LUND UNIVERSITY

Fragment screening using WAC towards new SMARCA4 inhibitors

By

ARTUR SAHAKJAN

Department of Chemistry
Lund University, Faculty of Engineering,
SE-221 00, Lund, Sweden

Artur Sahakjan

Examination Project work, Organic Chemistry 30 ECTS
Master of Science in Chemical Engineering

Supervisor:	Professor **Ulf Nilsson**	Lund University Centre for Analysis and Synthesis SE-221 00 Lund, Sweden
Co-Supervisor:	Associate Professor **Richard E. Johnsson**	Red Glead Discovery SE-223 81 Lund, Sweden
Examiner:	Professor **Ulf Ellervik**	Lund University Centre for Analysis and Synthesis SE-221 00 Lund, Sweden

To my mom, Shushanik

Abstract

The drug discovery and development process has changed dramatically during the last decades and the growth of more efficient techniques as high throughput screening combined with combinatorial synthesis using amino acids and carboxylic acids on solid support makes it possible to investigate the affinity of molecules towards a biological target. Fragment based drug discovery (FBDD) has been implemented into different techniques with purpose of discovering new drugs. In this study, a high throughput technique, weak affinity chromatography (WAC) has been used to screen small aromatics (<400 Da) with purpose of finding a new starting point for a SMARCA4 inhibitor. It is known that the protein is involved in the transcription and repair of DNA, hence the discovery of a new inhibitor is of great interest in medicinal chemistry. To find new inhibitors with high quality data, the need for a more neutral reference column is crucial. It has been observed that the current reference column which consists of ethanolamine have a charge impact on the results, where positively charged molecules being repelled and negatively charged molecules more attracted to the column, consequently the molecules elutes faster and slower respectively. Thus, a reference peptide column was developed by mimicking the isoelectrical point and the amino acid sequence of the protein with the purpose of having a reference column that better mimics the protein. During this study, a total of 216 molecules where synthesized, they were distributed in 13 different libraries and each mix was examined and evaluated after the screenings on WAC. WAC has been proven to be an efficient, sensitive, and a robust method, which also provides the possibility to calculate the dissociation constant (K_D). Due to its smoothness, flexibility, and productivity, the technique has very high potential to contribute and to establish more reliable FBDD research.

Populärvetenskaplig Sammanfattning

Syftet med avhandlingen var att undersöka och utvärdera bindningen av molekyler mot ett målprotein som är involverad i tillväxten av cancerceller. Undersökning genomfördes på Red Glead Discovery med WAC (weak affinity chromotography) som är en kromatografisk separationsmetod, där molekyler med olika vikt pumpas genom olika kolonner och retentionstiden för molekylerna i kolonnen detekteras. Molekyler som elueras senare från proteinkolonen än referenskolonnen, anses inneha bindning med proteinet. Större skillnad i eluerings tid tyder på starkare bindning. För att kunna erhålla mer precis och konkret bindningsdata bör referenskolonnen vara så lik proteinkolonnen som möjligt. Den nuvarande referenskolonnen påverkar resultatet genom att kolonnen främjar repulsion av molekyler med positiv laddning och tvärtom för negativt laddade molekyler. Baserat på det uppstådda problemet designades en peptid baserat på proteinets isoelektriska punkt och dess aminosyra sekvens. Peptiden immobiliserades i en kolonn och användes som en referenskolonn i undersökningen av olika molekyler.

Det unika med metoden framhävs av förmågan att kunna detektera upp till 25 olika molekyler som finns i ett bibliotek. Antalet molekyler i ett bibliotek är endast begränsade av vikten på molekylerna och startmaterialens reaktivitet vid fastfas syntesen. Totalt syntetiserades 216 olika molekyler som var distribuerade i tretton olika bibliotek och varje bibliotek innehöll mellan sex och 25 olika molekyler.

Proteinet som har kodnamnet SMARCA4 har ofta upptäckts vara muterad hos cancerpatienter. Genom att inhibera proteinet förhindras okontrollerat tillväxt av cancerceller, eftersom proteinet är involverad i transkriptionen av DNA. Därför valdes det proteinet till denna studie.

Resultaten visade minskad påverkan av laddning med referenspeptid kolonnen, där negativt laddade molekyler hade kortare retentionstid på referenskolonnen än den nuvarande kolonnen. Detta visar att sökandet efter en ny referenskolonn material borde fortsättas. Under utvärderingen av molekyler som upptäcktes ett par stycken som visade affinitet för proteinet, även dessa borde bli fortsatt studerade.

Abbreviation

Ala	Alanine
Asp	Aspartic acid
Aze	Azetidine
DCM	Dichloromethane
DIC	N,N'-Diisopropylcarbodiimde
DMF	Dimethylformamide
Fmoc	Flourenylmethyloxycarbonyl
Gly	Glycine
HPLC	High performance liquid chromatography
Hyp	Hydroxyproline
K_D	Dissociation constant
Nap	3-(2-naphtyl)-L-alanine
Oxyma	Ethyl cyanohydroxyiminoacetate
Pal	3-(3-pyridyl)-L-alanine
PFI-3	(2E)-1-(2-hydroxyphenyl)-3.3[(1R,4R)-5-(pyridine-2-yl)-2,5-diazabicyclo[2.2.1]heptan-2-yl]prp-2-en-1-one
Phe	Phenylalanine
pK_a	Equilibrium association constant
Pip	Pipecolic
Pro	Proline
TFA	Trifluoroacetic acid
Tic	1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid
Trp	Tryptophan
Ser	Serine
SPPS	Solid Phase Peptide Synthesis
WAC	Weak Affinity Chromatography

Acknowledgement

I would like to express my very great appreciation to Richard E. Johnsson for being my supervisor and giving me valuable advises and motivating challenges during this project. Your willingness to help and support me have shown how much you care. Thank you, for always being available when I needed help.

Secondly, I would like to thank the people who have been involved in proposing and creating this project, which opened an opportunity for me to learn more from very experienced and skilled people. Thanks Johan Evenäs (CEO). I would also like to thank Ulf Nilsson for being my supervisor and Ulf Ellervik my examiner.

Thirdly I would like to thank the peptide group, especially Maria Volkova for helping me in need. Thank you, Kirill Popov for being so involved in WAC and contributing with interesting ideas. Thank you, Stella Timpka, for helping me to prepare the columns and run on WAC.

I also want to thank Bo Svensson at SARomics for providing this project with the protein.

None of this would have happen without my family, friends, and colleagues. Thank you for your support and encouragement throughout my study.

Content

Background & Introduction

Fragment Based Drug discovery

Fragment based drug discovery (FBDD) is implemented in drug discovery to find new lead fragments or compounds based on their low molecular weight and small size. The concept is based on screening libraries containing many low molecular weight fragments (MW<400 Da) against a biological target. In order to identify hits, the fragments must bind with the protein, a certain necessity as shape complementary between the protein surface and ligand have to be reached. The complexity of binding increases with increasing size of the fragments, bigger molecules tend to clash and decrease the probability of the fragment to bind with the protein. Smaller molecules with the same library size tend to occupy the chemical space more efficiently and have a greater probability to bind with the protein, therefore they tend to have lower affinity and higher hit rates compared to larger sized molecules. The fragments might have weak affinity towards the biological target, but they should not be observed as poor binders, instead seen as having the potential to become a high-quality binding ligand. The approximate low affinity value of hits is about (K_d > 0.1-10 mM)[1]. Although the fragments have a low interaction with the protein, they form efficient, high quality binding with the protein architecture and they can be further grown, merged or linked into a high-affinity substance which can be a potential drug candidate[1] due to the fragments dynamic and influence on the binding site.

High quality data is required for FBDD to be successful and that comes from primary screening with fragment screening techniques. WAC is one of the FBDD techniques and it has been studied in this project.

Screening methodologies

There are several fragment screening methods in the pharmaceutical industry, one of them is surface plasmon resonance (SPR), which is mostly used for membrane protein drug targets, like G-protein coupled receptors. When binding between protein and ligand appears there is change in the refractive index at the surface interface, which a biosensor measure[2]. The sensor calculates interaction of the molecules immobilized on the biosensor and in solution in real time and provides dissociation constant (K_D) and association constant (K_a). [2] SPR can screen from 20 000 to 90 000 fragments in a time frame of one day to two weeks, but the disadvantage is that it cannot easily determine the difference between specific and nonspecific interaction.[3] Another method is nanoscale differential scanning fluorometry (nanoDSF), which takes advantage of measuring the melting point of folded proteins. Inhibited proteins have a higher melting point, that is measured and provides an inflection point of a protein ligand complex with increased temperature. Further the melting point is used to calculate K_D.[4] There is also a newly improved method called isothermal titration calorimetry (ITC). The instrument contains a reference cell and a sample cell, they are monitored with a fixed temperature and pressure. The heat detector measures the heat change when ligands are injected into the sample cell containing the protein. The heat change measured is in direct proportion of the amount of ligands binding. Since the technique can determine the binding constant (K_a) and the enthalpy

change (ΔH) of binding this relationship $\Delta G = -RT \ln K_a = \Delta H - T\Delta S$ can be used to calculate Gibbs free energy and the affinity constant. The technique has existed for two decades, but only been used for single measurements, which could take three to four hours, after improving the technique by increasing the titration rate the instrument is able to measure 30-40 titrations per day, and requires approximately ten milligram of protein.[5,6] There are a couple more methods like X-ray, NMR, CE capillary electrophoresis and finally, WAC which has been used in this study and is explained below.

WAC- Weak Affinity Chromatography & Reference Peptide Column

Weak affinity chromatography is based on weak zonal affinity separation of analytes and is an efficient and a robust method of screening small fragments towards the biological target. The selectivity of this methods appears from letting the fragments interact with the stationary phase, which separates the fragments based on their interaction with the column. Higher affinity correlates with longer retention time, when fragments interact strongly with the protein it remains longer in the protein column and elutes later.

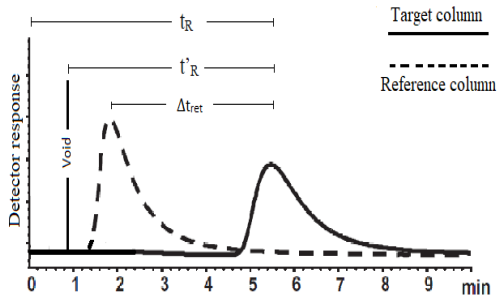


Figure 1. Chromatogram of two columns where the scattered curve is the reference column and solid curve is target column. The time is measured on the x-axis and the intensity (detector response) reflects on y-axis. Δt_{ret} is the net retention time and t'_R is the retention time without the dead volume.

The fragment screening is usually done on a standard analytical HPLC machine with mass spectrometry (MS) and UV detection. In order to see an interaction between the fragments and the protein on WAC-MS, there must be a reference column and a protein column, where the retention time difference explains how strongly they bind to the protein.

The reference column is usually deactivated silica (ethanolamine) but in this work we have also developed a new type of reference column based on mimicking the amino acid sequence of the protein with an eighteen amino acid long peptide immobilized on the silica column, in the same way as the protein was immobilized, the purpose of this peptide is to ensure the interaction between the protein and the fragments is specific interactions. The fragments that have been observed as hits has different retention time dependent on which column they have been screened on and K_D is calculated by the following equations,

$$K_D = \frac{B_{Tot}}{t\prime_R * F_R} \qquad \textit{equation 1}$$

B_{Tot} is the number of binding sits (nmol) on the column where the protein (in this work SMARCA4) is immobilized, t'_R is the retention time (min) difference between the protein and reference column and F_R is the flow rate (µL/min).[7] t'_R was calculated with *equation 2*, where

$t'_{protein}$ is the retention time for a selected fragment on the protein column and $t'_{Reference}$ is the retention time for the same fragment on the reference column.

$$t'_R = t'_{protein} - t'_{Reference} \qquad \textit{equation 2}$$

The method in this project has a limitation which can be explained by viewing the physiochemical part of the protein. Proteins often consist of long chains of amino acids and they are naturally charged. The hypothesis of a reference peptide column mimicking the pI (isoelectric point) value of the protein, is to have the same charge on the reference column and hence measure the retention time more fairly for charged fragments. By binding a peptide instead of ethanolamine to the reference column should make the data more accurate and reflects the true binding of the fragments.

Biological Target (SMARCA4)

The cause of cancer can vary in different ways dependent on how deep the biology is studied, but observing cancer from a top level, it either reflects on abnormal cell division (mitosis) or the malignant cells spread to other parts of the body. The target that have been studied in this project, which is believed to have a major impact on cancer, is a protein encoded by the SMARCA4 gene.

In order to create a protein, a gene needs to provide instructions, in this case the protein is called the same as the gene, SMARCA4, which gains the instructions from SMARCA4 gene. The protein forms a subunit of protein complexes called SWI/SNF. The complexes are involved in repairing damaged DNA, replicating DNA, and controlling the growth, division, and maturation of cells through their ability to regulate gene activity. The SWI/SNF complexes also performs a process called chromatin remodeling, and it controls the packaging of DNA into chromosomes. When DNA is loosely packed the gene expression is higher, than when DNA is more tightly packed.

Since the protein activates the transcription of the DNA, a mutation of the SMARCA4 gene might cause cancer cells growing uncontrolled.[8]

Cancer cells growing from a mutation of SMARCA4 appears in lungs, in the adrenal gland and it have also been discovered in the primary phase of brain cancer in children. SMARCA4 is the most frequently mutated chromatin remodeling ATPase in cancer.[9] The enzyme controls the ATP/ADP ratio which helps the cell metabolism to determine if it is oxidative or glycolytic.[10]

Previous researches have shown that by inhibiting the protein it is possible to prevent reduction of ATPase activity, which could lead to chromatin bridges and failure of topoisomerase IIa to bind DNA.[9]

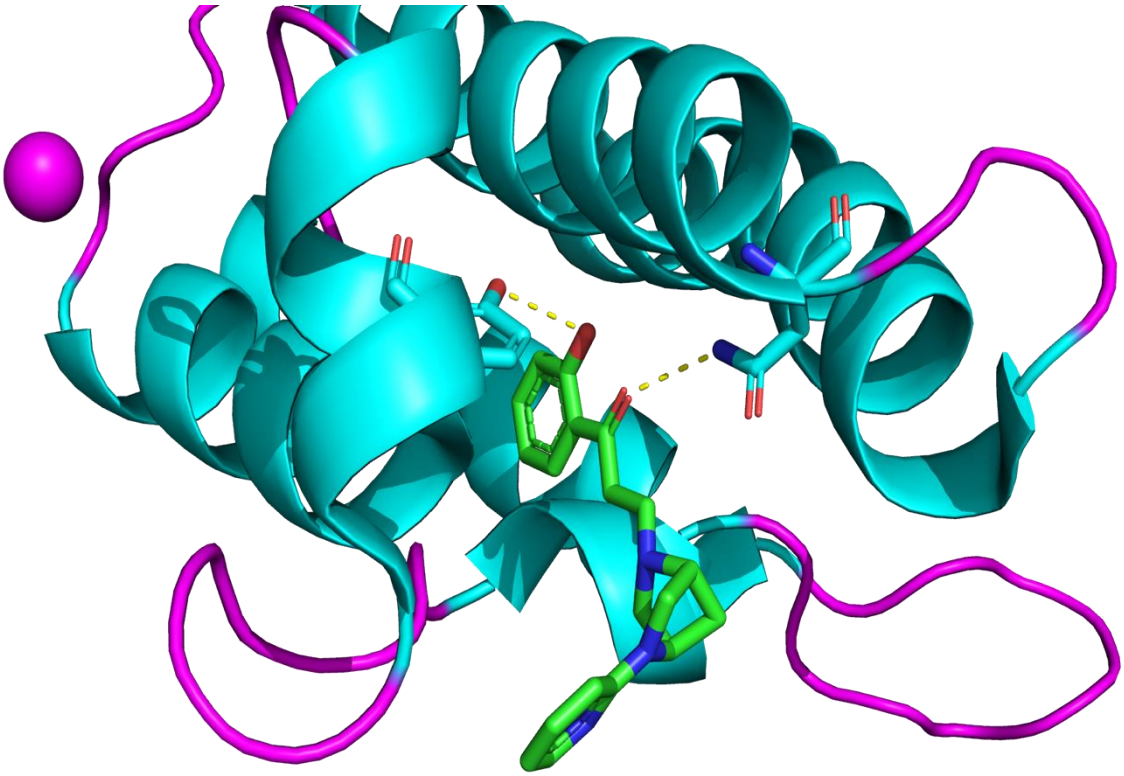


Figure 2. SMARCA4 protein in complex with PFI-3. The interactions of PFI-3 with TYR-497 and ASN-1540 appears as dotted lines, PDB code: 5DKD.

Designing fragments

The compounds have been designed based on a first screen using one of Red Glead Discovery's privilege libraries containing approximately 200 fragments and the simulation of SMARCA4 in FTMap[11], a fragment simulation program from Boston University. The previously screened fragments and the simulation program contributes with a better understanding about the binding pocket of the protein. The main approach of using previously screened fragment structure information was to improve the new fragments by studying their architecture, interaction, and variation of different functional

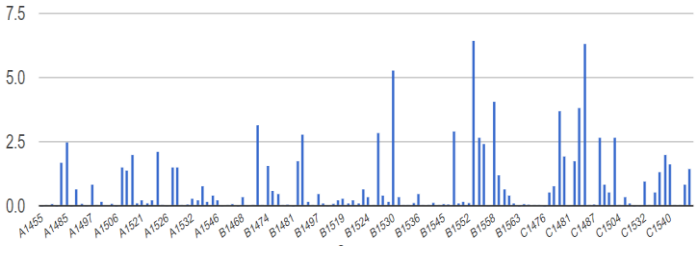


Figure 3. Non-bonded interaction graph between the functional groups and the protein. The intensity reflects on the y-axis and the x-axis shows the interacted amino acid.

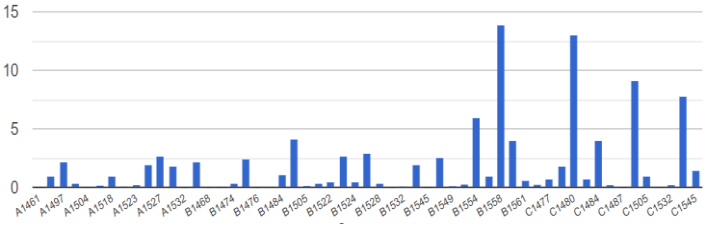


Figure 4. H-bonded interaction graph between the functional groups and the protein. The intensity reflects on the y-axis and the x-axis shows the interacted amino acid.

groups in different positions. The program simulates by docking 15 different type of probes/ functional groups into the protein pocket and gives a prediction of the probes binding orientation with the protein and a contact graph which describes the contact rate of the structure residues as a percentage of the total contacts with the protein, and provides two graphs. One graph *figure 3* shows non-bonded residue interaction and the second *figure 4* shows hydrogen bond interactions. The files where further transferred to PyMOL which made it possible to observe the bindings between residues and amino acids of the protein. *Figure 5* shows the groups that interacted most with glutamine GLN B1554 and these groups also showed the highest interaction rate overall.

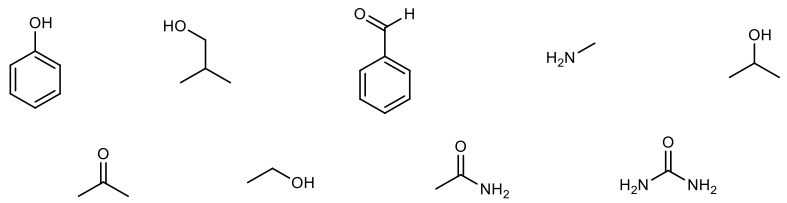


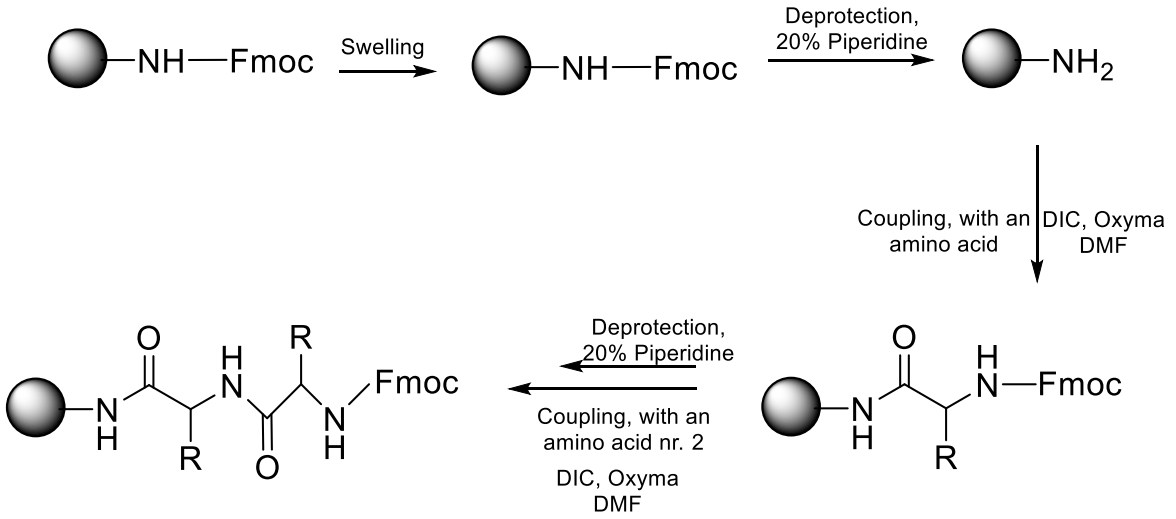
Figure 5. The functional groups that had highest overall rate of binding with GLN B1554.

Amino Acid-Carboxylic acid fragments

Solid phase peptide synthesis was developed by Robert Bruce Merrifield in the 1960's, the technique has been a huge contribution to the scientific community, which he also got rewarded for with the Nobel price. The method is based on sequential coupling of amino acids on a solid support (resin).[12]

In a previous master project by Tomas Laszlo Szakacs, tripeptides were synthesized and screened on neutrophil elastase in a similar fashion.[13] Tripeptides are quite polar and in the large end of fragments and that triggered part of the motivated to create a new project, where an amino acid coupled to an carboxylic acid and subsequently screened on WAC.

Resins are modified with a linker on the surface of the bead, the number of linkers decide the loading (mmol/g) of the resin- To start the synthesis the resin is swelled, deprotected and exposed to an amino acid that binds to the resin. The key feature in this technique is the controlled coupling of amino acids which consists of a carboxylic acid and an amine group. Proteins are naturally synthesized from the N-terminal, but in order to decrease racemization, synthetic peptides are synthesized from the C-terminal, utilizing amino acids protected on the nitrogen (Fmoc).



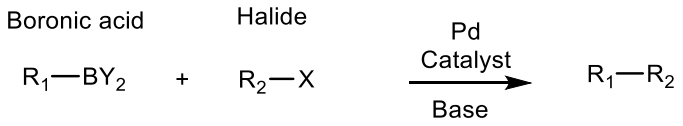
Scheme 1. Reaction procedure of solid phase peptide synthesis. The resin is illustrated by the grey sphere attached to the nitrogen.

The stepwise synthesis was done by swelling the resin with DMF, deprotecting by removing Fmoc with 20% piperidine, wash with DMF, and couple the desired amino acid with the coupling reagents DIC/ Oxyma. next step is to deprotect the previously added amino acid and couple the next one on the C-terminal, and the cycle is repeated until the desired peptide chain is synthesized. After the last deprotection of the amino acid the peptide is cleaved from the resin with TFA. This method has been used in this project to produce libraries containing many diamino acid mimicking molecules.

Suzuki coupled fragments

Since small aromatics are typical fragments in drugs, it is an advantage to create fragments containing polycyclic aromatics. In order to access them, a synthetic method called solid phase Suzuki coupling has been used in this project.[14] The coupling reaction is named after Akira Suzuki, who published his work 1979 for the first time and shared the Nobel price 2010 for the discovery and development of the palladium catalyzed reaction. Some advantages of the Suzuki coupling are the accessibility of boronic acids and the more environmentally friendly reaction conditions, compared to organozinc and organotin compounds.

We used the Suzuki coupling was to create di-aromatic fragments. The following scheme explains generally how they are formed.



Scheme 2. General reaction scheme where the R_x is the different groups attached either to a bohronic acid (BY_2) or a halide (x)

Aim

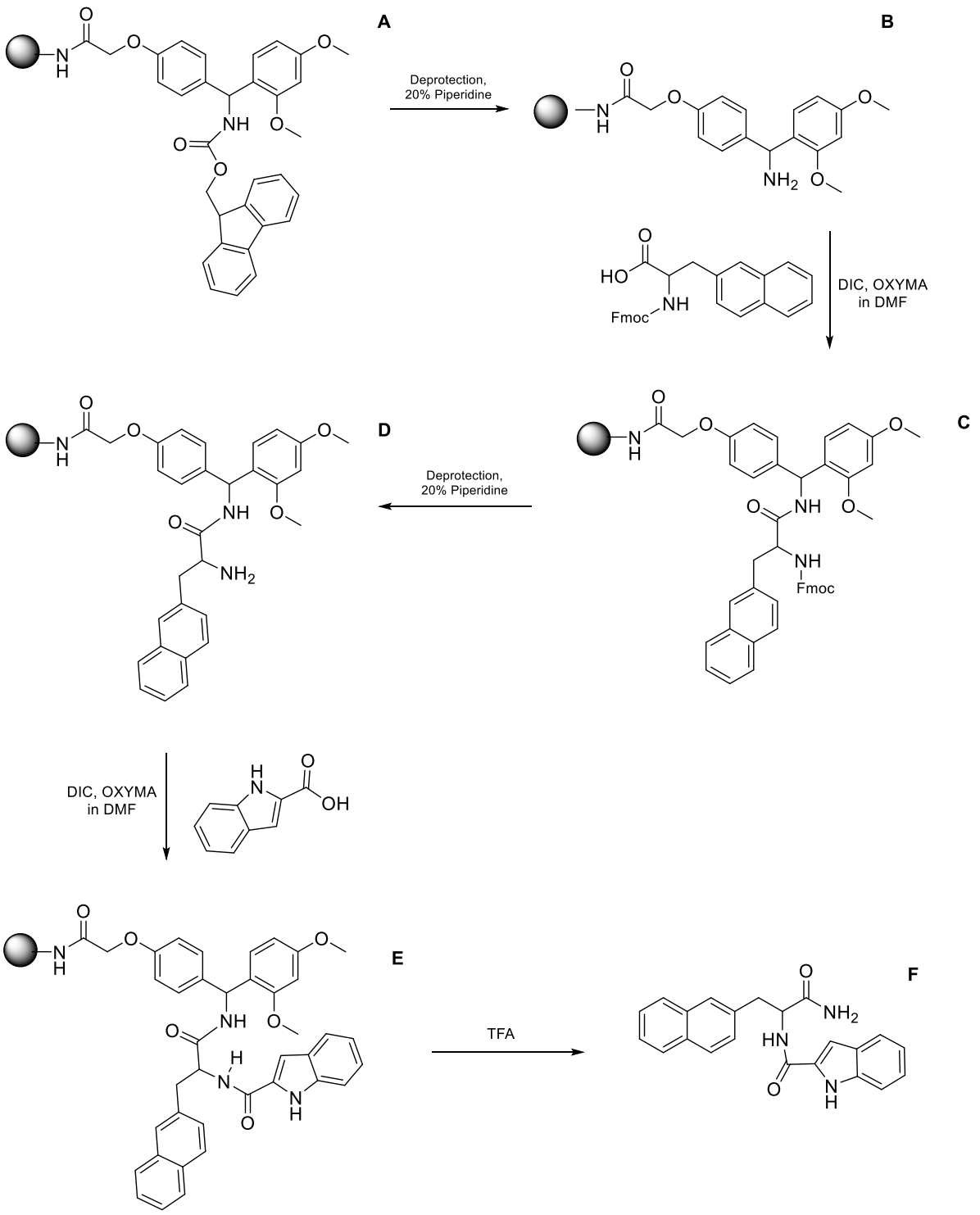
1. Hypothesis: A reference peptide column will give more true hits since it mimics the protein column better than the ethanolamine column does.
    - Designing and synthesizing a peptide and use as an alternative reference column in WAC, compare its performance with the ethanolamine and protein column.
2. Hypothesis: Synthesis of libraries is a great advantage in WAC and can be used to synthesize and screen many different compounds in a short time period.
    - Designing and synthesizing libraries of amino acid/carboxylic acid on solid support and screen them using WAC.
    - Designing and synthesizing libraries of Suzuki coupled di-aromatic compounds on solid support and screen them using WAC.
3. Hypothesis: WAC screening of libraries can give good starting points for a drug discovery program.
    - The synthesized libraries will be screened against a biologically relevant target, SMARCA4.

Results and Discussion

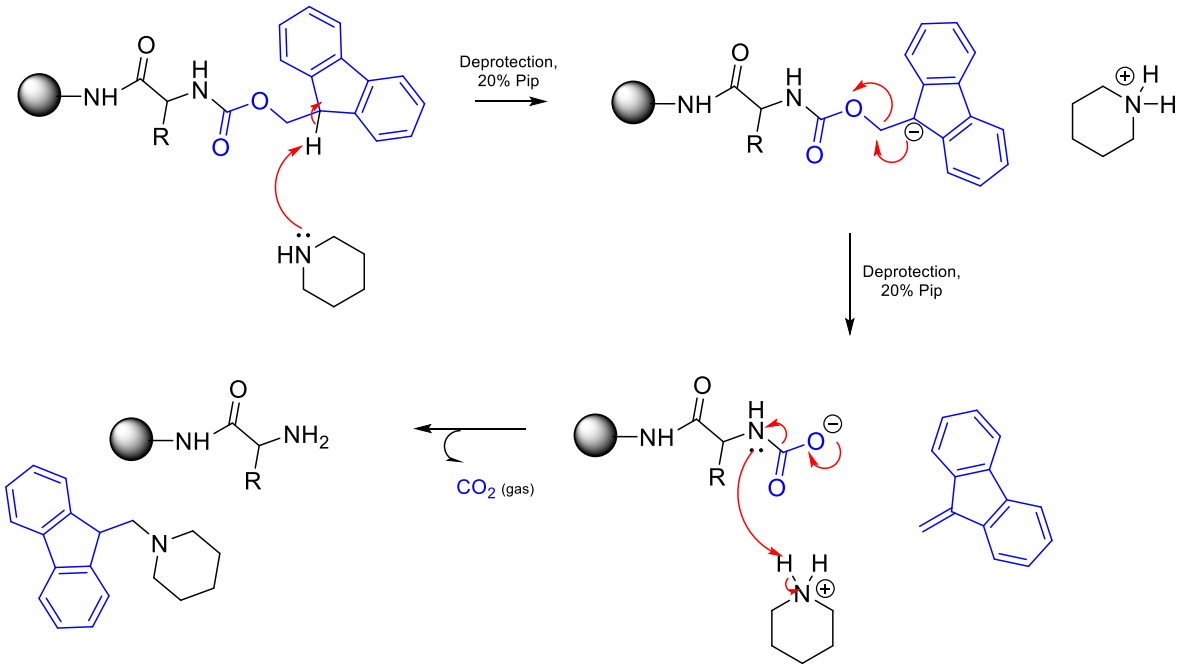
Preparation of libraries

The general synthetic strategy applied for the solid phase amino acids and carboxylic acid synthesis of all fragments are described in *scheme 1*. In *scheme 3*, the synthesis of N-(1-amino-3-(naphthalen-2-yl)-1-oxopropan-2-yl)-1H-indole-2-carboxamide is illustrated. All the syntheses follow the same strategy as in *scheme 3* with different amino acids and carboxylic acids.

Compound **B** in *scheme 3* was formed by swelling the resin in DMF and deprotecting it from flourenylmethyloxycarbonyl (Fmoc) with 20% piperidine in DMF the deprotection mechanism of Fmoc is illustrated in *scheme 4*. Compound **C** was formed when 3-(2-naphtyl)-L-alanine (Nap) was coupled with N,N'-diisopropylcarbodiimide (DIC) and ethyl cyanohydroxyiminoacetate (Oxyma). The amino acid reacts first with DIC to create O-acylisourea, and the intermediate is further attacked by Oxyma to activate the ester and DIC leaves as urea. Oxyma was preferred as a coupling reagent due to its property of being a better racemization suppressant and less explosive than other coupling reagents[15]. Compound **C** was deprotected of Fmoc with 20 % piperidine in DMF, forming compound **D**. Compound **E** was formed by coupling indole-2-carboxylic acid with DIC and Oxyma to compound **D.** The final step was achieved by cleaving compound **E** with trifluoracetic acid (TFA) from the resin beads.



Scheme 3. Reaction steps of coupling Nap with indole-2-carboxylic acid on TentaGel S Ram (resin), by using 20 % Piperidine for deprotection of Fmoc, DIC and Oxyma as coupling reagents, and TFA for cleavage from resin.



Scheme 4. Deprotection mechanism of Fmoc with 20 % Piperidine in DMF.

The carboxylic acids had different shapes and where also substituted differently which resulted in varied reactivity of the carbonyl carbon, the majority of the carboxylic acids took longer time to couple compared to amino acids. The longer coupling times was probably due to the substitution pattern and electronegativity of the carboxylic acids and thus, the reaction time was prolonged to make sure that all of the carboxylic acids could couple.

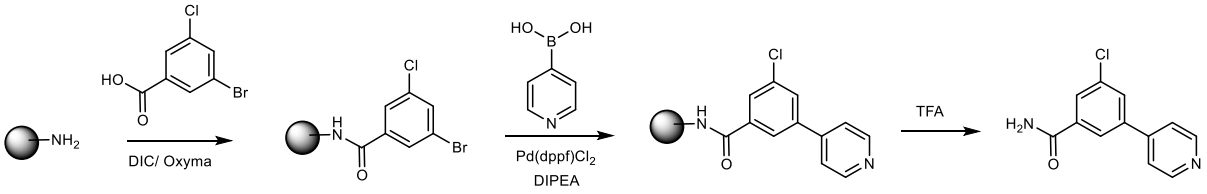
The synthesis of libraries containing 6-25 fragments, was carried out by following the general reaction procedure in *scheme 1* and modified the number of amino acids and carboxylic acid to 3-5 and 4-5 respectively in each library, to obtain most of the possible couplings between the amino acids and carboxylic acids.

Suzuki coupling

As an alternative to the amino acid/carboxylic acid libraries, we investigated Suzuki couplings on solid support. Considering the advantages of the Suzuki coupling and the usefulness of the di-aromatic fragments in drug discovery, and how well recognized the reaction is, it was no doubt of trying it. The idea was to attach several organohalides to the resin with DIC/ Oxyma, and couple it to several boronic acids in presence of a palladium catalyst and a base. Subsequently, the resin would be washed with DMF and the product cleaved and used without any further purification.

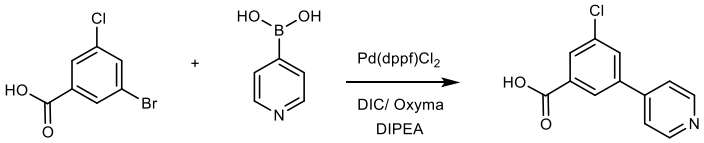
Previous studies have shown that the temperature, and the amount of time applied to the reaction, was crucial factors in a microwave assisted reactions. By using the required parameters, high yield reaction could be obtained. The Suzuki coupling in this study was both on solid phase with the peptide synthesizer and in solution phase with a microwave. The

syntheses were carried by following a published method.[14] The reaction was done by attaching 3-bromo-5-chlorobenzoic acid on the resin with DIC and Oxyma followed by coupling pyridine-4-ylboronic acid in presence of a palladium catalyst and DIPEA. Unfortunately, the product was not formed (confirmed by LCMS), and only the starting material could be detected. Different reaction conditions were tested , including changing the resin but without successful results.



Scheme 5. Solid phase Suzuki coupling reaction steps.

Since the solid phase synthesis did not work, solution phase Suzuki coupling were performed, by mixing an organohalide (3-bromo-5-chlorobenzoic acid), boronic acids (pyridine-4-ylboronic acid), $Pd(dppf)Cl_2$ in dioxane/water. The reaction was heated for 30 min at 70 °C in a microwave. The reaction worked, but besides the product other impurities were detected on LCMS. The crude was further filtrated through celite and analyzed, but without any improvement. Even though WAC could handle impure compounds and a completely pure product is not necessary, the major concern here was the residue palladium we could see, which would not be good for neither the column nor the MS. Hence, the challenging part was to remove palladium with a simple purification without using flash column chromatography. Considering that, the reaction was carried with a different catalyst, $Pd(PPh_3)_4$ and different ways to remove Pd residues was tested, such as a washing step with a dilute thiol solution and a thiol doped silica filtration. However, none of these improved the result.



Scheme 6. Solution phase Suzuki coupling reaction steps.

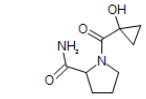
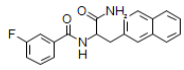
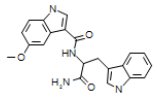
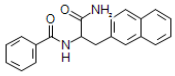
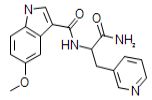
In the interest of time we decided to put the Suzuki couplings aside and focus on the amino acid/carboxylic acid libraries to generate more data from the screening.

SAR Study

After the first synthesis round we had seven libraries of in total 132 molecules that we screened against SMARCA4. All data in the screening is based on the ethanolamine reference column, for the results of the peptide reference column see below.

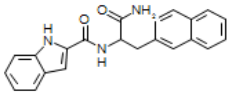
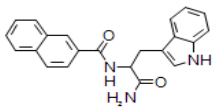
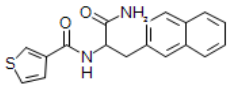
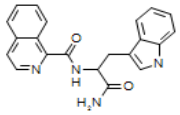
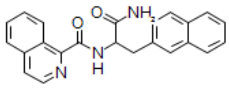
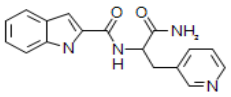
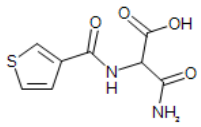
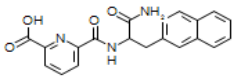
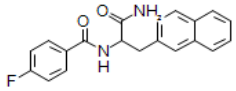
After the first run we identified five compounds as hits, i.e. a ΔRT >0.5 min. These compounds showed some structural similarities where the bicyclic aromatic compounds were more favored. Nap (Table 1, Compound **7** and **11**) and Trp (Table 1, compound **9**) showed promising enough results to further improve them, by coupling them with another similar carboxylic acids and see if we could get SAR data.

Table 1. First run screened hits, based on the retention time difference between the protein and ethanolamine column.

Structures	Compound number	RUN 1	LCMS [H]+	ΔRT Protein - Ethanolamine (min)	Charge
	3	1-Hydroxy-1-cyclopropanecarboxylic acid-Pro	197,13	1,06	zero
	7	3-fluorobenzoic acid-Nap	337,14	0,61	zero
	9	5-methoxyindole-2-carboxylic acid-Trp	377,16	0,54	neg
	11	Benzoic acid-Nap	319,15	0,53	zero
	12	5-methoxyindole-2-carboxylic acid-Pal	339,15	0,51	pos

Six more libraries of in total another 84 compounds were designed based on the results from the first screen where we tried to pick up on the hits we saw. In the second screening seven compounds were identified as hits.

Table 2. Second run screened hits, based on the retention time difference between the protein and ethanolamine column.

Structures	Compound number	RUN 2	LCMS [H]+	ΔRT Protein - Ethanolamine (min)	Charge
	1	Indole-2-carboxylic acid-Nap	358,16	2,14	zero
	2	naphthalene-2-carboxylic acid-Trp	358,15	2,14	zero
	4	3-Thiophenecarboxylic acid-Nap	325,10	0,84	zero
	5	Isoquinoline-1-carboxylic acid-Trp	359,15	0,66	zero
	6	Isoquinoline-1-carboxylic acid-Nap	370,16	0,64	zero
	8	Indole-2-carboxylic acid-Pal	309,14	0,55	pos
	10	3-Thiophenecarboxylic acid-Ser	215,05	0,54	neg
	13	6-carbamoylpyridine-2-carboxylic acid-Nap	363,14	0,45	neg
	14	4-Fluorobenzoic acid-Nap	337,14	0,40	zero

There are several comparisons between the fragments that could be observed. Starting with compound **7** in *table 1* and compound **14** in *table 2* which shows a slightly different ΔRT, 0,6116 min and 0,4 min, respectively, indicating that meta substituted benzoic acid coupled to Nap is more favorable.

Another comparison showed higher improvement of compounds containing Nap and Trp, based on compound **11** and **2** in *table 1*. Also compounds where Nap was coupled with bicyclic carboxylic acids in varying sizes showed higher affinity towards the protein, comparing compounds **1** and **8** in *table 2*.

Another comparison was between bicyclic carboxylic acid compounds containing one 5-membered ring or none, the results showed compounds containing one 5-membered ring had higher affinity for SMARCA4, comparing compounds **6** and **1**. Another comparison between mono and bicyclic compounds containing nitrogen, showed that bicyclic with nitrogen in a certain position had higher affinity for SMARCA4.

Estimation of dissociation constant for compound **1** is done by using *equation 1 & 2*. K_D=3,11mM.

The efficiency of new reference column

A reference peptide containing 18 amino acids was designed based on the amino acid sequence of SMARCA4 and its percentage composition. The reference peptide mimicked the isoelectrical point (pI) of the protein, which is the pH value when the protein does not carry any charge.[16] At a pH below the pI value generates positive charge on the protein and pH higher than the pI generates negatively charged protein. The current reference column which is covered with ethanolamine is positively charged, due to the secondary amine present, and it affects the results by promoting repulsion of molecules with the same charge and vice versa for negatively charged molecules.

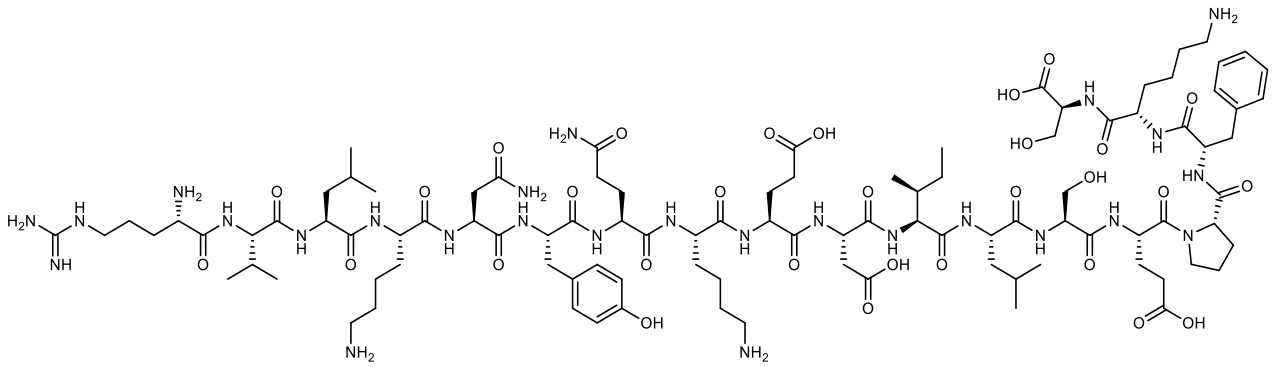
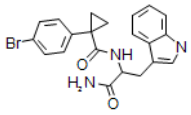
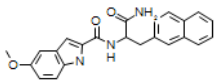
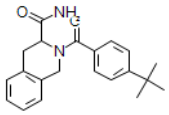
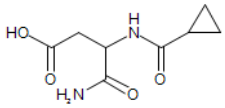
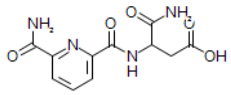
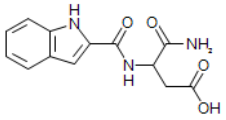
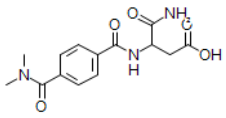
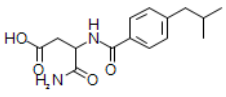
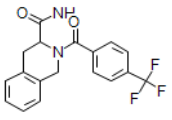
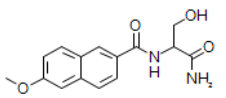
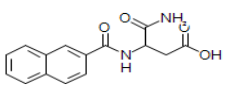


Figure 6. The reference peptide, containing 18 amino acids and their tree letter code.

The reference peptide had a pI value of 9.56 and the protein 9.12. To create similar conditions for the fragments in the column as human like as possible, a pH 6.8 is selected on the mobile phase, which resulted in approximately the same net charge being created between solution and protein in the column. Ammonium acetate was used as a buffer, due to its property of being protein and electroscopic friendly, and due to its property of being volatile it allows the mass spectrometer to operate in a high throughput manner.

A significant amount of data was processed to find trends between the different columns. As mentioned before the negatively charged compounds were often retained longer on the ethanolamine column, so also in this work. However, in the reference peptide column the retention time was shorter (see *table 3)*, which indicates that the negative charge is not affecting the reference peptide column as much as the ethanolamine column.

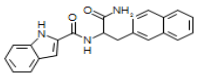
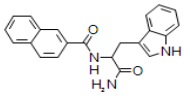
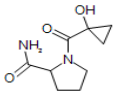
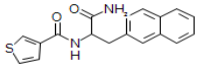
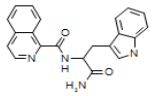
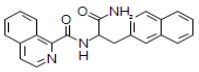
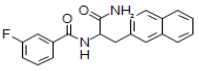
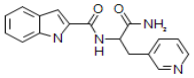
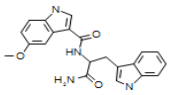
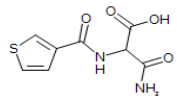
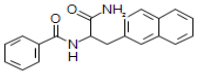
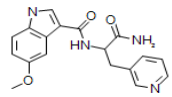
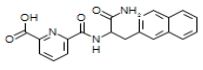
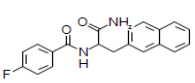
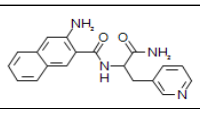
Table 3. *List of fragments that had higher retention time on the ethanolamine reference* column than the protein column.

Structures	Name	LCMS [H]+	ΔRT Protein - Ethanolamine (min)	ΔRT Protein - Custom Ref (min)	ΔRT Custom Ref - Ethanolamine (min)	Charge
	16	426,08	-0,07	-2,52	2,45	zero
	17	388,17	-0,08	-0,02	-0,06	zero
	18	337,33	-0,09	-1,54	1,45	zero
	19	201,09	-0,10	0,02	-0,11	neg
	20	281,09	-0,10	0,00	-0,10	neg
	21	276,10	-0,11	0,13	-0,24	neg
	22	308,13	-0,12	0,01	-0,13	neg
	23	293,15	-0,22	0,00	-0,22	neg
	24	349,12	-0,26	-0,11	-0,15	zero
	25	288,12	-0,27	0,04	-0,32	neg
	26	287,11	-0,38	0,07	-0,45	neg

By looking at the correlation between charge and retention time difference of the positive ΔRT peptide reference – ethanolamine column (see *table 4)*, shows that that positively charged or neutral compounds have less retention time on the ethanolamine column, as expected and opposite of negatively charged fragments. The hypothesis of the reference peptide column is almost fulfilled except that some compounds binds better on the reference peptide column than on the protein column. This could be due to other non-specific binding that could occur between the reference peptide and the screened molecules.

Since the reference peptide contain tree lysins, which were placed one at each end and one in the middle. This could affect the shape of the peptide in the stationary phase. Since only lysins were binding with the silica, the peptide could result either in a ''V'' shape or just a long chain. The shape variation might affect the binding.

Table 4. *List of fragments that are considered as ''hits'', which had a higher retention* time on the protein column than ethanolamine reference column.

Structures	Name	LCMS [H]+	ΔRT Protein - Ethanolamine (min)	ΔRT Protein - Custom Ref (min)	ΔRT Custom Ref - Ethanolamine (min)	Charge
	1	358,16	2,14	-2,87	5,00	zero
	2	358,15	2,10	-2,80	5,00	zero
	3	197,13	1,06	0,13	1,60	zero
	4	325,10	0,84	-0,41	1,25	zero
	5	359,15	0,66	-0,64	1,30	zero
	6	370,16	0,64	-3,00	3,64	zero
	7	337,14	0,61	-0,98	1,60	zero
	8	309,14	0,55	-0,03	0,57	pos
	9	377,16	0,54	-0,85	1,39	zero
	10	215,05	0,54	0,49	0,04	neg
	11	319,15	0,53	-0,84	1,37	zero
	12	339,15	0,51	-0,54	1,05	pos
	13	363,14	0,45	-0,51	0,97	neg
	14	337,14	0,44	-0,82	1,25	zero
	15	335,15	0,42	-0,35	0,77	pos

Conclusion and Future Perspectives

Negatively and positively charged compounds showed lower and higher retention time, respectively on the reference peptide column compared to the ethanolamine column. The reference peptide column decreased the charge impact on hit finding and made comparing compounds with opposite charge more feasible. This demonstrates that the search for a better reference column should continue and other reference peptides should be synthesized and tested.

216 different compounds were created with solid phase synthesis, containing an amino acid and a carboxylic acid. Even if some optimization of the reaction conditions was performed not all possible combinations were observed in the product mixture. Hence, more optimization should be done to ensure that all possible combinations are formed in the synthesis. Three thing that could be further examined is the coupling time, coupling temperature, and coupling reagents of the carboxylic acid coupling. Also, it might be interesting to sort the carboxylic acids into reactivity groups so that fast reacting acids are not mixed with much slower ones. In this way the competition between the acids would be less and there would be a more equal chance of all acids to react.

We did not detect any compounds with solid phase Suzuki coupling, but two compounds where formed from Suzuki coupling in solution phase. Unfortunately, the palladium content in the mix was too high, thus it could not be screened on WAC. This demonstrates that the reaction could be further experimented and improved, by, for example, with a solid phase bound palladium catalyst.

The binding data was evaluated by screening all molecules on WAC against SMARCA4 and we identified 12 hits in total, where 3-(2-naphtyl)-L-alanine (Nap) and indole-2-carboxylic acid was involved in most of them. The best had a $\Delta Rt=2.13$ min, this compound should be taken to the next step analysis, X-ray crystallography to verity it is hits in the active pocket.

There is a bright future for di-peptide mimicking on solid phase synthesis, it is an easy and fast method to achieve a large number of compounds in one library.

Experimental section

Standard Procedure for Amino and Carboxylic acid coupling

The reactions were carried in a peptide synthesizer (Biotage Alstra Initiator). 0.5 g TentaGel S Ram resin (0,24mmol/g) was weighed into a Biotage 30 mL reaction vial and employed to the synthesizer. The synthesizer was programmed to swell the resin in 5 mL DMF in 20 min at 50°C, next step was to filter of the solvent and perform a deprotection of fmoc in 5 min with 5 mL 20 % piperidine and 10 min with 5 mL with 20 % piperidine at r.t. The deprotection was followed by a washing step of piperidine with 4x5 mL DMF where every wash was carried by rotating the reaction vial containing DMF for 1 min and emptying it.

The coupling of amino acids was performed by adding 4 equivalents of 0.4 M amino acids(dissolved in DMF) , 0.5 M of DIC (dissolved in DMF) and 0.5 M of Oxyma (dissolved in DMF) to the resin, and rotate the vial for 60 min at r.t.

The amino acids where fmoc deprotected in 5 min with 5 mL 20 % piperidine and 10 min with 5 mL with 20 % piperidine at r.t. The deprotection was followed by a washing step of piperidine with 4x5 mL DMF where every wash was carried by rotating the reaction vial containing DMF for 1 min and emptying it. The coupling of carboxylic acids was performed by adding 4 equivalents of 0.4 M carboxylic acids (dissolved in DMF), 0.5 M of DIC (dissolved in DMF) and 0.5 M of Oxyma (dissolved in DMF) to the resin, and rotate the vial for 60 min at r.t. The final coupling was followed by a precleavage wash with iPrOH before drying in desiccator.

Cleavage of Fragments from Resin

The fragments were cleaved from the resin using a mixture of 10 mL 95% TFA, 3% H_2O and 2% TIPS for 40 min. The resin was filtered off, washed with TFA and DCM and the solution were concentrated in vacuum.

Running on WAC

The sample was solubilized in DMSO to 10 mM and further diluted in WAC buffer (ammonium acetate) to a final concentration of 20 µM.

The screening was performed on a LCMS, where the columns where changed after the collection of libraries was screened, every sample was analyzed two times on the protein (SMARCA4) column, one time on the peptide reference column and one time on the ethanolamine column. The mobile phase consisted of 100 % ammonium acetate buffer with pH 6,8 with a concentration of 20 mM, and the injection of each sample had a volume of 5 µL.

The instrument was set to scan all masses between 150-600 Da and at UV 280 nm. The detection of retention time was based on peaks from UV traces and extracted ion chromatography (EIC).

Library 1

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 1 are presented in ***Table 5***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 58.8 mg.

Table 5. Amino acids and carboxylic acids used in library 1.

Amino acids	**Carboxylic acids**
Aze	2,4-Dimethoxybenzoic Acid
Pro	1-Naphthoic acid
Pip	2-Bromo-4-chlorobenzoic Acid
	3,4,5-Trimethoxybenzoic Acid
	3-Sulfamoylbenzoic Acid

Library 2

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 2 are presented in ***Table 6***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 53.5 mg.

Table 6. Amino acids and carboxylic acids used in library 2.

Amino acids	**Carboxylic acids**
Phe	Benzoic acid
Pal	2-bromo-3,5-difluorobenzoic acid
Trp	3-fluorobenzoic acid
Hyp	5-methoxyindole-2-carboxylic acid
Nap	3,4,5-trimethoxybenzoic acid

Library 3

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 3 are presented in ***Table 7***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 25.9 mg.

Table 7. Amino acids and carboxylic acids used in library 3.

Amino acids	**Carboxylic acids**
Aze	Benzoic acid
Pro	4-Fluorobenzoic acid
Pip	4-Chlorobenzoic acid
Tic	4-tert-butylbenzoic acid
	4-(trifluoromethyl)benzoic acid

Library 4

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 4 are presented in ***Table 8***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 16.4 mg.

Table 8. Amino acids and carboxylic acids used in library 4.

Amino acids	**Carboxylic acids**
Gly	Indole-3-carboxylic acid
Ala	1-fluorocyclopropane-1-carboxylic acid
Trp	Cyclohexane-carboxylic acid
Nap	6-Cyano-1H-indole-3-carboxylic acid
	1-(4-Bromophenyl)cyclopropanecarboxylic acid

Library 5

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 5 are presented in ***Table 9***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 18.7 mg.

Table .9 Amino acids and carboxylic acids used in library 5.

Amino acids	**Carboxylic acids**
Aze	Indole-3-carboxylic acid
Pro	1-fluorocyclopropane-1-carboxylic acid
PIp	Cyclohexane-carboxylic acid
Pal	6-Cyano-1H-indole-3-carboxylic acid
	1-(4-Bromophenyl)cyclopropanecarboxylic acid

Library 6

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 6 are presented in ***Table 10***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 31.7 mg.

Table 10. Amino acids and carboxylic acids used in library 6.

Amino acids	**Carboxylic acids**
Gly	1-Hydroxy-1-cyclopropane-carboxylic acid
Ala	2-(4-PIperidinyl)propanoic acid hydrochloride
Trp	1-Phenyl-1-cyclopropane-carboxylic acid
Nap	3-aminonaphthalene-2-carboxylic acid
	1-Hydroxy-1-cyclopropane-carboxylic acid

Library 7

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 7 are presented in ***Table 11***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 39.9 mg.

Table 11. Amino acids and carboxylic acids used in library 7.

Amino acids	**Carboxylic acids**
Aze	1-Hydroxy-1-cyclopropane-carboxylic acid
Pro	2-(4-PIperidinyl)propanoic acid hydrochloride
PIp	1-Phenyl-1-cyclopropane-carboxylic acid
Pal	3-aminonaphthalene-2-carboxylic acid
	1-Hydroxy-1-cyclopropane-carboxylic acid

Library 8

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 8 are presented in ***Table 10***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 33.2 mg.

Table 12. Amino acids and carboxylic acids used in library 8.

Amino acids	**Carboxylic acids**
Nap	3,4-Difluorobenzoic acid
Pal	4-Fluorobenzoic acid
Pro	4-Chlorobenzoic acid
Trp	Isoquinoline-1-carboxylic acid
	6-methoxy-2-naphthoic acid

Library 9

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 9 are presented in ***Table 13***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 39.7 mg.

Table 13. Amino acids and carboxylic acids used in library 9.

Amino acids	**Carboxylic acids**
Nap	naphthalene-2-carboxylic acid
Pal	Indole-2-carboxylic acid
Pro	4-(propan-2-yl)benzoic acid
Trp	6-carbamoylpyridine-2-carboxylic acid
	3-Thiophenecarboxylic acid

Library 10

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 10 are presented in ***Table 14***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 90.4 mg.

Table 14. Amino acids and carboxylic acids used in library 10.

Amino acids	**Carboxylic acids**
Asp	3,4-Difluorobenzoic acid
Ser	4-Fluorobenzoic acid
Ala	4-Chlorobenzoic acid
	Isoquinoline-1-carboxylic acid
	6-methoxy-2-naphthoic acid

Library 11

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 11 are presented in ***Table 15***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 30.8 mg.

Table 15. Amino acids and carboxylic acids used in library 11.

Amino acids	**Carboxylic acids**
Asp	naphthalene-2-carboxylic acid
Ser	Indole-2-carboxylic acid
Ala	4-(propan-2-yl)benzoic acid
	6-carbamoylpyridine-2-carboxylic acid
	3-Thiophenecarboxylic acid

Library 12

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 12 are presented in ***Table 16***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 55.4 mg.

Table 16. Amino acids and carboxylic acids used in library 12.

Amino acids	**Carboxylic acids**
Nap	4-(dimethylcarbamoyl)benzoic acid
Pal	cyclopropanecarboxylic acid
Pro	
Trp	

Library 13

Was synthesized and cleaved according to the standard method described above. All the amino acids and carboxylic acids used in Library 13 are presented in ***Table 17***. The presence of obtained compounds was confirmed by extracting ion chromatogram on a LCMS. Attained amount of mix 32.6 mg.

Table 17. Amino acids and carboxylic acids used in library 13.

Amino acids	**Carboxylic acids**
Asp	4-(dimethylcarbamoyl)benzoic acid
Ser	cyclopropanecarboxylic acid
Ala	

Table 18. Shows all the fragments made, Libraries containing all the fragments with name, where the tree letter name is the amino acid. $[M+H]^+$ is the protonated mass of the fragments. The retention time differences between all the columns and the green marked rows have been detected as hits.

Compounds Name	LCMS $[M+H]^+$	ΔRT Protein - Ethanolamine (min)	ΔRT Protein - Custom Ref (min)	ΔRT Custom Ref - Ethanolamine (min)
Library 1				
2,4-Dimethoxybenzoic Acid-**Aze**	265,121	0,018	0,005	0,013
1-Naphthoic acid-**Aze**	255,115	0,101	0,023	0,078
2-Bromo-4-chlorobenzoic Acid-**Aze**	316,971	0,0835	0,0235	0,06
3,4,5-trimethoxybenzoic Acid-**Aze**	295,131	0,002	-0,013	0,015
3-Sulfamoylbenzoic Acid-**Aze**	284,263	Not Found	Not Found	Not Found
2,4-Dimethoxybenzoic Acid-**Pro**	279,137	0,022	-0,001	0,023
1-Naphthoic acid-**Pro**	269,131	0,083	-0,052	0,135
2-Bromo-4-chlorobenzoic Acid-**Pro**	330,987	0,0635	-0,0215	0,085
3,4,5-trimethoxybenzoic Acid-**Pro**	309,147	-0,0005	-0,0205	0,02
3-Sulfamoylbenzoic Acid-**Pro**	297,279	Not Found	Not Found	Not Found
2,4-Dimethoxybenzoic Acid-**Pip**	293,152	0,0505	0,0305	0,02
1-Naphthoic acid-**Pip**	283,146	Not Found	Not Found	Not Found
2-Bromo-4-chlorobenzoic Acid-**Pip**	345,002	Not Found	Not Found	Not Found
3,4,5-trimethoxybenzoic Acid-**Pip**	323,162	0,041	0,012	0,029
3-Sulfamoylbenzoic Acid-**Pip**	311,294	Not Found	Not Found	Not Found
Library 2				
Benzoic acid-**Phe**	292,08	0,058	-0,025	0,083
2-bromo-3,5-difluorobenzoic acid-**Phe**	405,972	-0,0545	-0,0235	-0,031
3-fluorobenzoic acid-**Phe**	310,071	0,055	-0,035	0,09
5-methoxyindole-2-carboxylic acid-**Phe**	361,102	0,063	-0,056	0,119
3,4,5-trimethoxybenzoic acid-**Phe**	382,112	0,065	-0,043	0,108
Benzoic acid-**Pal**	270,126	0,051	-0,032	0,083
2-bromo-3,5-difluorobenzoic acid-**Pal**	376,068	0,091	-0,207	0,298
3-fluorobenzoic acid-**Pal**	288,117	0,0475	-0,0335	0,081
5-methoxyindole-2-carboxylic acid-**Pal**	339,148	0,5065	-0,5395	1,046
3,4,5-trimethoxybenzoic acid-**Pal**	360,158	0,06	-0,051	0,111

Benzoic acid-**Trp**	308,142	0,2855	-0,1815	0,467
2-bromo-3,5-difluorobenzoic acid-**Trp**	414,084	Not Found	Not Found	Not Found
3-fluorobenzoic acid-**Trp**	326,133	0,2875	-0,2015	0,489
5-methoxyindole-2-carboxylic acid-**Trp**	377,164	0,541	-0,853	1,394
3,4,5-trimethoxybenzoic acid-**Trp**	398,174	0,2125	-0,2655	0,478
Benzoic acid-**Hyp**	235,11	-0,01	-0,008	-0,002
2-bromo-3,5-difluorobenzoic acid-**Hyp**	341,052	0,1345	-0,3325	0,467
3-fluorobenzoic acid-**Hyp**	253,101	0,0035	-0,0115	0,015
5-methoxyindole-2-carboxylic acid-**Hyp**	304,132	0,0715	-0,0945	0,166
3,4,5-trimethoxybenzoic acid-**Hyp**	325,142	-0,009	-0,143	0,134
Benzoic acid-**Nap**	319,146	0,5275	-0,8405	1,368
2-bromo-3,5-difluorobenzoic acid-**Nap**	425,088	0,214	-0,241	0,455
3-fluorobenzoic acid-**Nap**	337,137	0,6115	-0,9845	1,596
5-methoxyindole-2-carboxylic acid-**Nap**	388,168	-0,0785	-0,0165	-0,062
3,4,5-trimethoxybenzoic acid-**Nap**	409,178	0,2235	-0,9915	1,215
Library 3				
Benzoic acid-**Aze**	205,099	0,011	-0,009	0,02
4-Fluorobenzoic acid-**Aze**	223,09	0,0125	-0,0085	0,021
4-Chlorobenzoic acid-**Aze**	239,0608	0,026	-0,016	0,042
4-tert-butylbenzoic acid-**Aze**	261,294	0,0565	-0,0475	0,104
4-(trifluoromethyl)benzoic acid-**Aze**	273,087	0,023	-0,01	0,033
Benzoic acid-**Pro**	219,115	0,012	-0,013	0,025
4-Fluorobenzoic acid-**Pro**	237,106	0,0095	-0,0135	0,023
4-Chlorobenzoic acid-**Pro**	253,0768	0,03	-0,017	0,047
4-tert-butylbenzoic acid-**Pro**	275,31	0,041	-0,078	0,119
4-(trifluoromethyl)benzoic acid-**Pro**	287,103	0,022	-0,021	0,043
Benzoic acid-**Pip**	233,13	0,015	-0,012	0,027
4-Fluorobenzoic acid-**Pip**	251,121	Not Found	Not Found	Not Found
4-Chlorobenzoic acid-**Pip**	267,0918	Not Found	Not Found	Not Found
4-tert-butylbenzoic acid-**Pip**	289,325	Not Found	Not Found	0,048
4-(trifluoromethyl)benzoic acid-**Pip**	301,118	0,072	-0,195	0,267
Benzoic acid-**Tic**	281,131	0,0445	-0,0545	0,099
4-Fluorobenzoic acid-**Tic**	299,122	0,1225	0,0315	0,091

4-Chlorobenzoic acid-**Tic**	315,0928	0,082	-0,192	0,274
4-tert-butylbenzoic acid-**Tic**	337,326	-0,0915	-1,5435	1,452
4-(trifluoromethyl)benzoic acid-**Tic**	349,119	-0,26	-0,111	-0,149
Library 4				
Indole-3-carboxylic acid-**Gly**	218,098	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Gly**	159,088	-0,006	-0,004	-0,002
Cyclohexane-carboxylic acid-**Gly**	185,128	0,0025	-0,0125	0,015
6-Cyano-1H-indole-3-carboxylic acid-**Gly**	243,09	Not Found	Not Found	Not Found
1-(4-Bromophenyl)cyclopropanecarboxylic acid-**Gly**	297,026	0,029	-0,037	0,066
Indole-3-carboxylic acid-**Ala**	232,113	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Ala**	173,103	Not Found	Not Found	Not Found
Cyclohexane-carboxylic acid--**Ala**	199,143	0,014	-0,007	0,021
6-Cyano-1H-indole-3-carboxylic acid-**Ala**	257,105	0,1125	-0,1665	0,279
1-(4-Bromophenyl)cyclopropanecarboxylic acid-**Ala**	311,041	0,0445	-0,0455	0,09
Indole-3-carboxylic acid-**Trp**	347,15	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Trp**	288,14	Not Found	Not Found	Not Found
Cyclohexane-carboxylic acid-**Trp**	314,18	0,124	-0,14	0,264
6-Cyano-1H-indole-3-carboxylic acid-**Trp**	372,142	0,1225	-0,1235	0,246
1-(4-Bromophenyl)cyclopropanecarboxylic acid-**Trp**	426,078	-0,071	-2,522	2,451
Indole-3-carboxylic acid-**Nap**	358,16	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Nap**	299,15	0,035	-0,037	0,072
Cyclohexane-carboxylic acid-**Nap**	325,19	0,2105	-0,7725	0,983
6-Cyano-1H-indole-3-carboxylic acid-**Nap**	383,152	0,1975	-0,7325	0,93
1-(4-Bromophenyl)cyclopropanecarboxylic acid-**Nap**	437,088	0,0165	0,0165	0
Library 5				

Indole-3-carboxylic acid-**Aze**	244,113	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Aze**	185,103	-0,009	-0,003	-0,006
Cyclohexane-carboxylic acid-**Aze**	211,143	0,0005	-0,0145	0,015
6-Cyano-1H-indole-3-carboxylic acid-**Aze**	269,105	0,0345	-0,0015	0,036
1-(4-Bromophenyl)cyclopropane-carboxylic acid-**Aze**	323,041	0,0525	-0,0335	0,086
Indole-3-carboxylic acid-**Pro**	258,129	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Pro**	199,119			
Cyclohexane-carboxylic acid-**Pro**	225,159	0,011	0	0,011
6-Cyano-1H-indole-3-carboxylic acid-**Pro**	283,121	-0,011	-0,027	0,016
1-(4-Bromophenyl)cyclopropane-carboxylic acid-**Pro**	337,057	0,073	-0,061	0,134
Indole-3-carboxylic acid-**Pip**	272,144	Not Found	Not Found	Not Found
1-fluorocyclopropane-1--carboxylic acid-**Pip**	213,134	Not Found	Not Found	Not Found
Cyclohexane-carboxylic acid-**Pip**	239,174	0,0025	-0,0125	0,015
6-Cyano-1H-indole-3-carboxylic acid-**Pip**	297,136	Not Found	Not Found	0,032
1-(4-Bromophenyl)cyclopropane-carboxylic acid-**Pip**	351,072	Not Found	Not Found	Not Found
Indole-3-carboxylic acid-**Pal**	309,14	Not Found	Not Found	Not Found
1-fluorocyclopropane-1-carboxylic acid-**Pal**	250,13			
Cyclohexane-carboxylic acid-**Pal**	276,17	0,03	-0,015	0,045
6-Cyano-1H-indole-3-carboxylic acid-**Pal**	334,132	Not Found	Not Found	Not Found
1-(4-Bromophenyl)cyclopropane-carboxylic acid-**Pal**	388,068	0,1035	-0,1525	0,256
Library 6				
1-Hydroxy-1-cyclopropane-carboxylic acid-**Gly**	157,1	-0,011	-0,003	-0,008
2-(4-PIperidinyl)propanoic acid hydrochloride -**Gly**	214,158	0,048	0,006	0,042
1-Phenyl-1-cyclopropane-carboxylic acid-**Gly**	219,116	0,0005	-0,0155	0,016
3-aminonaphthalene-2-carboxylic acid-**Gly**	244,111	Not Found	Not Found	Not Found
1-Hydroxy-1-cyclopropane-carboxylic acid-**Ala**	171,115	0,0395	-0,0205	0,06
2-(4-PIperidinyl)propanoic acid hydrochloride-**Ala**	228,173	Not Found	Not Found	Not Found

1-Phenyl-1-cyclopropane-carboxylic acid-**Ala**	233,131	0	-0,016	0,016
3-aminonaphthalene-2-carboxylic acid-**Ala**	258,126	0,0885	-0,0845	0,173
1-Hydroxy-1-cyclopropane-carboxylic acid-**Trp**	286,152	Not Found	Not Found	Not Found
2-(4-PIperidinyl)propanoic acid hydrochloride-**Trp**	343,21	Not Found	-0,02	Not Found
1-Phenyl-1-cyclopropane-carboxylic acid-**Trp**	348,168	0,1585	-0,3875	0,546
3-aminonaphthalene-2-carboxylic acid-**Trp**	373,163	Not Found	-0,006	Not Found
1-Hydroxy-1-cyclopropane-carboxylic acid-**Nap**	297,162	Not Found	Not Found	Not Found
2-(4-PIperidinyl)propanoic acid hydrochloride-**Nap**	354,22	0,148	-0,054	0,202
1-Phenyl-1-cyclopropane-carboxylic acid-**Nap**	359,178	0,003	-2,732	2,735
3-aminonaphthalene-2-carboxylic acid-**Nap**	384,173	0,146	-0,009	0,155
Library 7				
1-Hydroxy-1-cyclopropanecarboxylic acid-**Aze**	183,115	0,0115	0,0035	0,008
2-(4-PIperidinyl)propanoic acid hydrochloride-**Aze**	240,173	0,06	-0,032	0,092
1-Phenyl-1-cyclopropanecarboxylic acid-**Aze**	245,131	0,005	-0,011	0,016
3-aminonaphthalene-2-carboxylic acid-**Aze**	270,126	Not Found	Not Found	Not Found
1-Hydroxy-1-cyclopropanecarboxylic acid-**Pro**	197,131	1,058	0,125	0,933
2-(4-PIperidinyl)propanoic acid hydrochloride-**Pro**	254,189	0,1815	-0,1465	0,328
1-Phenyl-1-cyclopropane-carboxylic acid-**Pro**	259,147	0,008	-0,013	0,021
3-aminonaphthalene-2-carboxylic acid-**Pro**	284,142	Not Found	Not Found	Not Found
1-Hydroxy-1-cyclopropanecar-boxylic acid-**Pip**	211,146	Not Found	Not Found	Not Found
2-(4-PIperidinyl)propanoic acid hydrochloride-**Pip**	268,204	0,0265	-0,0085	0,035
1-Phenyl-1-cyclopropane-boxylic acid-**Pip**	273,162	Not Found	Not Found	0,256
3-aminonaphthalene-2-boxylic acid-**Pip**	298,157	Not Found	Not Found	0,099

1-Hydroxy-1-cyclopropane-boxylic acid-**Pal**	248,142	0,0185		
2-(4-PIperidinyl)propanoic acid hydrochloride-**Pal**	305,2	0,0545	0,0135	0,041
1-Phenyl-1-cyclopropane-boxylic acid-**Pal**	310,158	0,0375	-0,0455	0,083
3-aminonaphthalene-2-boxylic acid-**Pal**	335,153	0,416	-0,352	0,768
Library 8				
3,4-Difluorobenzoic acid-**Nap**	355,128	0,0275	0,0125	0,015
4-Fluorobenzoic acid-**Nap**	337,137	0,435	-0,818	1,253
4-Chlorobenzoic acid-**Nap**	354,1078	Not Found	Not Found	Not Found
Isoquinoline-1-carboxylic acid-**Nap**	370,157	0,636	-3,004	3,64
6-methoxy-2-naphthoic acid-**Nap**	399,173			
3,4-Difluorobenzoic acid-**Pal**	306,108	0,051	-0,019	0,07
4-Fluorobenzoic acid-**Pal**	288,117	0,062	-0,008	0,07
4-Chlorobenzoic acid-**Pal**	305,0878	0,026	0,008	0,018
Isoquinoline-1-carboxylic acid-**Pal**	321,137	0,1435	-0,0345	0,178
6-methoxy-2-naphthoic acid-**Pal**	350,153	0,2185	-0,1555	0,374
3,4-Difluorobenzoic acid-**Pro**	255,097	0,0215	-0,0065	0,028
4-Fluorobenzoic acid-**Pro**	237,106	0,012	0,022	-0,01
4-Chlorobenzoic acid-**Pro**	254,0768			
Isoquinoline-1-carboxylic acid-**Pro**	270,126	0,072	0,013	0,059
6-methoxy-2-naphthoic acid-**Pro**	299,142	0,1175	-0,0355	0,153
3,4-Difluorobenzoic acid-**Trp**	344,118	Not Found	Not Found	Not Found
4-Fluorobenzoic acid-**Trp**	326,127	0,264	-0,092	0,356
4-Chlorobenzoic acid-**Trp**	343,0978	Not Found	-0,0465	Not Found
Isoquinoline-1-carboxylic acid-**Trp**	359,147	0,664	-0,636	1,3
6-methoxy-2-naphthoic acid-**Trp**	388,163	Not Found	Not Found	Not Found
Library 9				
naphthalene-2-carboxylic acid-**Nap**	369,162	Not Found	Not Found	Not Found
Indole-2-carboxylic acid-**Nap**	358,157	2,139	-2,869	5,008
4-(propan-2-yl)benzoic acid-**Nap**	375,209			
6-carbamoylpyridine-2-carboxylic acid-**Nap**	363,144	0,452	-0,514	0,966
3-Thiophenecarboxylic acid-**Nap**	325,1	0,8365	-0,4095	1,246
naphthalene-2-carboxylic acid-**Pal**	320,142	0,3565	-0,2645	0,621
Indole-2-carboxylic acid-**Pal**	309,137	0,545	-0,029	0,574

4-(propan-2-yl)benzoic acid-**Pal**	326,189	0,194	-0,425	0,619
6-carbamoylpyridine-2-carboxylic acid-**Pal**	314,124	0,077	0,001	0,076
3-Thiophenecarboxylic acid-**Pal**	276,08	0,061	-0,008	0,069
naphthalene-2-carboxylic acid-**Pro**	269,131	0,1035	-0,0385	0,142
Indole-2-carboxylic acid-**Pro**	258,126	0,1385	-0,0535	0,192
4-(propan-2-yl)benzoic acid-**Pro**	275,178	0,0565	-0,0865	0,143
6-carbamoylpyridine-2-carboxylic acid-**Pro**	263,113	0,0125	-0,0015	0,014
3-Thiophenecarboxylic acid-**Pro**	225,069	0,02	-0,004	0,024
naphthalene-2-carboxylic acid-**Trp**	358,152	2,139	-2,869	5,008
Indole-2-carboxylic acid-**Trp**	347,147	Not Found	Not Found	Not Found
4-(propan-2-yl)benzoic acid-**Trp**	364,199	0,371	-0,528	0,899
6-carbamoylpyridine-2-carboxylic acid-**Trp**	352,134	0,332	-0,029	0,361
3-Thiophenecarboxylic acid-**Trp**	314,09	0,077	0,001	0,076
Library 10				
3,4-Difluorobenzoic acid-**Asp**	272,071	Not Found	Not Found	Not Found
4-Fluorobenzoic acid-**Asp**	254,08	Not Found	Not Found	Not Found
4-Chlorobenzoic acid-**Asp**	271,0508	0,0865	0,0395	0,047
Isoquinoline-1-carboxylic acid-**Asp**	287,1	Not Found	Not Found	Not Found
6-methoxy-2-naphthoic acid-**Asp**	316,116	0,032	0,015	0,017
3,4-Difluorobenzoic acid-**Ser**	244,077	0,0315	-0,0085	0,04
4-Fluorobenzoic acid-**Ser**	226,086	Not Found	0,0055	Not Found
4-Chlorobenzoic acid-**Ser**	243,0568	0,142	0,006	0,136
Isoquinoline-1-carboxylic acid-**Ser**	259,106	0,0655	0,0235	0,042
6-methoxy-2-naphthoic acid-**Ser**	288,122	-0,2735	0,0445	-0,318
3,4-Difluorobenzoic acid-**Ala**	228,085	Not Found	Not Found	Not Found
4-Fluorobenzoic acid-**Ala**	210,094	Not Found	Not Found	Not Found
4-Chlorobenzoic acid-**Ala**	227,0648	Not Found	0,003	Not Found
Isoquinoline-1-carboxylic acid-**Ala**	243,114	Not Found	0,006	Not Found
6-methoxy-2-naphthoic acid-**Ala**	272,13	Not Found	Not Found	Not Found
Library 11				
naphthalene-2-carboxylic acid-**Asp**	287,105	-0,3765	0,0705	-0,447
Indole-2-carboxylic acid-**Asp**	276,1	-0,108	0,127	-0,235
4-(propan-2-yl)benzoic acid-**Asp**	293,152	-0,2225	-0,0035	-0,219
6-carbamoylpyridine-2-carboxylic acid-**Asp**	281,087	-0,098	0	-0,098

3-Thiophenecarboxylic acid-**Asp**	243,043	0,134	-0,041	0,175
naphthalene-2-carboxylic acid-**Ser**	259,111	0,275	-0,038	0,313
Indole-2-carboxylic acid-**Ser**	248,106	0,2825	-0,0075	0,29
4-(propan-2-yl)benzoic acid-**Ser**	265,158	0,0615	-0,0675	0,129
6-carbamoylpyridine-2-carboxylic acid-**Ser**	253,093			
3-Thiophenecarboxylic acid-**Ser**	215,049	0,536	0,492	0,044
naphthalene-2-carboxylic acid-**Ala**	243,119	0,134	-0,041	0,175
Indole-2-carboxylic acid-**Ala**	232,114	0,1715	0,0125	0,159
4-(propan-2-yl)benzoic acid-**Ala**	249,166	0,0915	-0,0945	0,186
6-carbamoylpyridine-2-carboxylic acid-**Ala**	237,101	0,049	0,02	0,029
3-Thiophenecarboxylic acid-**Ala**	199,057	0,006	-0,009	0,015
Library 12				
4-(dimethylcarbamoyl)benzoic acid-**Nap**	390,188	0,1995	-0,3495	0,549
cyclopropanecarboxylic acid-**Nap**	283,146	0,156	-0,065	0,221
4-(dimethylcarbamoyl)benzoic acid-**Pal**	341,168	0,047	0,008	0,039
cyclopropanecarboxylic acid-**Pal**	234,126	0,0145	0,0075	0,007
4-(dimethylcarbamoyl)benzoic acid-**Pro**	290,157	0,011	-0,002	0,013
cyclopropanecarboxylic acid-**Pro**	183,115	-0,012	0,002	-0,014
4-(dimethylcarbamoyl)benzoic acid-**Trp**	379,178	0,1435	-0,0645	0,208
cyclopropanecarboxylic acid-**Trp**	272,136	0,098	0,002	0,096
Library 13				
4-(dimethylcarbamoyl)benzoic acid-**Asp**	308,131	-0,1185	0,0075	-0,126
cyclopropanecarboxylic acid-**Asp**	201,089	-0,0965	0,0175	-0,114
4-(dimethylcarbamoyl)benzoic acid-**Ser**	280,137	0,0115	0,0015	0,01
cyclopropanecarboxylic acid-**Ser**	173,095	0,0255	0,0025	0,023
4-(dimethylcarbamoyl)benzoic acid-**Ala**	264,145	0,0195	0,0075	0,012
cyclopropanecarboxylic acid-**Ala**	157,103	0,0065	0,0065	0

WAC parameters

LCMS-set up	
Instrument	Titan
Detection	UV-220 and 254, and mass spectrometer
Mobile phase	Ammonium acetate buffer 20 mM 6.8
Run time	10 min ref. column, 15 min protein column
Injection volume	5 µL
Mix concentration	20 µM

References

1. Cons, A. J. P. S. H. and B. D. Fragment- based drug discovery and its application to challenging drug targets. *Portl. Press* (2017).

2. Shepherd, C. A., Hopkins, A. L. & Navratilova, I. Fragment screening by SPR and advanced application to GPCRs. *Prog. Biophys. Mol. Biol.* **116**, 113–123 (2014).

3. Ahmed, F. E., Wiley, J. E., Weidner, D. A., Bonnerup, C. & Mota, H. Surface plasmon resonance (SPR) spectrometry as a tool to analyze nucleic acid-protein interactions in crude cellular extracts. *Cancer Genomics and Proteomics* **7**, 303–310 (2010).

4. Magnusson, A. O. *et al.* nanoDSF as screening tool for enzyme libraries and biotechnology development. *FEBS J.* **286**, 184–204 (2019).

5. Torres(1), F. E., Michael I. Recht(1), Coyle(2), J. E., , Richard H. Bruce(1), and G. & Williams(2). Higher Throughput Calorimetry: Opportunities, Approaches and Challenges. *NIH* **23**, 1–7 (2008).

6. John E. Ladbury, G. K. and E. F. Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat. Rev.* **9**, 23–27 (2010).

7. Singh, P., Madhaiyan, K., Duong-Thi, M. D., Dymock, B. W. & Ohlson, S. Analysis of protein target interactions of synthetic mixtures by affinity-LC/MS. *SLAS Discov.* **22**, 440–446 (2017).

8. Medina, P. P. & Sanchez Cespedes, M. Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. *Epigenetics* **3**, 64–68 (2008).

9. Hodges, C., Kirkland, J. G. & Crabtree, G. R. The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb. Perspect. Med.* **6**, (2016).

10. Lemasters, E. N. M. and J. J. ATP/ADP Ratio, the Missed Connection between Mitochondria and the Warburg Effect. *Natl. Inst. Heal.* **23**, 1–7 (2014).

11. Kozakov, D. *et al.* The FTMap family of web servers for determining and characterizing ligand binding hot spots of proteins. *PMC* **10**, 733–755 (2016).

12. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *ACS Publ.* (1963) doi:2149-2154.

13. Szakacs, T. L. Peptide fragment screening towards a new inhibitor of neutrophil elastase By. (2019).

14. Larhed, M., Lindeberg, G. & Hallberg, A. Rapid microwave-assisted Suzuki coupling on solid-phase. *Tetrahedron Lett.* **37**, 8219–8222 (1996).

15. Subirós-Funosas, R., Prohens, R., Barbas, R., El-Faham, A. & Albericio, F. Oxyma: An efficient additive for peptide synthesis to replace the benzotriazole-based HOBt and HOAt with a lower risk of explosion. *Chem. - A Eur. J.* **15**, 9394–9403 (2009).

16. Serban C. Moldoveanu and Victor David. Chapter 5 - Properties of Analytes and Matrices Determining HPLC Selection. in *Selection of the HPLC Method in Chemical Analysis* 189–230 (2017).