A2** and **H2A4**, have moderate inhibition values, where **H1A1** indeed is having a lesser effect (Figure 6). Hydroxylamines **H1** – **H3** show weak inhibition of endothiapepsin's activity, similar to **H1A1** (supporting information, Figure S16).

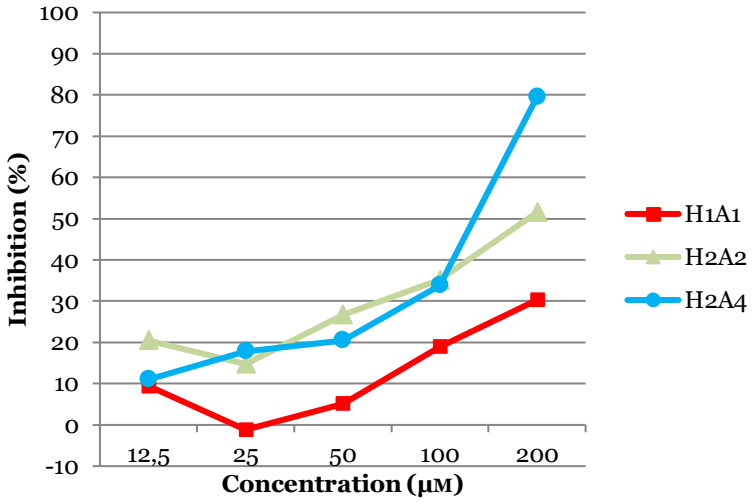


Figure 6. Endothiapepsin activity assay: showing that H1A1, H2A2 and H2A4 are moderate inhibitors. Data obtained from single experiments.

5.3 Conclusions

Based on the work of Hochgürtel and Sadownik, we extended the DCC of nitrones to a protein-templated setting. First, we showed that the nitrone reaction can take place in aqueous environment, with DMSO as co-solvent. Next, we determined the pH-range where nitrone formation is still reversible. Under slightly acidic conditions, ammonium acetate buffer (10 mM, pH 5.0), the equilibrium was reached within one day. We applied these conditions to ptDCC where endothiapepsin, an aspartic protease, was used as protein target. From this library, we obtained two hit compounds, which we then synthesised and analysed for their biochemical activities. The hit compounds, as well as the hydroxylamine building blocks, showed no toxicity at 100 μM on HepG2 cells. This founding is in line with the fact that nitrones are being used in clinical studies. From the activity assay we showed that **H1A1**, as negative control, hardly has any inhibitory effect. The two hit compounds, **H2A2** and **H2A4**, show moderate inhibitory activity. This finding is in line with the outcome of DCC experiment DCL-4. The hydroxylamines **H1** – **H3** showed similar weak activities as **H1A1**, indicating that the activity of the hit compounds are authentic. The work described herein, allows the medicinal chemist to choose from one more reversible linkage, when planning to apply DCC in drug-discovery processes.

5.4 Experimental

5.4.1 Materials and methods

See section 3.4.1.

5.4.2 DCC conditions

The corresponding hydroxylamines (each 0.2 μL, stock solution 100 mM in DMSO) and the aldehydes (each 2.0 μL, stock solution 100 mM in DMSO) were added to an ammonium acetate buffer (10 mM, pH 5.0). Endothiapepsin (169.2 μL, stock solution 0.591 mM) was added in the protein samples. DMSO was added to reach a final concentration of DMSO in the DCLs of 5%. The end-volume was 1 mL. Final concentrations in the DCLs are shown in Table 2. The DCL was left shaking at room temperature and was frequently monitored *via* UPLC-MS.

Table 2. Final concentrations in the DCLs.

	Endothiapepsin	Blank
DMSO	5%	5%
Aniline	10 mM	10 mM
Aldehyde	200 μM (each)	200 μM (each)
Hydroxylamine	20 μM (each)	20 μM (each)
Protein	100 μM	-

For monitoring *via* UPLC-MS, 10 µL of the corresponding library was diluted in 89 µL acetonitrile, the pH was raised to pH > 7 by adding NaOH (1 µL, 1 M) to freeze the reaction. The mixture was centrifuged at 10,000 rpm for 2 minutes, and the supernatant was analysed *via* UPLC-MS.

5.4.3 Cytotoxicity assay; determination of viable cell mass

HepG2 cells (2 x 10^5 cells per well) were seeded in 24-well, flat-bottomed plates. Culturing of cells, incubation and OD measurement were performed as described previously with small modifications.[13] Twenty-four hours after seeding the cells, the incubation was started by the addition of compounds in a final DMSO concentration of 1%. The living cell mass was determined after 48 hours in a PHERAstar microplate reader (BMG labtech, Ortenberg, Germany). Two independent measurements were performed for each compound.

5.4.4 Fluorescence-based Endothiapepsin inhibition assay

Endothiapepsin was purified as described recently.[14] Stock solutions were prepared for all compounds (10 mM DMSO). The final reaction volume was 200 µL containing 0.8 nM endothiapepsin, 1.8 µM substrate (Abz-Thr-Ile-Nle-p-nitro-Phe-Gln-Arg-NH_2 trifluoroacetate salt) (Bachem, Bubendorf, Switzerland) and 5% DMSO. The final concentration of inhibitors was between 0.1 and 200 µM. The assay was performed in flat bottom 96-well microplates (Greiner Bio-One, Frickenhausen, Germany) using a CLARIOstar microplate reader (BMG Labtech, Ortenberg, Germany) at an excitation wavelength of 337 nm and an emission wavelength of 414 nm. The assay buffer (0.1 M sodium acetate buffer, pH 4.6, containing 0.001% Tween 20) was premixed with the substrate and inhibitor; endothiapepsin was added directly before the measurement. The reaction was performed at 25 °C for 20 min and accompanied by detection of the fluorescence signal once per min. The resulting slopes were related to a DMSO control yielding % inhibition values. Each compound was measured at least at two different occasions. The final result represents the average of these measurements. IC_{50} values were calculated by using the software OriginPro 2017 (OriginLab Corporation, Northampton, MA).

5.4.5 Synthesis

General procedure for hydroxylamine formation GP1:

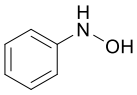
To a rapidly stirred mixture of ammonium chloride (1.2 eq.) and an aryl-nitro compound (1 eq.) in H_2O/ EtOH (2:5), was zinc dust (3 eq.) added portion-wise. The reaction mixture was stirred for one hour at room temperature before filtering through a pad of celite with warm water. After cooling to room temperature, the mixture was extracted using Et_2O (3 x 10 mL), the organic fractions were combined, dried over $MgSO_4$ and filtered. After the evaporation of

the solvent *in vacuo*, the hydroxylamine was obtained as a crystalline solid in 59 — 94% yield.[15]

General procedure for nitrone formation GP2:

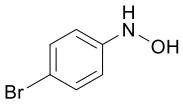
Nitrone (1 eq.) and aldehyde (1 eq.) were dissolved in minimal amount of EtOH in a brown-coloured glass-vial. The mixture was stirred at room temperature for 16 hours. The solvent was removed *in vacuo*, and the product was washed with ice-cold EtOH and purified using column chromatography if needed (DCM/ EtOac 1:0 → 0:1).[15]

N-Phenylhydroxylamine (H1)



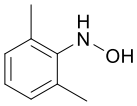
The hydroxylamine was synthesised at room temperature following the general procedure GP1, using nitrobenzene (61.6 mg, 0.5 mmol) in H_2O/ EtOH (2:5, 2.5 mL), zinc dust (98 mg, 1.5 mmol) and ammonium chloride (32 mg, 0.6 mmol). The hydroxylamine **H1** was obtained as a white, crystalline solid (32.1 mg, 59%). ^{1}H-NMR (500 MHz, DMSO-d_6) δ= 8.27 (s, 1H), 8.23 (s, 1H), 7.15 (t, J = 7.9 Hz, 2H), 6.82 (d, J = 7.9 Hz, 2H), 6.73 (t, J = 7.9 Hz, 1H). ^{13}C-NMR (126 MHz, DMSO-d_6) δ= 152.1, 128.4, 119.2, 112.9. NMR data were in accordance with those from the literature.[15]

N-(4-Bromophenyl)hydroxylamine (H2)



The hydroxylamine was synthesised following the general procedure GP1, using 1-bromo-4-nitrobenzene (101 mg, 0.5 mmol) in H_2O/ EtOH (2:5, 2.5 mL), zinc dust (98 mg, 1,5 mmol) and ammonium chloride (32 mg, 0.6 mmol). The hydroxylamine **H2** was obtained as a lightbrown, crystalline solid (88.6 mg, 94%). ^{1}H-NMR (500 MHz, DMSO-d_6) δ= 8.43 (s, 1H), 8.41 (s, 1H), 7.30 (d, J = 8.0 Hz, 2H), 6.77 (d, J = 8.0 Hz, 2H). ^{13}C-NMR (126 MHz, DMSO-d_6) δ= 151.4, 131.1, 114.8, 110.0. NMR data were in accordance with those from the literature. [16]

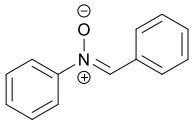
N-(2,6-Dimethylphenyl)hydroxylamine (H3)



The hydroxylamine was synthesised following the general procedure GP1, using 1,3-dimethyl-2-nitrobenzene (75.6 mg, 0.5 mmol) in H_2O/ EtOH (2:5, 2.5 mL), zinc dust (98 mg, 1.5 mmol) and ammonium chloride (32 mg, 0.6 mmol). The hydroxylamine **H3** was obtained as an off-white, crystalline solid (88.6 mg, 94%). ^{1}H-NMR (500 MHz, DMSO-d_6) δ= 8.07 (s, 1H), 7.15 (s, 1H), 6.92 (d, J = 7.4 Hz, 2H), 6.83 (t, J = 7.4 Hz, 1H), 2.27 (s, 6H). ^{13}C-NMR (126 MHz, DMSO) δ= 146.01,

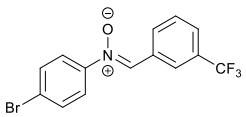
130.51, 128.24, 123.63, 18.19. NMR data were in accordance with those from the literature.[17]

(*Z*)-*N*,1-Diphenylmethanimine oxide (H1A1)



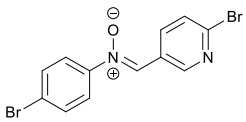
The nitrone **H1A1** was synthesised, using **H1** (10 mg, 0.09 mmol) and benzaldehyde (9.8 µL, 0.09 mmol), following the general procedure GP2. ^{1}H-NMR (500 MHz, $CDCl_3$) δ= 8.42 – 8.39 (m, 2H), 7.93 (s, 1H), 7.81 – 7.76 (m, 2H), 7.49 (m, 6H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ= 149.2, 135.0, 131.2, 130.8, 130.1, 129.4, 129.3, 128.8, 122.0. HRMS (ESI) calcd for $C_{13}H_{11}NO$ [*M*+H]$^+$: 198.0913, found 198.0900. NMR data were in accordance with those from the literature.[15]

(*Z*)-*N*-(4-Bromophenyl)-1-(3-(trifluoromethyl)phenyl)methanimine oxide (H2A2)



The nitrone **H2A2** was synthesised, using **H2** (15.0 mg, 0.08 mmol) and 3-(trifluoromethyl)benzaldehyde (10.7 µL, 0.08 mmol) following the general procedure GP2. After purification by column chromatography, the product **H2A2** was obtained as a white semi-solid (20.1 mg, 73%). ^{1}H-NMR (500 MHz, $CDCl_3$) δ = 8.70 (s, 1H), 8.55 (d, *J* = 8.0 Hz, 1H), 7.98 (s, 1H), 7.74 – 7.58 (m, 6H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ = 147.8, 133.9, 132.6, 131.9, 131.5, 131.2, 129.4, 127.6 (q, *J*=3.6), 125.7 (q, *J*=3.9), 125.0, 124.5, 123.3. HRMS (ESI) calcd for $C_{14}H_{10}BrF_3NO$ [*M*+H]$^+$: 343.9892, found 343.9868.

(*Z*)-*N*-(4-Bromophenyl)-1-(6-bromopyridin-3-yl)methanimine oxide (H2A4)



The nitrone **H2A4** was synthesised at room temperature using **H2** (15.0 mg, 0.08 mmol) and 6-bromo-3-pyridinecarboxaldehyde (14,8 mg, 0.08 mmol) following the general procedure GP2. After purification by column chromatography, the product **H2A4** was obtained as an off-white solid (20.4 mg, 72%). ^{1}H-NMR (500 MHz, $CDCl_3$) δ = 8.99 (dd, *J* = 8.5, 2.3 Hz, 1H), 8.89 (d, *J*=2.3, 1H), 7.94 (s, 1H), 7.69 – 7.58 (m, 5H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ = 150. 6, 147.5, 143.5, 137.0, 132.6, 130.5, 128.4, 126.4, 124.8, 123.2. HRMS (ESI) calcd for $C_{12}H_9Br_2N_2O$ [*M*+H]$^+$: 354.9076, found 354.9052.

5.5 References

[1] O. Ramström, J.-M. Lehn, *Nat. Rev. Drug Discov.* **2002**, *1*, 26–36.

[2] R. Van der Vlag, A. K. H. Hirsch, in *Compr. Supramol. Chem. 2*, Elsevier, **2017**, pp. 487–509.

[3] A. M. Hartman, R. M. Gierse, A. K. H. Hirsch, *European J. Org. Chem.* **2019**, DOI: 10.1002/ejoc.201900327.

[4] G. Durand, A. Polidori, J. P. Salles, B. Pucci, *Bioorganic Med. Chem. Lett.* **2003**, *13*, 859–862.

[5] S. Chen, K. Zhao, G. Chen, *J. Chem.* **2015**, *2015*, 1–6.

[6] Z. Zhang, G. Zhang, Y. Sun, S. S. W. Szeto, H. C. H. Law, Q. Quan, G. Li, P. Yu, E. Sho, M. K. W. Siu, et al., *Sci. Rep.* **2016**, *6*, 1–10.

[7] G. Zhang, T. Zhang, L. Wu, X. Zhou, J. Gu, C. Li, W. Liu, C. Long, X. Yang, L. Shan, et al., *NeuroMolecular Med.* **2018**, *20*, 97–111.

[8] K. R. Maples, F. Ma, Y. K. Zhang, *Free Radic. Res.* **2001**, *34*, 417–426.

[9] K. R. Maples, A. R. Green, R. A. Floyd, *CNS Drugs* **2004**, *18*, 1071–1084.

[10] M. Hochgürtel, *Eur. Pat. Appli.*, EP 1496038 A1, **2005**.

[11] J. W. Sadownik, PhD thesis Evolving Complex Systems from Simple Molecules, http://hdl.handle.net/10023/857, **2008**.

[12] S. M. Turega, C. Lorenz, J. W. Sadownik, D. Philp, *Chem. Commun.* **2008**, 4076–4078.

[13] J. Haupenthal, C. Baehr, S. Zeuzem, A. Piiper, *Int. J. Cancer* **2007**, *121*, 206–210.

[14] M. Mondal, M. Y. Unver, A. Pal, M. Bakker, S. P. Berrier, A. K. H. Hirsch, *Chem. - A Eur. J.* **2016**, *22*, 14826–14830.

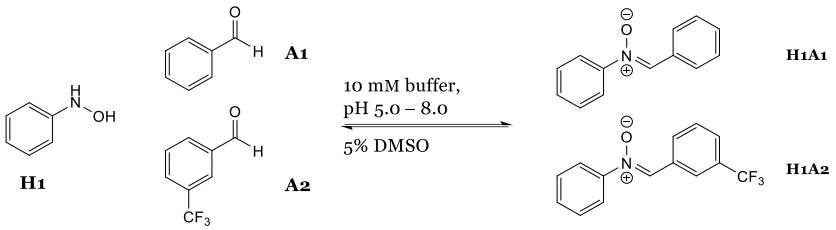
[15] D. A. Evans, H. J. Song, K. R. Fandrick, *Org. Lett.* **2006**, *8*, 3351–3354.

[16] K. N. Hojczyk, P. Feng, C. Zhan, M.-Y. Ngai, *Angew. Chemie Int. Ed.* **2014**, *53*, 14559–14563.

[17] F. Wang, O. Planas, J. Cornella, *J. Am. Chem. Soc.* **2019**, *141*, 4235–4240.

5.6 Supporting information

DCL-N1 pH range:



Scheme S1. DCL-N1, pH window, 20 μM H1, 200 μM **A1** and **A2**.

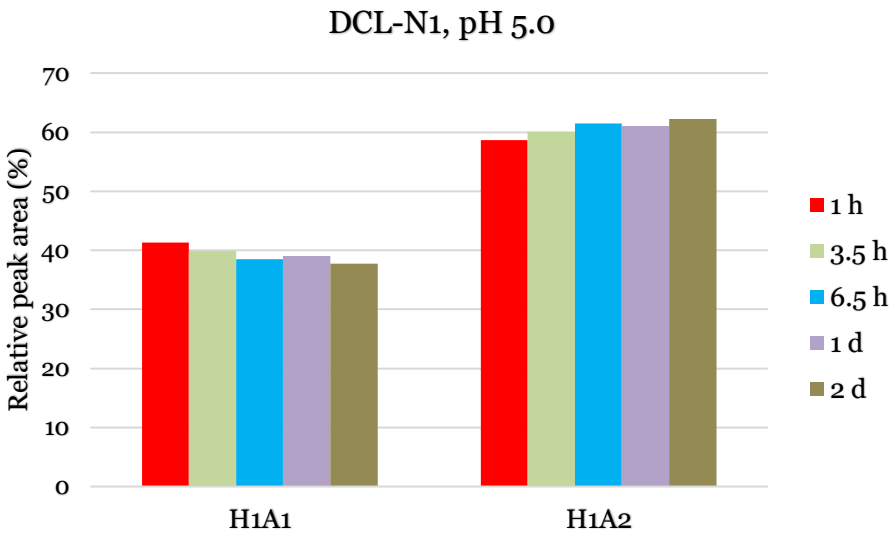


Figure S1. Dynamic combinatorial library-1 at pH 5.0. Data obtained from single experiment.

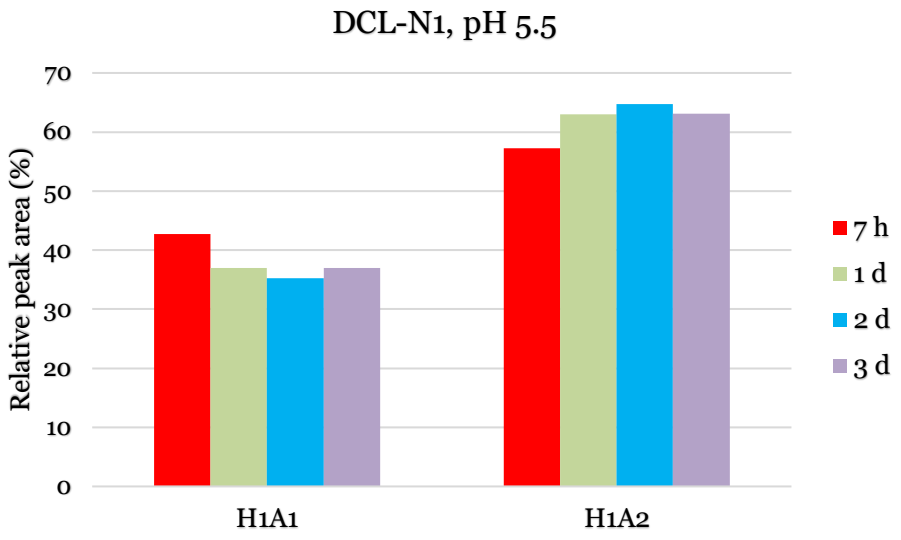


Figure S2. Dynamic combinatorial library-1 at pH 5.5. Data obtained from single experiment.

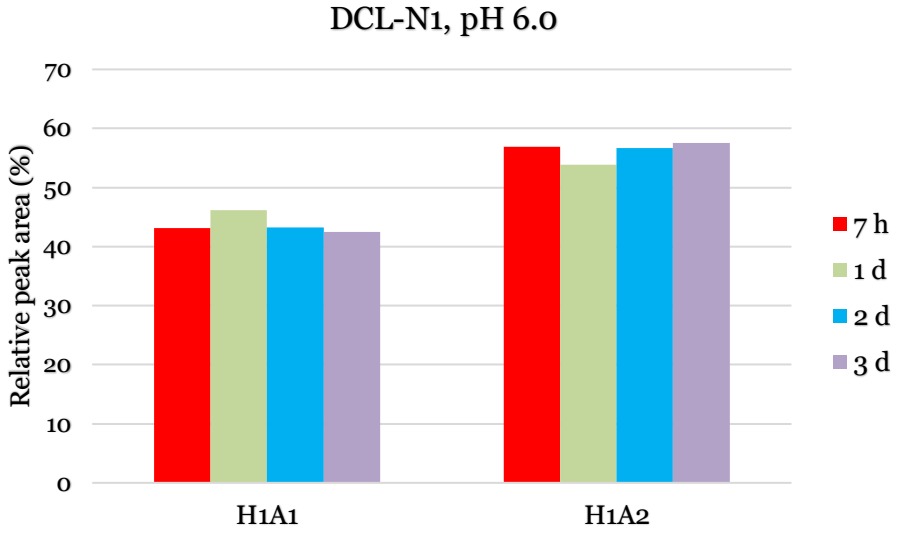


Figure S3. Dynamic combinatorial library-1 at pH 6.0. Data obtained from single experiment.

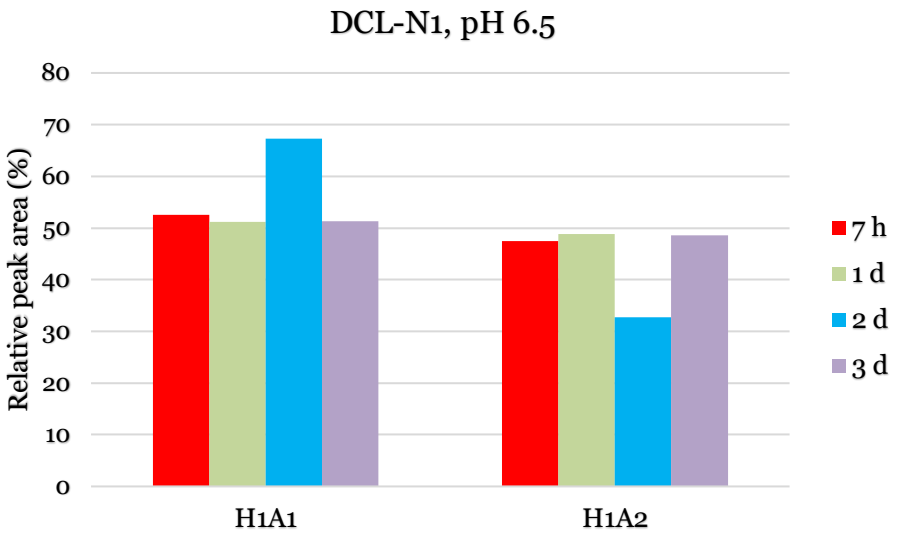
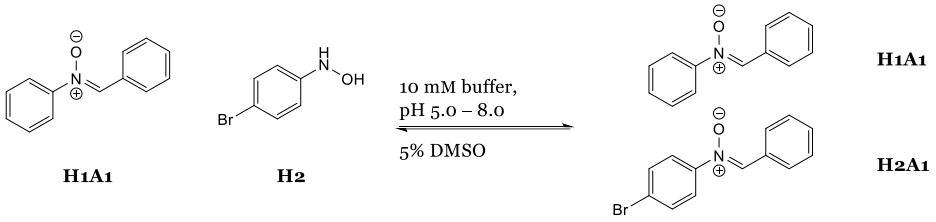


Figure S4. Dynamic combinatorial library-1 at pH 6.5. Data obtained from single experiment.

DCL-N2 pH range:



Scheme S2. DCL-N2, pH window, 20 μM **H1A1**, 200 μM **H2**.

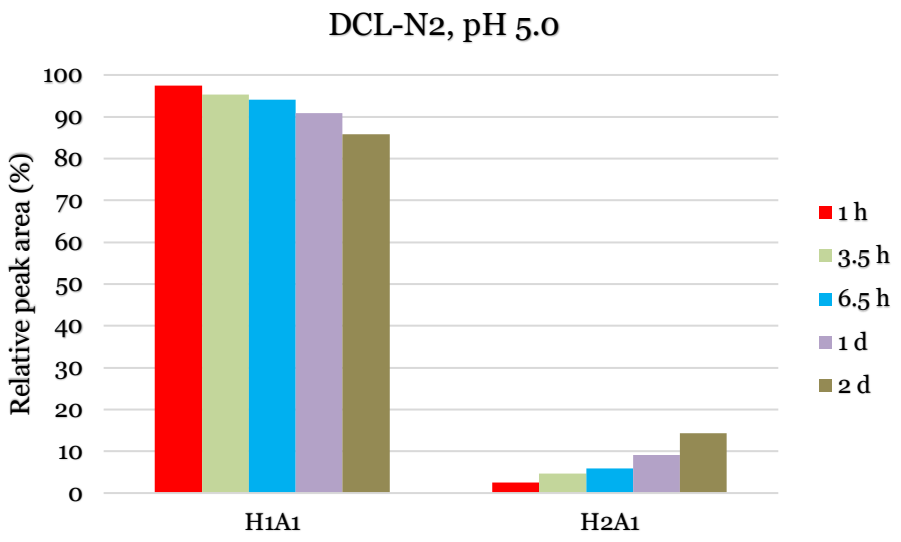


Figure S5. Dynamic combinatorial library-2 at pH 5.0. Data obtained from single experiment.

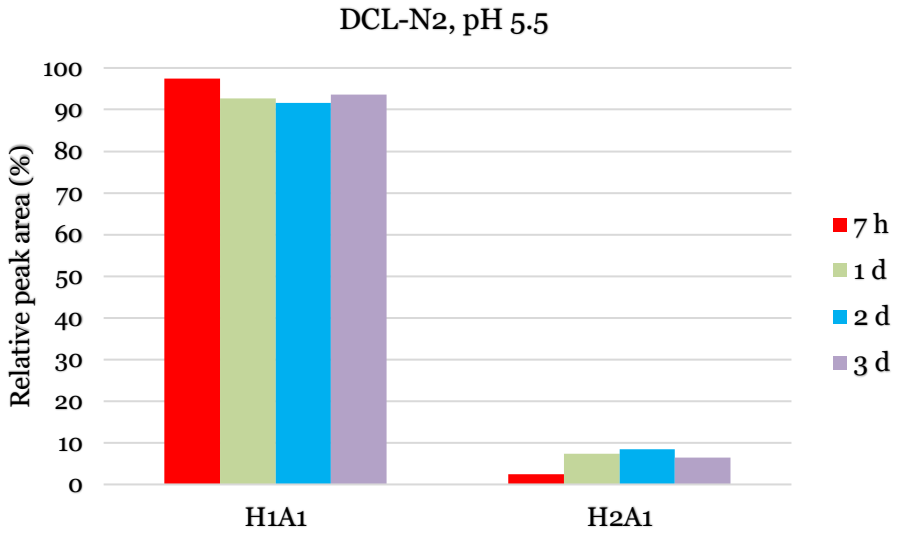


Figure S6. Dynamic combinatorial library-2 at pH 5.5. Data obtained from single experiment.

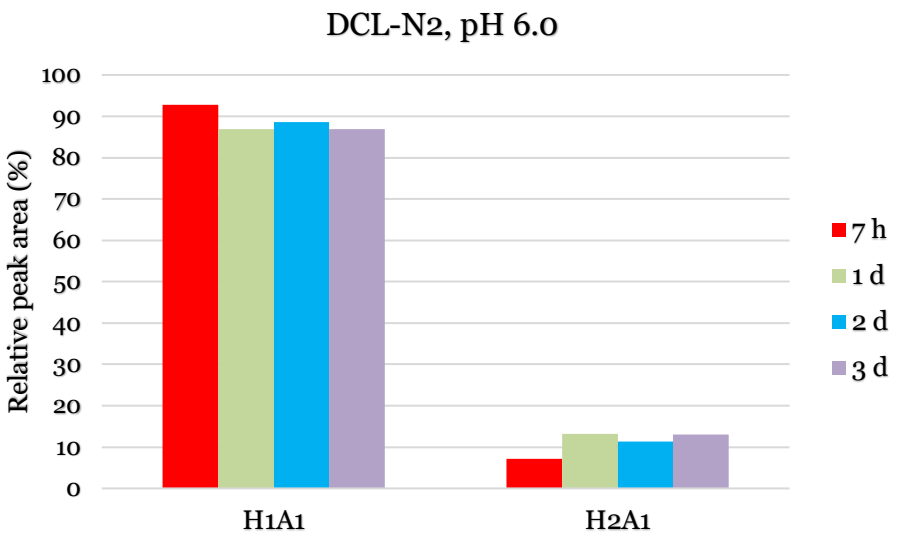


Figure S7. Dynamic combinatorial library-2 at pH 6.0. Data obtained from single experiment.

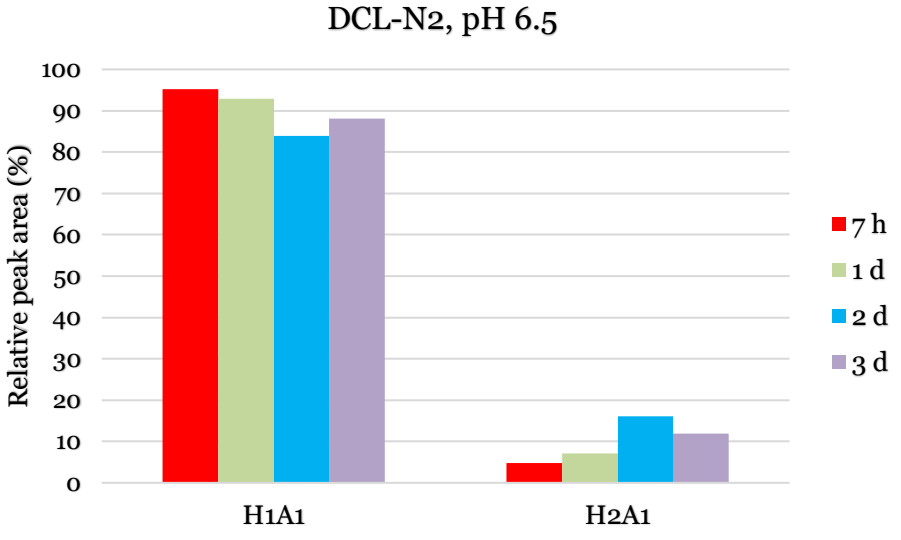


Figure S8. Dynamic combinatorial library-2 at pH 6.5. Data obtained from single experiment.

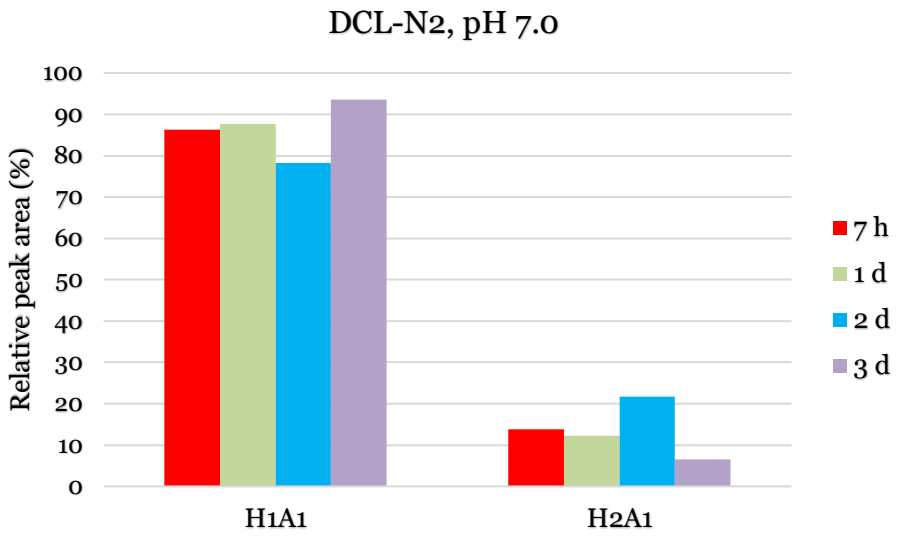


Figure S9. Dynamic combinatorial library-2 at pH 7.0. Data obtained from single experiment.

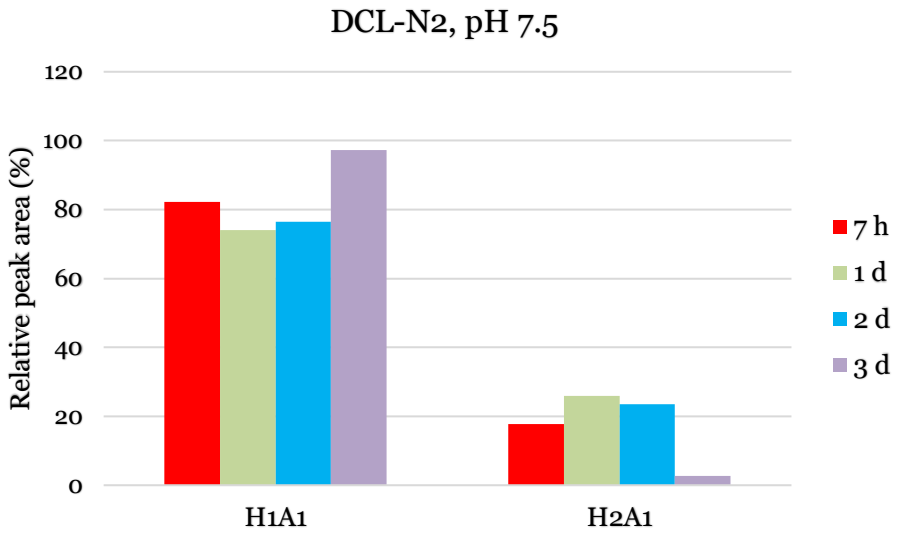


Figure S10. Dynamic combinatorial library-2 at pH 7.5. Data obtained from single experiment.

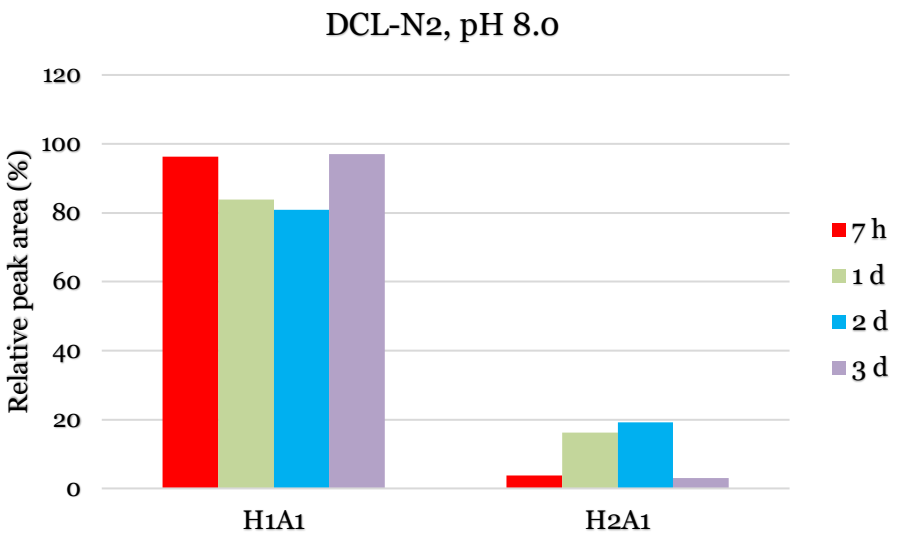
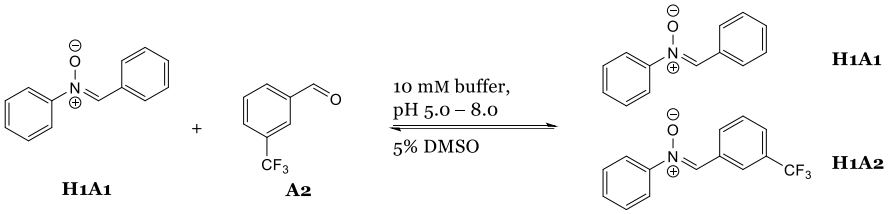


Figure S11. Dynamic combinatorial library-2 at pH 8.0. Data obtained from single experiment.

DCL-N3 pH range:



Scheme S3. DCL-N3, pH window. 20 μM **H1A1**, 200 μM **A2**.

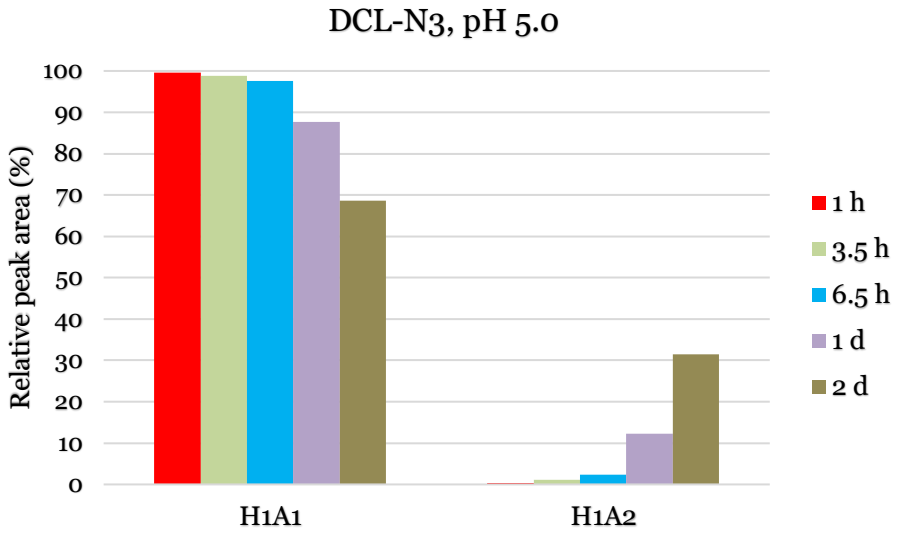


Figure S12. Dynamic combinatorial library-3 at pH 5.0. Data obtained from single experiment.

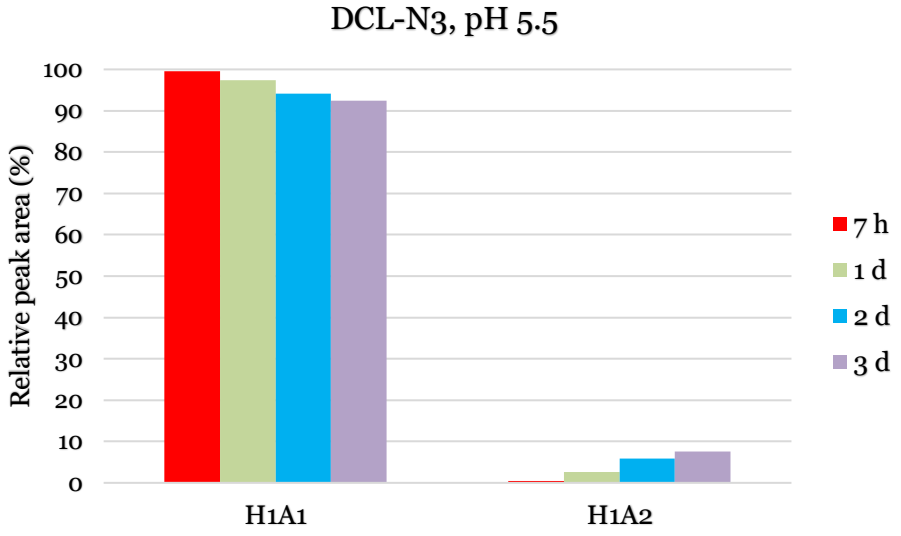


Figure S13. Dynamic combinatorial library-3 at pH 5.5. Data obtained from single experiment.

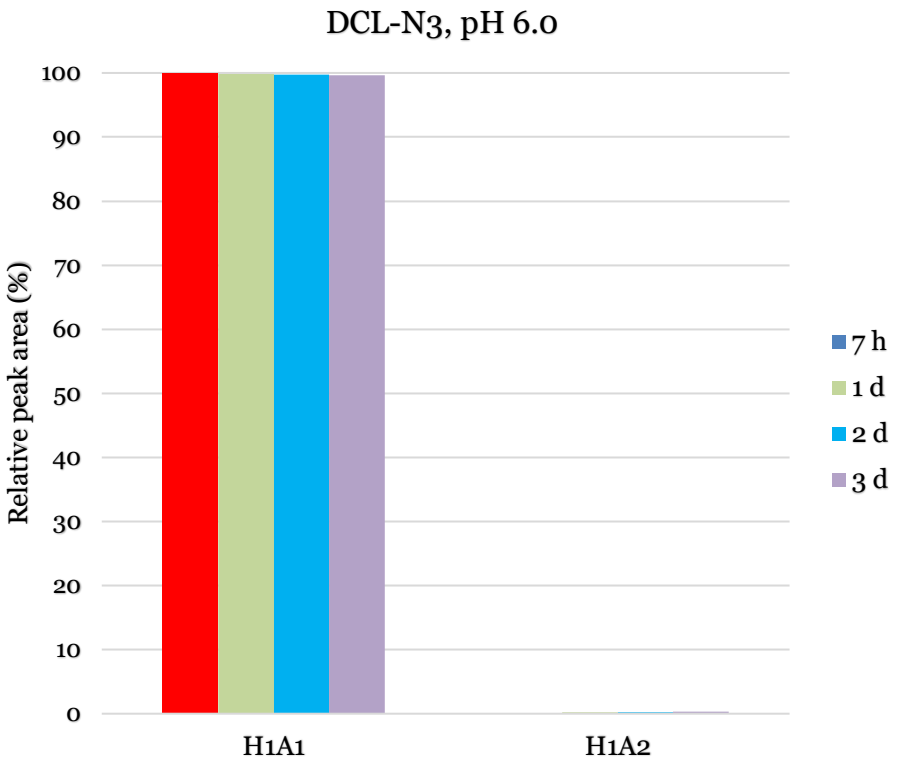


Figure S14. Dynamic combinatorial library-3 at pH 6.0. Data obtained from single experiment.

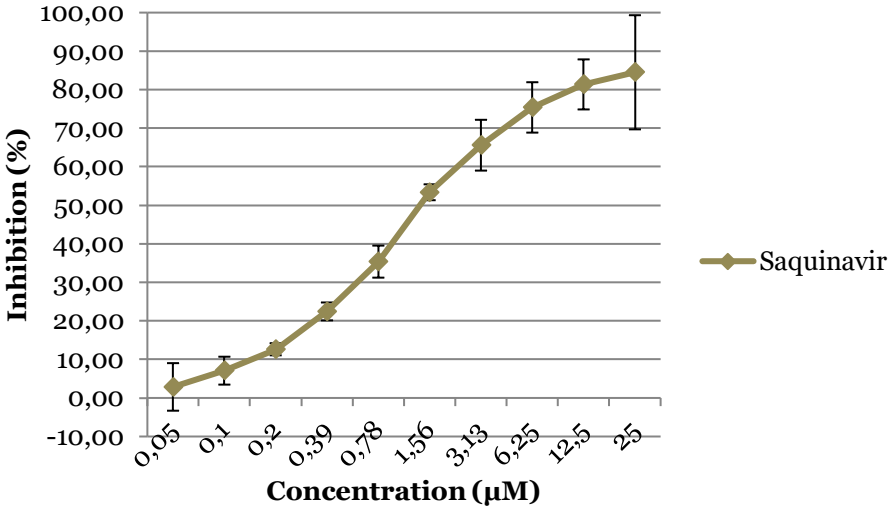


Figure S15. Endothiapepsin activity assay, using saquinavir as positive control. Data points are the average of two independent measurements, and standard deviation is given as error bars.

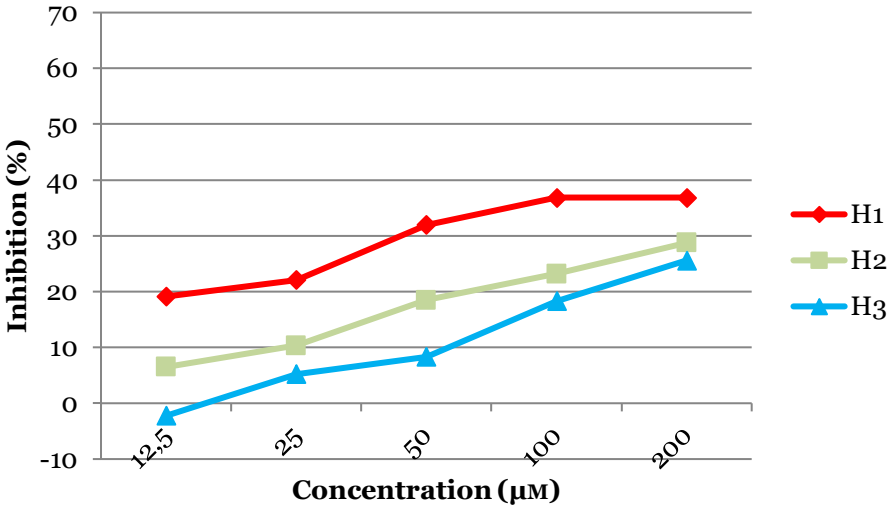


Figure S16. Endothiapepsin activity assay, using hydroxylamines H1 – H3. Partial inhibition was observed for these building blocks. Data obtained from single experiment.

Chapter 6

Thiazolidines in protein-templated Dynamic Combinatorial Chemistry

Expanding the reaction scope for dynamic combinatorial chemistry (DCC) is necessary to further improve the applicability of DCC on protein targets, allowing the medicinal chemist to choose from a diverse set of reversible reactions. In this chapter, we implement the reversible thiazolidine formation for the first time in a protein-templated (pt) DCC setting, using endothiapepsin as a target protein.

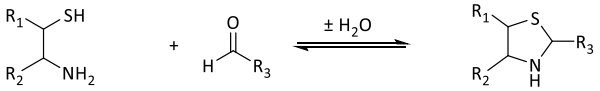
A. M. Hartman, B. Schroeder, A. K. H. Hirsch, *manuscript in preparation*.

A. M. Hartman and B. Schroeder contributed equally to this work. A. M. Hartman was involved in the design of the project, performing the DCC experiments, synthesis of compounds, and writing of the manuscript. B. Schroeder was involved in the DCC experiments, the synthesis of compounds, and writing of the manuscript, and A. K. H. Hirsch was involved in editing the manuscript and supervision of the project.

6.1 Introduction

In 1936, it was proposed that the thiazolidine motif could be of biological interest, due to its reversible formation and dissociation in neutral aqueous solution at room temperature. The thiazolidine motif can be found in natural compounds as well as in drug molecules, and it offers the possibility for multiple interactions in ligand–target binding. In 2009, it was shown that thiazolidines could be used in dynamic combinatorial chemistry (DCC). However, the application on a target was not yet published. In this present work, we present the application of protein-templated DCC based on this promising scaffold, using endothiapepsin as a target protein. Afterwards, the hit compound was analysed for its biochemical properties.

Dynamic combinatorial chemistry (DCC) has evolved into an ever more important technique to identify bioactive molecules.[1–10] Proteins, DNA and also RNA have been used in target-directed (td) DCC and it can be expected that the applicability of DCC will be expanded even further in the future.[6,11–18] Multiple reviews and book chapters on DCC have been written[19–21] and recently we published a minireview on protein-templated (pt) DCC, giving a brief overview and an experimental protocol (see Introduction chapter 1) .[22] The most commonly used reaction in DCC is based on acylhydrazone formation (see Chapters 3 and 4). This requires hydrazide and aldehyde building blocks which form acylhydrazones in a reversible manner. Upon addition of an external stimulus, for example a protein, the formation towards one or more favourable products can be templated. Other reactions used in DCC so far are based on alkene cross metathesis, (boronate) ester-, imine-, hemithioacetal-, disulfide- and thioether formation.[17,18,23–29] This limited scope needs to be extended by identifying other reversible reactions, which can take place in an aqueous environment. Ideally, the type of reaction scaffold would contribute to binding of the ligand to the target. An interesting scaffold, which can be found in natural products as well as in drugs is the thiazolidine motif. This five-membered heterocycle carries a thioether, an amine and can be reversibly formed by condensation of an aminothiol with an aldehyde (Scheme 1).



Scheme 1. Reversible thiazolidine formation by the reaction of an aldehyde and an aminothiol.

A few examples of thiazolidine derivatives in different molecules are given in Figure , illustrating the basic structure of Penicillin[30], the antiparasitic drug

Levamisole[31] and Rosiglitazone (Avandia) and Pioglitazone (Actos), which are both antidiabetic drugs[32].

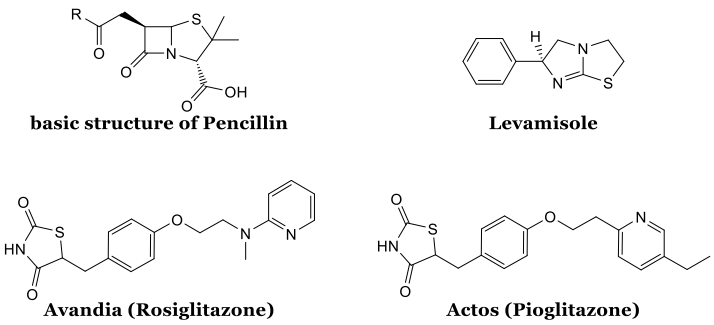


Figure 2. Example of natural compounds and drugs with a thiazolidine-like scaffold.

The structural motif of 2-arylthiazolidine-4-carboxylic acids was found as a class of cytotoxic agents for the treatment of prostate or ovarian cancer. Gududuru *et al.* first introduced a 4-thiazolidinone moiety as a phosphate mimic and then optimised it to 2-arylthiazolidine-4-carboxylic acid amides (Figure 2).[33]

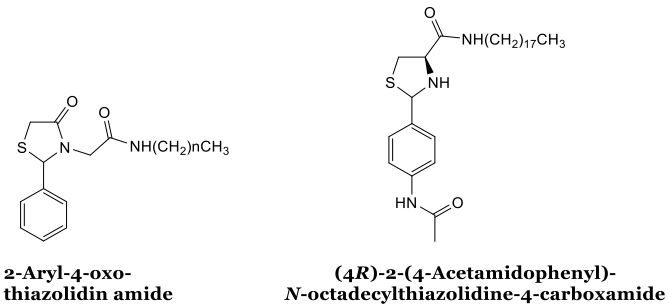


Figure 3. Thiazolidine motifs in a class of cytotoxic agents for prostate cancer.[33]

Although thiazolidines are a common motif in drugs and natural products, they have not been used in protein-templated DCC (ptDCC) until now. A lot of interesting properties make them a promising scaffold for ptDCC. First of all, the nitrogen atom is the only non-metal atom which, by protonation, carries a positive charge at physiological pH. Furthermore, the nitrogen atom is able to contribute to a lot of different supramolecular interactions: free electrons are able to make electrostatic interaction, cations can interact with π-systems and the protons which are bound to the nitrogen can act as H-bond donors. Sulfur atoms, on the other hand, have low-energy σ* orbitals, also called σ-hole.[34] It gives bivalent sulfur atoms the possibility to interact with electron donors like nitrogen- and oxygen atoms and π-systems. However, replacement of an

aromatic ring by a sulfur-containing heterocycle leads to a new substituent trajectory which can be used to optimise the fitting of ligands into a protein binding pocket. Taken together, these properties could lead to a new interesting scaffold for protein-templated DCC.

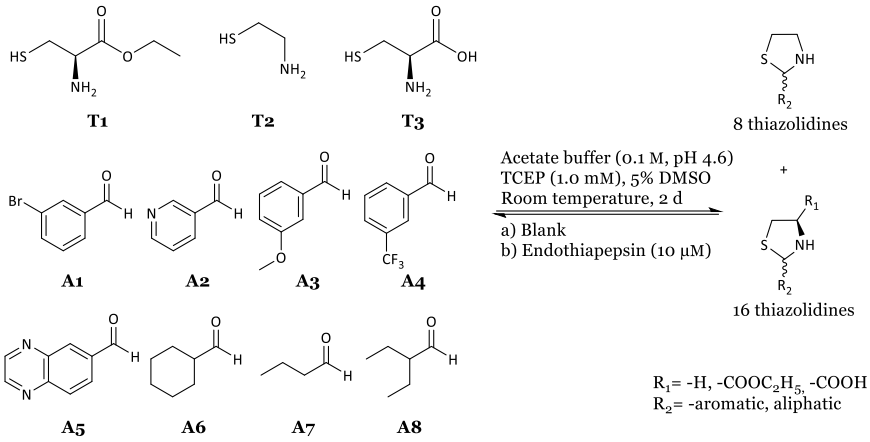
Already in 1936, the first work on thiazolidines was published by Schubert, namely the condenstioncysteine was condensed with various aldehydes. It was mentioned that due to their reversible formation and dissociation in neutral aqueous solution at room temperature, that these structures are of biochemical interest.[35] One year later it was shown that not only aldehydes, but also ketones can be used.[36] Then in 2009, Mahler and coworkers applied thiazolidines for the first time in a DCC experiment. Optimal pH values and temperatures were determined for the formation, equilibration and also for the exchange reaction between different products and building blocks. It was found that building blocks (1 mM) in an acetate buffer at pH 4–5 at room temperature reach equilibrium within two days. Methanol was used as cosolvent (acetate buffer / methanol, 2:1). To be able to analyse the dynamic combinatorial library (DCL), the pH was increased to 7 by the addition of $NaHCO_3$. This `freezes` the equilibrium by blocking its reversibility.[37]

In this work, we present the first application of ptDCC based on thiazolidines. Instead of using methanol as a cosolvent, which might not be tolerated by the protein, we chose to use DMSO in a final concentration of 5%. This allows the building blocks to stay solubilised. Our protein of interest is endothiapepsin, which is a member of the family of pepsin-like aspartic proteases. Aspartic proteases are involved in numerous biological processes; maturation of the HIV virus particle, implication in tumorigenesis by cathepsin D and they are involved in malaria, amyloid disease and other diseases.[38] We have used endothiapepsin as a model enzyme in DCC before, which was based on acylhydrazone formation. We showed that endothiapepsin is stable at pH 4.5 at room temperature for more than 20 days.[25,39,40] These conditions fit well with the conditions reported by Mahler and coworkers, which gave us an optimal starting point.

6.2 Results and Discussion

6.2.1 Design of the libraries.

The group of Mahler used L-cysteine ethyl ester and a variety of aldehydes to form their libraries, so we decided to follow their approach and applied this for the first time on a protein target. We designed dynamic combinatorial library-1 (DCL-1) and used UPLC-MS to monitor the reaction (Scheme 2).



Scheme 2. Design of dynamic combinatorial library-1. Final concentration of aldehydes and aminothiols, 100 μM each; endothiapepsin, 10 μM; TCEP, 1.0 mM.

With the knowledge that, under the applied conditions, the equilibrium would be reached within two days, we took samples at 4 and 8 hours and 1 and 2 days. Each product's mass was extracted from the chromatogram, and the resulting peak was integrated. The relative peak area was then compared between the samples in presence and in absence of protein. As can be seen from Figure 4, the equilibrium was reached around 1 day since no change was observed after this time point. Interestingly, the amplification folds of products with the aromatic aldehydes and L-cysteine ethyl ester (**T1**) decrease and an increase in these products with the same aromatic aldehydes with L-cysteine (**T3**) is observed. As an aspartic protease, endothiapepsin is also able to cleave esters. This is most likely the reason why an increase in concentration of compounds with L-cysteine (**T3**) is observed in the protein sample. Cysteamine (**T2**) products on the other hand cannot be hydrolysed, and accordingly no change in amplification fold was observed.

It must be noted that the concentrations of nearly all compounds decreased after 8 hours. This was seen in the blank as well as in the protein sample. Crashing out of the compounds could lead to this effect, however, no precipitation was observed (the solutions remained clear over time). Another explanation for a decrease in concentration could be the formation of disulfides, which would influence the equilibrium by taking away the thiolamine building blocks. However, hardly any disulfides were observed, presumably owing to the presence of tris(2-carboxylethyl)phosphine (TCEP). TCEP is a mild reducing agent, which is most commonly used in biochemistry for the selective cleavage of disulfide-crosslinked cysteines.[41] The mechanism for this disulfide cleavage is given in Scheme 3.

It starts with the nucleophilic attack by the phosphorus atom on one of the sulfur atoms.[42] The resulting thiolate is protonated and hydrolysis of the phosphonium cation leads to the phosphine oxide. Only the last step of the reaction is irreversible because of the high stability of the phosphorus-oxygen bond. The reducing agent can therefore not be regenerated, and sequential addition should provide the system with a constant reducing power.

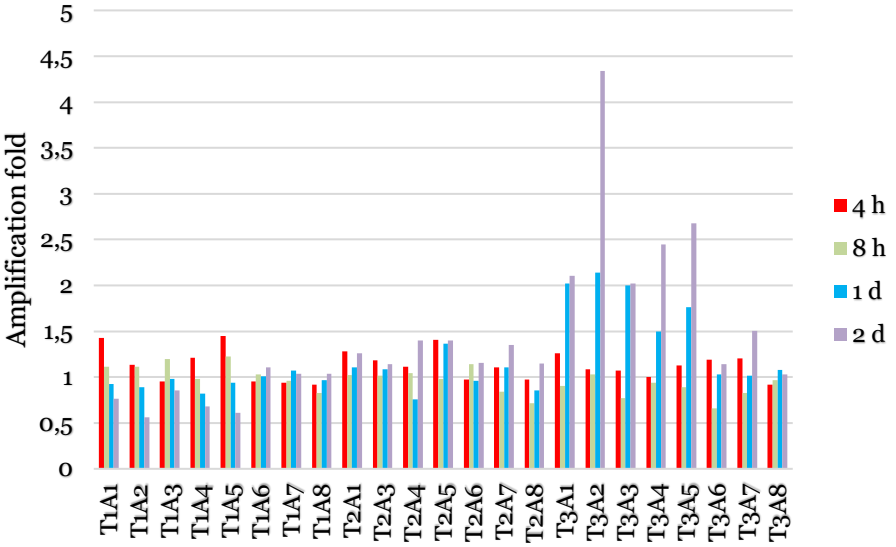
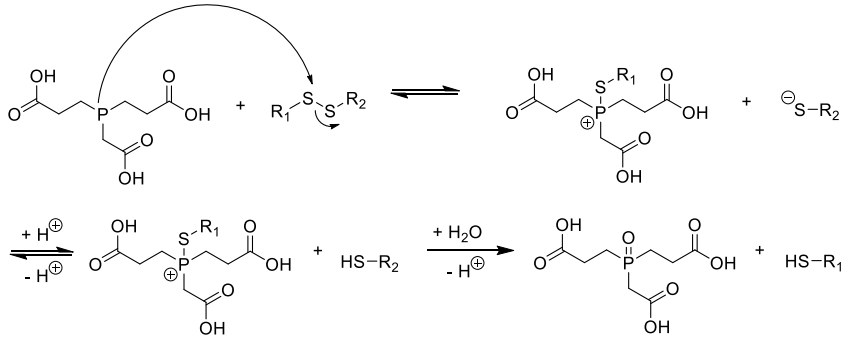
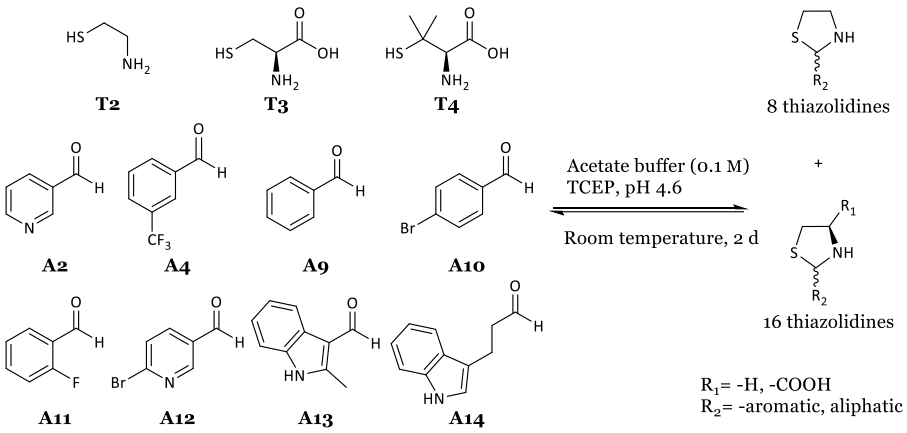


Figure 4. Amplification folds for the products of the dynamic combinatorial library-1. Data obtained from single experiment.



Scheme 3. Cleavage of the disulfide bond by TCEP.

Based on the observation that the ethyl ester functionality was cleaved during the reaction conditions, we designed a second library in which we omitted the ethyl ester (Scheme 4). All products could be observed, except thiazolidines resulting from aldehyde **A13**. This 2-methylindole-3-carboxaldehyde might be sterically hindering thiazolidine formation due to its 2-methyl group. Dynamic combinatorial library 2 was analysed at 7 hours, 1 day and 2 days. The equilibrium was most probably reached at around 1 day, the relative peak areas of the products do not change too much anymore (supporting information, Figure S1). The amplification folds of all products between day 1 and day 2 is approaching towards the value 1, indicating no significant difference between the protein and the blank sample (supporting information, Figure S2). The amplification of compound **T3A2** seems again to be the most, as was also observed in DCL-1. The amplification in DCL-2 is much less than in DCL-1, indicating that indeed in DCL-1 the hydrolysis of the ethyl ester was the major contributor to the amplification of **T3A2**. Even though the amplification of **T3A2** was much smaller in DCL-2, we set out to determine its biochemical properties.



Scheme 4. Design of dynamic combinatorial library-2.

6.3 Cytotoxicity assay

We were interested to see if hit compound **T3A2**, which was amplified in DCL-1 and DCL-2, would show toxicity against a human cell line. For this test, we chose the HepG2 cell line, which is the same as in section 5.2.3. This compound does not inhibit the cell growth and is therefore not cytotoxic.

6.4 Biochemical evaluation of hit T3A2 *via* a fluorescence-based inhibition assay

The product **T3A2** was evaluated for its inhibitory activity against endothiapepsin. The product **T3A2** is a very weak inhibitor, having only 14% inhibition at 500 μM (Figure 4). This finding is in line with DCL-2, where this product was present but not amplified.

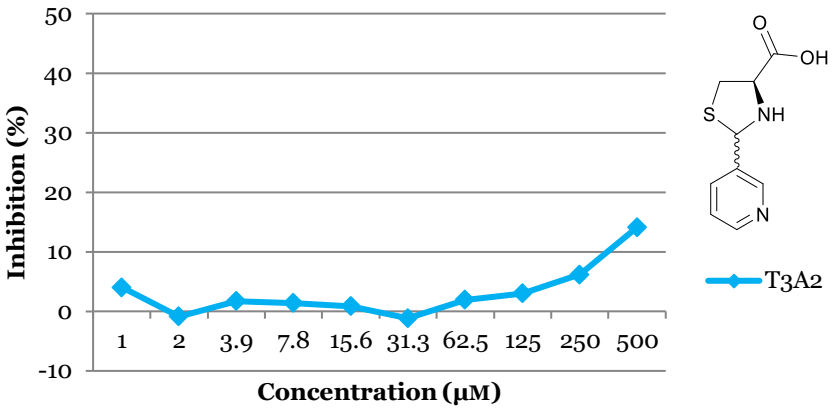
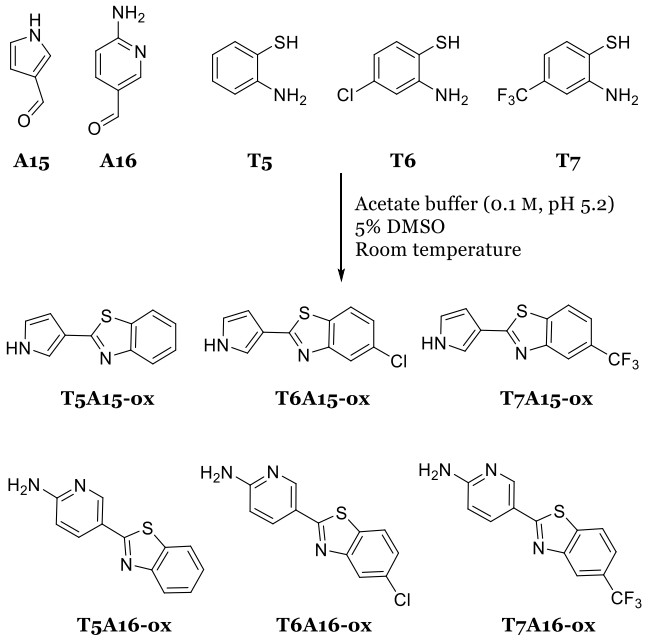


Figure 5. Endothiapepsin activity assay showing that **T3A2** is a very weak inhibitor. Data obtained from single experiment.

6.5 Expanding the reaction scope to aromatic aminothiols

We wanted to see if the thiazolidine reaction scope could be expanded to products which are based on aromatic aminothiols (**T5**—**T7**) (Scheme 5). After analysing the library, we predominately found the oxidised products, which have a conjugated aromatic system. Unfortunately, this makes the reaction irreversible and therefore not applicable for DCC.



Scheme 5. Initial library resulting in the formation of oxidised products.

6.6 Conclusions

From the first two libraries, it can be concluded that thiazolidine formation can be applied under the present conditions. Changing from 33% methanol, in the work of Mahler, to 5% DMSO allowed us to apply this reaction in a ptDCC setting. We illustrated the use of reversible thiazolidine formation with L-cysteine derivatives in presence of endothiapepsin, as a protein target. Since endothiapepsin belongs to the aspartic protease family, it can cleave peptides as well as esters, which we observed in DCL-1. Cleavage of the ester in **T1A2** resulted in **T3A2**, which was therefore amplified. By omitting the ester functionality we were hoping to see an amplification of one of the products. Unfortunately, no significant amplification was observed between the blank and the protein sample.

Biochemical evaluation of **T3A2** showed that it has very weak inhibitory activity of endothiapepsin. Therefore, this amplified product of DCL-1 is not a real hit, which was also seen in DCL-2, where no product was amplified. The work described herein shows that 2-arylthiazolidine-4-carboxylic acids and 2-arylthiazolidine-4-esters can be used in a ptDCC setting. This work paves the way to apply ptDCC on real drug targets as for example in the work of Gududuru *et al.*. These target would be very appropriate since already the activity of these thiazolidine structures is proven.[33]

As a second part of this work, we tried to expand the reaction scope towards aromatic aminothiols. However, this proved unsuccessful. These products oxidise rapidly, which makes the reaction irreversible and therefore incompatible for DCC. However, when such scaffolds are desired, then the formation of such diverse static libraries might still be of interest. MS-binding assays as performed by Wanner and coworkers might be a good approach to distinguish binders from non-binders.[44,45]

6.7 Experimental

6.7.1 Materials and methods

See section 3.4.1.

For high-resolution mass spectroscopy, a ThermoScientific Accucore Phenyl Hexyl column, 100 x 2.1 mm, was used for separation under the same conditions as described before.

6.7.2 DCC conditions

The corresponding cysteine derivatives (**T1** – **T3**, each 1 µL, stock solution 100 mM in DMSO) and the aldehydes (each 1.0 µL, stock solution 100 mM in DMSO) were added to an ammonium acetate buffer (100 mM, pH 4.6). Endothiapepsin (169.2 µL, stock solution 0.591 mM) was added accordingly. DMSO was added to reach a final concentration of DMSO in the DCL of 5%. TCEP was added as a disulfide reducing agent (10 µL, 100 mM in DMSO). The end-volume was 1 mL. Final concentrations in the DCLs are shown in Table 1. The DCL was left shaking at room temperature and was frequently monitored *via* UPLC-MS.

Table 6. Final concentrations in the DCLs.

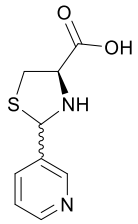
	Endothiapepsin	Blank
DMSO	5%	5%
Cysteine derivatives	100 µM (each)	100 µM (each)
Aldehyde	100 µM (each)	100 µM (each)
TCEP	1 mM	1 mM
Protein	100 µM	-

For monitoring *via* UPLC-MS, 10 µL of the corresponding library was diluted in 89 µL acetonitrile, the pH was raised to pH > 7 by adding NaOH (1 µL, 1.0 M) to freeze the reaction. The mixture was centrifuged at 10,000 rpm for 2 minutes, and the supernatant was analysed *via* UPLC-MS.

6.7.3 Cytotoxicity assay; determination of viable cell mass
See section 5.4.3.

6.7.4 Fluorescence-based endothiapepsin inhibition assay
See section 5.4.4.

6.7.5 Synthesis
(4*R*)-2-(Pyridin-3-yl)thiazolidine-4-carboxylic acid (T3A2)



L-Cysteine (1.5 mmol, 181.7 mg) was dissolved in 1 mL distilled water and to this a solution of nicotinaldehyde (1.0 mmol, 94 µL) in 1 mL EtOH and 0.5 mL EtOAc was added dropwise. The mixture was stirred over night at room temperature. Excess L-cysteine remained as solid in the flask and only the supernatant was taken up by pipette. The product was dried *in vacuo*, and afterwards washed with ice-cold water. Remaining solvent was evaporated *in vacuo* to obtain thiazolidine **T3A2** as a white, crystalline solid (60 mg, 29%) ^{1}H-NMR (500 MHz, DMSO-d$_6$) δ = 8.68 (d, J = 2.0 Hz, 0.5H), 8.62 (d, J = 2.1 Hz, 0.5H), 8.52 (dd, J = 4.8, 1.6 Hz, 0.5H), 8.47 (dd, J = 4.7, 1.5 Hz, 0.5H), 7.98 (dt, J = 7.9, 1.8 Hz, 0.5H), 7.84 (dt, J = 7.9, 1.8 Hz, 0.5H), 7.40 (dd, J = 7.9, 4.8 Hz, 0.5H), 7.37 (dd, J = 7.9, 4.8 Hz, 0.5H), 5.74 (s, 0.5H), 5.55 (s, 0.5H), 4.20 (dd, J = 6.9, 4.9 Hz, 0.5H), 3.92 (dd, J = 8.9, 6.9 Hz, 0.5H), 3.40 – 3.28 (m, 1H), 3.17 – 3.06 (m, 1H) HRMS (ESI) calcd for $C_9H_{10}N_2O_2S$ $[M+H]^+$:211.0536, found 211.0515. NMR data is in correspondence with those reported in the literature.[33]

6.8 References

[1] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84.

[2] R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* **1991**, *354*, 84–86.

[3] R. A. Houghten, *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 5131–5135.

[4] H. M. Geysen, R. H. Meloen, S. J. Barteling, *Proc. Natl. Acad. Sci.* **1984**, *81*, 3998 LP – 4002.

[5] R. Frank, W. Heikens, G. Heisterberg-Moutsis, H. Blöcker, *Nucl. Acid. Res.* **1983**, *11*, 4365–4377.

[6] I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci.* **1997**, *94*, 2106–2110.

[7] M. H. Ohlmeyer, R. N. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci.* **1993**, *90*, 10922 LP – 10926.

[8] V. A. Polyakov, M. I. Nelen, N. Nazarpack-Kandlousy, A. D. Ryabov, A. V Eliseev, *J. Phys. Org. Chem.* **1999**, *12*, 357–363.

[9] A. Ganesan, *Angew. Chemie Int. Ed.* **1998**, *37*, 2828–2831.

[10] C. Karan, B. L. Miller, *Drug Discov. Today* **2000**, *5*, 67–75.

[11] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* **2010**, *2*, 490–497.

[12] Z. Fang, W. He, X. Li, Z. Li, B. Chen, P. Ouyang, K. Guo, *Bioorganic Med. Chem. Lett.* **2013**, *23*, 5174–5177.

[13] M. Demetriades, I. K. H. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y. M. Tian, T. D. W. Claridge, P. J. Ratcliffe, E. C. Y. Woon, et al., *Angew. Chemie - Int. Ed.* **2012**, *51*, 6672–6675.

[14] S. Sakai, Y. Shigemasa, T. Sasaki, *Tetrahedron Lett.* **1997**, *38*, 8145–8148.

[15] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Angew. Chemie Int. Ed.* **2002**, *41*, 3405–3407.

[16] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Tetrahedron* **2004**, *60*, 771–780.

[17] B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* **2006**, *128*, 8459–8467.

[18] R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero, O. Ramström, *Angew. Chemie Int. Ed.* **2010**, *49*, 589–593.

[19] P. Frei, R. Hevey, B. Ernst, *Chem. Eur. J.* **2019**, *25*, 60–73.

[20] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455–2488.

[21] R. Van der Vlag, A. K. H. Hirsch, in *Compr. Supramol. Chem. 2*, Elsevier, **2017**, pp. 487–509.

[22] A. M. Hartman, R. M. Gierse, A. K. H. Hirsch, *European J. Org. Chem.* **2019**, DOI: 10.1002/ejoc.201900327.

[23] G. Artigas, P. López-Senín, C. González, N. Escaja, V. Marchán, *Org. Biomol. Chem.* **2015**, *13*, 452–464.

[24] S. Jana, D. Panda, P. Saha, G. D. Pantos, J. Dash, *J. Med. Chem.* **2019**, *62*, 762–773.

[25] M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch, *Angew. Chemie Int. Ed.* **2016**, *55*, 9422–9426.

[26] I. K. H. Leung, T. Brown, C. J. Schofield, T. D. W. Claridge, *Medchemcomm* **2011**, *2*, 390–395.

[27] W. He, Z. Fang, Z. Yang, D. Ji, K. Chen, K. Guo, *RSC Adv.* **2015**, *5*, 23224–23228.

[28] S.-A. Poulsen, L. F. Bornaghi, *Bioorg. Med. Chem.* **2006**, *14*, 3275–3284.

[29] B. M. R. Liénard, R. Hüting, P. Lassaux, M. Galleni, J.-M. Frère, C. J. Schofield, *J. Med. Chem.* **2008**, *51*, 684–688.

[30] A. J. Wright, *Mayo Clin. Proc.* **1999**, *74*, 290–307.

[31] N. Scheinfeld, J. D. Rosenberg, J. M. Weinberg, *Am. J. Clin. Dermatol.* **2004**, *5*, 97–104.

[32] M. T. Tran, M. D. Navar, M. B. Davidson, *Diabetes Care* **2006**, *29*, 1395–1396.

[33] V. Gududuru, E. Hurh, J. T. Dalton, D. D. Miller, *J. Med. Chem.* **2005**, *48*, 2584–2588.

[34] B. R. Beno, K.-S. Yeung, M. D. Bartberger, L. D. Pennington, N. A. Meanwell, *J. Med. Chem.* **2015**, *58*, 4383–4438.

[35] M. P. Schubert, *J. Biol. Chem.* **1936**, *114*, 341–350.

[36] G. E. Woodward, E. F. Schroeder, *J. Am. Chem. Soc.* **1937**, *59*, 1690–1694.

[37] C. Saiz, P. Wipf, E. Manta, G. Mahler, *Org. Lett.* **2009**, *4*, 3170–3173.

[38] J. B. Cooper, *Curr. Drug Targets* **2002**, *3*, 155–173.

[39] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259–3263.

[40] M. Mondal, D. E. Groothuis, A. K. H. Hirsch, *Medchemcomm* **2015**, *6*, 1267–1271.

[41] P. Liu, B. W. O'Mara, B. M. Warrack, W. Wu, Y. Huang, Y. Zhang, R. Zhao, M. Lin, M. S. Ackerman, P. K. Hocknell, et al., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 837–844.

[42] J. A. Burns, J. C. Butler, J. Moran, G. M. Whitesides, *J. Org. Chem.* **1991**, *56*, 2648–2650.

[43] M. Mondal, M. Y. Unver, A. Pal, M. Bakker, S. P. Berrier, A. K. H. Hirsch, *Chem. - A Eur. J.* **2016**, *22*, 14826–14830.

[44] P. Neiens, A. De Simone, A. Ramershoven, G. Höfner, L. Allmendinger, K. T. Wanner, *Biomed. Chromatogr.* **2018**, *32*, 1–18.

[45] F. Kern, K. T. Wanner, *Bioorg. Med. Chem.* **2019**, *27*, 1232–1245.

[46] J. Haupenthal, C. Baehr, S. Zeuzem, A. Piiper, *Int. J. Cancer* **2007**, *121*, 206–210.

6.9 Supporting information

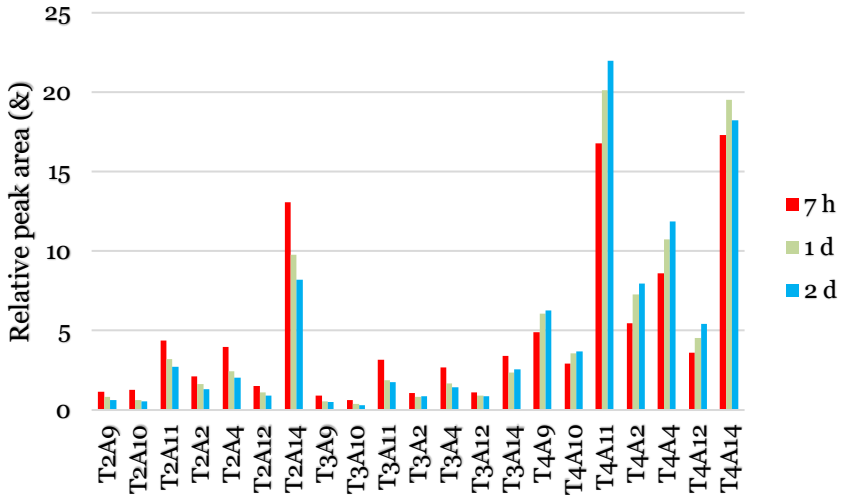


Figure S1. Dynamic combinatorial library-2: relative peak area of the ion counts of each individual product over time in the protein sample. Data obtained from single experiment.

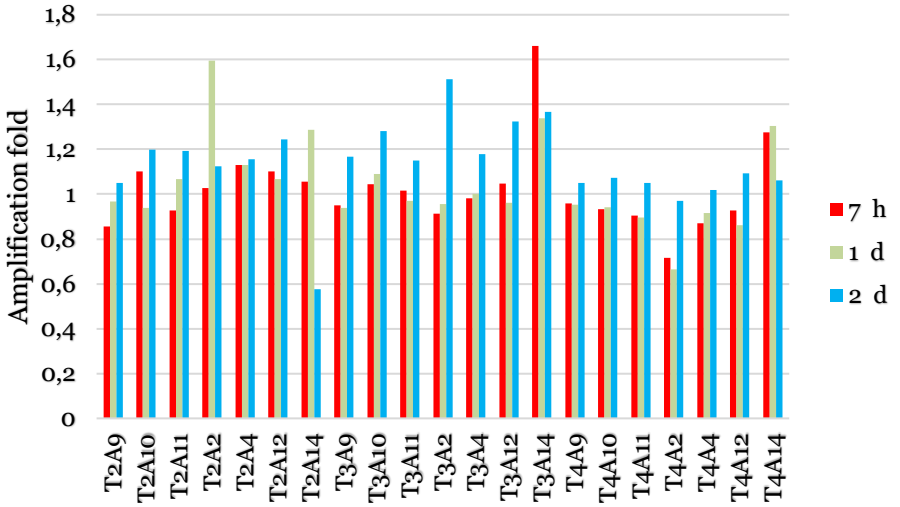


Figure S2. Dynamic combinatorial library-2: amplification fold (P/B); ion counts of individual relative peaks in the protein sample were divided by the relative peaks of the blank sample. Data obtained from single experiment.

Summary & Perspectives

7.1 Context and scope of this thesis

Applying dynamic combinatorial chemistry (DCC) to medicinal chemistry projects can be a helpful strategy for finding starting points in the drug-discovery process. Especially, when information of the target is known, structure-based or fragment-based drug design approaches combined with DCC could lead to potent compounds. As a first aim of this thesis, we have applied DCC on a number of biologically relevant targets, Chapters 3 and 4.

As relevant drug target, 14-3-3 proteins play a role in several diseases and many biological processes. Proteins of this family engage in protein-protein interactions (PPIs), and can do so with numerous different binding partners. By forming these PPIs, these proteins regulate their binding partner's activity. The activity can be both up- or down-regulated.

Another family of relevant targets are glucansucrases, which are important enzymes in the initiation and development of cariogenic dental biofilms, commonly known as dental plaque. Inhibiting these enzymes might therefore prevent dental caries.

The second aim of this thesis was to extend the list of reversible reactions applicable in a tdDCC setting. We have shown in Chapters 5 and 6 that nitrones and thiazolidines could be applied in DCC under physiological conditions. Based on this finding, we ran ptDCC experiments using these reversible linkages with endothiapepsin as a target protein. An important aspect in DCC is the stability of compounds, as noted for the nitrones and thiazolidines. Dynamic combinatorial libraries (DCLs) of these types of linkages should therefore be carefully monitored. Another important aspect in DCC is the analysis of the DCLs, which is best performed by determining the surface area under each peak in the UV chromatogram of the LC-MS. For the libraries in Chapters 5 and 6, analysis via UV was not possible, as most products were not UV-active under the screened wavelengths. We therefore analysed these libraries using the total ion counts of each product. To avoid possible problems of ionisation efficiency and the injection volume, which can deviate per run, we have taken the relative ion counts. This gives the ratio of products as a percentage, which allows the comparison of the libraries even when factors like injection volume or the ionisation might not be identical per run. One could consider adding an internal standard, which allows for a quantitative analysis.

In the last two chapters, endothiapepsin was used for protein-templated DCC (ptDCC). Endothiapepsin belongs to the family of the aspartic proteases. Aspartic proteases are involved in numerous biological processes and diseases, for example the maturation of the HIV virus particle.

Throughout this thesis, we focus on applying DCC to various projects. The main achievements are: 1) the description of the in-house protocol of DCC, in which aspects like solubility of building blocks and products, protein stability and more need to be taken in to account, 2) the application of acylhydrazone-based DCC to two attractive targets, a (PPI)-target and a glucansucrase, 3) the identification of small-molecules, which stabilise PPIs of 14-3-3/ synaptopodin, 4) expanding the reaction toolbox of ptDCC by two additional reactions: nitrone and thiazolidine formation.

Here, we summarise the main results reported in this thesis and we give an outlook on DCC for medicinal chemistry projects.

7.2 Summary

In **Chapter 1**, provides a clear and concise description of a comprehensive DCC protocol. This work focused on important aspects concerning the practical handling in ptDCC. The best biological and chemical conditions should be harmoniously combined for successful application in ptDCC. Moreover, analytical techniques used to determine which compounds are favoured were briefly discussed.

In **Chapter 2**, we present an overview of modulators of 14-3-3 proteins. We critically analysed the known modulators in terms of their structures and molecular recognition by 14-3-3 proteins. We proposed small modifications of certain modulators, based on reported cocrystal structures.

In **Chapter 3**, we describe the identification of modulators of 14-3-3 PPIs *via* DCC. We targeted a PPI-complex of 14-3-3(ζ) with synaptopodin, a 21 amino acid peptide chain. Hit compounds of the DCC experiments were analysed for their biochemical activities *via* surface plasmon resonance (SPR). These compounds show a stabilising effect on the PPI-complex. Co-crystallisation studies are ongoing to confirm the binding mode, setting the stage for SBDD to optimise the affinity.

In **Chapter 4**, we applied acylhydrazone-based DCC on a glucansucrase. Glucansucrases play an important role in cariogenic dental biofilms, which are causative for dental caries. Glucansucrase inhibitors should therefore prevent caries and may be added as additives in toothpaste. We described the syntheses, DCC experiments as well as biochemical characterisation of the hit compounds.

In **Chapter 5**, we describe the first application of nitrone-based ptDCC. We optimised conditions for ptDCC, keeping in mind both the target protein endothiapepsin as well as the applicability of the nitrone reaction. Besides the application to ptDCC, we determined the optimal pH window of this `new`

reaction. Synthesis of hit compounds from the DCC experiments enabled evaluation of their biochemical activity in a fluorescence-based activity assay. Preliminary results on the cytotoxicity of the nitrone linkage showed no cytotoxicity on the liver cell line, HepG2.

In **Chapter 6**, we portray another new reaction in ptDCC, using thiazolidines. Thiazoldines can be found in numerous natural compounds and are therefore very attractive scaffolds from a biochemical point of view. We described a number of DCC experiments, illustrating that this reaction can also be applied in medicinal chemistry projects. Expansion of the reaction scope from cysteine derivatives towards aromatic aminothiol building blocks resulted in the oxidation of the products. This oxidation makes the reaction irreversible, and therefore thiazolidine DCC is limited for now to cysteine derivatives.

7.3 Perspectives

As can be seen from this thesis, dynamic combinatorial chemistry can be applied to a variety of different targets. It opens up the possibility of finding starting points in medicinal-chemistry projects. It can be expected that the number of successful DCC projects will increase over time, as more and more research groups are showing interest in this technique. The applicability of DCC in a high-throughput manner is still limited but this might change in the near future, if the analysis *via* LC-MS can be automated. This would open-up the possibility of running library sizes over 100 compounds, especially when extracted ion counts of each individual product will be used instead of analysis *via* UV.

DCC will likely become also more important in optimisation rounds. Hit compounds of previous studies, obtained *via* DCC or *via* other hit-discovery strategies, can be valuable starting points. Molecules can for example also be grown by applying DCC. Drug discovery requires multiparameter optimisation, focusing on more than just affinity. Therefore, DCC will most likely stay most useful in the early phases of the drug-discovery process.

Since the chemical linkage formed in the reversible reactions used in DCC (for example: imine, hydrazone, thiazolidine) is part of the drug-molecule, it could also be involved in binding to the protein. Depending on the type of target, it would be helpful to have a variety of reversible reaction at hand. This would require a constant search for new reactions and the corresponding scaffolds which can be applied in a tdDCC setting.

All in all, we have given examples of the broad utility of DCC, showing why this technique is evolving over time!

Samenvatting

Het toepassen van dynamische combinatorische chemie (DCC) in medicinale-chemieprojecten kan een waardevolle strategie zijn om startpunten te vinden in het medicijn-ontdekkingsproces. Vooral wanneer informatie over het biologische doelwit bekend is, kan structuur-gebaseerd medicijnontwerp (SGMO) of fragment-gebaseerd medicijnontwerp (FGMO) tot potente verbindingen leiden.

Als relevante biologische doelwitten, spelen 14-3-3 eiwitten een rol in verschillende ziekten en biologische processen. Eiwitten in deze familie gaan eiwit-eiwit interacties (EEIs) aan, en kunnen dat doen met een groot aantal verschillende interactiepartners. Door het aangaan van deze EEIs, beïnvloeden ze de werking van hun interactiepartner. Deze werking kan zowel gestimuleerd als ook geremd worden.

Een andere biologisch relevante familie is die van de glucansucrasen, welke belangrijke enzymen zijn in de start en ontwikkeling van cariogene gebitsbiofilmen, ook wel bekend als tandplak.

In de laatste twee hoofdstukken wordt gebruik gemaakt van endothiapepsin in eiwit-gerichte (eg) DCC. Endothiapepsin behoort tot de familie van de aspartaatproteasen. Deze zijn betrokken bij veel biologische processen en ziekten, bijvoorbeeld de rijping van het hiv-virus.

In dit proefschrift focussen we op het toepassen van DCC in verschillende projecten. De belangrijkste resultaten zijn: 1) de beschrijving van ons eigen DCC-protocol, waarin rekening gehouden wordt met belangrijke aspecten als oplosbaarheid van bouwstenen en producten, eiwitstabiliteit en meer, 2) de toepassing van acylhydrazon-gebaseerde DCC op twee aantrekkelijke doelwitten, een EEI-doelwit en een glucansucrase, 3) de identificatie van kleine moleculen die de EEI van 14-3-3/ synaptopodin stabiliseren, 4) het uitbreiden van de gereedschapskist van egDCC middels twee aanvullende reacties: nitron en thiazolidin formatie.

In **Hoofdstuk 1** van dit proefschrift, geven we een heldere en bondige beschrijving van het uitgebreide DCC-protocol. Dit werk concentreert zich op belangrijke aspecten betreffende de praktische uitvoering in eiwit-gerichte dynamische combinatorische chemie (egDCC). Om een succesvolle toepassing van egDCC te bewerkstelligen, moeten de beste biologische en chemische condities harmonieus samengevoegd worden. Bovendien bediscussiëren we de analytische technieken die toegepast worden om te bepalen welke verbindingen geamplificeerd zijn.

In **Hoofdstuk 2** presenteren we een overzicht van regelaars van 14-3-3 eiwitten. We analyseren kritisch de bekende regelaars op hun structuur en de herkenning door 14-3-3 eiwitten. We stellen een aantal kleine veranderingen voor een aantal regelaars voor, die gebaseerd zijn op gepubliceerde co-kristalstructuren.

In **Hoofdstuk 3** beschrijven we de identificatie van regelaars van 14-3-3 EEIs *via* DCC. We richten ons op een EEI-complex van 14-3-3(ζ) met synaptopodin, een peptide van 21 aminozuren. Hit-verbindingen van de DCC-experimenten analyseren we op hun biochemische werking *via* oppervlakte plasmaresonantie (OPR). Deze verbindingen blijken een stabiliserende werking te hebben op het EEI-complex. Co-kristallisatiestudies worden uitgevoerd om de bindingswijze te bevestigen. En als basis voor SGMO om de affiniteit te verbeteren.

In **Hoofdstuk 4** passen we een acylhydrazon-gebaseerde DCC toe op een glucansucrase. Glucansucrasen spelen een belangrijke rol in tandplak, dat cariës veroorzaakt. Remmers van glucansucrase zouden daarom cariës kunnen voorkomen en aan tandpasta toegevoegd kunnen worden.

In **Hoofdstuk 5** beschrijven we de eerste toepassing van een nitron-gebaseerde egDCC. We hebben de condities voor de egDCC geoptimaliseerd, daarbij rekening gehouden met het eiwit alsook de toepasbaarheid van de nitronreactie. Tevens hebben we het optimale pH-venster bepaald voor deze 'nieuwe' reactie. De synthese van de hit-verbindingen maakte het evalueren van de biochemische activiteiten mogelijk *via* een fluorescentie-gebaseerde activiteits assay.

In **Hoofdstuk 6** laten we nog een nieuwe reactie in egDCC zien waarbij gebruik gemaakt wordt van thiazolidinen. Thiazolidinen worden in verscheidene natuurstoffen gevonden en zijn daarom zeer attractieve bouwstenen vanuit een biochemisch oogpunt. We beschrijven een aantal DCC-experimenten die laten zien dat deze reactie ook in medicinale-chemieprojecten toegepast kan worden. De uitbreiding van de reactiemogelijkheden van cysteïne derivaten naar aromatische aminothiol bouwstenen resulteerde in de oxidatie van de producten. De oxidatie maakt de reactie irreversibel waardoor thiazolidin DCC tot dusver gelimiteerd is tot cysteïne derivaten.

Zusammenfassung

Die Verwendung dynamisch kombinatorischer Chemie (Dynamic Combinatorial Chemistry, DCC) in medizinisch-chemischen Projekten kann eine sehr hilfreiche Strategie sein, um Anknüpfungspunkte für die Wirkstoffentdeckung zu finden. Insbesondere, wenn Targetinformationen bekannt sind, können mit DCC kombinierte struktur- oder fragmentbasierte Wirkstoffdesign-Ansätze zu wirksamen Verbindungen führen.

14-3-3 Proteine spielen als relevantes Target eine Rolle in verschiedenen Krankheiten und vielen biologischen Prozessen. Proteine dieser Familie beteiligen sich an Protein-Protein-Interaktionen (PPIs) und können diese mit zahlreichen Bindungspartnern ausbilden. Durch Bildung dieser PPIs regulieren diese Proteine die Aktivität der Bindungspartner. Die Aktivität kann sowohl hoch- als auch herabreguliert werden.

Eine andere Familie relevanter Targets sind die Glukansucrasen, welche wichtige Enzyme in der Initiierung und Entwicklung von kariogenen dentalen Biofilmen, allgemein bekannt als Plaque, sind. Durch Inhibition dieser Enzyme könnte deshalb Karies vorgebeugt werden.

In den letzten beiden Kapiteln wurde Endothiapepsin für Protein-vermittelte DCC (ptDCC) verwendet. Endothiapepsin gehört zur Familie der Aspartylproteasen. Aspartylproteasen sind an zahlreichen biologischen Prozessen und Krankheiten beteiligt, zum Beispiel an der Reifung des HIV Viruspartikels.

Im Verlauf dieser Arbeit fokussieren wir uns auf die Anwendung von DCC in verschiedenen Projekten. Die Hauptleistungen sind: 1) die Beschreibung des hausinternen DCC-Protokolls, in welchem Aspekte wie Löslichkeit von Bausteinen und Produkten, Proteinstabilität und weiteres wichtige zu beachten sind, 2) die Anwendung von Acylhydrazon-basierter DCC auf zwei attraktive Targets, eine Glukansucrase und ein PPI-Target, 3) die Identifikation kleiner Moleküle, die PPIs von 14-3-3/ Synaptopodin stabilisieren, 4) die Erweiterung des Reaktionsspielraums der ptDCC durch zwei zusätzliche Reaktionen: Nitron- und Thiazolidinbildung.

In **Kapitel 1** liefert eine klare und knappe Beschreibung eines umfassenden DCC-Protokolls. Diese Arbeit fokussierte sich auf wichtige Aspekte hinsichtlich der praktischen Handhabung in der ptDCC. Um ptDCC erfolgreich anzuwenden, sollten die besten biologischen und chemischen Bedingungen aufeinander abgestimmt werden. Darüber hinaus wurden analytische Techniken, die

verwendet wurden, um zu bestimmen, welche Verbindungen die favorisierten sind, kurz diskutiert.

In **Kapitel 2** zeigen wir einen Überblick der Modulatoren von 14-3-3 Proteinen. Kritisch analysierten wir in Hinblick die bekannten Modulatoren auf ihre Strukturen und molekulare Erkennung durch 14-3-3 Proteine. Basierend auf publizierten Cokristallstrukturen, schlugen wir kleine Modifikationen bestimmter Modulatoren vor.

In **Kapitel 3** beschreiben wir die Identifikation von Modulatoren der 14-3-3 PPIs mittels DCC. Wir adressierten einen PPI-Komplex der 14-3-3(ζ) mit Synaptopodin, ein aus 21 Aminosäuren bestehendes peptid. Hit-Verbindungen der DCC-Experimente wurden auf ihre biologischen Aktivitäten *via* Oberflächenplasmonenresonanz (SPR)-Spektroskopie untersucht. Diese Verbindungen zeigen einen stabilisierenden Effekt auf den PPI-Komplex. Cokristallisationsstudien laufen um den Bindungsmodus zu bestätigen, und SBDD zur Optimierung der Aktivität zu ermöglichen.

In **Kapitel 4** wendeten wir Acylhydrazon-basierte DCC auf eine Glukansucrase an. Glukansucrasen spielen in kariogenen dentalen Biofilmen eine wichtige Rolle, welche Karies verursachen. Glukansucrase-Inhibitoren sollten daher Karies vorbeugen und könnten als Additive in Zahnpasta zugesetzt werden. Wir beschrieben die Synthesen, DCC Experimente als auch die biochemische Charakterisierung der Hit-Verbindungen.

In **Kapitel 5** beschreiben wir die erste Anwendung von Nitron-basierter ptDCC. Wir optimierten die Bedingungen für ptDCC unter Berücksichtigung des Targetproteins Endothiapepsin sowie der Anwendbarkeit der Nitronreaktion. Neben der Anwendung auf ptDCC, bestimmten wir das optimale pH Fenster dieser „neuen“ Reaktion. Die Synthese der Hit-Verbindungen ermöglichte die Evaluation ihrer biochemischen Aktivität in einem Fluoreszens-basierten Aktivitätsassay. Vorläufige Ergebnisse hinsichtlich der Zytotoxizität des Nitron Linkers zeigte keine Zytotoxizität auf der Leberzelllinie HepG2.

In **Kapitel 6** porträtieren wir eine weitere neue Reaktion in ptDCC unter Einsatz von Thiazolidinen. Thiazolidine können in zahlreichen Naturstoffen gefunden werden und sind deshalb aus biochemischer Sicht sehr attraktive Gerüste. Wir beschrieben eine Zahl von DCC-Experimenten, die zeigten dass diese Reaktion ebenfalls in Medizinalchemieprojekten verwendet werden kann. Die Erweiterung des Reaktionsumfangs von Cysteinderivaten auf aromatischen Aminothiol-Bausteinen resultierte in der Oxidation der Produkte. Diese Oxidation macht die Reaktion irreversibel, und deshalb ist Thiazolidin-DCC vorerst auf Cysteinderivate begrenzt.

Acknowledgements

In the last few pages of this thesis, I would like to thank all the persons who have helped me over the past four years. None of this work could have been done without you. I would like to sincerely thank every single one of you for your support, trust, energy and optimism during my PhD.

Dear Anna, this amazing journey started not four years but a bit longer ago. It was during my Master's internship that you asked me what I was planning to do after I would finish. Honestly, I never thought about doing a PhD, since I wanted to do a second Master's degree in education. You offered me to start as a PhD student in your group, I happily accepted and I am forever grateful. I remember the moment well, when you told us that you had received a job offer from Saarbrücken. I was a bit shocked and overwhelmed at the beginning, as where more group members, but it didn't take me long to see the potential of continuing our research abroad. Thank you Anna, for guiding me during my PhD, for always supporting me and for being not only a great mentor but definitely also a friend.

Not only Anna was a great mentor, but also you, Adri, helped me a lot. During the time in Groningen you always showed great interest in my research and you were always available for questions and guidance. Even after moving to Saarbrücken, you kept supervising me and your feedback on my research progress was very valuable to me. Hartelijk dank dat u altijd voor me klaar stond, zeker ook toen we als werkweek in Hamburg waren en we wat problemen met het hotel hadden. Ik mag me erg gelukkig prijzen met u en Anna als mijn promotoren.

Dear Rolf, I would like to sincerely thank you for being my co-supervisor from Saarbrücken and for your support in all the events I organised at HIPS.

There are more professors whom I would like to thank. Beste Martin, dankjewel dat je je onvermoeid ingezet en nuttige feedback gegeven hebt tijdens de groepsbijeenkomsten en daarbuiten om. Lieber Alex und lieber Martin, ich möchte euch herzlich für euer Interesse an meiner Forschung und eure Tipps danken, die ich von euch bekommen habe. Lieber Prof. Hartmann, vielen Dank, dass Sie uns als Fachgruppe so herzlich in Saarbrücken aufgenommen haben.

I would like to thank all the members of my PhD reading committee, Prof. Frank Dekker, Prof. Rolf Müller, Prof. Sijbren Otto, Prof. Martin Witte, Prof. Anna Hirsch and Prof. Adri Minnaard, for their comments and feedback.

Many thanks to the collaborators with whom I had the pleasure to work with. Beste Sebastian, dankjewel voor je enorme hulp bij het 14-3-3 project. Zonder

jouw inzet was het überhaupt niet mogelijk geweest. Lieber Christian, vielen Dank dass du uns die Möglichkeit gegeben hast um DCC zu verwenden mit 14-3-3. Beste Lubbert en Evelien, ik wil jullie hartelijk danken voor jullie samenwerking bij het glucansucrase project. Also at this stage I would like to thank you, Varsha, for helping and supporting me with the glucansucrase project. main aapake dhairy aur dostee kee bahut saraahana karata hoon. After thanking Varsha, I immediately thought of you, Ravi. You are an amazing person, I enjoy spending time with you, scientifically and also outside of the labs. I cannot thank you enough for reading my manuscripts and also the chapters of this thesis, and giving me very useful feedback. ek adbhut dost hone ke lie dhanyavaad.

My dear Walid, I know it is your pleasure helping your colleagues, and I definitely enjoyed working together! I cannot thank you enough for your help with the SPR-measurements, teaching and guiding me. shukraan litaeawunikum aljamil. I would also like to thank you Samir, for always being available to help with the LC-MS and questions related to synthesis. Merci pour ton aide précieuse.

One year ago, you joined our group Henni, and I am really happy you did. You are calm, capable, extremely intelligent and someone you can trust! Kiitos paljon siitä, että autat minua hakemuskirjeteni kanssa, ja ystäväksi. My dear Jelena, also you joined our group one year ago. And it has been a really nice time since then. Hvala vam što ste mi bili prijatelj i što ste bili tako ljubazni!

My dear office mates from Saarbrücken: Andreas, Philine, Isabell, Melissa and Christina, you guys have made my time in Saarbrücken very pleasant, joyfull and valuable. Lieber Andreas, du bist jemand, auf den man sich verlassen kann, danke für deine Hilfsbereitschaft, wenn ich Fragen hatte, für die tolle Zeit im Büro und deine Freundschaft. Liebe Christina, du hast mir unglaublich gut mit der Übersetzung der Zusammenfassung meiner Doktorarbeit und allen anderen Briefen sowie E-mails geholfen. Danke für dein Interesse an dem, was ich mache und für deine Freundschaft. Meine liebe Philine, wie schön war die Zeit mit dir im Büro. Ich schätze dich sehr und bin sehr froh, dass wir zusammen etwas publizieren werden, und somit auch unsere Zusammenarbeit verewigt ist. Ich bin dir sehr dankbar, dass du mich in euer Büro eingeladen hast. Ciao Melissa, ho anche trascorso molto tempo con te in ufficio. Sei una persona meravigliosa, grazie per la tua amicizia.

Another dear Italian, whom I would like to thank, is you Massimo. Grazie mille per la tua gentilezza, ospitalità, amicizia e i tuoi deliziosi vini.

There are more persons I would like to thank from Saarbrücken. Lieber Chris, dich konnte ich immer etwas zu Synthesen fragen. Du bist ein genauso guter Mensch wie Chemiker. Danke für die coole Zeit mit dir. Lieber Dirk, wir haben

viel mit einander gekämpft, aber das hat uns beiden auch viel Spaß gemacht. Ich habe mich immer über unsere Squash-Spiele gefreut. Es tut mir leid, dass ich bei dir gequetschte Rippen verursacht habe, ich werde versuchen weniger knochig zu werden. Am Anfang warst du auch dabei, Roman. Nachdem ich nach Saarbrücken gezogen bin, bist du ziemlich schnell nach Groningen umgezogen. Ich hoffe, dass es euch gut geht und bin sicher, dass euch die Stadt gefällt!

Dear Italian gang, Eleonora, Federica, Melissa, Serena, Vittoria, Saverio and of course Claudia, you guys are also very good friends of mine. Grazie per le serate di pizza e il tempo divertente insieme.

Dear Jörg, Mostafa, Ahmed-Saad, Ahmed Merabet, Stefanie, Sandra, Eike, Joscha, Olga, Ghamdan, Varvara, Eva, Ali, Alaa, Valentin, Vlad, Dominik, Aylin, Julia, Cansu, Alex and most likely many more, thank you all for being great colleagues and making my time at HIPS pleasant!

Dear Robin, Spyros, Judy, Eleonora, Yagiz, Varsha, Ravi and Anna. I think we can look back on a successful move from Groningen to Saarbrücken. We have supported each other a lot and found our way around. I am happy that I was part of the team!

My dear Ramon, you did not move with us to Saarbrücken, but you did stay an important part of the Hirsch-group. At the beginning in Saarbrücken, we had our frequent problem sessions, which you joined *via* Skype. Dankjewel voor jouw vriendschap, hulp en dat je altijd geïnteresseerd was in mijn onderzoek en hoe het met mij ging. Het was jammer dat je niet meegegaan bent naar Saarbrücken, maar ik begrijp het heel goed dat je fijn samen met Joyce in Groningen gebleven bent. Ik wens jullie samen viel geluk en plezier, en hopelijk kun jij binnenkort ook je PhD afronden.

My dearest Simone, Jeannine, Tabea, Dennis, Teresa and Andreas, thanks to all of you for the great lunch and coffee breaks we had. Vielen Dank, dass Ihr immer für mich da wart, und meine Zeit am HIPS richtig schön gemacht habt.

Many, many thanks to the `Hausmeister‘, Christian and Frank and to the IT-guys, Mark and Michael. Ohne euch kann keiner am HIPS arbeiten, ihr seid die stillen Kräfte, auf die man sich verlassen kann. I would also like to thank the people at DDEL and MINS, for being great colleagues and friends, especially Olga, Rebekka, Justus, Patrick and Benedikt.

Also many people from Groningen were unmissable. Dankjewel Nabil, voor de potjes snooker die we gespeeld hebben en je hulp bij de laatste stapjes voor mijn thesis. Het was een genot om tijd met je doorgebracht te hebben in Groningen!

Datzelfde geldt voor Mark, met wie ik ook meerdere potjes snooker heb kunnen spelen, en tevens het kantoor gedeeld te hebben in de Linnaeusborg. Dear Kaja, thanks a ton for your company in the office and for making the illustration for my first review. Dear Leticia, Mira, Jonas, Paul, Niek, Manuel, Jeffrey, Ashmir, Ruben, Wiktor, Ilse, Mickel, Michael, Dusan, Friederike, Stella and many more, I wish to thank you for the good time we spent in and outside of the Linnaeusborg.

Dear Yagiz and Milon, you both have been an example for me during my PhD. You both were there for me, when I had questions or wanted to discuss other things. Dear Yagiz, it was really special, being at your wedding party, I wish you and Nour all the best.

A very special thanks goes to my students: Stef, Barbara, Nathalie and Claudia. It was my pleasure being part of your academic journeys, and you have all helped me to become a better mentor and scientist myself.

A big thank you goes to Theodora, Monique, Hilda and Alphons, who make sure that everyone can work properly in Groningen!

Without family and friends, I could not have achieved any of this. I could write another book, on what you all mean to me, but I will keep it concise. My dearest Corinna, liebe `Schwester`. Wir kennen uns schon ziemlich lange, und du bist eine meiner besten Freundinnen. Du warst immer ein Beispiel für mich, gute Arbeit zu leisten und Einfluss in Forschung zu haben. Auch danke an deine Eltern, die sind auch awesome. Lieve Bram en Tineke, jullie zijn twee geweldige vrienden en hebben drie super leuke dochters! Dankjulliewel, ook aan jullie familie, voor alle steun en interesse! Beste Jeffrey, samen naar Groningen gegaan om te studeren, en dan ga jij naar Duitsland. Toen je terug kwam ging ik bijna naar Duitsland. Ik hoop dat ook jij binnenkort je PhD succesvol kunt gaan afronden. Lieve Marcella, dankjewel voor je interesse in mijn onderzoek en de vriendschap die we al vele jaren hebben. Lieve Renate en Derek, jullie zijn ook twee onmisbare personen voor mij, dankjulliewel voor alles. Lieve Marianne & Annyttsje, Frits & Willeke en Jolande & Richard, Jasper & Stephany en Melissa, ook jullie wil ik hartelijk danken voor jullie steun en interesse!

Last but certainly not least comes my family. Lieve papa, en zeker ook mama, lieve Stefan & Kirsten + Julian, lieve Lysette & Maarten, ik wil jullie heel erg bedanken voor alles wat jullie voor mij betekenen. Jullie hebben mij altijd gesteund, en me geholpen waar nodig. Lieve zus, dankjewel voor de fantastische cover van mijn thesis, dat heb je echt super mooi gedaan!

Lieber Tobias, dankeschön, dass du mein Freund bist, dass du zuverlässig bist und für dein Interesse an mir als Person und Wissenschaftler. Danke, dass wir

uns jeden Tag wieder lieb haben. Ich hoffe, dass wir noch lange zusammen bleiben werden. Auch würde ich hier gerne deiner Mutter, Christine, deiner Schwester, Alexandra, ihrem Mann, Johannes, und den Kinder herzlich für euer Interesse an mir danken.

For all of you, who I might have missed to thank by name, I am grateful for your help, support and trust in me over the past years.